

## **Statistical Analysis of Soft Commodities Returns in the Period 2007-2016**

**Abstract.** Soft commodities, often referred to as tropics, constitute a significant element of international trade and are also important to the Polish economy. Moreover, Polish investors may invest in foreign markets of soft commodities through commodity-linked ETFs, offered by some brokerage houses. Obviously, each investment decision should be preceded by an analysis of asset performance. This paper provides results of statistical analysis of soft commodities returns over the period January 2007 to December 2016. They reveal the existence of weak positive correlation between the returns, non-normal distributions, negative trends, and serial autocorrelation.

**Keywords:** soft commodities, distributional characteristics, correlation, trend, autocorrelation

**JEL Classification:** C10, C12, G12

### **Introduction**

Soft commodities (softs) constitute an important part of the world trade. Their role is reflected in construction of commodity indices, such as the Tomson Reuters/Core Commodity CRB Index that is a broad indicator of commodity price movement and the most widely followed index of commodities futures. The index, originally composed of 28 commodities, currently is made up of 19 commodities and soft commodities cover more than 20% of its portfolio.

Basic soft commodities are: coffee, cocoa, sugar, cotton, orange juice, and rubber<sup>3</sup>. With no doubt, the most important of them is coffee - the most valuable agriculture item on the world market. According to the FAIRTRADE Poland, the actual value of the world coffee market exceeds the amount of 81 billion USD ([www.fairtrade.org.pl](http://www.fairtrade.org.pl)). The coffee bean had its origin in Ethiopia, where it was roasted for the first time in the 14th century and then brewed into coffee (Eller and Sagerer, 2008). Today, this tropical crop is grown in many different varieties, but the basic classification establishes two main coffee types: the higher-quality African coffee Arabica and Latin American coffee Robusta. Geman (2007) notes that the USA almost exclusively import Robusta coffee, probably because the US consumers have gotten to its taste. The largest coffee producer is Brazil, followed by Colombia, Ecuador, Costa Rica, and Vietnam. Eller and Sagerer (2008) claim that since the trend in Asian regions has been to plant less tea crops and more coffee crops, demand for this agro-sector is only expected to increase in the future.

---

<sup>1</sup> PhD, Faculty of Economic Sciences, Nowoursynowska 166, 02-787 Warsaw, e-mail: [anna\\_gorska@sggw.pl](mailto:anna_gorska@sggw.pl), <https://orcid.org/0000-0003-0763-0550>

<sup>2</sup> PhD, Faculty of Applied Informatics and Mathematics, Nowoursynowska 159, 02-776 Warsaw, e-mail: [krawiec.monika@gmail.com](mailto:krawiec.monika@gmail.com)

<sup>3</sup> According to Fabozzi, Füss and Kaiser (2008), there are three other commodities that may fall into the group of softs. These are: silk, timber, and wool.

On the contrary, cocoa is mainly cropped at the Ivory Coast. It is characterized by a long production cycle as new trees take approximately seven years to mature and fruit (for new coffee bushes it takes 3 – 5 years). The typical cocoa flavor occurs after fermentation of the bean semen. More than 90% of the world harvest is used in the sweets industry, the rest in cosmetics and pharmaceuticals industries.

The agro-raw material sugar is manufactured from sugarcane and from sugar beets. Sugarcane is grown in tropical sub-regions such as Brazil, India, and Cuba, whereas sugar beets come from moderate climates such as in Europe, Australia or China. Sugar is used to improve flavor and preserve goods, but it has also been used increasingly as gasoline competitor in the form of ethanol fuel. Brazil is the largest sugar producer, but half of the Brazilian sugar harvest is currently processed into biofuel. The second largest producer is India followed by European Union<sup>4</sup>, and China.

Cotton is another hot-weather crop that has been in use for more than 5 000 years in India. Also in China, Egypt and the United States it has been in use for 1 000 years. Today, cotton is grown in more than seventy countries worldwide, of which China and the United States are the most important producers (Eller and Sagerer, 2008). In the US it is grown from Georgia through Arizona and California. According to Geman (2007), Californian cotton is among the best quality cotton in the world. After the harvest, the cotton is dried, cleaned, and packed into the bales. Cotton fibers are used in the textile industry.

