Journal of
Risk and Financial Management

MDPI

Article

The Distribution of Commodity Futures Prices and Risk Estimates

Hector O. Zapata * and Raghav Goyal

Department of Agricultural Economics & Agribusiness, Louisiana State University & LSU AgCenter, Baton Rouge, LA 70803, USA; rgoyal@agcenter.lsu.edu
* Correspondence: hzapata@lsu.edu

Abstract

This article compares ARMA-GARCH models under alternative statistical distributions to identify the best fit for risk assessment in selected U.S. agricultural commodity markets. The Kupiec proportion-of-failures test assesses whether actual losses exceed expected losses from each model for rice, soybeans, wheat, corn, cotton, orange juice and live cattle in daily futures prices from 1994 to 2026. In sample, ARMA-eGARCH models with skewed Student-t innovations have the lowest AIC and give correct 5% Value-at-Risk coverage for all seven commodities, whereas models with normal innovations are rejected for five of the seven. Out of sample (March 2020 to June 2026), the selected models pass the Kupiec test in 26 of 28 cases, while a normal GARCH(1,1) produces too many violations at the 1% level. A GARCH(1,1) with skewed Student-t innovations forecasts Value-at-Risk about as well as the selected models, so the distributional assumption, rather than the complexity of the volatility equation, accounts for most of the gain in accuracy. The study finds, similarly to other studies, that the skewed Student-t distribution works best in fitting kurtosis and skewness for risk analysis purposes.

Keywords: financial markets; risk management and analysis; Value-at-Risk; Expected Shortfall; eGARCH; skewed Student-t; agricultural futures

1. Introduction

The distribution of price changes has been of continual interest to researchers and market participants in commodity futures markets. Knowledge of best-fit distributions is useful in risk and financial management applications. For example, the pricing of commodity futures options depends directly on the normality assumption. Similarly, accurate measurement of Value-at-Risk (VaR) in futures trading depends critically on estimates of best-fit distributions. In applications involving futures trading, the estimation of the distribution of futures and also cash prices is needed for estimating optimal hedge ratios. Empirical evidence on the distribution of futures price changes often reports the existence of higher moments and nonnormal price changes (Hudson et al., 1987; Hall et al., 1989). When viewing futures prices as non-stationary random walks with independent and identically distributed price changes (e.g., Fama, 1965; Stevenson & Bear, 1970; Campbell et al., 1997), the implication is that markets are operating efficiently, that is, price changes reflect new information and follow random variations. In practice, daily commodity futures prices may have a more complex structure than that of a pure random walk. One frequent empirical finding in previous work is that the mean function of daily price changes follows an autocorrelated moving average (ARMA) process and that volatility in price changes

Check for updates

Academic Editor: Thanasis Stengos

Received: 30 September 2026
Revised: 29 October 2026
Accepted: 30 October 2026
Published: 4 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

is non-constant and follows a generalized autoregressive conditional heteroskedasticity (GARCH) model (R. Engle, 1982; Bollerslev, 1987; Poon & Granger, 2003; Hansen & Lunde, 2005); GARCH models have been useful in explaining the distribution of common stock and commodity futures prices. When the purpose of building a model for price changes is to estimate risk, an ARMA-GARCH structure is fitting, given the autocorrelated nature of commodity price changes, and also allows volatility to be predicted, which is another empirical regularity in commodity futures prices and a key component of risk-measurement formulas (Manfredo & Leuthold, 1999). As discussed in the literature review section below, assumptions about the distribution of model errors must be consistent with the empirical properties of actual prices, particularly when estimating risk because risk measures use quantile functions that depend critically on estimates of a best-fit distribution.

While empirical literature in commodity markets has periodically analyzed some of these questions, research on the distribution of U.S. agricultural futures price changes has received less attention than research on equity markets over the past three decades, a period during which U.S. agricultural commodity markets have experienced booms and busts in futures prices, alongside the advancement of new statistical distributions that accommodate higher moments and non-normality. Since efficient price discovery and risk management functions of commodity futures markets are essential for the performance of commercial trade in commodities for both cash and futures markets, finding a best-fit distribution of futures price changes using recent statistical advances can improve accuracy in estimates of price risk and the ensuing value investors in commodity futures trading are putting at risk.

Several studies have nevertheless examined tail risk in commodity markets. Giot and Laurent (2003) estimate VaR for long and short commodity positions with skewed Student APARCH models and evaluate them in and out of sample, and Morgan et al. (2012) and van Oordt et al. (2021) study the extreme price risk of agricultural futures. That skewed, fat-tailed innovations fit daily returns better than the normal distribution is also an established result (Fernández & Steel, 1998; Kuester et al., 2006). Relative to these studies, this paper adds three things: (i) a long and recent sample (1994–2026) for seven U.S. agricultural futures markets, covering the 2006–2008 boom, the biofuel mandates, the pandemic and the 2022 supply shocks; (ii) a comparison of VaR and ES across innovation distributions and volatility structures for both long and short positions; and (iii) an out-of-sample evaluation with coverage, independence and Expected Shortfall backtests.

Recent research on stock prices and market indexes (e.g., Levy & Duchin, 2004; Corlu et al., 2016) has reported evidence of non-normality in price changes, with measures of skewness and kurtosis exceeding values of a normal distribution (0 and 3, respectively); research with U.S. agricultural commodity futures prices (e.g., Hall et al., 1989) has found similar results. The magnitude of skewness and kurtosis estimates varies among commodity markets, reflecting commodity-specific impacts of changes in market fundamentals, weather conditions, agricultural policy, and geopolitical events. The boom-and-bust cycles in commodity futures markets (Irwin et al., 2009) are often accompanied by extreme price changes (volatility clusters) which often drive market participants to respond symmetrically or asymmetrically to market events.

In commodity futures markets, Value-at-Risk (VaR) and Expected Shortfall (ES) are useful metrics to assess how much investment value is put at risk when hedging or speculating (e.g., Manfredo & Leuthold, 1999). Standard practice in the investment community is to model mean returns using autoregressive-moving average (ARMA) and to account for volatility with a generalized autoregressive conditional heteroskedasticity (GARCH) model. What remains of research interest in U.S. agricultural commodity futures markets literature is finding the estimation of best-fit statistical distributions for the errors in an ARMA-GARCH process. How much variation in estimates of measures of VaR and ES

exists when modeling a best-fit distribution? This paper jointly estimates ARMA–GARCH volatility models using various goodness-of-fit tests and applies recent statistical methods to identify the best-fit empirical distribution to measure price risk in agricultural commodity futures markets. The paper shows that, for risk estimates in agricultural commodity futures, the choice of innovation distribution matters more for risk estimates than the choice of the volatility model.

2. Review of Literature

Numerous empirical studies have shown that commodity futures price changes do not follow a normal distribution. In particular, returns in commodity futures markets tend to display excess kurtosis, skewness, volatility clustering, and occasional extreme movements. This implies that the probability of large price changes is significantly higher than what would have been predicted by the Gaussian distribution. However, comprehensive literature has documented these distributional features for financial assets, including equities (see, e.g., Levy & Duchin, 2004). For agricultural futures markets, the reasons for this non-normality might include production seasonality, weather effects, inventory levels, lags in biological production, policy, geo-political events, and arrival of market-specific information.

Mann and Heifner (1976) analyze daily futures prices of U.S. corn, soybeans, wheat, soybean oil, soybean meal, shell eggs, live cattle, frozen pork bellies, and Maine potatoes. In their study, they examine whether the short-run behavior of futures prices is consistent with normality and a random walk. They find no significant evidence of normality and also reject random walk in several commodities. The findings in Mann and Heifner imply that models that assume normally distributed and independently generated price changes may not be a good fit for agricultural futures markets.

Gordon (1985) expanded the literature by utilizing daily closing prices on U.S. corn, cotton, soybeans, wheat, orange juice, live cattle, live hogs, and rough rice, from January 1979 to May 1984. Gordon considered the assumptions of independence, normality, and equal variance, which are some of the basic assumptions underlying most models of futures prices. The author showed that the variance of agricultural futures price changes does not remain constant over the lifespan of the futures contract. Variance tends to be relatively high during the growing seasons but relatively low after the harvest period; however, the variance is more or less constant during two-month periods within the season. Therefore, at least part of the nonnormality found in futures returns can be attributed to the fact that the observations are taken from periods with different levels of volatility rather than one constant-variance distribution.

The question of best-empirical fit has obvious practical significance. According to Gordon (1985), option pricing models based on constant variance assumptions will undervalue options during high-volatility situations and over-value them in low-volatility situations. Similarly, time-series models built on the basis of the covariance stationarity assumption should take into consideration time-varying variance before applying them to agricultural futures prices. A forecast evaluation can be misleading if there is a discrepancy between the forecasts and actuals due to the change in seasonality. Here, the relevance of Gordon's work goes beyond just rejecting the normality assumption as it reveals the economically significant cause of non-normal distribution of agricultural futures returns.

Likewise, Helms and Martell (1985) studied the distribution of price changes in futures markets and found additional support for the conclusion that Gaussian distributions cannot sufficiently describe empirical deviations from normality. This study is part of a group of early studies that investigated whether fat tails could be attributed to other sources such as stochastic processes or information diffusion. Hudson et al. (1987), with a focus on the commodity options markets of U.S. wheat, soybeans, and live cattle, continued the

examination of normality and its consequences for agricultural option pricing. The findings of these researchers were used extensively as support for the presence of leptokurtosis in futures price changes.

An important contribution to our understanding of where fat tails come from was made by Hall et al. (1989). Rather than testing for normality per se, they compared two alternative theories about the distribution of futures price changes: one based on the stable-Paretian assumption and another based on the mixture-of-normals theory. Utilizing 20 long-time-series of daily closing futures price changes, they tested the stability characteristics of the distributions on both the original and randomized data. They found evidence supporting the mixture-of-normals theory. In economic terms, their findings support the idea that the return distribution is approximately normal, given the information or volatility state, but is fat-tailed overall because the returns are observed in different volatility states.

It was becoming more common during the 1990s and the 2000s for researchers to investigate this conditional form of volatility rather than taking the unconditional return distribution to be constant. For instance, the contribution of Manfredo and Leuthold (1999) centers upon the implications of non-normal price risk distributions in agriculture for VaR modeling and agricultural risk management. Their contribution is thus best understood as an illustration of the implications of distributional assumptions for measuring agricultural market risk rather than a further test of normality. Similarly, Goodwin and Schnepf (2000) shifted the emphasis onto the factors that determine endogenous volatility of corn and wheat futures markets. This research investigates how market and fundamental factors affect volatility of futures prices, explaining why volatility cannot be considered to be time-invariant.

