Detecting Chaotic Behavior in Agricultural Exchange-traded Funds

Jo-Hui Chen¹, Tushigaa Batsukh², Yu-Fang Huang³

¹ Dept. of Finance, Chung Yuan Christian University, Chung Li, Taiwan, R.O.C. johui@cyu.edu.tw
² Dept. of Business Administration, Chung Yuan Christian University, Chung Li, Taiwan, R.O.C. Tushigeet_2010@yahoo.com
³ Fu Jen Catholic University, Taipei, Taiwan, R.O.C. 014447@mail.fju.edu.tw

Abstract: This study investigates the chaos effect of agricultural exchange-traded funds (ETFs) using Brock, Dechert, and Scheinkman test, rescaled range analysis, and correlation dimension analysis. The standardized residuals from generalized autoregressive conditional heteroskedasticity models are fitted into eight ETFs and examined in each case for evidence of chaotic behavior. This study also examines whether or not the ETFs are consistent with the chaos effect based on the underlying random data with trend-reinforcing series. Research results outline the financial insights for the agricultural ETF field of investment forecasting to eliminate trading emotions, while pursuing considerable profitable experience for investors.

Keywords: agricultural ETF, chaos effect, brock dechert scheinkman test, rescaled range analysis, correlation dimension analysis

1. Introduction

In recent years, agricultural commodities have become popular choices because of the increasing number of people seeking anti-inflationary investments and because of land scarcity. Inamura et al.[6] and Tang and Xiong[7] asserted that in evaluating the extent of global commodity markets to examine the sensitivity of portfolio investment, commodity prices are significantly correlated with one another and greatly influence stock prices.

Exchange-traded funds (ETFs), which trace the advent of financial futures, are one of the most successful financial innovations introduced in the early 1990s as an investment fund traded on the U.S. and Canadian exchanges. Since its establishment in 2004 by ETF providers, a commodity ETF has become a useful tool for constructing diversification in investment portfolios and has been hedged against economic downturns (Ilan et al.[8]). A new investment opportunity of agricultural ETFs is a topic that has recently received substantial attention from economists. Agricultural ETFs are one of the most effective means to diversify or stabilize a portfolio, hedge risk to invest in a broad agriculture fund, or focus on a specific commodity such as sugar, coffee, or grains. Using the returns of agricultural ETFs and futures relationship, Tse[9] posited that agricultural ETFs are significantly related to the stock market. Stock market sentiment is seen as a key factor driving price behaviors.

The possible existence of nonlinear dynamic causes problems in forecasting financial and economic time series when linear models are applied. The literature has documented that chaotic series refer to the linear models (e.g., time series or regressions) that cannot capture regularities in such a series (Kohzadi et al.[10]). Chaotic systems based on generally nonlinear response systems composed of erratic behavior, amplified events, and discontinuities. Benhabib and Nishimura[11], Grandmont[12], Day[13], Wolff[14], and Guegan and Leroux[15] have utilized the dynamical chaotic systems in
economics and finance. These experts suggested that the character of noise is a distraction in a real data set by using chaos theory in practice, and that perceived development of robust deconvolution techniques is required. LeBaron\cite{11} argued that in financial markets, nonlinear forecasting may vary from actual chaotic dynamics.

In related studies, DeCoste et al.\cite{12} applied correlation dimension to daily futures prices of sugar, silver, copper, and coffee. Yang and Brorsen\cite{13} employed correlation dimension and Brock, Dechert, and Scheinkman (BDS) test on futures markets to examine the price changes for soybean, corn, and wheat. Bacsil\cite{14} applied a linear cobweb model along with Lyapunov models to analyze the weekly price of potatoes. Based on a deterministic price model unrelated to any stochastic component, the results observed irregular oscillations for fluctuations in real price series. Chatrath et al.\cite{15} estimated the correlation dimension and BDS test on the daily prices of futures contracts for soybean, corn, cotton, and wheat. The results revealed strong evidence of nonlinear dependence, which is not consistent with a long-lasting chaotic structure. Chappell and Panagiotidis\cite{16} and Mahmoodzadeh et al.\cite{17} used the correlation dimension to examine the nonlinear dynamics of the stock index and exchange rates of Athens. However, the findings contradicted the chaotic behavior and indicated the presence of low chaotic dimensions.

Substantial literature has explored chaotic behavior in food commodity. By contrast, no study has looked into whether agricultural ETFs exhibit chaotic dynamics. Therefore, this study aims to examine chaos structure in the daily price of agricultural ETFs by employing three different methods (i.e., BDS test, rescaled range (R/S) analysis, and correlation dimension analysis). The BDS test for nonlinearity detection and rescaled range (R/S) analysis confirm the existence of chaotic behavior in price. Correlation dimension verifies the results of R/S analysis. Through the analysis of chaotic effect for agricultural ETFs, investors can effectively use to predict tools to form investment decisions on agricultural futures and instruments.