The most important producers of concentrated orange juice are: Brazil (the state of Sao Paulo) and the United States (the state of Florida), and then Italy and Mexico. Frozen concentrated orange juice (FCOJ) was first traded in an organized way around 1947. Innovations in packaging and transportation permitted the emergence of an international market. In 1966 the New York Board of Trade (NYBOT), launched FCOJ futures contracts, whereas options on futures appeared in 1985. Both instruments allow for hedging against the impact of the weather on commodity price as production in Florida is very sensitive to Frost while orange trees in Brazil languish in a dry climate (Geman, 2007).

The last one of the soft commodities under consideration is rubber. Natural rubber is mainly produced in South and Southern-West Asia (in Thailand, Indonesia, and Vietnam), but also in India, China, and Malesia. The world centers for rubber trade are Tokyo Commodity Exchange and Singapore Commodity Exchange. Natural rubber is used extensively in many applications and products, either alone or in combination with other materials. In most of its useful forms, it has a large stretch ratio and high resilience, and is also waterproof.

As all these soft commodities are underlying assets for derivatives, not only producers and consumers, but also institutional and individual investors are interested in recognizing some patterns, trends and anomalies in performance of their prices and returns. Thus, this paper is aimed at the statistical analysis of soft commodity returns in the period January 2007 through December 2016. Data under consideration covers daily spot prices of the commodities available at [www.bloomberg.com](http://www.bloomberg.com).

---

<sup>4</sup> The sugar market in the EU is among most regulated food markets in the world. This regulation is based on production limits and foreign trade regulations (Hamulczuk and Szajner 2015).

## Methods of analysis

Following Chevallier and Ielpo (2013) we consider commodities as any other financial asset and work with returns on commodities:

$$r_t = p_t - p_{t-1}, \quad (1)$$

where  $p_t$  is the logarithm of commodity price at time  $t$  and  $p_{t-1}$  is the logarithm of commodity price in the previous period.

Then, we calculate several descriptive characteristics.

1. Mean return is an indicator for the location of the return distribution. It is:

$$\bar{r} = \frac{r_1 + r_2 + \dots + r_n}{n}. \quad (2)$$

2. Standard deviation:

$$\sigma = \left( \frac{1}{n-1} \sum_{i=1}^n (r_i - \bar{r})^2 \right)^{0,5}. \quad (3)$$

The standard deviation indicates the dispersion of the return distribution. It accounts for both the positive and negative deviations from the average value and is thus a measure for the total risk of an investment.

3. Skewness:

$$s = \frac{\frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})^3}{\sigma^3}. \quad (4)$$

The skewness describes the asymmetry of a probability distribution. If the distribution has a longer tail on the right (left) side then the distribution is referred to as positively (negatively) skewed. The skewness of a symmetrical probability distribution is equal to zero.

4. Kurtosis:

$$K = \frac{\frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})^4}{\sigma^4}. \quad (5)$$

Kurtosis describes whether a probability distribution is more acute or wider in comparison to normal distribution, which has kurtosis of three. If the probability is more acute (wider) and has more (less) returns at tails than a normal distribution, then the distribution is referred to as leptokurtic (platykurtic). Having thick distribution ends ("fat tails") is referred to as having positive excess kurtosis; having thinner tails when compared to the normal distribution indicates negative excess kurtosis.

In order to assess linear dependence between soft commodities returns, Pearson correlation measure is considered. The Pearson correlation coefficient  $k_{XY}$  of commodities X and Y is:

$$k_{XY} = \frac{\sum_{i=1}^n (r_{Xi} - \bar{r}_X)(r_{Yi} - \bar{r}_Y)}{(n-1)S_X S_Y}. \quad (6)$$

Its value ranges between  $<-1, 1>$ . The higher the value of its modulus, the stronger the linear dependence between the investigated time series.

