The sharp changes in the global food commodity prices in the recent years have raised concerns to the governments in developing countries as the low income groups became more vulnerable to higher inflation rates. Increasing volatility in the agriculture commodity prices creates uncertainties to farmers to meet the rising demand for agricultural food commodities, and to consumers to manage their future spending plans.

Analysts attribute the rising volatility in the food commodity prices to a number of factors, among them the speculations in future commodity markets (FAO 2008); crude oil price changes and its ramification on the bio-fuel commodity markets (Institute for Agriculture and Trade Policy 2008); Jeffrey Frankel (2008) attributes the soaring prices in food commodities to structural change in the global demand for food items, mainly due to the high and rapid economic growth in the countries like China and India. Whatever would be the prime cause behind the soaring food commodity prices, it is important to point out that the volatility modeling can help capturing the empirical regularities that characterize the commodity markets. While the literature on the volatility of food commodity markets in general is scarce, compared to the literature on the financial asset markets, a number of authors investigated the volatility in food commodity markets from the perspective of the spillover effect of the crude oil price (Babula and Somwaru 1992; Uri 1996; Du et al. 2009). Broadly speaking, the literature on the volatility forecast in commodity markets includes two main approaches, the implied volatility models which are based on the option pricing formulas, and the conditional volatility models of the time series.
analysis. An option is a contract that allows the holder, without requiring, to sell (put option) or buy (call option) underlying commodity at a pre-specified price (strike or exercise price). The more volatile the price of the underlying commodity, the higher the option’s price. Given that the option market is efficient, the option price reflects the future expected volatility. More specifically, option prices are functions of four observable variables (i.e., the price of the underlying asset, the exercise price, the time to the maturity of the option, and the risk-free rate of interest), and one non-observable variable which is the expected volatility of the underlying commodity price. Since the option price is observable, and it is a monotonically increasing function of the expected volatility, then via the option pricing formula it is possible to derive the expected volatility for the remaining period of the option maturity. Such expected volatility which is based on the option pricing formula is usually known as the implied volatility.

However, it should be noted that the implied volatility approach of the expected volatility has a number of drawbacks. Among them, the standard option pricing formula included in Black and Scholes (1973) applies only to the European type option, but it has no closed-form solution for the more popular American and exotic options. In addition, as noted by Kroner et al. (1993), the volatility forecast based on the implied volatility approach may be more appropriate for the short-term forecast, but it may not yield a reliable long-term forecast since usually trading is thin in the options that are far away from their maturity dates.

As a result, in this paper we employed the time series modeling approach in forecasting the conditional volatility in the food commodity market. The conditional volatility models include the ARCH/GARCH models developed by Engle (1982); and the stochastic volatility (SV) models, proposed by Taylor (1994). While the conditional volatility models of the ARCH/GARCH-type define volatility as a deterministic function of the past innovations, the SV models treat volatility as a stochastic process. The ARCH/GARCH-type models are relatively easy to estimate compared to the SV models which are directly connected to the diffusion process, and thus involve the volatility process that does not depend on the observable variables, and are therefore relatively more difficult to estimate (Shephard 2005).

A question which needs to be addressed is, why do we need to investigate volatility in the food commodity markets? In the light of the option pricing formula of Black and Scholes discussed above, robust estimates of volatility of food prices enhance a better option pricing mechanism in future commodity markets. Moreover, volatility estimates allow the investigation of empirical regularities that characterize the food commodity markets. Among the empirical regularities that characterize asset prices, there are the fat-tailedness and volatility persistence. It is well documented (Bollerslev et al., 2003) that the fat-tailedness in asset markets is intimately related to the so-called volatility clustering, which describes the phenomena that large changes in asset prices, in either sign, tend to be followed by large changes, and small changes are followed by small changes, reflecting market irregularities. Thus, volatility modeling can reveal the market imperfection in the global food commodity markets.

Bollerslev et al. (2003) indicated that the normality assumption is at odds when price changes exhibit the fat-tailedness (leptokurtosis behavior). It has been evidenced recently by a number of authors (Brooks and Persand (2003), Vilasuso (2002), and Hansen and Launda (2003), the standard GARCH models which use the normality assumption has an inferior forecasting performance compared to the models that reflect skewness and kurtosis in innovations.

The paper is divided into four sections. Section two includes the methodology of the research. Section three deals with the estimation procedure and discussion of the results. In the final section, we conclude the research findings.