Jin (2004) modeled agricultural cash price changes with stable distributions, which allow for heavier tails than the normal distribution.

Wilhelmsson (2006) formally assessed the forecasting performance of GARCH(1,1) models under various distributional assumptions on the Standard and Poor's 500 Index Futures returns; while this application was not on commodity futures markets, the results identified a GARCH model estimated with Student-t distribution as the best performing model under alternative forecast evaluation criteria. This finding is consistent with empirical evidence of modeling commodity futures price changes.

Additional studies have extended the focus beyond the marginal return distribution into that of the transmission of volatility between related agricultural markets. For instance, Hernandez et al. (2014) investigated volatility transmission in the major agricultural futures markets in the US and Europe. Zhu and Tansuchat (2019) document extreme risk spillovers between the U.S. and Chinese agricultural futures markets. This body of literature does not constitute a normal versus non-normal test per se, but it adds more weight to the argument for the use of dynamic multivariate volatility models because agricultural market volatility is neither constant nor confined to particular commodity markets. Shocks in one market may change the conditional risk profile of another market.

Existing literature has also focused on examining the tails of the return distribution. Junkus (1991) documented systematic skewness in futures contract returns. In a recent paper, Kateregga et al. (2017) used stable distributions to model the commodity futures log returns and compared them to Student-t location-scale distribution. The findings clearly indicate the practical value of using flexible stable distributions to describe the futures return distributions. Most importantly, their study demonstrates that imposing symmetry in the distribution, by assuming it to be a Student-t distribution, does not necessarily imply that the model accounts for the skewness in the commodity futures data[1].

When considered together, empirical literature offers much more support to our conclusions than just the lack of Gaussian returns. Several empirical studies, for example,

Mann and Heifner (1976), Gordon (1985), Helms and Martell (1985), Hudson et al. (1987), and Hall et al. (1989), have shown that the behavior of agricultural futures prices can hardly be explained by the simple model of normal distribution with constant variance. Further research has uncovered several economic and statistical properties which explain such non-conformity of the returns: seasonal and stochastic variance, volatility clustering, volatility regime mixture, long-term volatility, skewness, jumps, and tail correlation.

The implications are clear for empirical risk measurement. The normal distribution understates the probability in the tails when there is high kurtosis in the data and thus may seriously underestimate the probability and size of large price fluctuations. Furthermore, assumptions about constant variance do not allow for the well-known variability of volatility in agricultural futures observed in the literature. Therefore, to accurately measure downside risk—Value-at-Risk and Expected Shortfall in particular—more flexible distributional forms need to be employed that capture heavy tails, skewness, and time-varying volatility.

3. Materials and Methods

3.1. Futures Price Data and Construction of the Nearby Series

The empirical analysis uses daily closing prices for seven nearby agricultural futures series obtained from Bloomberg: the commodities and ticker symbols (Table 1) are corn (C1 Comdty), soybeans (S1 Comdty), hard red winter wheat (KW1 Comdty), rough rice (RR1 Comdty), Cotton No. 2 (CT1 Comdty), live cattle (LC1 Comdty), and frozen concentrated orange juice (JO1 Comdty). The sample extends from 9 December 1994, through 17 June 2026, which gives 7903 daily returns for each commodity. Each commodity is analyzed separately so that differences in trading calendars, contract specifications, price behavior, and underlying market fundamentals do not impact our empirical findings.

Table 1. Bloomberg nearby futures series, quotation units, and roll method.

Commodity	Bloomberg Ticker	Exchange	Contract Size	Price Quotation	Roll Method
Corn	C1 Comdty	CBOT	5000 bushels	Cents per bushel	Expiration
Soybeans	S1 Comdty	CBOT	5000 bushels	Cents per bushel	Expiration
Hard red winter wheat	KW1 Comdty	CBOT	5000 bushels	Cents per bushel	Expiration
Rough rice	RR1 Comdty	CBOT	2000 hundredweight	Dollars per hundredweight	Expiration
Cotton No. 2	CT1 Comdty	ICE Futures U.S.	50,000 pounds	Cents per pound	Expiration
Live cattle	LC1 Comdty	CME	40,000 pounds	Cents per pound	Expiration
Frozen concentrated orange juice	JO1 Comdty	ICE Futures U.S.	15,000 pounds	Cents per pound	Expiration

Note: The numerical quotations for corn, soybeans, and wheat are interpreted as cents per bushel, and the quotations for cotton, live cattle, and orange juice are interpreted as cents per pound. This interpretation is consistent with the contract values displayed by Bloomberg.

The nearby series are constructed using Bloomberg's roll method. Under this method the series rolls to the next contract at expiration, so the return on a roll date compares the prices of two different contracts; Bloomberg provides one continuous, automatically rolled series and does not report the roll dates, so we approximate them from the last-trading-day rules in the CME and ICE contract calendars. The returns around these dates are examined in Section 4 and Appendix A (Tables A2 and A3), and the implications are discussed in Section 7. We screen the data for nontrading days, missing values, nonpositive prices, and apparent recording errors. Consecutive observations refer to consecutive trading days rather than consecutive calendar days. Because trading calendars may differ across exchanges and commodities, the individual commodity series are not forced onto a common daily calendar.

3.2. Measurement of Daily Futures Price Changes

Let P_t denote the closing price of the nearby futures series for a particular commodity on trading day t. Following Campbell et al. (1997), Markowitz and Usmen (1996), and Levy and Duchin (2004), the continuously compounded daily price change is defined as

$$r_t = ln(P_t/P_{t-1}) = ln(P_t) - ln(P_{t-1}). \quad (1)$$

Throughout the analysis, r_t is referred to as the daily futures return or daily price change. Multiplying r_t by 100 provides the return in percentage terms. Logarithmic price changes are used because they are stationary over time and provide a consistent measure of proportional price variation across commodities with different price levels and units.

Let $p_t = ln(P_t)$. A benchmark random-walk model with independently and identically distributed innovations can be written as

$$p_t = \mu + p_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim iid\, N(0, \sigma^2). \quad (2)$$

Subtracting p_{t-1} from both sides of Equation (2) gives

$$r_t = \mu + \varepsilon_t\,. \quad (3)$$

This specification assumes that logarithmic futures prices are nonstationary, whereas daily returns have a constant conditional mean, constant conditional variance, and serially independent innovations. The empirical analysis relaxes these restrictions by allowing both the conditional mean and the conditional variance of returns to vary over time.

3.3. Descriptive Analysis and Tests of Stationarity

The empirical analysis begins with descriptive statistics for futures price levels and daily returns. For each commodity, we report key distributional measures, including central tendency, skewness, and kurtosis. Deviations from normal-distribution benchmarks—zero skewness and Pearson kurtosis of three—provide preliminary evidence of asymmetry and tail thickness.

Time-series plots of price levels and returns are used to identify broad price movements, unusually large changes, volatility clustering, possible structural breaks, and potential data irregularities. Histograms and kernel-density plots provide additional evidence concerning asymmetry, heavy tails, and departures from normality. The Jarque–Bera test is used to test the joint null hypothesis that sample skewness and kurtosis are consistent with a normal distribution.

The autocorrelation function and partial autocorrelation function are examined for logarithmic price levels and returns. Persistent autocorrelation in logarithmic price levels provides an initial indication of nonstationarity, but formal inference is based on the augmented Dickey–Fuller test. The augmented Dickey–Fuller test is applied to p_t and r_t, with the lag length set by the default rule. The Phillips–Perron test and the Kwiatkowski–Phillips–Schmidt–Shin test are also reported as complementary tests; the latter is included because its null hypothesis is stationarity, whereas the augmented Dickey–Fuller test has a unit root as its null hypothesis. The expected empirical pattern is that logarithmic price levels contain a unit root while daily returns are stationary. The ADF test results are reported in Appendix A (Table A1).

3.4. Conditional Mean Specification

Serial dependence in daily returns is modeled using an autoregressive-moving-average process. To distinguish the orders of the conditional mean model from those of the volatility

model, lower-case letters are used for the orders of the mean and upper-case letters for the orders of the variance, so the mean equation is denoted ARMA(p, q):

$$r_t = c + \Sigma_{i=1}^{p} \phi_i r_{t-i} + \Sigma_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \quad (4)$$

Here, c is the intercept, ϕ_i are autoregressive parameters, θ_j are moving-average parameters, and ϵ_t is the return innovation. The conditional mean of r_t, based on the information set available through the trading day $t - 1$, is therefore:

$$\mu_t = E(r_t \mid \mathcal{F}_{t-1}) = c + \Sigma_{i=1}^{p} \phi_i r_{t-i} + \Sigma_{j=1}^{q} \theta_j \epsilon_{t-j} \quad (5)$$

Candidate ARMA orders are evaluated over $p, q \in \{0, 1, \ldots, 5\}$, subject to successful convergence. Stationarity and invertibility are not imposed during estimation; they are checked for the selected models in Table 7.

3.5. Conditional Volatility Models

Commodity futures price changes commonly exhibit volatility clustering, in which large price changes tend to be followed by additional large price changes and relatively calm periods tend to be followed by subsequent calm periods. The conditional variance is modeled using the generalized autoregressive conditional heteroskedasticity framework. The return innovation is written as

$$\epsilon_t = \sqrt{h_t} z_t, \quad E(Z_t) = 0, \quad Var(Z_t) = 1, \quad (6)$$

where $h_t = \mathrm{Var}(\epsilon_t \mid \mathcal{F}_{t-1})$ is the conditional variance and z_t is a standardized innovation. The standard GARCH(P, Q) model is specified as

$$h_t = \omega + \Sigma_{i=1}^{P} \alpha_i \epsilon^2_{t-i} + \Sigma_{j=1}^{Q} \beta_j h_{t-j}, \quad (7)$$

where $\omega > 0$, $\alpha_i \geq 0$, and $\beta_j \geq 0$. Here P is the number of ARCH terms and Q the number of GARCH (β) terms. In a GARCH(1,1) model, α_1 measures the immediate response of volatility to a new return shock, β_1 measures the persistence of past volatility, and $\alpha_1 + \beta_1$ summarizes overall volatility persistence. A value close to one indicates that volatility shocks dissipate slowly.

