FUTURES PRICE VOLATILITY: MODELING NON-CONSTANT VARIANCE

PAUL L. FACKLER

1986

Futures trading

Paul L. Fackler is a doctoral candidate and research assistant in the Department of Agricultural & Applied Economics at the University of Minnesota.

AAEA paper, 1986

FUTURES PRICE VOLATILITY: MODELING NON-CONSTANT VARIANCE

PAUL L. FACKLER

ABSTRACT

A model is developed in which price variance is treated as a function of a deterministic seasonal component and a stochastic component which is conditional on past price changes. When applied to the corn futures market, both components are found to be significant. Implications for option pricing models are discussed.

FUTURES PRICE VOLATILITY: MODELING NON-CONSTANT VARIANCE

The stochastic nature of speculative prices has been the object of considerable interest and inquiry. Numerous probability models have been suggested and empirically investigated. For the most part, however, these models have concentrated on the properties of the marginal price distribution, with changes in prices being taken to be independent realizations from an assumed family of distributions.

The opening of markets in options has focused attention on the issue of non-constancy of price distributions when conditioned on currently available information. In most option pricing models, the volatility of price movements is a key factor which must be estimated by market participants. Empirical evidence suggests that this volatility changes over time. Futures prices for seasonally produced commodities, for example, exhibit large seasonal differences in volatility (Anderson, Gordon). It has also been observed (Mandelbrot, Engle) that uncertainty about the future is a cyclic phenomenon, with prices being relatively stable in some periods and quite volatile in others.

In this paper a model of futures price volatility will be developed that incorporates a deterministic seasonal component and a stochastic component which conditions current variance on past price changes. In the next section a brief review of the past work on which this model is based is given. Following this, the model of futures price movements is developed and is then used to examine price behavior in the corn futures market during the period 1974-1982. The concluding section contains comments on the implications of this study for the pricing of options.

Previous Modeling

It has long been noted that speculative prices exhibit both periods of relative calm and of large, often wild, movement. In particular, large movements in price tend to be followed by large movements, though not necessarily in the same direction. In part this can be explained by lumpiness in the arrival of information relevant to the future level of price. In agricultural commodity markets the well known "weather markets" of the summer months, in which prices can be extremely volatile, can be attributed to the continuous arrival of new information about the upcoming harvest. Anderson found, for example, that the monthly variance of futures price changes for corn, wheat, and soybeans, tended to be highest in July and lowest in February.

For commodities that are not produced seasonally and for other types of traded assets, the seasonal explanation is not adequate to explain the kinds of cyclic behavior that price volatility exhibits. The causes of such uncertainty are many, complex, and difficult to quantify. The result, however, is that the conditional uncertainty about the near future is closely linked to the uncertainty that existed in the immediate past. Therefore, without knowing the causes of the uncertainty about the future, it is nonetheless possible to build plausible models of that uncertainty.

Engle has provided such a model, which he calls the ARCH model, an acronym for Autoregressive Conditional Heteroskedasticity. A generalized version, called GARCH, has been developed by Bollerslev. A discrete stochastic process, e_t, which fits the GARCH framework is one that is distributed

$$e_t \sim N(0,h_t),$$

where

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i e^2_{t-i} + \sum_{j=1}^{p} \beta_j h_{t-j}$$

and

$$\alpha_i \geq 0,\ \beta_j \geq 0, \quad \text{for all } i \text{ and } j.$$

Such a process can be denoted as GARCH(p,q). The ARCH model takes the β_j terms to be identically zero, i.e. it is GARCH(0,q). GARCH processes can be thought of as having variances which behave like Autoregressive-Moving Average (ARMA) processes, with the α terms associated with the MA components and the β terms with the AR components.

Bollerslev provides a number of results concerning the properties of these processes. The necessary and sufficient condition for the wide sense stationarity of GARCH processes is that

$$\Sigma_i \alpha_i + \Sigma_j \beta_j \leq 1.$$

The marginal variance of such a process is given by:

$$Var(e_t) = \alpha_0/(1- \Sigma_i \alpha_i - \Sigma_j \beta_j).$$

It also can be shown that GARCH processes, though conditionally normal, have marginal distributions that are more kurtotic (fatter tailed) than is the normal distribution.