**METHODOLOGY**

Given that \( p_t \) is the commodity price at time \( t \), and \( I_{t-1} \) is the information set at time \( t - 1 \), then the standard GARCH(1,1) model specified on normal distributed and Student \( t \)-distributed error terms is defined as:

\[
\ln \left( \frac{p_t}{p_{t-1}} \right) = y_t = \mu + \epsilon_t
\]

or

\[
f(\epsilon_t | I_{t-1}) \sim N(0, \sigma_t^2)
\]

where \( \eta \) is degrees of freedom, and

\[
\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\]

Given an initial value for \( \sigma_0^2 \) (the conditional volatility), the estimated values for \( \omega, \alpha \) and \( \beta \) in equation (3) can be used for estimating expected volatility at

---

any given horizon time. Using equation (3), the expected volatility can be set (see Engle and Bollerslev 1986, equation 22) as:

\[
E(\sigma^2_{t+k} \mid I_t) = \begin{cases} 
  w + \alpha \varepsilon_t + \beta \sigma^2_{t-1} & \text{if } k = 1 \\
  w + (\alpha + \beta)E(h_{t,k-1} \mid I_t) & \text{if } k \geq 2 
\end{cases}
\]  

(4)

Alternatively, using the recursive substitution of equation (3) we get

\[
E(\sigma^2_{t+k} \mid I_t) = \begin{cases} 
  w + \alpha \varepsilon_t + \beta \sigma^2_{t-1} & \text{if } k = 1 \\
  w[1 + (\alpha + \beta) + \ldots (\alpha + \beta)^{k-2}] + (\alpha + \beta)^{k-1}(w + \alpha \varepsilon_t + \beta \sigma_{t-1}^2) & \text{if } k \geq 2 
\end{cases}
\]  

(5)

Equations (4) and (5) yield the forecast of conditional volatility at the horizons 1, 2, ..., \(k\). Bollerslev et al. (2003), discusses the usefulness of the GARCH models in the short-run volatility forecast.

It is well documented that the standard GARCH specification as stated in equation (1) fail to fully account for the leptokurtosis of the high frequency time series when they are assumed to follow the normal distribution. Bollerslev et al. (2003) indicate the ARCH models with conditional normal errors; the result in a leptokurtic unconditional distribution. However, the degree of the leptokurtosis induced by the time-varying conditional variance often does not capture all of the leptokurtosis present in the high frequency speculative price data. To circumvent this problem, Bollerslev et al. (2003) suggest the use of the Student t-distribution with the degrees of freedom greater than two.

When the residual errors in (3) distributed Student t-distribution the density function in equation (3) can be specified as:

\[
f(\varepsilon \mid \eta) = \frac{\Gamma(\eta + 1)/2}{\sqrt{\pi} \Gamma(\eta/2)\left(\eta + \varepsilon^2\right)^{\eta/2}} 
\]  

for \(-\infty < \varepsilon < \infty\)  

(6)

where \(\Gamma(.)\) denotes the gamma function, and \(\eta\) is the degrees of freedom. Now, we have two competing models, (equations 1 and 6), for the expected conditional volatility specification in equation (5).

Given there is no common single conventional model selection criteria, to assess the goodness-of-fit for the two models, we employed the predictive power performance criteria, and four other criterias including the log-likelihood function and the Akaike information criteria (AIC), the Schwarz criteria (SC), and the Hannan-Quinn (HQ), as indicated below:

\[
AIC = -2(\log(T)/T + 2k/T)
\]

\[
SC = -2(l/T) + k \log(T)/T
\]

\[
HQ = -2(l/T) + 2k \log(\log(T))/T
\]

where \(k\) and \((T)\) are respectively the number of parameters and the sample size, and \(l\) is the lag length. The model that minimizes the above information criteria is considered the best fit, given that the model also yield the highest log likelihood value.

**VOLATILITY PERSISTENCE**

The ARFIMA\((p, d, q)\) process

\[
\psi(L)(1-L)^d (y_t - \mu) = \theta(L) \varepsilon_t
\]  

(7)

where

\[
\psi(L) = \sum_{j=0}^{p} \phi_j L^j, \quad \theta(L) = \sum_{j=1}^{q} \theta_j L^j
\]

\[
(1-L)^d = \sum_{j=0}^{\infty} \binom{d}{j} \left(-1\right)^j L^j = 1 - dL + \frac{d(d-1)L^2}{2} - ..... 
\]

and \(L\) is the lag operator, \(d\) is the fractional differencing parameter, all roots of \(\psi(L)\) and \(\theta(L)\) assumed to lie outside the unit circle, and \(\varepsilon_t\) is the white noise.

