GARCH 101

Robert Engle

Robert Engle is Visiting Professor of Finance, Stern School of Business, New York University, New York, New York, and Professor of Economics, University of California at San Diego, La Jolla, California.

The great workhorse of applied econometrics is the least squares model. The basic version of the model assumes that, the expected value of all error terms, in absolute value, is the same at any given point. Thus, the expected value of any given error term, squared, is equal to the variance of all the error terms taken together. This assumption is called homoskedasticity. Conversely, data in which the expected value of the error terms is not equal, in which the error terms may reasonably be expected to be larger for some points or ranges iof the data than for others, is said to suffer from heteroskedasticity.

It has long been recognized that heteroskedasticity can pose problems in ordinary least squares analysis. The standard warning is that in the presence of heteroskedasticity, the regression coefficients for an ordinary least squares regression are still unbiased, but the standard errors and confidence intervals estimated by conventional procedures will be too narrow, giving a false sense of precision. However, the warnings about heteroskedasticity have usually been applied only to cross sectional models, not to time series models. For example, if one looked at the cross-section relationship between income and consumption in household data, one might expect to find that the consumption of low-income households is more closely tied to income than that of high-income households, because poor households are more likely to consume all of their income and to be liquidity-constrained.

In a cross-section regression of household consumption on income, the error terms seem likely to be systematically larger for high-income than for low-income households, and the assumption of homoskedasticity seems implausible. In contrast, if one looked at an aggregate time series consumption function, comparing national income to consumption, it seems more plausible to assume that the variance of the error terms doesn't changed much over time.

A recent developments in estimation of standard errors, known as "robust standard errors," has also reduced the concern over heteroskedasticity. If the sample size is large, then robust standard errors give quite a good estimate of standard errors even with heteroskedasticity. Even if the sample is small, the need for a heteroskedasticity correction that doesn't affect the coefficients, but only narrows the standard errors somewhat, can be debated.

However, sometimes the key issue is the variance of the error terms itself. This question often arises in financial applications where the dependent variable is the return on an asset or portfolio and the variance of the return represents the risk level of those returns. These are time series applications, but it is nonetheless likely that heteroskedasticity is an issue. Even a cursory look at financial data suggests that some time periods are riskier than others; that is, the expected value of error terms at some times is greater than at others. Moreover, these risky times are not scattered

randomly across quarterly or annual data. Instead, there is a degree of autocorrelation in the riskiness of financial returns. ARCH and GARCH models, which stand for autoregressive conditional heteroskedasticity and *generalized* autoregressive conditional heterosjedasticity, have become widespread tools for dealing with time series heteroskedastic models such as ARCH and GARCH. The goal of such models is to provide a volatility measure – like a standard deviation -- that can be used in financial decisions concerning risk analysis, portfolio selection and derivative pricing.

ARCH/GARCH Models

 Because this paper will focus on financial applications, we will use financial notation. Let the dependent variable be labeled r_t, which could be the return on an asset or portfolio. The mean value m and the variance h will be defined relative to a past information set. Then, the return r in the present will be equal to the mean value of r (that is, the expected value of r based on past information) plus the standard deviation of r (that is, the square root of the variance) times the error term for the present period.

    The econometric challenge is to specify how the information is used to forecast the mean and variance of the return, conditional on the past information. While many specifications have been considered for the mean return and have been used in efforts to forecast future returns, rather simple

4

specifications have proven surprisingly successful in predicting conditional variances. The most widely used specification is the GARCH(1,1) model introduced by Bollerslev (1986) as a generalization of Engle(1982). The (1,1) in parentheses is a standard notation in which the first number refers to how many autoregressive lags appear in the equation, while the second number refers to how many lags are included in the moving average component of a variable. Thus, a GARCH (1,1) model for variance looks like this:

$$h_t = \omega + \alpha h_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1}.$$

This model forecasts the variance of date t return as a weighted average of a constant, yesterday's forecast, and yesterday's squared error. Of course, if the mean is zero, then from the surprise is simply r_{t-1}^2.

Thus the GARCH models are conditionally heteroskedastic but have a constant unconditional variance.

Possibly the most important aspect of the ARCH/GARCH model is the recognition that volatility can be estimated based on historical data and that a bad model can be detected directly using conventional econometric techniques. A variety of statistical software packages like Eview **and others?** are available for implementing GARCH and ARCH approaches.