A Model for Futures Prices

The GARCH model, combined with a seasonal component, can be used to describe the time series behavior of the changes in the logs of daily futures prices, which will be denoted y_t. Specifically, this discrete random process is modeled as follows:

$$y_t \sim N(0,h_t),$$

with

$$h_t = s_t + v_t,$$

where s_t is a deterministic seasonal component and v_t is a stochastic GARCH component. Before describing these two components, a few comments on general assumptions are in order.

First, the process y_t is assumed to have zero mean. In a number of markets, notably the grain markets, previous studies have provided support for this assumption, and it should, therefore, cause few problems. The assumption allows attention to be focused on the volatility of the process. It would not be difficult, however, to model the mean of the process as an ARIMA process and/or as a function of some set of regressors.

In keeping with the GARCH model, the assumption is made that the process is conditionally normal. As has already been noted, this does not imply that y_t is marginally normal, but rather it is more kurtotic than is the normal distribution. This property is consistent with conventional wisdom about the nature of price distributions.

The seasonal component used in this study has two principle features. First it is deterministic, which allows it to be incorporated into a GARCH model without changing any of the essential features of that model. The second feature is that it is a smooth function of time, which contrasts with such work as Anderson's, in which monthly dummy variables were used. Specifically, s_t is taken to be a sum of trigonometric functions of the day of the year. Letting d_t be the day of the year on which observation t falls (thus on Jan. 1 $d_t=1$, on Jan. 2 $d_t=2$, etc.), this component can be written

$$s_t = \sum_{i=1}^{k} [\theta_i \cos(2\pi i d_t/365) + \phi_i \sin(2\pi i d_t/365)].$$

This representation not only ensures that the seasonal component is a periodic function with a period of one year, but can be justified as a kth order Fourier approximation to an arbitrary seasonal component. Furthermore, because each of the terms in the summation is bounded on [-1,1], computational problems that can plague Taylor approximations are avoided.

Two alternative specifications of the GARCH component, v_t, are examined. The first alternative is a restricted ARCH formulation. The exact specification of this model is intended to address an issue that exists in the options pricing literature. It is common to use so called "historical volatilities", which amount to moving averages of past squared realizations, to provide forecasts of future variance (Chiras & Manaster). There are two problems with this practice. First, it is not clear how many observations to use and second, and more importantly, there is a presumption that simple unweighted averages provide good forecasts of future volatility. This presumption may have some validity in a constant variance model, in which case, however, the forecasting problem essentially ceases to exist. In any model in which the variance changes over time there will be better forecasting methods. Furthermore, as a general rule, the more recent past is more useful for forecasting than is the more distant past. This suggests that a weighted moving average, with weights that decline with the lag length, would provide better forecasts than an unweighted sample.

To shed light on these issues, three moving average components were computed as follows:

$$MA_{kt} = \sum_i (y_{t-i})^2 / n_k,$$

where i=1,...,5 for k=1, i=6,...,10 for k=2, and i=11,...,20 for k=3. The divisors, n_k, are thus 5, 5, and 10, respectively. The v_t component is written as a function of these three variables:

$$v_t = \alpha_0 + \sum_{k=1}^{3} \alpha_k \, MA_{kt}.$$

This is equivelent to an ARCH(20) process with equality restrictions imposed on groups of the coefficents. With such a specification the question of how many lags to use in generating forecasts of variance and whether equal weight should be given to each lag used can be addressed in a formal manner.

The second alternative uses the simple GARCH(1,1) process. Bollerslev points out that, if stationary, this model can be written as an infinitely ordered ARCH model. It can, therefore, provide a parsimonious way of representing a system with declining weights on past realizations and avoids the need to specify an exact lag length. Thus, the ARCH model treats the variance as a function of a weighted sample variance, with weights given by the α_{t-i}, while the GARCH(1,1) model uses weights equal to $\alpha_1 \beta_1^{t-i+1}$.