GARCH\((p, q)\) models are often used for modeling the volatility persistence, which has the features of the relatively fast decaying persistence. However, in some cases the volatility shows a very long temporal dependence, i.e., the autocorrelation function decays very slowly. This motivates considering the Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity (FIGARCH) process (Baillie 1996) defined as:

\[
\varphi(L)(1-L)^d \varepsilon_t^2 = w + \beta(L)\nu_t
\]  

(8)

where \(\varphi(L)\) and \(\beta(L)\) are respectively the AR\((p)\) and MA\((q)\) vector coefficients and \(\nu_t = \varepsilon_t^2 - \sigma_t^2\).

Following Baillie et al. (1996), Bollerslev and Mikkelsen (1996), Granger and Ding (1996), the parameters in the ARFIMA\((p, d, q)\) and FIGARCH\((p, d, q)\) models in (7) and (8) estimated using the quasi-maximum likelihood (QMLE) method. In the ARFIMA models, the short-run behavior of the data series is represented by the conventional ARMA parameters, while the long-run dependence can be captured by the fractional differencing parameter, \(d\). A similar result also applies when modeling the conditional variance, as in equation (8). While for the covariance stationary GARCH\((p, q)\) model, a shock to the forecast of the
future conditional variance dies out at an exponential rate, for the FIGARCH(p,d,q) model, the effect of a shock to the future conditional variance decay at low hyperbolic rate. As a result, the fractional differencing parameter, $d$, in the equations (7) and (8) can be regarded the decay rate of a shock to the conditional variance (Bollerslev and Mikkelsen 1996).

In general, allowing for the values of $d$ in the range between zero and unity (or, $0 < d < 1$) adds a flexibility that plays an important role in modeling of the long-run dependence in time series.$^3$

Bollerslev (1996), indicates that if $d = 0$, the series is covariance stationary and possesses a short memory process, whereas in the case of $d = 1$, the series is non-stationary. However, in the case of $0 < d < 0.5$, the series, even though covariance stationary, its auto-covariance decays much more slowly than the ARMA process. If $d$ is $0.5 < d < 1$, the series is no longer covariance stationary, but still mean reverting with the effect of a shock persist for a long period of time, and in that case, the process is said to have a long memory. Given a discrete time series, $y_t$, with autocorrelation function, $\rho_j$, at lag $j$, Mcleod and Hipel (1978) define long memory as a process:

$$\sum_{j=-n}^{n} \rho_j | \quad n \to \infty$$

characterized as non-finite. In the non-stationary and in the long memory process, the shock $e_t$ at time $t$, continues to influence the future $y_{t+k}$ for a longer horizon, $k$, than would be the case for the standard stationary ARMA process. While there are varieties of ways to estimate the parameters of (3) and (4), in this paper we employed the maximum likelihood estimator.

**EMPIRICAL RESULTS**

**Estimation of parameters**

The analysis in this research is based on monthly data for six food commodity prices, during the sample period from October 1984 to September 2009. The food commodities include wheat, rice, sugar, beef, coffee, and groundnut. All price series were collected from the Index Mundi website, which in turn was extracted from the the IMF, Primary Commodity Price Tables$^4$. We employed the maximum likelihood estimation method to estimate the parameters in equations (2)–(5). The graphical exposition of price changes (see the Figures 1–6) indicates the evidence of volatility clustering, which is the phenomenon that large changes in asset prices, in either sign, tend to be followed by large changes, and small changes are followed by small changes. Table 1 presents the estimation results of the parameters in equation (3) under both the normal and the $t$-distribution errors. Results of GARCH(1,1) parameters show the evidence of stationarity of the conditional volatility; whereas the sample autocorrelation statistic indicated by the squared values of Ljung-Box, Q(5) suggests that the conditional homoskedasticity can be rejected for all six commodities. Also the results of the LM statistics for ARCH(5) errors confirm the significance of the ARCH effects in the data. The log likelihood and the information criteria test results overwhelmingly support the $t$-distribution specification of the innovations in the AR(1) model in equation (1). This is consistent with the existing literature on asset markets, which indicates the evidences of the conditional leptokurtosis in the high and medium frequency data analysis (Bai

![Figure 1. Groundnut price change](http://www.indexmundi.com/)

---

$^3$See Diebold and Rudebuch (1989); Cunado et al (2005); and Ding and Granger (1996) for a detailed discussion about the importance of allowing for non-integer values of integration when modeling long-run dependence in the conditional mean of time series data.