For the purpose of analysis, there is also tested normality of distribution of time series under consideration. The normal distribution can be statistically verified using respective Jarque-Bera test.

The null hypothesis:

$H_0$ : the returns follow a normal distribution

is tested against the alternative hypothesis

$H_1$ : the returns do not follow a normal distribution.

The test statistic  $JB$  is:

$$JB = \frac{n}{6} \left[ S^2 + \frac{1}{4}(K-3)^2 \right]. \quad (7)$$

The statistic follows an asymptotic chi-squared distribution with two degrees of freedom.

Commodity returns often exhibit trends. A trend in the commodity returns can be represented in the following way:

$$p_t = \alpha_0 + \alpha_1 t + \varepsilon_t, \quad (8)$$

where  $\varepsilon_t$  captures the deviation from the trend. The sign of  $\alpha_1$  estimator ( $a_1$ ) decides the nature of the trend. Moving back to returns, this equation turns out to be:

$$r_t = a_1 + e_t - e_{t-1}, \quad (9)$$

where  $e_t$  - residuals from equation (8).

When  $a_1$  is positive (respectively negative), the commodity is driven by a positive (negative) trend.

Trends can be understood as a natural tendency for returns to have one particular sign, regardless of the period considered or of the frequency used to compute it. This is measured in equation (8) by the estimator of  $\alpha_1$ . The question is whether the trends observed are persistent. A first simple and intuitive approach to gather the evidence that commodities exhibit persistent trends is to compute the autocorrelation of order  $h$ , that is:

$$\hat{\rho}_h = \frac{\sum_{i=h+1}^n (r_i - \bar{r})(r_{i-h} - \bar{r})}{\sum_{i=1}^n (r_i - \bar{r})^2} . \quad (10)$$

If  $\hat{\rho}_h$  is positive and significant at 5% risk level, this provides evidence that a given commodity exhibits persistence. Conversely, when  $\hat{\rho}_h$  is negative and significant, the commodity under consideration has a mean-reverting behavior, rapidly correcting its trajectory in case of a large positive or negative return (Chevallier and Ielpo, 2013).

If observed trends are persistent, then investors applying trend-following strategies can benefit from commodity markets (see Lukac, Brorsen and Irwin (1988), Irwin et al. (1997), Erb and Harvey (2006), Miffre and Rallis (2007), Fuertes, Miffre and Rallis (2008), Szakamary, Shen and Dharma (2010), Górska and Krawiec (2011), Zaremba (2015)).

Instead of testing the statistical significance of any individual autocorrelation coefficient, we can test the joint hypothesis that all the  $\hat{\rho}_h$  up to certain lags are simultaneously equal to zero. This can be done by using the statistic developed by Box and Pierce (1970). In 1978 this statistic was modified by Ljung and Box. The Ljung-Box (LB) statistic is defined as:

$$LB = n(n+2) \sum_{h=1}^m \frac{\hat{\rho}(h)^2}{n-h} , \quad (11)$$

where:  $n$  – sample size,  $m$  – lag length.

In large samples LB statistic follows chi-square distribution with  $m$  degrees of freedom (Ljung and Box 1978). According to Gujarati (2003), the LB statistic has been found to have better (more powerful in the statistical sense) small-sample properties than the Box-Pierce statistic.

## Empirical results

Following the methodology described in the previous section, in the first step of research there are calculated logarithmic returns of soft commodities (displayed in Figure 1) and their basic descriptive characteristics (given in Table 1): maximum and minimum, range (that is the difference between them), standard deviation, kurtosis, skewness, the JB statistics, trend slope  $a_1$ , and coefficient of autocorrelation of the first order ( $\hat{\rho}_1$ ).