Estimation of these models is fairly straightforward. The likelihood function for a sample can be defined as the product of the conditional likelihoods and maximum likelihood techniques can be used to estimate model coefficients. The conditional log likelihood function, apart from a constant, is given by

$$l(h_t) = -0.5[\ln(h_t) + y_t^2/h_t],$$

where h_t is given above. In this study maximum likelihood estimates were obtained using the modified method of scoring (Berndt, et al.) with numerical derivatives and using half step squeezes to determine the step length. Initial coefficient values were obtained by setting the coefficient on the constant term, α_0, to the variance of the total sample and all other coefficients to zero.

Hypothesis testing can be carried out using any or all of the likelihood ratio, Wald, or Lagrange multiplier tests (Harvey). These tests can be used to help determine whether either the seasonal or the GARCH components are useful in modeling volatility. They can also be used to help assess the appropriate order of the Fourier terms and, in the ARCH(20) specification, which of the moving average terms are significant and whether they can be pooled.

Empirical Results

The data used in this study consists of the daily Chicago Board of Trade futures prices for December corn for the Jan.-Nov. 1974-1982 period. The first twenty business days of each year are not used in the analysis except in the computation of the variables involving lags. These include the three moving average terms used in the restricted ARCH(20) specification, as well as the lagged squared term and the h_{t-1} term in the GARCH(1,1) specification.

Due to the discontinuities in the data arising from the maturing of contracts, the GARCH(1,1) model requires a small modification. In general the starting value used for h_{t-1} is somewhat arbitrary and in this study starting values are required for each contract. The sample variance for the first 20 business days of each year provide reasonable starting values. It is doubtful that any other reasonable selection would have a large impact on the results given that the sample used is composed of 1897 observations.

Estimation results are given in Table 1 for each of the two model specifications. A third order seasonal component is shown in the presented results. This order was chosen after examination of likelihood ratio statistics for second and fourth order models, from which it was determined that significant gains in fit were to be had by going from second to third order, but not from third to fourth order models.

The results provide clear support for the model. In particular, the coefficients associated with the GARCH components are highly significant in both of the specifications examined. Furthermore, the two models seem to be similar in their ability to capture the essential features of the variance. While they are not nested models, it is worth noting that their two log-likelihoods are quite close, being, apart from a constant, -1295.73 and -1292.98 for the ARCH(20) and the GARCH(1,1) models, respectively. This can be contrasted to the constant variance model, for which the comparable value is -1505.97.

The seasonal components in the models are also statistically significant, though all of the individual coefficients are not. Wald test statistics for the hypothesis that all of the trigonometric terms have zero coefficients have values of 21.7044 and 19.8475, with associated asymptotic p-values of 0.0014 and 0.0029.

The seasonal patterns for the two models are illustrated in Figure 1. Both have similar patterns of peaks and troughs, with the highest peak being in July, as expected. Other smaller peaks occur in early March and November. These may, in part, reflect the release of planting intentions and farm program specifications in March and of final crop reports in November.

There is an interesting contrast between the seasonal components, however. The ARCH(20) model exhibits far greater variation than does the GARCH(1,1) model, with the former having a range of about -0.3 to 0.5, as compared to about -0.02 to 0.05 for the latter. It is possible that this can be explained in part by the difference in the length of memory in the two models.

Despite the relatively small size of the seasonal component for the GARCH(1,1) model, however, there is still a non-zero probability that it could yield a negative predicted variance. This can be seen by noting that the seasonal component is negative and absolutely greater than the constant for some time periods. While empirically this is not a problem, it is nonetheless a bothersome feature and suggests that an alternate, less constraining specification of either the seasonal or the GARCH component might yield better results. The ARCH(20) model does not share this defect.