A Value at Risk Example

Applications of the ARCH/GARCH approach are widespread in situations where volatility of returns is a central issue. Many banks and other financial institutions use the idea of "value at risk" as a way to measure the risks faced by their portfolios. The 1% value at risk is defined as the number of dollars that one can be 99 percent certain exceeds any losses for the next day. Let's use the GARCH (1,1) tools to estimate the 1 percent value at risk of a $1,000,000 portfolio on March 23, 2000. This portfolio consists of 50 percent Nasdaq, 30 percent Dow Jones, and 20 percent long bonds. This date is chosen to be just before the big market slide at the end of March and April. It is a time of high volatility and great anxiety.

First, we construct the hypothetical historical portfolio. (All calculations in this example were done with the Eviews software program.) Figure 1 shows the pattern of the Nasdaq, Dow Jones, and long bonds. In Table 1, we present some illustrative statistics for each of these three investments separately, and for the portfolio as a whole in the final column.

Then we forecast the standard deviation of the portfolio and its 1 percent quantile. We carry out this calculation over several different time

frames: the entire 10 years of the sample up to March 23, 2000; the year before March 23, 2000; and from January 1, 2000 to March 23, 2000.

Consider first the quantiles of the historical portfolio at these three different time horizons. Over the full ten-year sample, the 1 percent quantile times $1,000,000 produces a value at risk of $22,477. Over the last year the calculation produces a value at risk of #24, 653 – somewhat higher, but not enormously so. However, if the first quantile is calculated based on the data from January 1, 2000 to March 23, 2000, the value at risk is $35,159. Thus, the level of risk has increased dramatically over the last quarter.

The basic GARCH(1,1) results are given in Table 2 below. Notice that the coefficients sum up to a number slightly less than one. The forecast standard deviation for the next day is 0.014605, which is almost double the average standard deviation of .0083 presented in the last column of Table 1. If the residuals were normally distributed, then this would be multiplied by 2.326348 giving a VaR=$33,977. As it turns out, the standardized residuals, which are the estimated values of $\{\varepsilon_t\}$, have a 1% quantile of 2.8437, which is well above the normal quantile. The estimated 1% VaR is $39,996. Notice that this VaR has risen to reflect the increased risk in 2000.

Extensions and Modifications of GARCH

The GARCH(1,1) is the simplest and most robust of the family of volatility models. However, the model can be extended and modified in many ways. We will briefly mention three modifications.

The GARCH (1,1) model can be generalized to a GARCH(p,q) model; that is, a model with additional lag terms. Such higher order models are often useful when a long span of data is used, like several decades of daily data or a year of hourly data. With additional lags, such models allow both fast and slow decay of information. A particular specification of the GARCH(2,2) by Engle and Lee(1999), sometimes called the component model, is a useful starting point to this approach.

Another version of GARCH models takes an asymmetric view by estimating positive and negative returns separately. Typically, higher volatilities follow negative returns than positive returns of the same magnitude. Two models which take this asymmetric approach are the TARCH model – threshhold ARCH -- attributed to Zakoian() and Glosten Jaganathan and Runkle (1993), and the EGARCH model of Nelson(1991

It is also possible to incorporate exogenous variables into the GARCH equation. **Like what variables?**

Software packages like Eviews offer a variety of tests to check specifications of ARCH/GARCH models or to choose between models.