Because positive and negative price shocks may have different effects on commodity-market volatility, the exponential GARCH model of Nelson (1991) is also considered. The EGARCH(P, Q) specification is given by

$$\ln(h_t) = \omega + \Sigma_{i=1}^{P} [\gamma_i (|z_{t-i}| - E|z_{t-i}|) + \alpha_i z_{t-i}] + \Sigma_{j=1}^{Q} \beta_j \ln\left(h_{t-j}\right) \quad (8)$$

The parameter α_i captures the sign (asymmetry) effect in the volatility response, and γ_i captures the size effect. A statistically significant value indicates that positive and negative innovations of equal magnitude have different effects on subsequent volatility[2]. Candidate GARCH and EGARCH models are estimated for $P, Q \in \{1, 2, 3, 4, 5\}$. In both equations P is the number of ARCH terms (α and γ) and Q the number of β terms, matching the order notation (e)GARCH(P,Q) used in the tables. The model-selection procedure has four steps:

(1) For each commodity, ARMA(p,q)-eGARCH(P,Q) models are estimated by maximum likelihood for all p, q $\in$ {0,...,5} and P, Q $\in$ {1,...,5}, that is, 900 specifications, with different innovations[3]. Specifications that fail to converge are excluded. The sGARCH grid is searched in the same way, and the volatility model with the lower AIC is retained.
(2) The specification with the lowest AIC is retained.
(3) This specification is re-estimated under normal, Student-t, generalized error and skewed Student-t innovations, and the distributions are compared by AIC and BIC.
(4) VaR and ES are computed from the retained model and backtested in sample and out of sample (Section 3.8).

3.6. Innovation Distributions, Model Selection, and Diagnostics

The standardized innovation z_t does not have to be restricted to follow a normal distribution. The analysis considers distributions capable of representing skewness, excess kurtosis, and heavy tails: the normal distribution is the basis of comparison and Student-t distribution, generalized error distribution, and skewed Student-t distribution are alternatives (as pointed out in the literature review). The same set of candidate distributions is applied to every commodity to preserve comparability across markets.

Each candidate ARMA–GARCH and ARMA–EGARCH model is estimated by maximum likelihood. Model fit is compared using the maximized log-likelihood, AIC, and BIC. As described in step (3), the distribution is chosen after the best ARMA-eGARCH order has been found: the model is re-estimated under each candidate distribution, and the fits are compared by AIC and BIC. The adequacy of the selected distribution is further assessed through the Kupiec coverage of the resulting Value-at-Risk estimates. All models are estimated with the rugarch package in R (version 4.5.3) (Ghalanos, 2024). The skewed Student-t is the Fernández and Steel (1998) skewed Student-t as implemented in rugarch, in which a skewness parameter equal to one corresponds to a symmetric distribution and values below (above) one indicate negative (positive) skewness.

The final model for each commodity is the specification with the lowest AIC; residual and variance diagnostics on standardized residuals are then reported to assess its adequacy. The standardized residuals are calculated as

$$\hat{z}_t = \frac{\hat{\epsilon}_t}{\sqrt{\hat{h}_t}} \tag{9}$$

The weighted Ljung–Box test (Fisher & Gallagher, 2012) is applied separately to the standardized residuals and their squares. The test on standardized residuals evaluates whether the conditional mean model has removed serial dependence in mean returns, whereas the test on squared standardized residuals evaluates whether the volatility model has removed remaining volatility clustering. A weighted ARCH-LM test is also applied to the standardized residuals. These diagnostics are reported for each selected model, and remaining serial dependence is noted where it appears; models are not discarded on this basis.

3.7. Estimation of Value-at-Risk and Expected Shortfall

The best selected ARMA–GARCH or ARMA–eGARCH model for each commodity is used to generate one-trading-day-ahead forecasts of the conditional mean, $\widehat{\mu_{t+1|t}}$, and conditional variance, $\widehat{h_{t+1|t}}$. Let $q_{\alpha,D}$ denote the lower-tail α-quantile of the selected standardized innovation distribution D. For a long futures position, one-day conditional Value-at-Risk, expressed as a positive return loss, is calculated as

$$VAR_{\alpha,t+1} = -\left[\widehat{\mu_{t+1|t}} + \sqrt{\hat{h}_{t+1|t}}\, q_{\alpha,D}\right] \tag{10}$$

Expected Shortfall measures the expected loss conditional on the loss exceeding Value-at-Risk. Let $m_{\alpha,D} = E(z|\, z \le q_{\alpha,\,D})$ denote the expected standardized innovation in the lower α tail. Conditional Expected Shortfall is calculated as

$$ES_{\alpha,\,t+1} = -\left[\widehat{\mu_{t+1|t}} + \sqrt{\hat{h}_{t+1|t}}\, m_{\alpha,D}\right] \tag{11}$$

Value-at-Risk and Expected Shortfall are estimated at the 95 and 99 percent confidence levels, corresponding to lower-tail probabilities of $\alpha = 0.05$ and $\alpha = 0.01$ respectively. To isolate the importance of the distributional assumption, Value-at-Risk and Expected Shortfall from the selected best-fitting distribution are compared with estimates from an otherwise identical conditional mean and volatility model with normally distributed innovations,

and with a standard ARMA-GARCH(1,1) model with normal innovations given that the GARCH(1,1) model is often reported to be hard to beat in forecasting experiments (e.g., Hansen & Lunde, 2005). In the out-of-sample evaluation, GARCH(1,1) models with skewed Student-t innovations are added as benchmarks: one with a constant mean, one with an AR(1) mean and one with the same ARMA order as the selected model. VaR and ES are reported as positive losses, consistent with Equations (10) and (11).

3.8. Backtesting of Value-at-Risk and Expected Shortfall

Risk-model performance is evaluated by backtesting the one-day-ahead Value-at-Risk estimates. Backtests are carried out both in sample and out of sample. In sample, the models are estimated once on the full sample and the conditional Value-at-Risk at each date is compared with the realized return. Because a distribution with estimated tail and skewness parameters can match the exceedance frequency of its own sample almost by construction, the in-sample results measure fit rather than forecasting accuracy. Out of sample, each model is estimated on the first 80% of the sample (6322 trading days) and re-estimated every 21 trading days on an expanding window, with the ARMA and eGARCH orders held at those selected on the full sample (Section 3.5), and one-day-ahead VaR and ES forecasts are evaluated over the remaining 1581 trading days (6 March 2020 to 17 June 2026). For each confidence level, a Value-at-Risk violation occurs when the realized return falls below the estimated lower-tail return threshold. The observed violation rate is compared with the theoretical rate using the Kupiec (1995) unconditional-coverage test. Out of sample, we also apply the Christoffersen (1998) independence and conditional-coverage tests and the dynamic quantile (DQ) test of R. F. Engle and Manganelli (2004), which detect clustering of violations. A correctly specified model should produce approximately 5 percent violations at the 95 percent confidence level and approximately 1 percent violations at the 99 percent confidence level. Violations are counted in both tails: the lower tail for long positions and the upper tail for short positions. Expected Shortfall is backtested with the McNeil and Frey (2000) test, which checks whether the exceedance residuals (the difference between the realized return and the predicted ES on violation days, scaled by the conditional standard deviation) have mean zero; Acerbi and Székely (2014) discuss alternative ES backtests.

4. Results

Time plots of prices and price changes, descriptive statistics (Table 2), ACFs and PACFs, and an empirical distribution function were analyzed for each commodity. A plot of daily prices for soybeans, one of the major oilseed crops in the United States, is illustrated in Figure 1. Soybean prices (top row, left panel) follow a typical random walk pattern in the sense that the mean and the variance change over time, and the magnitude of the fluctuations varies with the length of the time period selected. The top-right panel shows ACFs of price levels with autocorrelations decaying very slowly, a pattern typical of nonstationary prices, validating the use of Equation (1) to estimate price changes. Descriptive statistics of price changes in Table 2 show low daily price returns, for example, 0.007% daily returns for rice and 0.008% daily returns for soybeans. Skewness values are mostly negative (except for orange juice), and range from 0.0865 to −1.3056. We do not read this skewness as a response to negative news. Bloomberg rolls the nearby series at expiration, and most of the ten largest absolute returns fall within one trading day of an approximate roll date: all ten for rice and soybeans, nine for corn, eight for live cattle, and seven for wheat and cotton, but only three for orange juice (Table A2). When roll-day returns are excluded (Table A3), skewness lies between −0.22 and 0.01 for rice, soybeans, wheat and corn, and is −0.57 for cotton and −0.47 for live cattle; excess kurtosis falls from 40.5 to 4.2 for rice, from 20.1 to 2.6 for corn, from 18.7 to 9.9 for cotton and from 13.9 to 4.9 for live cattle. Much of the negative skewness

and kurtosis in Table 2 therefore comes from the roll; orange juice, whose largest returns are mostly not on roll dates, is little affected. Excess kurtosis is also prevalent, with rice, corn, cotton, and live cattle being more leptokurtic and thicker tails. For all crops, the distribution of price changes is not normal based on the Jarque–Bera (JB) test; all high values of the JB test on the last column of Table 2 have *p*-values much smaller than 0.001.

Table 2. Descriptive Statistics and Normality Test of Daily U.S. Futures Price Changes, 1994–2026.

Commodity	Mean	sd	Skew	Excess Kurtosis	JB
Rice	0.00007241906	0.0171137	−1.17098022	40.52623	542,921.33
Soybeans	0.00008826623	0.0152078	−0.93368992	7.74678	20,924.23
Wheat	0.00006003866	0.0182177	−0.00958245	2.15919	1537.17
Corn	0.00008174989	0.0176851	−1.24500010	20.06665	134,715.65
Cotton	−0.00000957504	0.0183485	−0.98947279	18.66423	116,067.34
Orange Juice	0.00004102388	0.0226810	0.08654139	4.86411	7807.13
Live Cattle	0.00016449451	0.0113307	−1.30555196	13.87744	65,701.05

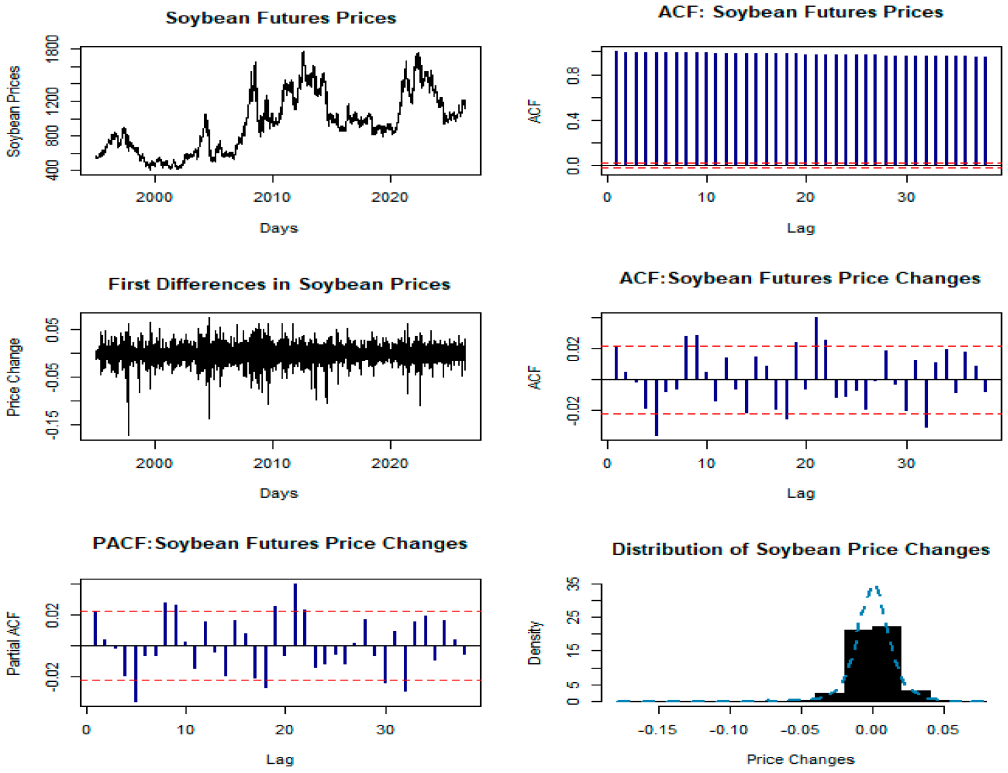


Figure 1. Temporal behavior in U.S. soybean futures prices, 1994–2026.