As was noted earlier, the GARCH aspect of these models can be thought of as taking the current variance to be a weighted sample variance of past realizations. While the stationary GARCH(1,1) framework constrains these weights to decline with the lag length, the estimated coefficients on the moving average terms in the ARCH(20) model also imply this feature. These weight structures are illustrated in Figure 2, where it can be seen that the weights on the first 20 lags, except the 11th, are larger for the ARCH(20) than for the GARCH(1,1), while the latter has a longer memory.

While it is not suggested that either of these models provides the optimal weighting structure, it can be shown that the weights implied by the ARCH(20) model are superior to an equally weighted sample of the previous 20 observations. The equal weights hypothesis is equivelent to the hypothesis that $\alpha_1=\alpha_2=\alpha_3/2$, where the α_i are the coefficients on the moving average terms. The Wald test statistic for this hypothesis has a value of 6.9483, with an associated asymptotic p-value of 0.0310, and, thus, the equal weights hypothesis can be rejected. This suggests that studies, such as Chiras & Manaster, which compared variance forecasts from "historical volatilities" to those implied by option prices, biased their results in favor of the latter by using a suboptimal forecasting method.

Implications for Option Pricing Models

The GARCH framework with a deterministic seasonal component provides a plausible model of the conditional variance of futures prices for seasonally produced commodities. The specific models used here assume that the changes in the logs of prices form a conditionally normal, zero-mean process, the variance of which is composed of a deterministic seasonal component and a component conditional on past realizations of the process. Both of these components were found to be significant in explaining variance in the corn market.

The nature of futures price volatility has important implications for option pricing models. Black's model, which is the most widely known and used, requires a forecast of future variance. This study suggests that forecasts of future variance based on a weighted sample of past realizations, with weights declining with the lag length, and with a seasonal adjustment, are an improvement over the unweighted, fixed-size samples variances that are often used. While only the one-step ahead variances have been discussed here, it is possible to derive multistep variance forecasts. An examination of the nature of such forecasts would be an interesting extension to the current study.

Another implication of the GARCH framework is that the assumptions on which current option pricing models are based may not yield accurate descriptions of the nature of price movements. A common assumption of such models is that the variance component is a known function of, at most, time and current and past realizations of price. This allows options to be valued using only arbitrage considerations and is a feature of the constant elasticity of variance (CEV) model of Cox & Ross, and of Black's model, which is a special case of the CEV model.

The addition of a deterministic seasonal component can be incorporated into these models with little difficulty; the GARCH model, however, adds an element not found in the specific models mentioned. In the CEV model price enters into determination of the variance through the price level, while the GARCH model specifies that it is the size of price changes that is important.

These two assumptions, of course, are not incompatible. Indeed, an interesting topic for further research is the exploration of GARCH type models which include the price level as an explanatory variable. The further issue that naturally arises from this study is whether an explicit option pricing formula could be developed that is based on a GARCH like process, with the variance being a function of the size of price changes. Given the empirical results presented, such a model would be a useful addition to the study of option pricing.

Table 1. Estimation Results for Corn Futures Prices, 1974-1982

Variable[1]	ARCH(20) Coefficient Value	ARCH(20) Asymptotic p-value	GARCH(1,1) Coefficent Value	GARCH(1,1) Asymptotic p-value
Constant	0.3595	0.0000	0.0191	0.0107
MA_1	0.3002	0.0000		
MA_2	0.2108	0.0003		
MA_3	0.2587	0.0000		
y_{t-1}^2			0.0565	0.0000
h_{t-1}			0.9297	0.0000
COS_1	-0.2537	0.0004	-0.0245	0.0072
SIN_1	-0.1263	0.0132	0.0017	0.7214
COS_2	0.0690	0.2755	0.0067	0.3874
SIN_2	0.0070	0.8836	-0.0057	0.1887
COS_3	-0.1402	0.0144	-0.0199	0.0113
SIN_3	-0.0741	0.1374	0.0025	0.6631

1. MA_i refer to the 3 moving average terms described in the text. COS_i and SIN_i are taken with respect to $2\pi i d_t/365$, where d_t is the day of the year.