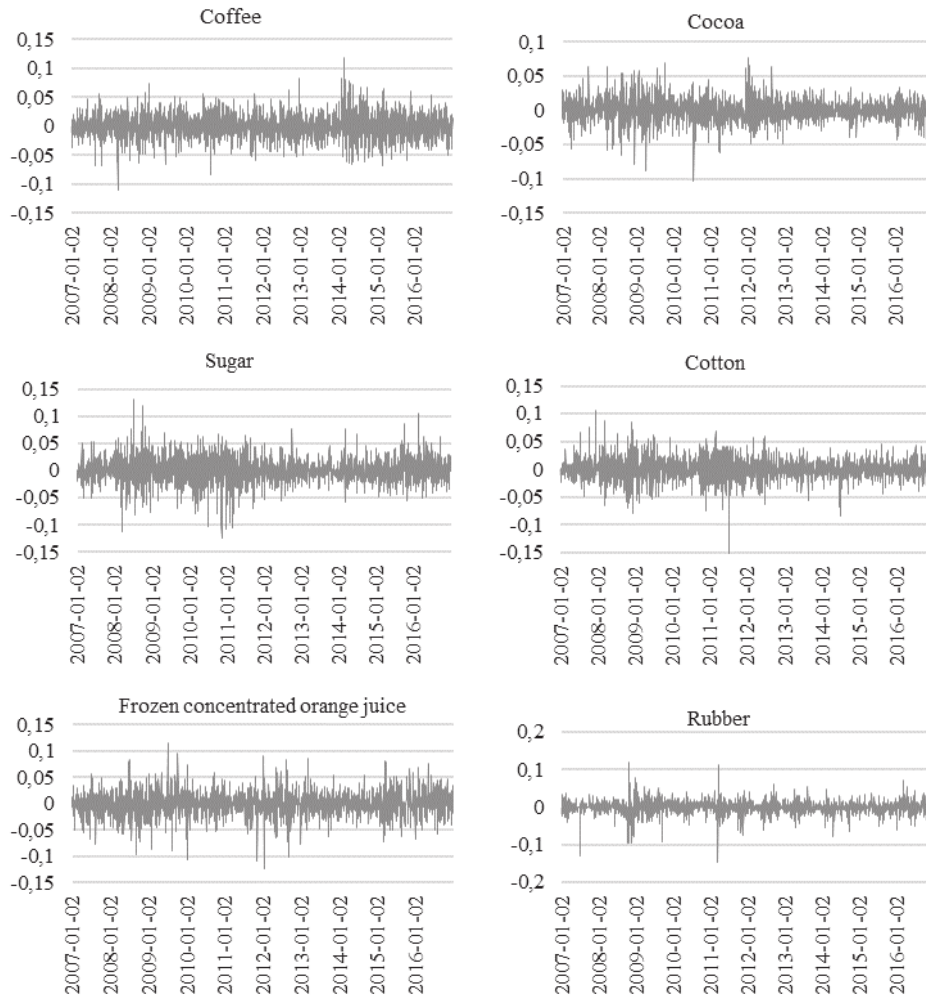


Fig.1. Soft commodities daily logarithmic returns in the period 2007-2016

Source: own elaboration.

On the base of results given in table 1 we can conclude that almost all soft commodities exhibit positive mean returns, the highest one - cocoa. The only exception is frozen concentrated orange juice with negative mean return. All commodities (except for coffee) are described by negative skewness. Moreover, kurtosis is higher than 3 for four out of six soft commodities (cocoa, sugar, cotton, rubber) whose distributions are leptokurtic. The Jarque-Bera (JB) statistic of normality confirms the rejection of null hypothesis for all return series at 5% level of significance. All commodities, with the exception of frozen concentrated orange juice, are driven by negative trends. We also find a statistically significant positive autocorrelation of order 1 for four soft commodities (cocoa, cotton, frozen concentrated orange juice, rubber), which is the evidence, although limited, of trend persistence.

