Estimation of Parameters of Marginal Distributions of Return Rates of Selected Agricultural Products Listed on the Chicago Commodity Exchange

Gabriela Malik PhD
College of Economics and Computer Science in Krakow (Poland)

Abstract

The article aims to estimate and verify the correctness of the specification of tail distributions of return rates of agricultural products listed on the Chicago Commodity Exchange from 1987 to 2022. The selected products include corn, soybeans, and wheat. The analysis determined a model describing the examined series concerning mean and variance values, considering the relationships between the series. To describe the tail distributions, a GARCH model was utilized. In the subsequent stage, the quality of fit of the estimated model was assessed.

During the conducted research, it was found that the model best describing the analyzed series of prices for agricultural futures contracts is the AR (1)-GARCH (1,1) model with a conditional t-Student distribution.

The conducted analysis is crucial for correctly determining the proper forms of estimated models. It is essential to emphasize that errors in the correctness of specifying tail distributions can consequently lead to incorrect parameter estimation in further studies conducted for agricultural products.

Keywords: agricultural products, estimation, tail distributions, GARCH model
1. Introduction

Time series of financial instrument return rates usually belong to a group of stationary or non-stationary processes, the degree of integration of which typically does not exceed unity. Autocorrelation is also a common feature, although it rapidly diminishes for higher lags [see Brzeszczyński and Kelm (2002), p. 82; Weron and Weron (2009), p. 300]. To describe the dynamics of conditional mean return rates, both univariate ARMA models can be applied [cf. Patton (2009), pp. 767-786; Chollete et al. (2009); Manner and Reznikova (2012); Liu et al. (2019); Moon et al. (2021)], as well as a multivariate Vector Autoregression (VAR) model [cf. Candelon and Manner (2010); Rossi and de Magistris (2013); Gautam and Kanoujiya (2022); Kilian and Zhou (2023)]. In this paper, the GARCH model will be used to describe the tail distributions, which was introduced into the global literature in 1986 by T. Bollerslev [see Bollerslev (1986)]. One of the undeniable advantages of GARCH models is the time-varying conditional variance, while the unconditional variance remains constant. Modeling daily variance based on the estimated GARCH-class model, whose best parameterization will be determined based on the Akaike Information Criterion, results in the highest flexibility. The choice of the model was also influenced by the analysis of the properties of the agricultural product return series as well as its popularity in the literature.

This article aims to estimate and verify the correctness of the specification of tail distributions of return rates of agricultural products listed on the Chicago Commodity Exchange, specifically for selected products: corn, soybeans, and wheat. The analysis will determine a model describing the examined series in terms of mean and variance values, considering the mutual relationships between the series. The goodness of fit of the estimated model will also be checked.

The article is divided into five parts. The second part presents the characteristics of the research sample. The third part describes the statistical properties of the distribution of agricultural product prices in the futures market. It includes descriptive statistics and the results of normality tests of return rates of the examined agricultural products, along with their graphical representation. The estimation of parameters of tail distributions of agricultural product prices, along with the assessment of their goodness of fit, is in the fourth part. Finally, the last, fifth part provides an overview of the most important conclusions formulated during the conducted study.

2. Description of the Research Sample

The data used for analysis in the empirical part of the study were sourced from the Chicago Mercantile Exchange (CME). The research sample consists of daily quotes of nominal prices of futures contracts for three agricultural products: corn, wheat, and soybeans. The contract value is expressed as the price per bushel of the respective commodity unit in US dollars. The data include the closing price of contracts with the shortest expiration term, allowing the series of quotations to be treated as futures prices with the shortest possible realization term. The selection of products was justified by their significance in the futures market and the availability:
of sufficiently long time series. The empirical data cover the years 1987-2022, and each individual time series comprises 9,090 observations. These data were checked for potential discontinuities and errors. To minimize the impact of arbitrary interventions on the obtained results, no procedures for data correction or supplementation were applied.**1**