Figure 1. Seasonal Components for the ARCH(20) and GARCH(1,1) Models

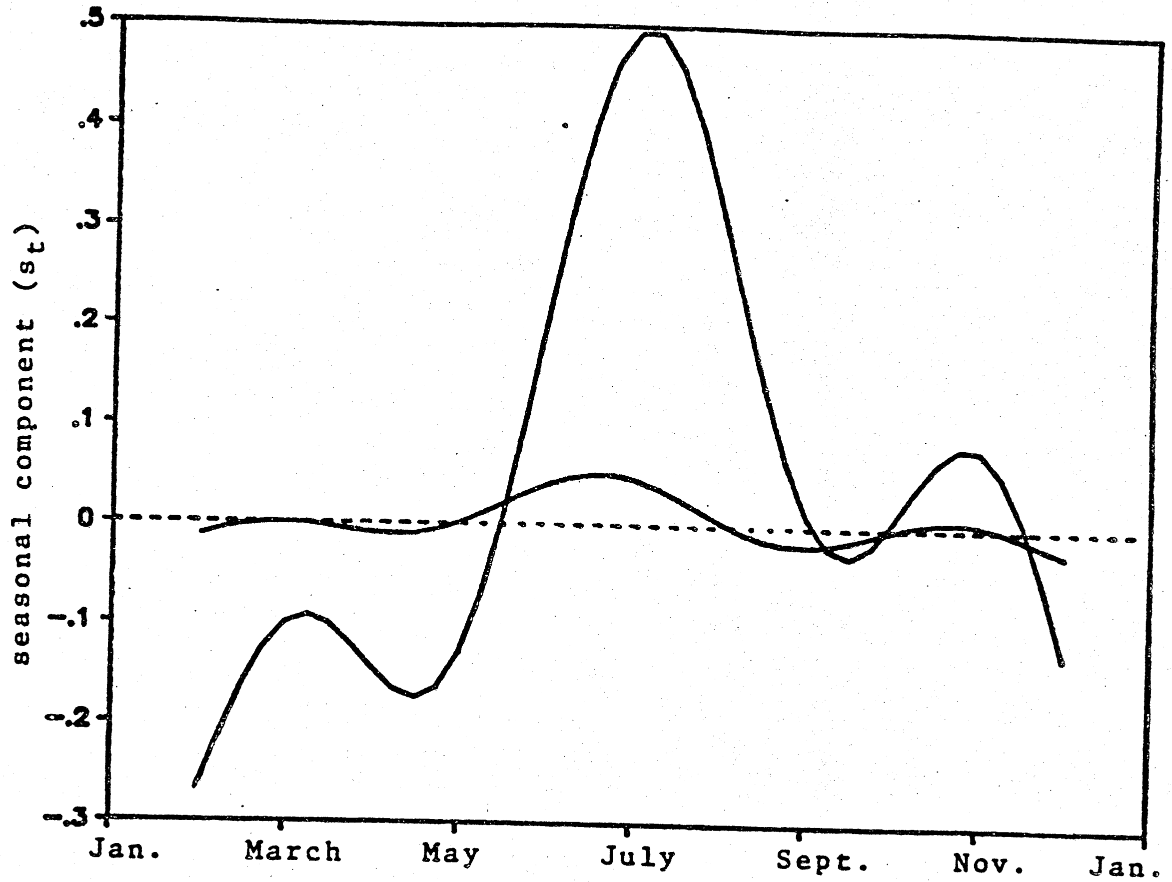
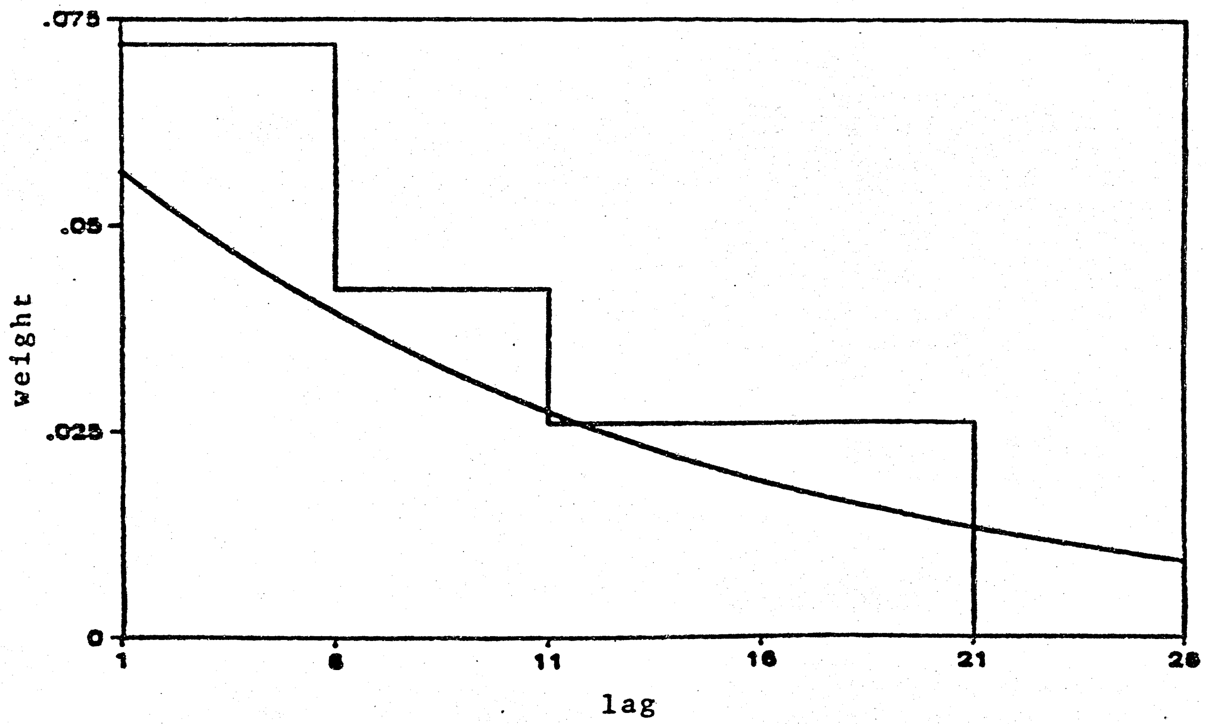