Table 1. Basic characteristics for logarithmic returns of soft commodities

Characteristic	Commodity					
	Coffee	Cocoa	Sugar	Cotton	Frozen concentrated orange juice	Rubber
Maximum	0.11789	0.07810	0.13062	0.10527	0.11474	0.11879
Minimum	-0.11089	-0.10401	-0.12366	-0.15555	-0.12314	-0.14774
Range	0.22878	0.18211	0.25428	0.26082	0.23789	0.26653
Mean	0.00003	0.00026	0.00019	0.00009	-0.00001	0.00002
Standard deviation	0.01982	0.01550	0.02190	0.01888	0.02163	0.01711
Kurtosis	2.08055	3.48264	3.26530	4.47554	2.74104	8.78294
Skewness	0.09118	-0.14439	-0.07872	-0.31238	-0.13873	-0.67168
Trend slope	-2.2e-07	-5.7e-07	-1.7e-07	-2.8e-07	6.8e-07	-3.1e-07
Autocorrel. of order 1	-0.0336	<b>0.0464</b>	-0.0099	<b>0.1126</b>	<b>0.0874</b>	<b>0.1233</b>
JB	471.33*	1320.77*	1155.69*	2209.30*	820.62*	8544.24*

Note: bold type denotes statistical significance at 0.05 level; \* denotes rejection of null hypothesis at 0.05 level

Source: own calculations.

In Table 2, there are presented values of Pearson correlation coefficient between soft commodities return series. Almost all of them are characterized by weak positive linear correlation. The strongest correlation is between coffee and sugar (0.283).

Table 2. Pearson correlation coefficients for logarithmic returns of soft commodities

Commodity	Coffee	Cocoa	Sugar	Cotton	Frozen concentrated orange juice	Rubber
Coffee	<b>1</b>					
Cocoa	<b>0.137</b>	<b>1</b>				
Sugar	<b>0.283</b>	<b>0.124</b>	<b>1</b>			
Cotton	<b>0.209</b>	<b>0.111</b>	<b>0.220</b>	<b>1</b>		
Frozen concentrated orange juice	<b>0.063</b>	<b>0.026</b>	<b>0.064</b>	<b>0.096</b>	<b>1</b>	
Rubber	<b>0.106</b>	<b>0.067</b>	<b>0.072</b>	<b>0.133</b>	0.048	<b>1</b>

Note: bold type denotes statistical significance at 0.05 level

Source: own calculations.

Additionally, to test whether soft commodities returns are independent, we plot the sample autocorrelation functions (ACFs) in Figure 2. Although some autocorrelations are statistically different from zero at the 5% level (e.g. at lag 3, 18, 20, 21, 23 and 24 for coffee; at lag 1, 25, 28, 31 and 34 for cocoa; etc.), there is no systematic pattern of autocorrelations. To investigate it further, we compute the Ljung-Box test statistic of the joint null hypothesis that all of the first 5, 10, 20, and 30 autocorrelations are zero. The results obtained are given in Table 3. They show that regardless the number of lags, we find statistically significant autocorrelation for cotton, frozen concentrated orange juice, and rubber returns. Only for cocoa returns, there is no evidence of significant autocorrelation. This finding is consistent with informationally efficient market, where price changes must be unpredictable if they are properly anticipated by market participants.