Figure 2.1 illustrates the behavior of prices for the analyzed agricultural products during the examined period. "It is visible that in mid-2008, there was a sharp increase in prices for each of the agricultural products, especially for soybeans, where the maximum price reached 1618.5 cents per bushel, while the maximum prices for wheat and corn were 1195 and 711 cents per bushel, respectively. The direct cause of such a significant increase in prices for the studied agricultural products was the emerging financial crisis in the USA, considered to have started in September 2008 when the American investment bank Lehman Brothers collapsed. The increase in the prices of the analyzed agricultural products can also be observed in 2010. It was not as sharp as two years earlier, and its main cause was the drought and fires that affected Russia."**2** However, the most spectacular increase in prices for the analyzed agricultural products was observed in 2022 when the price of soybeans stood at 1772.2 cents per bushel, and wheat and corn were priced at 1299.9 and 816.5 cents per bushel, respectively. This significant rise in agricultural product prices in global markets was and still is, albeit to a somewhat lesser extent, due to the Russian aggression against Ukraine.

![Figure 2.1. Daily Changes in Futures Contract Prices for Corn, Soybeans, and Wheat Traded on the Chicago Commodity Exchange from 1987 to 2022.](image)

Source: The author

---

1 Quote from Malik G., “Extreme value distribution of prices of chosen agricultural products listed on futures market”, International Journal of Applied Technology & Leadership, vol. 1, iss. 2 (July 2022), s. 2

2 Quote from Malik G., “Extreme value distribution of prices of chosen agricultural products listed on futures market”, International Journal of Applied Technology & Leadership, vol. 1, iss. 2 (July 2022), s. 3
3. Statistical Properties of Agricultural Commodity Prices in the Futures Market

Price series of financial instruments belong to the group of non-stationary processes. Therefore, to conduct statistical analysis between agricultural commodity prices, daily continuous return rates were calculated based on the price series. The use of logarithmic return rates is significant for the properties of the analysed series, as logarithm transformation, one of the Box-Cox transformations, stabilizes the series’ variance.

Hence, daily closing prices were used to compute logarithmic return rates for individual agricultural products, according to the formula presented below:

\[
R_t = \ln \left( \frac{X_t}{X_{t-1}} \right),
\]

where \(X_t\) represents the futures contract value on day \(t\).\(^3\)

In this article, graphical illustrations of daily return rates for futures contracts of corn, soybeans, and wheat listed on the Chicago Commodity Exchange during the analysed period, i.e., from 1987 to 2022, were not presented. This extension of the research period from 2010 to 2022 did not fundamentally alter the appearance of the charts and their characteristics. Analogous conclusions can be drawn regarding the behaviour of return rates for futures contracts of individual agricultural products [cf. G. Malik (2022)].

To conduct a detailed analysis of the empirical data, basic descriptive statistics for the return rates of the examined agricultural products are presented in Table 3.1. The last two rows of the table contain the results of normality testing using the Shapiro-Wilk and Jarque-Bera tests.

As the results presented in Table 3.1 show, for each of the examined agricultural products, the average daily return rates are close to zero, as expected. The comparison of the minimum and maximum values with the first and third quartiles respectively clearly indicates a strong tendency to extreme values, especially on the negative side of the return rate distribution.

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybeans</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>-0,207</td>
<td>-0,135</td>
<td>-0,229</td>
</tr>
<tr>
<td>Quartile 1.</td>
<td>-0,008</td>
<td>-0,007</td>
<td>-0,011</td>
</tr>
<tr>
<td>Median</td>
<td>0,000</td>
<td>0,000</td>
<td>0,000</td>
</tr>
<tr>
<td>Average</td>
<td>0,000</td>
<td>0,000</td>
<td>0,000</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0,017</td>
<td>0,015</td>
<td>0,021</td>
</tr>
<tr>
<td>Quartile 3.</td>
<td>0,009</td>
<td>0,008</td>
<td>0,011</td>
</tr>
</tbody>
</table>

\(^3\) Quote from Malik G., "Extrem value distribution of prices of chosen agricultural products listed on futures market", International Journal of Applied Technology & Leadership, vol. 1, iss. 2 (July 2022), s. 3
The minimum daily return rate ranges from -22.9% for wheat to -13.5% for soybeans. On the other hand, the highest maximum return rate was achieved for corn, reaching 23.3%, while the lowest was for soybeans at 12.5%. The standard deviation takes small values, ranging from 0.015 for soybeans to 0.021 for wheat. Since the standard deviation is greater than the mean, it can be concluded that the daily return rates of individual agricultural products exhibit high volatility. Additionally, the slightly negative skewness value and high kurtosis confirm a well-known fact in the literature, indicating the limited suitability of the normal distribution to describe return rates. They also suggest a slight leftward asymmetry in the examined series. For each agricultural product, the empirical distributions of return rates exhibit characteristics significantly deviating from normality. This is reflected in rejecting the null hypotheses of the empirical distributions conforming to the normal distribution in both applied normality tests at a very stringent level of significance [cf. G. Malik, 2016].