Returning to the middle two panels for futures price changes in Figure 1, the pattern of positive and negative price changes clearly illustrates that negative futures price responses in soybeans are of larger absolute magnitude than positive ones (consistent with the negative skewness observed earlier, although all ten of the largest soybean returns fall on approximate roll dates; Table A2); volatility also clusters around certain periods, with high and negative volatility in the first half of the data sample and much lower towards 2025–2026, all pointing toward the presence of GARCH effects. The ACFs and PACFs (middle right and bottom left panels in Figure 1) signal autoregressive and moving average (ARMA) effects in soybeans.

The joint estimation of the ARMA and eGARCH effects in multiple model combinations generated the best ARMA-eGARCH models reported in Table 3, selected by the AIC; the Weighted Ljung–Box tests reported in Table 3 show mostly no remaining GARCH effects

in the standardized squared residuals. The estimation is based on skewed Student's t-distribution given the skewness in all price changes discussed above.

Table 3. Best ARMA-eGARCH models under the skewed Student-t distribution, U.S. Commodity Futures Price Changes, 1994–2026.

Commodity	ARMA(p,q)-eGARCH(P,Q) Best Fit	AIC	WLB1 *p*-Values	WLB2 *p*-Values	Largest Inverse β Root
Rice	(2,5)–(5,3)	−5.67 sstd	0.13[1]/0.04[34]	0.41[1]/0.56[39]	0.997
Soybeans	(4,4)–(2,2)	−5.80 sstd	0.06[1]/0.00[39]	0.04[1]/0.37[19]	0.987
Wheat	(3,4)–(5,5)	−5.31 sstd	0.57[1]/0.94[34]	0.48[1]/0.85[49]	0.999
Corn	(5,5)–(4,3)	−5.53 sstd	0.14[1]/0.06[49]	0.72[1]/0.95[34]	0.997
Cotton	(5,4)–(1,3)	−5.42 sstd	0.03[1]/0.50[44]	0.61[1]/1.00[19]	0.992
Orange Juice	(5,5)–(5,4)	−4.98 sstd	0.11[1]/0.00[49]	0.63[1]/0.52[44]	0.999
Live Cattle	(4,4)–(3,4)	−6.45 sstd	0.11[1]/1.00[39]	1.00[1]/0.99[34]	0.988

WLB is the Weighted Ljung–Box statistic on standardized residuals (WLB1) and standardized squared residuals (WLB2), reported as *p*-values at lag 1 and at the lag in brackets, $[4(p+q)+(p+q)-1]$ for WLB1 and $[4(P+Q)+(P+Q)-1]$ for WLB2. The last column is the largest inverse root of the eGARCH β polynomial; values below one imply a stationary log-variance. Models are selected over ARMA(p,q) and eGARCH(P,Q) orders up to lag 5 by the AIC under skewed Student-t innovations; sstd = skewed Student-t distribution, which is the best fit among the normal, Student-t, and the generalized error distribution.

Table 3 reports the best ARMA-eGARCH model for each commodity, selected over ARMA(p,q) orders with $p, q \leq 5$ and eGARCH(P,Q) orders with $P, Q \leq 5$ as the combination that minimizes the AIC under skewed Student-t innovations (Section 3.5). When this order is re-estimated under each distribution, the skewed Student-t distribution has the lowest AIC among the normal, Student-t, skewed Student-t, and generalized error distributions for every commodity, and that is what we present here (for cotton the Student-t and skewed Student-t AIC values are equal to four decimals and the BIC prefers the Student-t). The AIC and BIC of the four distributions are reported in Table A4, and the parameter estimates with robust standard errors in Table A5. The selected mean-return equations are long, ranging from ARMA(2,5) for rice to ARMA(5,5) for corn and orange juice, and the volatility equations range from eGARCH(1,3) for cotton to eGARCH(5,5) for wheat.

Because several β coefficients are negative, volatility persistence is measured by the largest inverse root of the β polynomial rather than by the sum of the coefficients; the log-variance is stationary when all inverse roots lie inside the unit circle. The largest inverse root is below one for all seven commodities, ranging from 0.987 for soybeans to 0.999 for wheat and orange juice, which points to a stationary but highly persistent conditional variance. The estimated shape parameters, from 3.88 for live cattle to 8.87 for wheat, indicate heavy tails, and the skewness parameters differ from one in both directions (0.95 for soybeans and live cattle, and 1.08 for wheat)[4].

Table A6 reports the sign (asymmetry) coefficients α of the selected eGARCH models; in the rugarch parameterization of Equation (8), $\alpha > 0$ means that a positive return shock raises volatility more than a negative shock of the same size. At the first lag, α is positive and significant for rice (0.030) and wheat (0.041) and negative and significant for corn

(−0.077) and live cattle (−0.052); it is not significant for soybeans, cotton or orange juice. The higher-order α coefficients often have the opposite sign, so that their sum is small, from −0.047 for live cattle to 0.039 for soybeans, compared with the sum of the size coefficients γ (0.016 to 0.521). The sign effect in these markets is therefore modest, and its direction differs across commodities: there is no general tendency for positive price shocks to raise volatility more than negative ones. On parameter stability, the Nyblom test rejects stability for the GARCH(1,1) models of soybeans, wheat, orange juice and live cattle at the 1% level[5], which is consistent with the lower volatility of recent years visible in Figure 1. For the selected eGARCH models, the joint Nyblom statistic does not reject stability for soybeans (3.73) or cotton (3.99) at the 5% level (critical value 4.14); rugarch does not compute the joint statistic for the other five models, which have more than 20 parameters. Because parameters may drift over 1994–2026, the out-of-sample models in Section 6 are re-estimated every 21 trading days.

Table 3 also presents the weighted Ljung–Box (WLB) diagnostics. The tests on squared standardized residuals (WLB2) show no remaining conditional heteroskedasticity for any commodity at the reported lags, with the single exception of soybeans at lag 1 ($p = 0.042$). The tests on standardized residuals (WLB1) show some remaining serial correlation for soybeans and orange juice at the long lags ($p = 0.00$ at lags 39 and 49) and for cotton at lag 1 ($p = 0.03$). This correlation is small: the sample autocorrelations of soybean returns, for example, are within about ±0.03, so it changes the conditional mean, and hence the VaR, by only a few basis points. Adding ARMA lags does not remove it, because the AR and MA roots of the selected models already nearly cancel (Table 7), and a GARCH(1,1) with a constant mean forecasts VaR as accurately out of sample (Section 6). The weighted ARCH-LM tests (Fisher & Gallagher, 2012) do not reject the null of no remaining ARCH effects at the 5% level for any commodity, with p-values between 0.11 and 0.98.

5. Discussion

This article reported the best ARMA-GARCH models under alternative statistical distributions such as the normal, the generalized error distribution, Student-t and Skewed Student-t. The focus of the article is to assess the impact on estimates of risk, using Value-at-Risk (VaR) and Expected Shortfall (ES), under alternative distributions. This application is similar in spirit to that of Hansen and Lunde (2005) who compared volatility models in a forecasting experiment, and to Wilhelmsson (2009) who incorporated skewness, kurtosis, and second moment dependence for returns in the S&P 500, using VaR comparisons. Table 4 summarizes the findings encompassing the hard-to-beat GARCH(1,1) model as in Hansen and Lunde, and VaR estimates as in Wilhelmsson, as well as ES 5% measures in the seven commodity futures markets. Note that the best empirical fit appears in the last column of Table 4, with the models in the previous columns reported for comparison purposes. The results in Table 4 are in sample and therefore measure fit; Section 6 evaluates the models out of sample.

Table 4 compares, for each commodity, a standard GARCH(1,1) model with normal innovations (column 2) and the best ARMA-eGARCH model under normal, Student-t, and skewed Student-t innovations (columns 3–5). The GARCH(1,1) model uses the same ARMA specification as in the best models. The Kupiec test rejects correct 5% VaR coverage for the normal-distribution models for rice, live cattle, wheat, corn and cotton, under both the GARCH(1,1) and the best-model volatility equations, but not for soybeans or orange juice. Under Student-t innovations the wheat and corn models are rejected at the 5% level, but the other five are not, although live cattle are borderline ($p = 0.054$). Under skewed Student-t innovations, the best in-sample fit, coverage is not rejected for any of the seven commodities ($p = 0.953$, 0.750 and 0.924 for rice, soybeans and live cattle, and 0.074, 0.115,

0.446 and 0.750 for wheat, corn, cotton and orange juice). Relative to the best skewed Student-t model, the normal GARCH(1,1) VaR is larger for all seven commodities (for example, rice 2.64% versus 2.31%, soybeans 2.37% versus 2.29%, live cattle 1.82% versus 1.62%, and corn 2.79% versus 2.53%), whereas its ES is smaller for soybeans (2.97% versus 3.23%), live cattle (2.29% versus 2.50%), corn (3.50% versus 3.54%), cotton (3.68% versus 3.84%) and orange juice (4.51% versus 5.04%), and slightly larger for rice (3.32% versus 3.26%) and wheat (3.69% versus 3.65%).

Table 4. Value-at-Risk and Expected Shortfall at the 5% level: GARCH(1,1) versus best model, by innovation distribution, seven commodity futures (in sample).