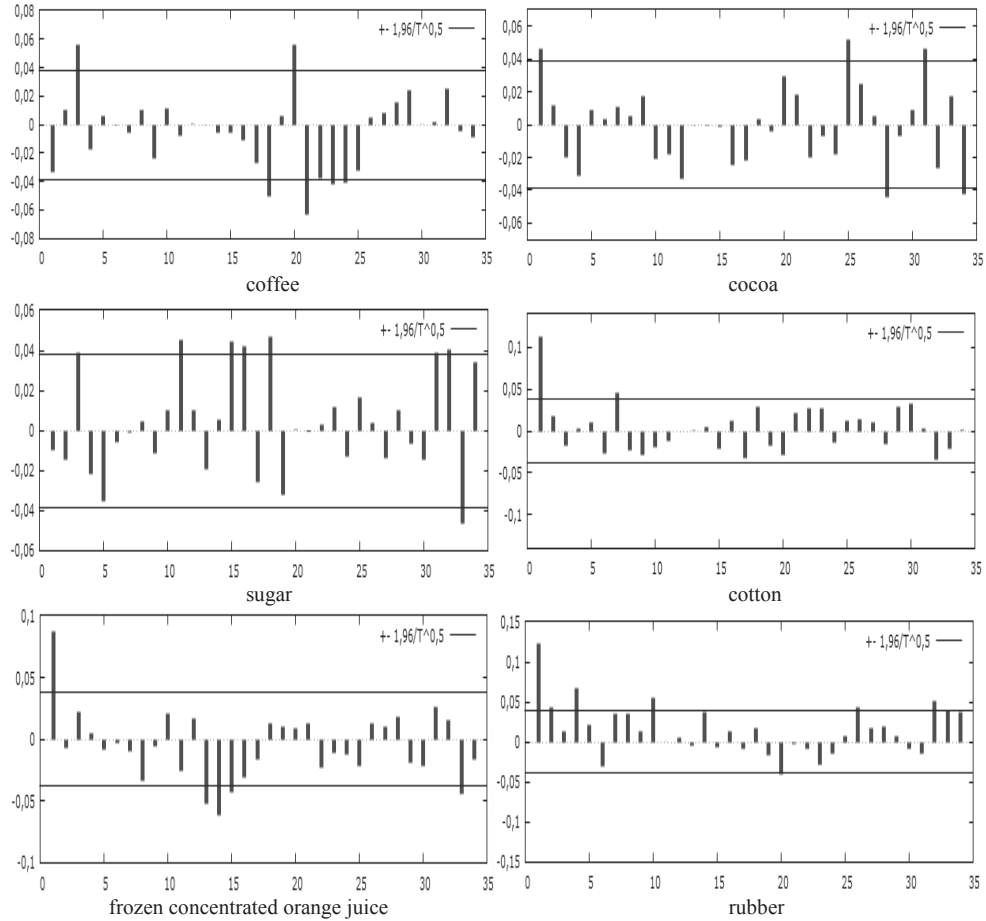


Fig.2. Sample autocorrelation functions (ACFs) for logarithmic returns of soft commodities

Source: own elaboration.

Table 3. Values of LB statistic for logarithmic returns of soft commodities

Commodity	Number of lags			
	5	10	20	30
Coffee	12.340*	14.564	32.304*	60.903*
Cocoa	9.794	12.088	21.215	38.321
Sugar	9.493	10.266	37.215*	40.347
Cotton	35.172*	46.966*	56.997*	69.494*
Frozen concentrated orange juice	21.612*	26.111*	55.013*	62.654*
Rubber	58.007*	75.274*	85.707*	96.172*

Note: \* denotes rejection of null hypothesis at 0.05 level

Source: own calculations.



## Concluding remarks

Soft commodities are often referred to as „tropics”, because these commodities are grown primarily in tropical or subtropical regions. In most cases they are grown in developing countries. The most important of soft commodities are: coffee, cocoa, sugar, cotton, orange juice, and rubber. They constitute a significant element of international trade balance, and are also important to Polish economy. According to the data released by the Central Statistical Office of Poland: “*Foreign Trade Turnover By Main Commodities 2016*” (Foreign..., 2017), in 2016 import of coffee to Poland reached the value of 459 924 000 USD and quantity of 148 394 tons, whereas import of cocoa (cocoa beans) – the value of 38 832 000 USD and quantity of 10 443 tons, cane or beet sugar, chemically pure sucrose in solid form – value of 125 353 000 USD (246 351 tons); cotton, carded or combed – 12 347 000 USD (6 689 tons); cotton yarn – 81 723 000 USD (30 218 tons), and woven fabrics of cotton - 302 123 000 USD (38 335 tons). Moreover, Polish investors have an opportunity to invest in foreign markets of soft commodities through commodity-linked exchange traded funds (ETFs), as some brokerage houses, for example Pekao or mBank, offer these instruments to their clients.