It is important to note that as early as the 1960s, it was demonstrated that the normal distribution is of limited use in describing the empirical distribution of prices and return rates in futures markets. Both E. Fama [cf. E. Fama, 1965] and P. Clark [cf. P. Clark, 1973] presented significant evidence in their works showing the dissonance between what was observable and what the normal distribution postulated. Similarly, B. Mandelbrot [cf. B. Mandelbrot, 1963], analyzing cotton prices in the commodity futures market, confirmed that they could not be described using a normal distribution.

Although the rejection of the normality hypothesis is a consensus in the literature, the open question remains regarding the choice of the best distribution to replace the discredited normal distribution, both in the context of financial markets and futures markets. Deng et al. (2002) as well as Jin (2007) emphasize the importance of α-stable distributions in describing prices of commodities listed on commodity exchanges. The extensive use of the scaled t-Student distribution to describe return rates has also been demonstrated in the global literature [cf. e.g., Praetz (1972); Blattberg and Gonedes (1974); Gray and French (1990); Peiro (1994); Aparicio and Estrada (2001); Broca (2002); Dritsaki (2019)]. Similar conclusions were reached by Malik (2011) when analysing the price distributions of agricultural products listed on the Chicago Commodity Exchange.

Table 3.1. Descriptive Statistics and Normality Test Results for Daily Return Rates of Examined Agricultural Products

| Source: The author |

* In parentheses, critical probability values (p-values) are provided.
4. Estimation of Parameters for Marginal Distributions of Agricultural Product Return Rates

The GARCH \((p, q)\) model is described by the following formula:

\[
\sigma_i^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 ,
\]

where: \(p > 0\), \(q \geq 0\) and \(\alpha_0 > 0\), \(\alpha_i > 0\) for \(i = 1, \ldots, p\), however \(\beta_j > 0\) for \(j = 1, \ldots, q\), are the conditions to ensure the positivity of conditional variance.

In equation (4.1) the conditional variance is explained using lagged squared returns and lagged conditional variance. The parameters \(\alpha\) are therefore responsible for the effect on volatility of the information contained in the \(\epsilon_{t-1}^2\), while the \(\beta\) parameters illustrate the dynamics characterizing market expectations, i.e. volatility not susceptible to changes and progressing into the future similarly to the past.

To assess the degree of integration of prices of the considered agricultural products, the generalized DF test was applied [see W. W. Charemza, D. F. Deadman, 1997, s. 114-117; D. A. Dickey, W. A. Fuller, 1979; D. A. Dickey, W. A. Fuller, 1981]. The selection procedure involved starting the test with the maximum lag and then reducing the lag by one in each subsequent round until rejecting the null hypothesis of the existence of a unit root. The testing allowed determining the integration degree of return rate series for futures contracts of agricultural products at \(d = 0\), indicating stationarity. The specific parametrization of the model occurred in a two-step manner. General guidelines regarding the number of autoregressive and moving average parameters were derived from the autocorrelation and partial autocorrelation function plots [por. G. Box, G. Jenkins, 1983, s. 93]. The final criterion was the Akaike Information Criterion value.

Consequently, in the estimation process, AR (1) models were used to describe the conditional mean. For describing conditional variances, one-dimensional GARCH models were applied, described by Equation (4.1) in their most popular parametrization, GARCH (1,1), with a conditional normal distribution and a conditional \(t\)-Student distribution. The choice of conditional distributions was motivated by results commonly presented in the literature, recommending their usage [por. J. Osiewalski, M. Pipień, 1999; J. Osiewalski, A. Pajor, M. Pipień, 2004] and by the author's scientific research [cf. G. Malik, 2011; 2013]. The standardized residuals from these models were then transformed using appropriate theoretical distribution quantiles into a series of observations, which are realizations of random variables uniformly distributed on the interval \([0,1]\).