$^4$[http://www.indexmundi.com/](http://www.indexmundi.com/)
Figure 2. Sugar price change

Figure 3. Rice price change

Figure 4. Coffee price change

Figure 5. Beef price change
To investigate further the robustness of the $t$-distribution model, we conducted, using the out-of-sample forecast analysis, the predictive power of the two models. Diebold and Mariano (1995) (DM) test was employed to compare the accuracy of forecast results. When comparing the forecasts from two competing models; model A, and model B, an important question that needs to be taken into account.

Table 1. Estimation of parameters

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Rice</th>
<th>Beef</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>normal</td>
<td>$t$-dist.</td>
<td>normal</td>
</tr>
<tr>
<td>$w$ (p-value)</td>
<td>0.068 (0.24)</td>
<td>0.01 (0.05)</td>
<td>0.58 (0.12)</td>
</tr>
<tr>
<td>$\alpha$ (p-value)</td>
<td>0.00 (0.94)</td>
<td>0.74 (0.00)</td>
<td>0.00 (0.93)</td>
</tr>
<tr>
<td>$\beta$ (p-value)</td>
<td>0.63 (0.00)</td>
<td>0.10 (0.01)</td>
<td>0.003 (0.95)</td>
</tr>
<tr>
<td>$Q^2$ (5) (p-value)</td>
<td>7.99 (0.15)</td>
<td>286 (0.00)</td>
<td>73.8 (0.00)</td>
</tr>
<tr>
<td>LM (5)</td>
<td>144*</td>
<td>6.83*</td>
<td>96.1*</td>
</tr>
<tr>
<td>Log-likelihood function</td>
<td>177</td>
<td>13 746</td>
<td>–798</td>
</tr>
<tr>
<td>AIC</td>
<td>0.18E-1</td>
<td>0.27E-41</td>
<td>13.12</td>
</tr>
<tr>
<td>SC</td>
<td>0.19E-1</td>
<td>0.28E-41</td>
<td>13.62</td>
</tr>
<tr>
<td>HQ</td>
<td>0.18E-1</td>
<td>0.27E-41</td>
<td>13.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Groundnut</th>
<th>Sugar</th>
<th>Coffee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>normal</td>
<td>$t$-dist.</td>
<td>normal</td>
</tr>
<tr>
<td>$w$ (p-value)</td>
<td>0.00 (0.94)</td>
<td>0.01 (0.00)</td>
<td>0.068 (0.00)</td>
</tr>
<tr>
<td>$\alpha$ (p-value)</td>
<td>0.28 (0.24)</td>
<td>0.00 (0.81)</td>
<td>0.12 (0.22)</td>
</tr>
<tr>
<td>$\beta$ (p-value)</td>
<td>0.59 (0.02)</td>
<td>0.25 (0.00)</td>
<td>0.16 (0.01)</td>
</tr>
<tr>
<td>$Q^2$ (5) (p-value)</td>
<td>103 (0.00)</td>
<td>76.6 (0.00)</td>
<td>15.72 (0.01)</td>
</tr>
<tr>
<td>LM (5)</td>
<td>43.6*</td>
<td>40.62*</td>
<td>5.25*</td>
</tr>
<tr>
<td>Log-likelihood function</td>
<td>844</td>
<td>1 2939</td>
<td>907</td>
</tr>
<tr>
<td>AIC</td>
<td>0.19E-3</td>
<td>0.63E-39</td>
<td>0.13E-3</td>
</tr>
<tr>
<td>SC</td>
<td>0.20E-3</td>
<td>0.66E-39</td>
<td>0.14E-3</td>
</tr>
<tr>
<td>HQ</td>
<td>0.20E-3</td>
<td>0.64E-39</td>
<td>0.13E-3</td>
</tr>
</tbody>
</table>

Note: Estimated values of parameters rounded into two decimals; terms in parenthesis are $p$-values

*significant at 5% significance level
is, whether the prediction of model A is significantly more accurate, in terms of a loss function, than the prediction of model B. The Diebold and Mariano test aims to test the null hypothesis of the equality of forecast accuracy against the alternative of the different forecasts across models. Table 2 reports the predictive power of the two models (the normal distribution and the $t$-distribution innovations) using the Root Mean Squared Error (RMSE) of forecast values of the conditional volatility.

Results in Table 3 report the FIGARCH(1, $d$, 1) results, and indicate a strong evidence of the stationary short memory process for the four of the six commodities, as only beef and coffee exhibit the long memory behavior (covariance non-stationary, but mean reverting). This result implies that for the food commodities exhibiting the short memory process, a shock is not likely to persist for a long period.