The rest of this paper is organized into four sections. Following the Introduction, Section II explains the data as well as BDS test, R/S analysis, and correlation dimension analysis. Section III discusses the empirical results, and Section IV presents the conclusions of the study.

2. Data and Methodology

This research uses daily closing prices of eight agricultural ETFs (i.e., ADZ, AGA, UAG, RJA, JJG, AGGP, COTN, and COW) obtained from Yahoo! Finance. The agricultural ETFs data set is tabulated in Table 1. ADZ and AGA ETFs are issued by PowerShares Capital Management LLC. These indices are examined by tracing the performance of several agricultural futures contracts (e.g., corn, wheat, soybean, and sugar) and financial futures for the returns from investing in U.S. Treasury Bills for three months. The UAG issued by the Union Bank of Switzerland is adopted to measure the from futures contracts for collateralized returns representing the agricultural sector. RJA represents the value of a basket of 20 agricultural commodity futures contracts and is a subindex of the Rogers International Commodity Index issued by Merrill Lynch, Pierce, Fenner & Smith, a subsidiary of the Bank of America. JJG is composed of three futures contracts on grains traded on the U.S. exchange and used by Barclays iPath. All these preceding ETFs are listed on the New York Stock Exchange. AGGP and COTN are listed on the London Stock Exchange and are issued by ETF Securities, which tracks the Dow Jones-AIG (DJ-AIG) Grains Subindex and DJ-AIG Cotton Subindex and pay a capitalized interest return. COW is listed on the Canadian Toronto Stock Exchange and is issued by the Claymore Investments, Inc., which traces the performance of the MFC Global Agriculture Index.
The data used in this study are those that fall between April 2008 and January 2011. As previously mentioned, this research applies three different approaches (i.e., BDS test, R/S analysis, and correlation dimension analysis) to explore chaotic behavior. Empirical research on finance and economics uses chaos theory to examine the nonlinear effects in deterministic systems transformed into a usable predictor of certain behavior (Djavanshir and Khorramshahgol[18]).

2.1. BDS statistics

The BDS test developed by Brock et al.[19][20] is superior to the alternative of deterministic chaos or stochastic nonlinear models. The null hypothesis states that certain variables have the independently and identically distribution (i.i.d.) against the unspecified alternative. The calculations of BDS test begin with a given time series \(\{x_t\}\) containing \(N\) observations, which should be the first difference of the natural logarithms of raw data in the time series to set up \(m\)-dimensional vectors called \(m\)-histories, \(x^m_i = (x_{i},x_{i+1},...,x_{i+m-1})\). These \(m\)-histories, collectively known as the embedding dimension, can be used to describe a correlation integral.

\[
C_{m,T}(\epsilon) = 2 \sum_{i=1}^{T-m+1} \sum_{t=1}^{T-m} I_\epsilon(x^m_i , x^m_u) / (T_m (T_m - 1)),
\]

where \(T_m = T - m + 1\), and \(I_\epsilon(x^m_i , x^m_u)\) is an indicator function of the event

\[
\left\|x^m_i - x^m_u\right\| = \max_{i=0,1,...,m-1} \left|x_{i+1} - x_{i+1}\right| \leq \epsilon.
\]

To calculate the pairs of the correlation integral lying within the tolerance distance \(\epsilon\) for the embedding dimension \(m\) and the time series of length \(T\), the BDS statistic is defined as

\[
BDS_{m,T}(\epsilon) = T^{\frac{1}{2}} \left[ C_{m,T}(\epsilon) - C_{1,T}(\epsilon) \right] / \sigma_{m,T}(\epsilon).
\]

2.2 R/S analysis

R/S analysis, which was invented by Hurst[21], is used mainly to estimate the repeated patterns in a time series of financial data. Given a time series, \(t\), and all observed values, \(u\), the estimation process is interpreted as follows:
\[ X_{t,N} = \sum_{n=1}^{N} \left( e_{u} - M_{N} \right), \]  

where \( X_{t,N} \) is the cumulative deviation over \( N \) periods, \( e_{u} \) stand for the influx in year \( u \), and \( M_{N} \) represents the average \( e_{u} \) over \( N \) periods. The range becomes the difference between maximum and minimum levels obtained from Equation (4).

\[ R = \max \left( X_{t,N} \right) - \min \left( X_{t,N} \right), \quad 1 \leq j \leq n, \]  

where \( R \) is the range of \( X \), \( \max(X) \) is the maximum value of \( X \), and \( \min(X) \) is the minimum value of \( X \).