The estimation results of the considered model are presented in Tables 4.1 and 4.2. Besides the parameter values and their standard errors given in parentheses, the tables also contain the values of the log-likelihood function (LLF) and the Akaike Information Criterion (AIC), as well as values of the Bayesian, Shibata, and Hannan-Quinn criteria included for comparative purposes.
All parameters related to the variance equation in the one-dimensional GARCH models presented in Tables 4.1 - 4.2 turned out to be statistically significant. Meanwhile, the parameters in the mean equations are statistically insignificant, confirming the absence of autocorrelation between the returns of the analyzed series. A detailed analysis of the estimated coefficients shows that the sum of the parameters in the one-dimensional GARCH models is close to unity, indicating a strong conditional variance dependency at time $t$ on the variance at the time $t-1$. This confirms the frequently occurring phenomenon called volatility clustering, which was already observed during the interpretation of the daily return charts of futures contracts in Chapter 3 of this article. On the other hand, the high value of the $\beta_1$ coefficient obtained in the estimation process in the variance equations confirms a clear GARCH effect in the model. Additionally, the noticeably lower than thirty shape parameter value in the case of the conditional $t$-Student distribution indicates the occurrence of extreme observations much more frequently than in the case of the normal distribution.
Analysing the results presented in Tables 4.1 - 4.2, it can also be observed that the lowest Akaike criterion value and the highest log-likelihood function value were obtained when applying the conditional $t$-Student distribution. This indicates a better fit of the conditional $t$-Student distribution to the estimated model. It is important to note here that the slight difference compared to the values obtained for the conditional normal distribution is a result of computations conducted on series with a very large number of observations (over 9,000 observations). Moreover, the use of the AR (1)-GARCH (1,1) model with a conditional $t$-Student distribution for a long series of observations is consistent with the results commonly presented in the literature [por. J. Osiewalski, M. Pipień, 1999; J. Osiewalski, A. Pajor, M. Pipień, 2004].

The next step in the analysis is to assess the quality of fit of the estimated model. For this purpose, Table 4.3 presents comprehensive results of tests conducted on the standardized residuals of the AR (1)-GARCH (1,1) model for both assumed marginal distributions.

<table>
<thead>
<tr>
<th>Normal marginal distribution</th>
<th>Marginal distribution $t$-Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>corn</td>
<td>soybeans</td>
</tr>
<tr>
<td>corn</td>
<td>soybeans</td>
</tr>
<tr>
<td>soybeans</td>
<td>corn</td>
</tr>
<tr>
<td>soybeans</td>
<td>corn</td>
</tr>
</tbody>
</table>

Box-Pierce Q-statistic for standardized residuals squared

| $Q [1]$ | 0.293 | 0.986 | 0.595 | 0.669 | 4.478 | 0.458 |

Table 4.2. Estimation Results of AR (1)-GARCH (1,1) Model with Conditional $t$-Student Distribution
Source: The author
The results presented in Table 4.3 show that the AR (1)-GARCH (1,1) model describes the relationship between returns for corn very well, both when assuming a normal marginal distribution and a t-Student marginal distribution. In both cases, there is no autocorrelation or ARCH effect in the standardized residual series. Moreover, for all agricultural products, there is no asymmetry effect in the residuals series. However, for soybeans and wheat, the results are not as clear-cut as for corn. Specifically, for the standardized residuals of soybeans and the squared standardized residuals of wheat for higher lags, the null hypothesis was rejected, both for the normal marginal distribution and the t-Student marginal distribution, albeit to a lesser extent. A similar situation applies to Engle's statistic for soybeans and wheat for both assumed marginal distributions. The ultimate conclusion from the analysis of the results presented in Table 4.3 is that the AR (1)-GARCH (1,1) model describes the relationships of returns very well for corn, quite well for soybeans and wheat, with the application of the t-Student distribution as the marginal distribution proving to be more useful.