Using commodities as a financial investment opens a whole new universe of potential assets. However, investors should consider carefully certain aspects of each asset such as returns distribution, the correlation to other assets or autocorrelation. This paper was aimed at statistical analysis of soft commodities returns in the period January 2007 through December 2016. The results show that in the period under consideration soft commodities returns exhibited weak positive correlation. What is more, the returns were not normally distributed. In most cases they exhibited negative skewness and excess kurtosis. It is worth to remember that negative skewness and positive excess kurtosis are distribution properties that investors do not appreciate, because they imply more overall large returns (positive and negative) compared to the normal distribution. The larger negative returns are generally not compensated for by larger positive returns. We also find the statistically significant autocorrelation in soft commodities returns (with the exception of cocoa), which could be the evidence for weak-form informational inefficiency of their markets.

## Literature

- Box, G.E.P., Pierce, D.A. (1970). Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, 65, 1509-15.
- Chevallier, J., Ielpo, F. (2013). *The Economics of Commodity Markets*. John Wiley & Sons Inc., Chichester, West Sussex.
- Eller, R., Sagerer, Ch. (2008). An Overview of Commodity Sectors. *The Handbook of Commodity Investing*. John Wiley & Sons Inc., Chichester, West Sussex, 681-711.
- Erb, C.B., Harvey, C.R. (2006). The Strategic and Tactical Value of Commodity Futures. *Financial Analyst Journal*, 62(2), 69-97.
- Fabozzi, J., Füss, R., Kaiser, D.G. (2008). A Primer on Commodity Investing. *The Handbook of Commodity Investing*. John Wiley & Sons Inc., Chichester, West Sussex, 3-37.
- Foreign Trade Turnover By Main Commodities 2016 (2017). CSO, Warsaw, August.
- Fuertes, A.M., Miffre, J., Rallis, G. (2008). Tactical Allocation in Commodity Futures Markets: Combining Momentum and Term Structure Signals. *EDHEC Business School Working Paper*, Lille.
- Geman, H. (2007). *Commodities and Commodity Derivatives*. John Wiley & Sons Inc., Chichester, West Sussex.
- Górska, A., Krawiec, M. (2011). Zastosowanie narzędzi analizy technicznej w inwestowaniu na rynkach metali szlachetnych. *Metody Ilościowe w Badaniach Ekonomicznych*, 12(2), 148-157.

- Gujarati, D.N.. (2003). *Basic Econometrics*. Mc-Graw Hill, New York.
- Hamulczuk, M., Szajner, P. (2015). Ceny cukru w Polsce i ich determinanty. *Zagadnienia Ekonomiki Rolnej*, 4, 59-79.
- Irwin, S.H., Zulauf, C.R., Gerlow, M.E., Tinker, J.N. (1997). A Performance Comparison of a Technical Trading Systems with ARIMA Models for Soybean Complex Prices. *Advances in Investment Analysis and Portfolio Management*, 4, 193-203.
- Ljung, G.M., Box, G.P.E. (1978). On a Measure of Lack Fit in Time Series Models. *Biometrika*, 66, 66-72.
- Lukac, L.P., Brorsen, B.W., Irwin, S.H. (1988). A Test of Futures Market Disequilibrium Using Twelve Different Trading Systems. *Applied Economics*, 20(5), 623-639.
- Miffre, J., Rallis, G. (2007). Momentum Strategies in Commodity Futures Markets. *Journal of Banking and Finance*, 31(6), 1863-1886.
- Szakamary, A., Shen, Q., Dharma, S. (2010). Trend Following Trading Strategies in Commodity Futures: A Reexamination. *Journal of Banking and Finance*, 34, 409-426.
- [www.bloomberg.com](http://www.bloomberg.com).
- [www.fairtrade.org.pl](http://www.fairtrade.org.pl).
- [www.stat.gov.pl](http://www.stat.gov.pl).
- Zaremba, A. (2015). Strategie oparte na momentum i strukturze terminowej w warunkach finansjalizacji rynku surowców. *Studia Oeconomica Posnaniensia*, 3(2), 29-54.