	GARCH(1,1)	Best Model		
Statistic	**Normal**	**Normal**	**Student-t**	**Skewed Student-t**
Rice				
ARMA order	ARMA(2,5)	ARMA(2,5)	ARMA(2,5)	ARMA(2,5)
GARCH order	sGARCH(1,1)	eGARCH(5,3)	eGARCH(5,3)	eGARCH(5,3)
Kupiec *p*-value	0.000	0.000	0.093	0.953
Reject at 5%	**Yes**	**Yes**	No	No
VaR 5%	2.64%	2.58%	2.41%	2.31%
ES 5%	3.32%	3.24%	3.45%	3.26%
Soybeans				
ARMA order	ARMA(4,4)	ARMA(4,4)	ARMA(4,4)	ARMA(4,4)
GARCH order	sGARCH(1,1)	eGARCH(2,2)	eGARCH(2,2)	eGARCH(2,2)
Kupiec *p*-value	0.190	0.294	0.104	0.750
Reject at 5%	No	No	No	No
VaR 5%	2.37%	2.33%	2.23%	2.29%
ES 5%	2.97%	2.93%	3.13%	3.23%
Live Cattle				
ARMA order	ARMA(4,4)	ARMA(4,4)	ARMA(4,4)	ARMA(4,4)
GARCH order	sGARCH(1,1)	eGARCH(3,4)	eGARCH(3,4)	eGARCH(3,4)
Kupiec *p*-value	0.000	0.000	0.054	0.924
Reject at 5%	**Yes**	**Yes**	No	No
VaR 5%	1.82%	1.81%	1.56%	1.62%
ES 5%	2.29%	2.27%	2.40%	2.50%
Wheat				
ARMA order	ARMA(3,4)	ARMA(3,4)	ARMA(3,4)	ARMA(3,4)
GARCH order	sGARCH(1,1)	eGARCH(5,5)	eGARCH(5,5)	eGARCH(5,5)
Kupiec *p*-value	0.000	0.000	0.000	0.074
Reject at 5%	**Yes**	**Yes**	**Yes**	No
VaR 5%	2.94%	2.88%	2.86%	2.75%
ES 5%	3.69%	3.62%	3.83%	3.65%
Corn				
ARMA order	ARMA(5,5)	ARMA(5,5)	ARMA(5,5)	ARMA(5,5)
GARCH order	sGARCH(1,1)	eGARCH(4,3)	eGARCH(4,3)	eGARCH(4,3)
Kupiec *p*-value	0.000	0.000	0.020	0.115
Reject at 5%	**Yes**	**Yes**	**Yes**	No
VaR 5%	2.79%	2.74%	2.59%	2.53%
ES 5%	3.50%	3.43%	3.68%	3.54%
Cotton				
ARMA order	ARMA(5,4)	ARMA(5,4)	ARMA(5,4)	ARMA(5,4)
GARCH order	sGARCH(1,1)	eGARCH(1,3)	eGARCH(1,3)	eGARCH(1,3)
Kupiec *p*-value	0.005	0.001	0.649	0.446
Reject at 5%	**Yes**	**Yes**	No	No
VaR 5%	2.93%	2.90%	2.73%	2.70%
ES 5%	3.68%	3.64%	3.90%	3.84%

Table 4. *Cont.*

	GARCH(1,1)	Best Model		
Statistic	**Normal**	**Normal**	**Student-t**	**Skewed Student-t**
Orange Juice				
ARMA order	ARMA(5,5)	ARMA(5,5)	ARMA(5,5)	ARMA(5,5)
GARCH order	sGARCH(1,1)	eGARCH(5,4)	eGARCH(5,4)	eGARCH(5,4)
Kupiec *p*-value	0.093	0.103	0.446	0.750
Reject at 5%	No	No	No	No
VaR 5%	3.59%	3.56%	3.40%	3.45%
ES 5%	4.51%	4.47%	4.94%	5.04%

Note: Left tail (long position). VaR 5% and ES 5% are averages of the one-step-ahead conditional estimates, in daily log-return percent, reported as positive losses. The Kupiec *p*-value tests correct unconditional coverage of the 5% VaR (in-sample). The GARCH(1,1) column is a standard sGARCH(1,1) with normal innovations and the same ARMA(p,q) as the best model. The best model is the lower-AIC of sGARCH and eGARCH (eGARCH for all seven commodities), each at its own best order selected with skewed Student-t innovations; the same order is used for all three distributions. The yellow shaded column is the best in-sample fit. Out of sample, a GARCH(1,1) with skewed Student-t innovations forecasts VaR about as well (Section 6) and is the recommended starting point for forecasting.

In the rice panel in Table 4 (top panel) it is observed that the standard GARCH(1,1) volatility model with an ARMA(2,5) mean model is rejected as an adequate model for rice based on a Likelihood Ratio test (Kupiec test) at the 5% level of significance; the best model under normal innovations in the third column [ARMA(2,5)-eGARCH(5,3)] is also rejected by the Kupiec test suggesting that alternative statistical distributions may produce a better model; this is consistent with the rejection of normality reported earlier in Table 2, due to the presence of kurtosis and skewness. These latter two properties are modeled using Student's t distribution (for kurtosis) and by the skewed Student-t (for kurtosis and skewness). Using the AIC and Ljung–Box tests on standardized residuals and standardized-squared residuals, the best-fit model favors the skewed t-distribution in this application which for rice is an ARMA(2,5)-eGARCH(5,3). Similar results are found for live cattle, even though the model structure changed to an ARMA(4,4)-eGARCH(3,4). Note that for soybean and orange juice futures, however, all models perform adequately based on the Kupiec test.

6. Out-of-Sample Evaluation

The in-sample results above measure fit. To evaluate forecasting accuracy, Table 5 reports out-of-sample backtests for all seven commodities. Each model is estimated on the first 80% of the sample with an expanding window re-estimated every 21 trading days, keeping the ARMA and eGARCH orders selected on the full sample, and one-day-ahead VaR and ES forecasts are evaluated over the remaining 1581 trading days (6 March 2020 to 17 June 2026), at the 1% and 5% levels and for both long (lower-tail) and short (upper-tail) positions. In addition to the selected ARMA-eGARCH skewed Student-t models and the normal GARCH(1,1), the comparison includes GARCH(1,1) models with skewed Student-t innovations and a constant mean, an AR(1) mean, or the same ARMA order as the selected model. Table 6 summarizes the number of rejections by test.

Out of sample, the selected eGARCH skewed Student-t models pass the Kupiec test in 26 of 28 cases, and the normal GARCH(1,1) is rejected far more often. At the 5% level the normal GARCH(1,1) tends to produce too few long-position violations (for example, 54 against 79 expected for rice and 56 for wheat), so its VaR is conservative, whereas at the 1% level it produces too many (36 against 16 expected for live cattle, and 33 for the short tail of wheat). The finding that the normal model overestimates VaR therefore applies to

the 5% level only. The McNeil and Frey test rejects the ES forecasts of the selected models in 5 of 28 cases, against 14 of 20 for the normal GARCH(1,1).

Table 5. Out-of-sample VaR and ES backtests, all seven commodities (6 March 2020–17 June 2026).

Commodity	Model	Long 1%	Short 1%	Long 5%	Short 5%
Rice	eGARCH skew-t, best ARMA	17	**20 (CD)**	**74 (CD)**	**92 (CD)**
	GARCH(1,1) skew-t, constant	18	14	**74 (CD)**	**88 (CD)**
	GARCH(1,1) skew-t, AR(1)	19	15	**72 (CD)**	**87 (CD)**
	GARCH(1,1) skew-t, best ARMA	19	16	**73 (CD)**	**86 (CD)**
	GARCH(1,1) normal, best ARMA	**20 (E)**	**25 (KCD)**	**54 (KCDE)**	**78 (CDE)**
Soybean	eGARCH skew-t, best ARMA	14	**14 (E)**	66	76
	GARCH(1,1) skew-t, constant	12	**11 (E)**	70	**85 (E)**
	GARCH(1,1) skew-t, AR(1)	12	**11 (E)**	67	**84 (E)**
	GARCH(1,1) skew-t, best ARMA	12	**11 (E)**	69	**85 (E)**
	GARCH(1,1) normal, best ARMA	**20 (E)**	18	**68 (E)**	66
Wheat	eGARCH skew-t, best ARMA	12	16	70	**97 (K)**
	GARCH(1,1) skew-t, constant	12	16	71	**106 (KCD)**
	GARCH(1,1) skew-t, AR(1)	11	17	69	**102 (KCD)**
	GARCH(1,1) skew-t, best ARMA	11	17	70	**104 (KCD)**
	GARCH(1,1) normal, best ARMA	15	**33 (KCD)**	**56 (KC)**	**104 (KCDE)**
Corn	eGARCH skew-t, best ARMA	**15 (D)**	13	**75 (D)**	72
	GARCH(1,1) skew-t, constant	14	14	**79 (D)**	67
	GARCH(1,1) skew-t, AR(1)	14	14	**80 (D)**	68
	GARCH(1,1) skew-t, best ARMA	14	14	**78 (D)**	72
	GARCH(1,1) normal, best ARMA	**22 (E)**	**21 (E)**	**65 (E)**	**57 (KCE)**
Cotton	eGARCH skew-t, best ARMA	12	**12 (E)**	80	**88 (E)**
	GARCH(1,1) skew-t, constant	13	**10 (E)**	78	**87 (E)**
	GARCH(1,1) skew-t, AR(1)	13	**10 (E)**	76	**86 (E)**
	GARCH(1,1) skew-t, best ARMA	15	**11 (E)**	75	**87 (E)**
Orange Juice	eGARCH skew-t, best ARMA	15	**7 (KCE)**	90	**82 (E)**
	GARCH(1,1) skew-t, constant	18	**12 (E)**	**103 (KCD)**	**96 (CDE)**
	GARCH(1,1) skew-t, AR(1)	18	**12 (E)**	**109 (KCDE)**	**100 (KCDE)**
	GARCH(1,1) skew-t, best ARMA	17	**13 (E)**	**104 (KCD)**	**100 (KCDE)**
Live Cattle	eGARCH skew-t, best ARMA	21	11	**82 (CD)**	81
	GARCH(1,1) skew-t, constant	24	10	**78 (CDE)**	69
	GARCH(1,1) skew-t, AR(1)	24	10	**76 (CDE)**	69
	GARCH(1,1) skew-t, best ARMA	24	10	**80 (CD)**	71
	GARCH(1,1) normal, best ARMA	**36 (KCDE)**	**18 (DE)**	**71 (CDE)**	**52 (KCDE)**

Note: Entries are the number of VaR violations over 1581 out-of-sample days (expected 15.8 at 1% and 79.1 at 5%). Models are estimated on the first 80% of the sample with an expanding window and re-estimated every 21 trading days. Letters mark rejection at the 5% level: K = Kupiec (1995) unconditional coverage, C = Christoffersen (1998) conditional coverage, D = R. F. Engle and Manganelli (2004) dynamic quantile, E = McNeil and Frey (2000) ES test; rejected cells are in bold. The GARCH(1,1)-normal model could not be estimated over all rolling windows for cotton and orange juice.