### Table 4.3. Results of Tests Conducted on Standardized Residuals of AR (1)-GARCH (1,1) Model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0,911</td>
<td>2,080</td>
<td>0,904</td>
<td>1,385</td>
<td>5,433</td>
<td>0,897</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q [15]</td>
<td>6,703</td>
<td>7,567</td>
<td>5,920</td>
<td>6,967</td>
<td>11,177</td>
<td>5,186</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Box-Pierce Q-statistic for standardized residuals squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q [1]</td>
<td>3,426 0,054 0,001 1,364 0,287 0,901</td>
</tr>
<tr>
<td>Q [5]</td>
<td>4,143 2,684 6,642 2,428 2,477 8,488</td>
</tr>
<tr>
<td>Q [15]</td>
<td>6,689 15,555 22,599 5,370 17,451 23,482</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Engle test statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCH [1-2]</td>
<td>3,427 1,138 2,945 1,428 1,061 6,679</td>
</tr>
<tr>
<td>ARCH [1-5]</td>
<td>5,231 15,131 7,591 3,384 17,009 8,801</td>
</tr>
<tr>
<td>ARCH [1-10]</td>
<td>7,377 24,420 24,857 6,541 26,121 24,774</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Joint Sign Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2,435 3,977 7,250 3,193 3,904 6,068</td>
</tr>
</tbody>
</table>

Source: The author
The figures below present estimates of volatility for individual agricultural products according to the specification of the AR (1)-GARCH (1,1) model for marginal distributions with a conditional normal distribution and a conditional t-Student distribution. Analyzing the charts shown in Figure 4.1, it can be observed that the estimates of conditional variance for each of the studied agricultural products are almost identical, regardless of the applied marginal distribution. Additionally, it is evident that the highest volatility of corn exhibits the largest deviations in conditional variance, while soybeans have lowest volatility.

![Conditional variance charts for different agricultural products](image)

**Figure 4.1.** Conditional variance of agricultural products’ returns for conditional normal distribution [Fig. A, B, C] and conditional t-Student distribution [Fig. D, E, F] obtained using the AR (1)-GARCH (1,1) model.

Source: The author

A crucial aspect in the course of the conducted research is the correctness of specifying marginal distributions, especially since errors in this area can lead to incorrect estimation in subsequent studies. Therefore, the next step of the analysis was to verify the conformity of the marginal distributions with the uniform distribution on the interval [0,1] using the Kolmogorov-Smirnov test and to check for autocorrelation using the Ljung-Box test. Additionally, to assess the uniformity of marginal distributions, the Anderson-Darling test was employed due to its high sensitivity to extreme observations. The results of the conducted tests are summarized in the tables below. The critical probability values in the tests for the correctness of specifying marginal distributions are provided in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Corn</th>
<th>Soybeans</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test K-S</td>
<td>0.074</td>
<td>0.055</td>
<td>0.069</td>
</tr>
</tbody>
</table>

© 2024 Journal of Applied Technology and Leadership
The results presented in Tables 4.4-4.5 show that all applied tests do not provide grounds to reject the null hypotheses regarding both the lack of autocorrelation and the conformity of distributions with the uniform distribution.

### 5. Conclusions

In the process of studying relationships in the futures market of agricultural products, it is crucial from a modeling perspective to determine a model describing the analyzed series concerning the mean and variance values, taking into account the mutual relationships between the series. The model that best describes the analyzed series of futures contract prices for agricultural products is the AR (1)-GARCH (1,1) model with a conditional t-Student distribution. All parameters related to the variance equation proved to be statistically significant, whereas parameters in the mean equations are statistically insignificant, confirming the lack of autocorrelation between the returns of the analyzed series. A detailed analysis of the estimated coefficients confirmed the phenomenon commonly known as volatility clustering, a clear GARCH effect, and the occurrence of extreme observations much more frequently than in the case of the normal distribution. The analysis of the fit quality showed that this model particularly well describes the relationship between the returns for corn. However, for soybeans and wheat, the results were not as clearly definitive. The ultimate conclusion from the analysis of the fit quality results is that the AR (1)-GARCH (1,1) model with a conditional t-Student distribution accurately describes the relationships in the returns for corn, and fairly well for soybeans and wheat.
References