This range is divided by Hurst to measure the standard deviation, \( S_{0} \), of the original observation to compare the different types of time series. The “rescaled range” should increase with time, and can be determined with the following formula:

\[ R/S = (an)H, \]  

where \( R/S \) is the rescaled range, \( N \) is the number of observations, \( a \) is a constant, and \( H \) is the Hurst exponent.

Based on the series in a consistency of random walk, \( H \) is equal to 0.5. However, the observations are dependent when \( H \) is not 0.5. Each observation theoretically carries a long-term “memory” process of all the events that precede and last forever. Hurst statistic is interconnected through a long stream of events. To estimate the Hurst exponent, the natural logarithm is rewritten as follows:

\[ \ln \left( \frac{R(n)}{S(n)} \right) = H \ln(an), \quad H = \ln \left( \frac{R}{S} \right) / \ln(an). \]  

If \( H \) is 0.5, then the underlying time series has no correlation; 0 ≤ \( H \) < 0.5. In this case, the time series increments represent a related coefficient smaller than 0, and therefore, are negatively correlated with one another. If 0.5 < \( H \) < 1, then the time series increments reveal a related coefficient which is larger than 0, indicating a positive correlation.

### 2.3 Correlation dimension analysis

The correlation dimension analysis was devised by Grassberger and Procaccia[22]. This technique measures the random time series from the chaotic one. In this study, the correlation dimension algorithm is realized through four procedures explained below.

1. Residual analysis is performed to remove autocorrelation and conditional heteroskedasticity.
2. The information embedded dimension \( n \) estimates only the process of embedding the dimension value, which is not greater than the length of the original sequence, \( V \), but is not less than 1. 1-history is expressed as \( x_{t}^{1} = x_{t} \), 2-history is \( x_{t}^{2} = \left( x_{t-1} , x_{t} \right) \), and \( n \)-history is \( x_{t}^{n} = \left( x_{t-n+1}, \ldots, x_{t} \right) \), which signifies a point in \( n \)-dimensional space (\( n \) is the “embedding dimensions”).
3. The correlation integral is determined. Grassberger and Procaccia[22] used the integration process as

\[ C_{n}(\varepsilon) = \lim_{T \to \infty} \frac{\# \{ (t,s) : 0 < t, s, < T : \|x_{t}^{n} - x_{s}^{n}\| < \varepsilon \} \cdot T^{2}}{T^{2}}. \]  

where \( \# \) is the number of elements in the set, and the measurement of \( (x_{t}^{n}, x_{s}^{n}) \) distance implies that

\[ \max_{i=0,\ldots,n-1} \left\{ \left| x_{t-i} - x_{s-i} \right| \right\} < \varepsilon. \]
(4) The minimum value of $\varepsilon$ denotes the slope for the calculation of demand, and $\|l\|$ is calculated with the following formula:

$$v_{n} = \lim_{\varepsilon \to 0} \log C_{n}(\varepsilon)/\log \varepsilon .$$

When $n$ increases, the correlation dimension $\left( v_{n} \right)$ increases continuously, revealing that the data do not have a fractal structure without a compliance characteristic of chaotic systems (Chappell and Panagiotidis[16]).

3. Empirical Results

3.1 BDS test results

Price series and futures prices contain mostly unit roots, and are nonstationary (Moore and Cullen[23]; Beck[24]). Hence, stationarity should be tested first before taking the nonlinearity test. In this study, the stationarity test is performed using the augmented Dickey–Fuller unit root test proposed by Dickey and Fuller[25] to examine the presence of unit root in the series (difference stationarity). The agricultural ETF series is filtered with autoregressive integrated moving average (ARIMA) model and autoregressive conditional heteroskedasticity (ARCH) models to delete linear dependence and to eliminate the conditional heteroskedasticity, respectively. The generalized autoregressive conditional heteroskedasticity (GARCH) model is also used to reduce conditional heteroskedasticity. The standardized residuals of GARCH are likewise employed.

After the ETF series is filtered with GARCH model, BDS test statistics are estimated for ARIMA and GARCH filtered residuals as well as for the series compared before filtering. The null hypothesis of the BDS test based on underlying data following i.i.d is rejected. This condition suggests that the data are not i.i.d, because a nonlinear dependence is observed in the time series after the first differences of the natural logarithm are obtained. Table 2 compiles the BDS test results and shows that the JG series is insignificant when the m embedding dimension is equal to 2. When the $\sigma$ standard deviation ranges from 1 to 2, the BDS statistic results are insignificant. The results of the ARIMA filtered residuals indicate that AGGP, COW, ADZ, and COTN all have 1% significance level. The estimation of GARCH filtered residuals shows that the results of JG, GOTN, and RJA are insignificant, whereas those of AGGP series are significant. However, the remaining four series are determined to be insignificant when $\sigma$ standard deviation increases to 1, 1.5, and 2.