The volatility structure matters much less than the distribution. At the 1% level the GARCH(1,1) skewed Student-t models are never rejected by the Kupiec, Christoffersen or DQ tests, while the selected eGARCH models are rejected once, twice and twice. At the 5% level the eGARCH models perform somewhat better (1, 3 and 4 rejections against 2, 6 and 7 for the constant-mean GARCH(1,1)), with the clearest gains for orange juice and the short tail of wheat. Adding an AR(1) term or the full ARMA order to the GARCH(1,1) makes little difference. Some clustering of violations remains for rice and live cattle at the 5% level, where the conditional-coverage and DQ tests reject for most models, which indicates that violations in these markets cluster in turbulent periods.

Table 6. Number of out-of-sample rejections at the 5% significance level.

Level	Model	Cases	Kupiec	Christoffersen	DQ	McNeil–Frey
1% VaR	eGARCH skew-t, best ARMA	14	1	2	2	3
	GARCH(1,1) skew-t, constant	14	0	0	0	3
	GARCH(1,1) skew-t, AR(1)	14	0	0	0	3
	GARCH(1,1) skew-t, best ARMA	14	0	0	0	3
	GARCH(1,1) normal, best ARMA	10	3	3	4	6
5% VaR	eGARCH skew-t, best ARMA	14	1	3	4	2
	GARCH(1,1) skew-t, constant	14	2	6	7	4
	GARCH(1,1) skew-t, AR(1)	14	3	6	7	5
	GARCH(1,1) skew-t, best ARMA	14	3	6	7	3
	GARCH(1,1) normal, best ARMA	10	5	7	5	8

Note: Cases = commodities × tails (7 × 2; 5 × 2 for the normal GARCH(1,1), which could not be estimated over all rolling windows for cotton and orange juice).

Table 7 helps explain why the parsimonious models do as well. The AR and MA roots of every selected model lie close to the unit circle and nearly cancel (the closest AR–MA root pair is between 0.001 and 0.027 apart), which signals parameter redundancy in the high-order mean equations. A parsimonious GARCH(1,1) with a constant mean and skewed Student-t innovations therefore delivers VaR forecasts of similar accuracy to the selected models. The skewness parameter differs significantly from one, the value for a symmetric distribution, for soybeans, wheat and live cattle, but not for rice, corn, cotton or orange juice. For rice, the AR and MA roots with the smallest moduli (0.99998 and 0.999) lie marginally inside the unit circle, so the selected rice model does not strictly satisfy the stationarity and invertibility conditions. These two root pairs are only 0.001 apart and cancel, which leaves a mean equation close to white noise; the constant-mean GARCH(1,1) with skewed Student-t innovations, which has no ARMA terms, gives similar rice VaR forecasts out of sample (Table 5). Finally, the ARMA and eGARCH orders were selected on the full sample, which includes the evaluation period. This look-ahead can only favor the selected eGARCH models; the GARCH(1,1) benchmarks with a constant or AR(1) mean involve no order selection, and they forecast about as well.

Table 7. Roots and skewness test for the selected ARMA–eGARCH skewed Student-t models.

Commodity	Smallest |AR Root|	Smallest |MA Root|	Closest AR–MA Root Pair	Sum of β	Largest Inverse β Root	*p*-Value, Skew = 1
Rice	1.000	0.999	0.001	0.957	0.997	0.458
Soybean	1.004	1.003	0.002	0.978	0.987	0.002
Wheat	1.093	1.083	0.011	0.999	0.999	0.000
Corn	1.015	1.018	0.003	0.939	0.997	0.949
Cotton	1.004	1.002	0.002	0.989	0.992	0.166
Orange Juice	1.021	1.017	0.006	0.999	0.999	0.192
Live Cattle	1.127	1.165	0.027	0.970	0.988	0.001

Note: Full-sample estimates. Stationarity of the eGARCH log-variance requires all inverse roots of the β polynomial to lie inside the unit circle. The last column is a Wald test, with robust standard errors, of the null that the skewed Student-t is symmetric (skew = 1). For rice the smallest AR root is 0.99998 and the smallest MA root 0.99905, both marginally inside the unit circle (see text).

7. Conclusions

Overall, given that normality of returns in commodity futures prices is strongly rejected by the Jarque–Bera test, and that the skewed Student-t distribution captures departures from normality such as kurtosis and skewness, the skewed Student-t models in the last column of Table 4 give the best in-sample fit to the properties of these commodity futures prices. In sample and at the 5% level, the normal GARCH(1,1) gives a larger VaR than

the skewed Student-t model for all seven commodities but a smaller ES for five of them; for rice and wheat its ES is slightly larger. Out of sample, the normal GARCH(1,1) VaR is conservative at the 5% level but too low at the 1% level, where it produces too many violations. When investing in commodity futures, the VaR and ES estimates reported here can result in considerable dollars at risk. One futures contract for soybeans is 5000 bushels, which, when multiplied by the VaR95% of 2.29% for the best-fit model in Table 4, at a price of $12/bu, the daily loss exceeded on 5% of trading days would be $1,374 per contract; using the ARMA(4,4)-standard GARCH(1,1) model, the corresponding loss would be $1422 per contract. The difference of $48 for one contract is modest at the 5% level; the out-of-sample results suggest that the choice of distribution matters more at the 1% level, where the normal model understates risk.

These findings carry an implication for risk management in agricultural commodity markets, although in-sample fit should be distinguished from forecasting performance. In sample, models that account for skewness and heavy tails fit the data better than a normal-distribution GARCH(1,1) model, which is still the default in many trading and hedging applications. Out of sample, the gain in forecasting accuracy comes mainly from the distributional assumption: a GARCH(1,1) with skewed Student-t innovations forecasts VaR about as well as the selected ARMA-eGARCH models, while the normal GARCH(1,1) can misstate tail risk, particularly at the 1% level. For practical risk measurement in these markets, a GARCH(1,1) with skewed Student-t innovations is therefore a reasonable starting point.

Our study leaves some open areas for future researchers to explore. First, the nearby futures series are rolled at contract expiration, so the return on each roll date compares the prices of two different contracts. For storable crops the price gap between the expiring (old-crop) contract and the next (new-crop) contract can be large, and holding the contract to expiration also brings delivery-month effects and thin trading into the series. Most of the largest returns fall on approximate roll dates, and excluding them removes much of the negative skewness and kurtosis for rice, soybeans, corn, cotton and live cattle (Tables A2 and A3), so for these markets the sample skewness reflects the roll more than the response of prices to news. Bloomberg does not report the roll dates of its continuous series, and ours are approximated from the exchange calendars; back-adjusted series, or series that roll before the delivery month, would address this. Second, daily price limits censor returns on some days. He and Serra (2022) find that price limits do not reduce volatility and that volatility is higher and liquidity lower when trading resumes; intraday data would allow a closer analysis of limit moves. Third, the variance equation has no seasonal or time-to-maturity component. Gupta and Rajib (2018) document seasonality in VaR exceptions in spot markets, but we are not aware of a method that captures seasonality well in daily futures returns. While this paper provides initial evidence on the effect of GARCH models under various distributional assumptions for price changes in U.S. agricultural commodity futures markets, future work will focus on the existence of time-varying higher moments (i.e., skewness and kurtosis) in volatility models of financial risk in the context of out-of-sample density forecasting (e.g., Wilhelmsson, 2013). If time-dependence in higher moments is found, the new models will improve not only model specification for all commodities but also risk estimates reported in this article.

Author Contributions: Conceptualization, H.O.Z. and R.G.; Methodology, H.O.Z. and R.G.; Software, H.O.Z. and R.G.; Validation, H.O.Z. and R.G.; Formal Analysis, H.O.Z. and R.G.; Investigation, H.O.Z. and R.G.; Resources, H.O.Z. and R.G.; Data Curation, R.G.; Writing—Original Draft Preparation, H.O.Z. and R.G.; Writing—Review & Editing, H.O.Z. and R.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the USDA National Institute of Food and Agriculture, Hatch Project, grant number LAB94695.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The price data were obtained from Bloomberg and cannot be redistributed under its license terms. The R code used for estimation and backtesting is available from the authors on request.

Acknowledgments: We thank the anonymous reviewers for substantially contributing via comments. The content is the responsibility of the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Augmented Dickey–Fuller unit-root tests on futures price levels, 1994–2026.

Commodity	ADF Statistic	Lag	*p*-Value
Rice	−2.674	19	0.293
Soybeans	−2.744	19	0.263
Wheat	−3.031	19	0.142
Corn	−2.827	19	0.228
Cotton	−3.356	19	0.061
Orange Juice	−2.603	19	0.323
Live Cattle	−0.655	19	0.975

Note: The null hypothesis is a unit root; the lag length is set by the default rule. The Phillips–Perron and KPSS tests lead to the same conclusion. For daily returns, all three tests indicate stationarity; these results are available from the authors on request.

Table A2. Ten largest absolute daily returns (%) by commodity, with approximate roll dates marked, 1994–2026.