### CONCLUDING REMARKS

The paper employs two competing models, including the thin tailed the normal distribution and the fat-tailed Student $t$-distribution models, to explore the volatility in the global food commodity prices of wheat, rice, beef, groundnut, sugar, and coffee. The sample period in the study includes monthly data covering the period from October 1984 to September 2009. Using the predictive power of the volatility forecast and other goodness of fit measures, the performance of each model was assessed. The analysis in the paper indicates that the $t$-distribution model outperforms the normal distribution model, revealing the evidence of leptokurtosis in the volatility of food commodity prices. This result implies that if such leptokurtic behavior is not taken into account when estimating the conditional volatility, the standard option pricing formula of Black and Scholes, which depends on the expected volatility parameter, could lead into unreliable results when pricing the future option contracts in the commodity markets. The paper also shows that the volatility of the future

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Wheat</th>
<th>Rice</th>
<th>Beef</th>
<th>Groundnut</th>
<th>Sugar</th>
<th>Coffee</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$ (standard error)</td>
<td>0.48 (8.04)</td>
<td>0.44 (7.2)</td>
<td>0.57* (10.2)</td>
<td>0.42 (6.0)</td>
<td>0.47 (6.6)</td>
<td>0.56* (10.1)</td>
</tr>
<tr>
<td>$\phi_1$ (standard error)</td>
<td>$-0.49$ ($-7.6$)</td>
<td>$-0.45$ ($-7.0$)</td>
<td>$-0.58$ ($-10.3$)</td>
<td>$-0.17$ ($-2.1$)</td>
<td>$-0.16$ ($-1.9$)</td>
<td>$-0.58$ ($-10.4$)</td>
</tr>
<tr>
<td>$\theta_1$ (standard error)</td>
<td>0.01 (0.01)</td>
<td>0.01 (0.002)</td>
<td>0.4 (0.01)</td>
<td>0.01 (0.01)</td>
<td>0.02 (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Log-likelihood function</td>
<td>201</td>
<td>$-676$</td>
<td>475</td>
<td>837</td>
<td>886</td>
<td>$-205$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Wheat</th>
<th>Rice</th>
<th>Beef</th>
<th>Groundnut</th>
<th>Sugar</th>
<th>Coffee</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$ (standard error)</td>
<td>0.75* (8.6)</td>
<td>0.55* (9.9)</td>
<td>0.59* (9.4)</td>
<td>0.41 (5.01)</td>
<td>0.48 (7.1)</td>
<td>0.63* (11.87)</td>
</tr>
<tr>
<td>$\phi_1$ (standard error)</td>
<td>$-0.39$ ($-4.36$)</td>
<td>$-0.59$ ($-10.4$)</td>
<td>$-0.44$ ($-6.3$)</td>
<td>$-0.05$ ($-2.1$)</td>
<td>$-0.19$ ($-2.4$)</td>
<td>$-0.44$ ($-7.7$)</td>
</tr>
<tr>
<td>$\theta_1$ (standard error)</td>
<td>0.82 (0.62)</td>
<td>9.7 (0.47)</td>
<td>$-14.0$ (1.79)</td>
<td>8.28 (0.59)</td>
<td>$-6.2$ ($-2.25$)</td>
<td>52.5 (8.47)</td>
</tr>
<tr>
<td>Log-likelihood function</td>
<td>97.1</td>
<td>$-683$</td>
<td>438</td>
<td>850</td>
<td>884</td>
<td>$-152$</td>
</tr>
</tbody>
</table>

*The loss functions are based on three days ahead forecast errors. Under the null hypothesis of equal forecast accuracy, DM is asymptotically normally distributed. *The loss functions are based on three days ahead forecast errors. Under the null hypothesis of equal forecast accuracy, DM is asymptotically normally distributed.
commodity prices is mean reverting, and exhibits the short memory behavior for all commodities, except for beef and coffee which show intermediate memory behavior. The short term persistence implies that the effect of shocks on the future conditional variance persist only for short periods.

REFERENCES


Arrived on 9th March 2010

Contact address:

Ibrahim A. Onour, Arab Planning Institute, Kuwait, P.O. Box 5834, Safat 13059 State of Kuwait
Bruno S. Sergi, University of Messina, Via T. Cannizzaro, 9, 1-98122 Messina, Italy
e-mails: onour@api.org.kw, ibonour@hotmail.com