Figure 2. Weights on y_{t-i}^2 for the ARCH(20) and GARCH(1,1) Models



REFERENCES

Anderson, R.W. "Some Determinants of the Volatility of Futures Prices." Journal of Futures Markets. 5(1985): 331-48.

Berndt, E.K., B.H. Hall, R.E. Hall, & J.A. Hausman. "Estimation and Inference in Nonlinear Structural Models." Annals of Economic and Social Measurement. 4(1974): 653-65.

Black, F. "The Pricing of Commodity Contracts." Journal of Financial Economics. 3(1976): 167-79.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroscedasticity." Discussion Paper 85-16, Department of Economics, University of California, San Diego. 1985.

Chiras, D.P. & S. Manaster. "The Information Content of Option Prices and a Test for Market Efficiency." Journal of Financial Economics. 6(1978): 213-34.

Cox, J. & S. Ross. "The Valuation of Options for Alternative Stochastic Processes." Journal of Financial Economics. 3(1976): 145-66.

Engle, R.F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." Econometrica. 50(1982): 987-1007.

Harvey, A.C. The Econometric Analysis of Time Series. John Wiley & Sons. New York. 1981.

Gordon, J.D. "The Distribution of Daily Changes in Commodity Futures Prices." USDA-ERS Technical Bulletin #1702. Washington, D.C. 1985.

Mandelbrot, B. "The Variation of Certain Speculative Prices." Journal of Business. 36(1963): 394-419.