Rank	Rice	Soybeans	Wheat	Corn	Cotton	Orange Juice	Live Cattle
1	2020-07-15 −29.97 *	1997-09-22 −17.43 *	1996-12-20 −12.59 *	1996-07-23 −27.62 *	1995-07-10 −30.44 *	1998-10-09 22.72	2018-05-01 −15.65 *
2	2024-07-15 −28.51 *	2004-07-15 −13.84 *	2005-05-13 −11.15 *	2013-07-15 −26.86 *	2022-06-24 −27.29	1999-08-12 18.32	2017-05-01 −10.60 *
3	2000-07-21 28.08 *	2004-08-16 −12.48 *	1997-05-21 −10.92 *	2021-07-15 −19.10 *	2011-07-08 −15.55 *	2006-10-12 15.43	1996-10-25 −10.38
4	2004-07-22 −24.45 *	2022-08-15 −11.09 *	2009-01-12 −8.99	2023-07-17 −18.34 *	2003-03-10 13.62 *	2026-01-12 −13.32 *	2005-05-02 −10.04 *
5	2003-07-23 16.25 *	2009-08-17 −10.74 *	1995-05-22 −8.55 *	2022-07-15 −13.99 *	2011-07-07 −13.19 *	1997-08-14 −12.91	2003-11-03 −9.69 *
6	2023-07-17 −14.29 *	2014-08-15 −10.50 *	2008-03-20 −8.45 *	2009-09-15 12.76 *	2001-06-25 11.35	2012-01-11 −12.31 *	2003-05-01 −9.24 *
7	1999-07-21 −12.97 *	1997-08-21 −10.41 *	2024-07-15 −8.37 *	1996-09-20 −12.12 *	2017-07-10 −10.60 *	2025-07-08 12.12	1997-04-24 −8.97
8	1999-05-20 −12.58 *	2013-09-16 −9.91 *	2022-02-25 −8.25	2011-06-30 −10.41	2007-12-07 10.53 *	1999-07-13 −11.73 *	2019-05-01 −8.32 *
9	2008-07-15 −11.35 *	2014-09-15 −9.77 *	2010-10-08 8.10	1997-07-23 −9.17 *	2002-03-07 9.54 *	2009-07-06 11.47	2016-09-01 −7.78 *
10	2014-07-15 −11.09 *	2022-07-15 −9.37 *	2024-03-15 −8.01 *	1999-12-22 9.13 *	2008-03-03 8.68	2026-02-02 −11.43	2000-05-01 −7.73 *
Near roll	10 of 10	10 of 10	7 of 10	9 of 10	7 of 10	3 of 10	8 of 10

Note: Entries are the date and the log return in percent, ranked by absolute value. * = within one trading day of an approximate roll date. Bloomberg does not report the roll dates of its continuous series; roll dates are approximated from the last-trading-day rules of the CME (corn, soybeans, wheat, rough rice, live cattle) and ICE (cotton, orange juice) contract calendars, so a date may be off by a day where a rule changed or a holiday intervened.

Table A3. Descriptive statistics of daily returns with and without roll-date returns, 1994–2026.

Commodity	Sample	N	Mean (%)	SD (%)	Skewness	Excess Kurtosis	JB
Rice	All returns	7903	0.0072	1.711	−1.17	40.54	542,921
	Excl. roll days	7526	−0.0164	1.515	−0.02	4.18	5488
	Excl. roll window	6778	−0.0176	1.515	0.05	3.41	3295
Soybeans	All returns	7903	0.0088	1.521	−0.93	7.75	20,924
	Excl. roll days	7464	0.0371	1.418	−0.22	2.92	2718
	Excl. roll window	6592	0.0502	1.405	−0.20	2.84	2254
Wheat	All returns	7903	0.0060	1.822	−0.01	2.16	1537
	Excl. roll days	7587	−0.0127	1.774	0.01	1.90	1147
	Excl. roll window	6956	−0.0091	1.771	0.12	1.41	594
Corn	All returns	7903	0.0082	1.769	−1.25	20.07	134,716
	Excl. roll days	7587	−0.0021	1.613	−0.00	2.59	2118
	Excl. roll window	6956	0.0039	1.607	−0.04	2.57	1914
Cotton	All returns	7903	−0.0010	1.835	−0.99	18.67	116,067
	Excl. roll days	7747	−0.0229	1.734	−0.57	9.94	32,331
	Excl. roll window	7435	−0.0134	1.715	−0.54	10.20	32,575
Orange Juice	All returns	7903	0.0041	2.268	0.09	4.87	7807
	Excl. roll days	7714	0.0126	2.199	0.24	4.72	7232
	Excl. roll window	7336	0.0132	2.188	0.24	4.94	7519
Live Cattle	All returns	7903	0.0164	1.133	−1.31	13.88	65,701
	Excl. roll days	7525	0.0362	0.985	−0.47	4.90	7809
	Excl. roll window	6768	0.0383	0.977	−0.50	5.32	8255

Note: Roll days are the first trading day on the new contract under the approximate roll dates of Table A2; the roll window adds the trading day before and after. Mean and SD are in percent. JB is the Jarque–Bera statistic; all p-values are below 0.001.

Table A4. AIC and BIC of the selected ARMA–eGARCH order under each innovation distribution.

Commodity	AIC: Normal	Student-t	GED	Skewed t	BIC: Normal	Student-t	GED	Skewed t
Rice	−5.5471	−5.6623	−5.6401	**−5.6694**	−5.5276	−5.6420	−5.6198	**−5.6482**
Soybeans	−5.7360	−5.7963	−5.7838	**−5.7973**	−5.7219	−5.7813	−5.7688	**−5.7814**
Wheat	−5.2764	−5.3071	−5.2952	**−5.3102**	−5.2552	−5.2851	−5.2731	**−5.2872**
Corn	−5.4366	−5.5317	−5.5114	**−5.5344**	−5.4163	−5.5105	−5.4902	**−5.5123**
Cotton	−5.3066	**−5.4181**	−5.4072	**−5.4181**	−5.2925	**−5.4031**	−5.3922	−5.4022
Orange Juice	−4.8772	−4.9767	−4.9675	**−4.9778**	−4.8542	−4.9529	−4.9437	**−4.9531**
Live Cattle	−6.2126	−6.4490	−6.4093	**−6.4501**	−6.1950	−6.4305	−6.3908	**−6.4307**

Note: Information criteria per observation (rugarch convention); lower is better, and the lowest value in each block is in bold. Each commodity uses the order in Table 3. The skewed Student-t values are those of Table 3; the other distributions are re-estimated at the same order.

Table A5. Parameter estimates of the selected ARMA–eGARCH skewed Student-t models (robust standard errors in parentheses).

Parameter	Rice	Soybeans	Wheat	Corn	Cotton	Orange Juice	Live Cattle
c	0.0004 (0.0002)	0.0002 (0.0002)	0.0002 (0.0001)	0.0002 (0.0001)	0.0001 (0.0002)	0.0000 (0.0003)	0.0004 (0.0001)
ϕ1	0.3860 (0.0003)	0.1842 (0.0041)	−1.2326 (0.0010)	−0.1253 (0.0017)	1.4991 (0.0065)	0.4923 (0.0129)	1.8997 (0.0138)
ϕ2	−1.0000 (0.0013)	0.9927 (0.0168)	−1.2095 (0.0006)	−0.0279 (0.0001)	−0.0009 (0.0046)	−0.9209 (0.0249)	−1.8574 (0.0270)
ϕ3	–	0.1951 (0.0095)	−0.4413 (0.0020)	0.0576 (0.0006)	−1.4879 (0.0034)	1.0246 (0.0231)	1.4121 (0.0258)

https://doi.org/10.3390/jrfm19110891

Table A5. *Cont.*

Parameter	Rice	Soybeans	Wheat	Corn	Cotton	Orange Juice	Live Cattle
ϕ4	–	−0.9774 (0.0032)	–	−0.2124 (0.0001)	0.9552 (0.0023)	−0.1465 (0.0173)	−0.5814 (0.0105)
ϕ5	–	–	–	−0.8762 (0.0008)	−0.0152 (0.0046)	0.2913 (0.0274)	–
θ1	−0.3202 (0.0000)	−0.1908 (0.0007)	1.2639 (0.0046)	0.1379 (0.0056)	−1.4870 (0.0000)	−0.4363 (0.0004)	−1.8641 (0.0002)
θ2	0.9459 (0.0000)	−0.9956 (0.0119)	1.2399 (0.0038)	0.0178 (0.0007)	−0.0264 (0.0001)	0.8620 (0.0000)	1.7815 (0.0025)
θ3	0.0719 (0.0001)	−0.1954 (0.0086)	0.4735 (0.0038)	−0.0527 (0.0005)	1.5025 (0.0000)	−0.9835 (0.0000)	−1.3234 (0.0036)
θ4	−0.0286 (0.0001)	0.9866 (0.0000)	−0.0061 (0.0285)	0.1989 (0.0001)	−0.9432 (0.0004)	0.0661 (0.0006)	0.5204 (0.0000)
θ5	−0.0057 (0.0003)	–	–	0.8722 (0.0009)	–	−0.2734 (0.0002)	–
ω	−0.3637 (0.0457)	−0.1908 (0.0835)	−0.0095 (0.0019)	−0.5083 (0.0008)	−0.0934 (0.0176)	−0.0105 (0.0029)	−0.2775 (0.0098)
α1	0.0296 (0.0110)	−0.0105 (0.0160)	0.0411 (0.0017)	−0.0768 (0.0001)	−0.0009 (0.0108)	0.0048 (0.0584)	−0.0524 (0.0151)
α2	0.0021 (0.0010)	0.0497 (0.0163)	0.0041 (0.0169)	−0.0271 (0.0000)	–	0.0247 (0.0255)	−0.0045 (0.0166)
α3	0.0161 (0.0124)	–	0.0116 (0.0045)	0.0332 (0.0000)	–	0.0251 (0.0361)	0.0103 (0.0157)
α4	−0.0248 (0.0057)	–	−0.0184 (0.0016)	0.0818 (0.0001)	–	−0.0203 (0.0507)	–
α5	−0.0025 (0.0008)	–	−0.0322 (0.0017)	–	–	−0.0488 (0.0201)	–
γ1	0.2196 (0.0038)	0.0970 (0.0253)	0.1330 (0.0015)	0.1635 (0.0002)	0.1438 (0.0129)	0.2536 (0.0305)	0.1394 (0.0221)
γ2	0.1135 (0.0005)	0.1296 (0.0257)	−0.0063 (0.0077)	0.2203 (0.0003)	–	0.0798 (0.0321)	−0.0357 (0.0240)
γ3	0.1185 (0.0038)	–	0.0281 (0.0059)	0.1357 (0.0007)	–	−0.1782 (0.1154)	0.0993 (0.0213)
γ4	−0.0799 (0.0014)	–	−0.0502 (0.0022)	0.0010 (0.0002)	–	−0.1376 (0.0675)	–
γ5	−0.0423 (0.0039)	–	−0.0886 (0.0005)	–	–	0.0298 (0.2508)	–
β1	0.3066 (0.0002)	0.3220 (0.0047)	0.8741 (0.0000)	−0.4097 (0.0001)	0.4291 (0.0011)	0.2881 (0.0007)	0.6458 (0.0005)
β2	−0.3270 (0.0000)	0.6557 (0.0050)	−0.0017 (0.0001)	0.3718 (0.0000)	0.6142 (0.0005)	0.8407 (0.0000)	−0.4771 (0.0005)
β3	0.9772 (0.0000)	–	0.1335 (0.0003)	0.9771 (0.0001)	−0.0548 (0.0035)	0.4622 (0.0001)	0.4338 (0.0004)
β4	–	–	0.8644 (0.0007)	–	–	−0.5924 (0.0012)	0.3672 (0.0005)
β5	–	–	−0.8715 (0.0000)	–	–	–	–
skew	1.0745 (0.1004)	0.9511 (0.0160)	1.0788 (0.0220)	1.0302 (0.4721)	1.0217 (0.0156)	0.9757 (0.0186)	0.9505 (0.0148)
shape	5.0774 (3.9918)	6.0449 (0.4226)	8.9539 (35.1049)	5.7208 (0.3274)	5.1127 (0.1775)	4.6646 (1.6704)	3.8787 (0.1859)
Log-lik.	22,430.9	22,926.0	21,007.3	21,895.8	21,422.3	19,694.2	25,509.5

Note: Parameters follow Equations (4) and (8); c is the mean, ϕ and θ the AR and MA coefficients, ω the log-variance constant, α the sign and γ the size coefficients, β the log-variance persistence coefficients; skew = 1 is a symmetric distribution. Robust (sandwich) standard errors. Where the AR and MA roots nearly cancel (Table 7), the likelihood is almost flat along the canceling directions and the Hessian is nearly singular, so the very small standard errors of some ARMA and β coefficients should not be read as precise estimates. These estimates are from the re-estimation used for Table 7; for rice, wheat, corn, cotton and orange juice their log-likelihood differs from the fits in Table 3 by between 1.7 and 5.5 points because the optimizer started from different values.