3.2 R/S analysis results

An H value of 0.5 designates a random series. An H value of more than 0.5 indicates a trend-reinforcing series, whereas large H values signify strong trends. Table 3 shows that the original values of H are mostly lower than 0.5. After the data are disordered, all values of H are determined to be significantly asymptotic to 0.5. The results reveal that the underlying data have a persistent and trend-reinforcing series. This finding illustrates that the upward or downward trend of the selected agricultural ETFs in the past would continue to be either positive or negative in the subsequent periods.

3.3 Correlation dimension analysis results

The results of the correlation dimension test are presented in Figure 1. The embedding and correlation dimensions are
located on the horizontal and vertical axes, respectively. The eight-series correlation dimension evidently becomes convergent when the embedding dimension is at least greater than 4. This condition implies that the primary data of the agricultural ETFs may be chaotic.

Table 2 Empirical result of BDS test
<table>
<thead>
<tr>
<th>j</th>
<th>Embedding dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before filtering</td>
</tr>
<tr>
<td>0.5</td>
<td>0.003*** 0.004*** 0.002*** 0.001*** 0.001***</td>
</tr>
<tr>
<td>1.0</td>
<td>0.005*** 0.015*** 0.018*** 0.015*** 0.018***</td>
</tr>
<tr>
<td>1.5</td>
<td>0.004 0.017*** 0.039*** 0.049*** 0.064***</td>
</tr>
<tr>
<td>2.0</td>
<td>0.002 0.010*** 0.021*** 0.038*** 0.052***</td>
</tr>
</tbody>
</table>

**Note:** * Significance level of $\alpha = 10\%$; ** Significance level $\alpha = 5\%$; *** Significance level $\alpha = 1\%$.
Table 3 Empirical result of rescales range analysis test

<table>
<thead>
<tr>
<th>Variable names</th>
<th>Before filtering</th>
<th>ARIMA filtered residuals</th>
<th>GARCH filtered residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original Hurst index</td>
<td>Hurst index after scrambled</td>
<td>Original Hurst index</td>
</tr>
<tr>
<td>ADZ</td>
<td>-0.0067</td>
<td>0.4772*</td>
<td>0.0017</td>
</tr>
<tr>
<td>AGA</td>
<td>0.0023</td>
<td>0.5227*</td>
<td>0.0022</td>
</tr>
<tr>
<td>AGGP</td>
<td>-0.0049</td>
<td>0.4175</td>
<td>0.0056</td>
</tr>
<tr>
<td>COTN</td>
<td>0.4205</td>
<td>0.9911</td>
<td>-0.0064</td>
</tr>
<tr>
<td>COW</td>
<td>-0.0144</td>
<td>0.5017*</td>
<td>-0.0118</td>
</tr>
<tr>
<td>JJJG</td>
<td>-0.0021</td>
<td>0.5389*</td>
<td>0.0006</td>
</tr>
<tr>
<td>RJA</td>
<td>0.5363</td>
<td>0.9957</td>
<td>0.0008</td>
</tr>
<tr>
<td>UAG</td>
<td>-0.0007</td>
<td>0.5017*</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

Note: * denotes $H = 0.5$; ** denotes $0 \leq H < 0.5$; *** denotes $0.5 < H < 1$. 

Figure 1 Correlation dimension results of GARCH filtered series
4. Conclusion

This study applies BDS test, R/S analysis, and correlation dimension analysis, as well as with ARIMA and GARCH residual tests to examine the chaotic behavior of eight agricultural ETFs. The BDS results demonstrate that all agricultural ETFs are significant for the data before being filtered by ARIMA and GARCH models. The ARIMA residuals are also identified to be significant. If the series of the GARCH filtered residuals is verified to be non i.i.d., then the presence of nonlinear dependence may be observed. However, deterministic chaos could probably exist if the null hypothesis is not accepted. The GARCH filtered results are merely significant for the series of AGGP ETF. The BDS results contradict the expectation of this study, but are consistent with previous research that BDS may not be strong enough to examine weak nonlinearity (Wang et al.26). For R/S analysis, the underlying data are determined to have a random and trend-reinforcing series. This observation specifies that the returns of agricultural ETFs are fractal and chaotic. The correlation dimension values are calculated against the corresponding embedding dimension, and all time series converge as expected. This finding demonstrates that the underlying data of agricultural ETFs are consistent with chaos and that conventional linear models are not appropriate for use in analyzing such ETFs.

In sum, this study verifies that a nonlinear and possibly chaotic structure may exist in the price series of all analyzed agricultural ETFs. Among all methods applied in this study, the R/S analysis provides the strongest results and verifies the uncertainty in the overall empirical results. Therefore, the findings presented in this paper can be adopted by future researchers or investors to forecast or form investment decisions.

References

16. Chappell D and Panagiotidis T, Using the Correlation Dimension to Detect Nonlinear Dynamics: Evidence from the Athens