Table A6. Sign (asymmetry) coefficients α of the selected eGARCH models.

Commodity	Model	α1	α2	α3	α4	α5	Σα	Σγ
Rice	(2,5)–(5,3)	0.030 (0.007)	0.002 (0.037)	0.016 (0.194)	−0.025 (0.000)	−0.002 (0.002)	0.021	0.329
Soybeans	(4,4)–(2,2)	−0.011 (0.509)	0.050 (0.002)	–	–	–	0.039	0.227
Wheat	(3,4)–(5,5)	0.041 (0.000)	0.004 (0.807)	0.012 (0.010)	−0.018 (0.000)	−0.032 (0.000)	0.006	0.016
Corn	(5,5)–(4,3)	−0.077 (0.000)	−0.027 (0.000)	0.033 (0.000)	0.082 (0.000)	–	0.011	0.520
Cotton	(5,4)–(1,3)	−0.001 (0.931)	–	–	–	–	−0.001	0.144
Orange Juice	(5,5)–(5,4)	0.005 (0.934)	0.025 (0.332)	0.025 (0.488)	−0.020 (0.689)	−0.049 (0.015)	−0.014	0.047
Live Cattle	(4,4)–(3,4)	−0.052 (0.001)	−0.004 (0.788)	0.010 (0.511)	–	–	−0.047	0.203

Note: Estimates with robust p-values in parentheses. $\alpha > 0$: a positive shock raises next-period log-variance more than a negative shock of equal size; $\alpha < 0$: a negative shock raises it more. $\Sigma\alpha$ and $\Sigma\gamma$ are the sums of the sign and size coefficients.

Notes

1 More recent work on volatility forecasting and tail risk offers alternatives to the GARCH family used here. Heterogeneous autoregressive models of realized volatility (Corsi, 2009) exploit intraday data, score-driven (GAS) models let volatility and higher moments respond to the score of the conditional density (Creal et al., 2013), and Markov-switching GARCH models allow the parameters to change across volatility regimes (Haas et al., 2004; Ardia et al., 2019). Giot and Laurent (2003) and Kuester et al. (2006) show that GARCH models with skewed, fat-tailed innovations forecast VaR well. The ARMA-eGARCH models used here need only daily closing prices, which are available over a long sample for all seven markets, and they nest the standard GARCH(1,1) benchmark; regime-switching and score-driven specifications are a natural extension.

2 The asymmetry coefficients α are significant in most cases. For example, in an EGARCH(1,1) model with skewed Student-t innovations, the asymmetry coefficient is significant at the 1% level for every commodity except cotton ($p = 0.3151$). The sign coefficients of the selected models are reported in Table A6, and all parameter estimates with robust standard errors in Table A5.

3 As a robustness check, the grid was extended to seven lags. The longer lags give only marginal improvements in fit and do not change the preferred distribution or the risk estimates, so the search is kept at five lags.

4 Parameter stability was tested for the best ARMA-GARCH(1,1) models with the Nyblom test, which rejects stability at the 1% level for soybeans, wheat, orange juice and live cattle, and for the ARMA-eGARCH models in Table 3 with the cusum test of Lee and Song (2008) and Song and Kang (2018), which does not reject stability for any commodity. All test results are available from the authors on request.

5 See Note 4 above.

References

Acerbi, C., & Székely, B. (2014). Backtesting expected shortfall. *Risk*, *27*(11), 76–81.

Ardia, D., Bluteau, K., Boudt, K., Catania, L., & Trottier, D.-A. (2019). Markov-switching GARCH models in R: The MSGARCH package. *Journal of Statistical Software*, *91*(4), 1–38. [CrossRef]

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, *69*, 542–547. [CrossRef]

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, *39*(4), 841–862. [CrossRef]

Corlu, C., Meterelliyoz, M., & Tinic, M. (2016). Empirical distributions of daily equity index returns: A comparison. *Expert Systems with Applications*, *54*, 170–192. [CrossRef]

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

Creal, D., Koopman, S. J., & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, *28*(5), 777–795. [CrossRef]

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50*, 987–1007. [CrossRef]
Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, *22*(4), 367–381.
Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, *21*(5), 55–59. [CrossRef]
Fernández, C., & Steel, M. F. J. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, *93*(441), 359–371. [CrossRef]
Fisher, T. J., & Gallagher, C. M. (2012). New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Association*, *107*(498), 777–787. [CrossRef]
Ghalanos, A. (2024). *rugarch: Univariate GARCH models. R package*. Available online: https://CRAN.R-project.org/package=rugarch (accessed on 29 October 2026).
Giot, P., & Laurent, S. (2003). Market risk in commodity markets: A var approach. *Energy Economics*, *25*(5), 435–457. [CrossRef]
Goodwin, B. K., & Schnepf, R. (2000). Determinants of endogenous price risk in corn and wheat futures markets. *Journal of Futures Markets*, *20*(8), 753–774. [CrossRef]
Gordon, J. (1985). *The distribution of daily changes in commodity futures prices*. Technical Bulletin Number 1702. Economic Research Service, USDA.
Gupta, A., & Rajib, P. (2018). Do var exceptions have seasonality? An empirical study on Indian commodity spot prices. *IIMB Management Review*, *30*(4), 369–384. [CrossRef]
Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to markov-switching GARCH models. *Journal of Financial Econometrics*, *2*(4), 493–530. [CrossRef]
Hall, J., Brorsen, B., & Irwin, S. (1989). The distribution of futures prices: A test of the stable paretian and mixture of normals hypotheses. *Journal of Financial and Quantitative Analysis*, *24*(1), 105–116. [CrossRef]
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. [CrossRef]
He, X., & Serra, T. (2022). Are price limits cooling off agricultural futures markets? *American Journal of Agricultural Economics*, *104*(5), 1724–1746. [CrossRef]
Helms, B. P., & Martell, T. F. (1985). An examination of the distribution of futures price changes. *Journal of Futures Markets*, *5*(2), 259. [CrossRef]
Hernandez, M. A., Ibarra, R., & Trupkin, D. R. (2014). How far do shocks move across borders? Examining volatility transmission in major agricultural futures markets. *European Review of Agricultural Economics*, *41*(2), 301–325. [CrossRef]
Hudson, M., Leuthold, R., & Sarassoro, G. (1987). Commodity futures price changes: Recent evidence for wheat, soybeans and live cattle. *The Journal of Futures Markets (1986–1998)*, *7*(3), 287–301. [CrossRef]
Irwin, S., Sanders, D., & Merrin, R. (2009). Devil or angel? The role of speculation in the recent commodity price boom (and bust). *Journal of Agricultural and Applied Economics*, *41*(2), 377–391. [CrossRef]
Jin, H. (2004). The stable distribution of agricultural cash prices changes. *Journal of Rural Development/Nongchon-Gyeongje*, *27*(4), 177–195.
Junkus, J. C. (1991). Systematic skewness in futures contracts. *Journal of Futures Markets*, *11*(1), 9–24. [CrossRef]
Kateregga, M., Mataramvura, S., & Taylor, D. (2017). Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics & Finance*, *5*(1), 1318813. [CrossRef]
Kuester, K., Mittnik, S., & Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, *4*(1), 53–89. [CrossRef]
Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, *3*(2), 73–84. [CrossRef]
Lee, S., & Song, J. (2008). Test for parameter change in ARMA models with GARCH innovations. *Statistics & Probability Letters*, *78*(13), 1990–1998. [CrossRef]
Levy, H., & Duchin, R. (2004). Asset return distributions and the investment horizon. *Journal of Portfolio Management*, *30*(3), 47–62. [CrossRef]
Manfredo, M. R., & Leuthold, R. M. (1999). Value-at-Risk analysis: A review and the potential for agricultural applications. *Applied Economic Perspectives and Policy*, *21*(1), 99–111. [CrossRef]
Mann, J. S., & Heifner, R. G. (1976). *The distribution of shortrun commodity price movements*. Technical Bulletin No. 1536. U.S. Department of Agriculture, Economic Research Service.
Markowitz, H., & Usmen, N. (1996). The likelihood of various stock market return distributions, part 1: Principles of inference. *Journal of Risk and Uncertainty*, *13*, 207–219. [CrossRef]
McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3–4), 271–300. [CrossRef]

Morgan, W., Cotter, J., & Dowd, K. (2012). Extreme measures of agricultural financial risk. *Journal of Agricultural Economics*, *63*(1), 65–82. [CrossRef]

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*(2), 347–370. [CrossRef]

Poon, S., & Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478–539. [CrossRef]

Song, J., & Kang, J. (2018). Parameter change tests for ARMA–GARCH models. *Computational Statistics & Data Analysis*, *121*, 41–56. [CrossRef]

Stevenson, R., & Bear, R. (1970). Commodity futures: Trends or random walks? *The Journal of Finance*, *25*, 65–81. [CrossRef]

van Oordt, M. R. C., Stork, P. A., & de Vries, C. G. (2021). On agricultural commodities' extreme price risk. *Extremes*, *24*(3), 531–563. [CrossRef]

Wilhelmsson, A. (2006). GARCH forecasting performance under different distribution assumptions. *Journal of Forecasting*, *25*(8), 561–578. [CrossRef]

Wilhelmsson, A. (2009). value at risk with time varying variance, skewness and kurtosis—the NIG-ACD model. *The Econometrics Journal*, *12*, 82–104. [CrossRef]

Wilhelmsson, A. (2013). Density forecasting with time-varying higher moments: A model confidence set approach. *Journal of Forecasting*, *32*, 19–31. [CrossRef]

Zhu, Q., & Tansuchat, R. (2019). The extreme risk spillovers between the US and china's agricultural commodity futures markets. *Journal of Physics: Conference Series*, *1324*(1), 012085. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.