Conclusion

Volatility models have been applied in a wide variety of applications. In most cases, volatility is itself an interesting aspect of the problem. In some cases, volatility is an input used for purposes of measurement, like in the example of estimating value at risk given earlier. In other cases, volatility may be a causal variable, as in models expected volatility is a determinant of expected returns.
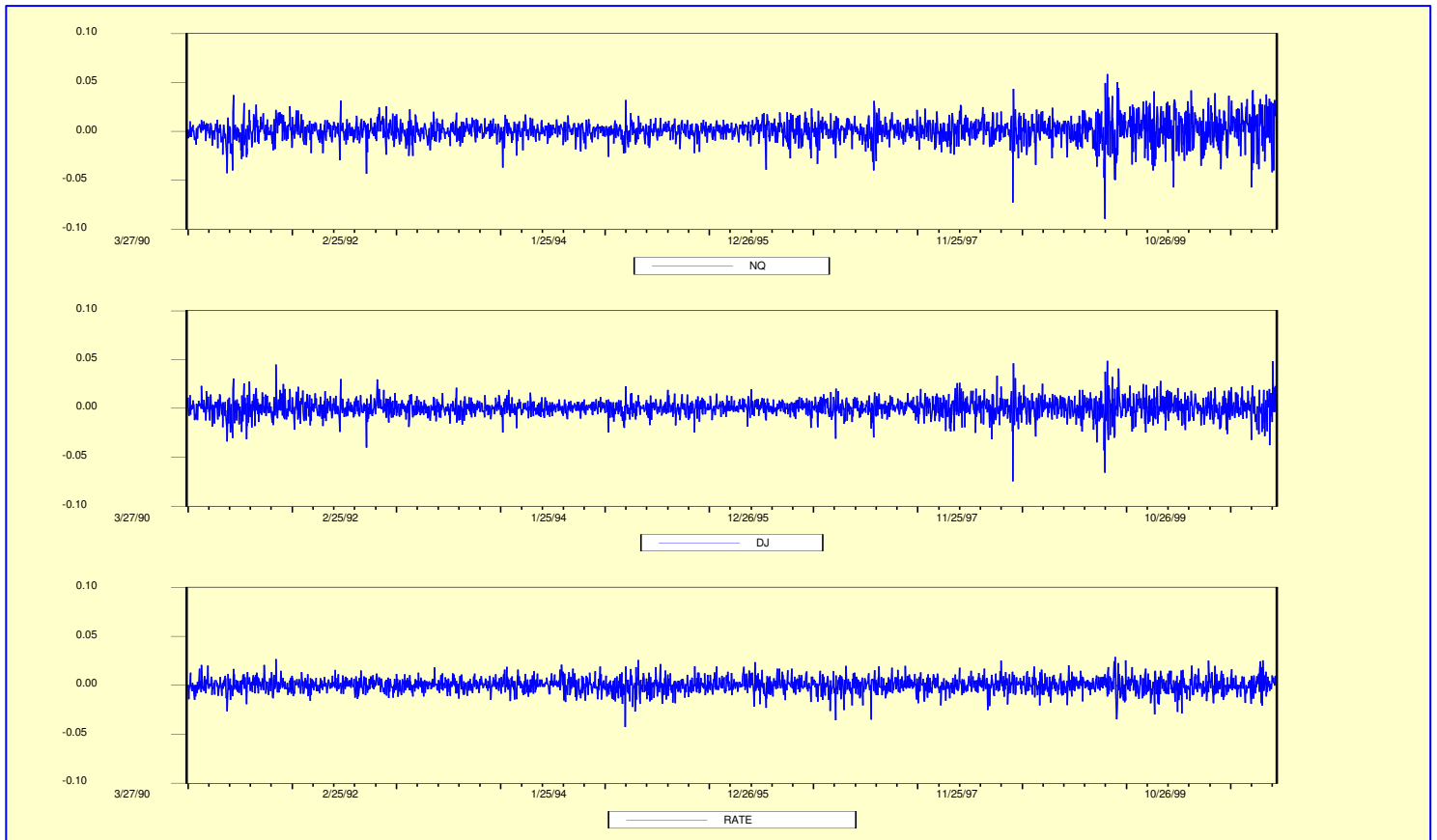
Nasdaq,Dow Jones, and Bond Returns

Figure 1

Table 1

Portfolio Data

Sample: 3/23/1990 3/23/2000				
	NQ	DJ	RATE	PORT
Mean	0.0009	0.0005	0.0001	0.0007
Std. Dev.	0.0115	0.0090	0.0073	0.0083
Skewness	-0.5310	-0.3593	-0.2031	-0.4738
Kurtosis	7.4936	8.3288	4.9579	7.0026

Table 2

GARCH(1,1)

Dependent Variable: PORT
Sample(adjusted): 3/26/1990 3/23/2000
Convergence achieved after 16 iterations
Bollerslev-Wooldrige robust standard errors & covariance

Variance Equation			
C	0.0000	0.0000	3.1210	0.0018
ARCH(1)	0.0772	0.0179	4.3046	0.0000
GARCH(1)	0.9046	0.0196	46.1474	0.0000

S.E. of regression	0.0083	Akaike info criterion	-6.9186
Sum squared resid	0.1791	Schwarz criterion	-6.9118
Log likelihood	9028.2809	Durbin-Watson stat	1.8413

References

Bollerslev, Tim, 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

Bollerslev, Tim and Wooldridge, Jeffrey M., 1992, Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances, *Econometric Reviews*, 11(2), 143-172.

Engle, Robert F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987-1007.

Engle, Robert F., and Manganelli, Simone, 1999, CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles, University of California, San Diego, Department of Economics Working Paper 99-20.

Engle, Robert F., and Mezrich, Joseph, 1996, GARCH for Groups, *RISK*, 9(8), 36-40.

Engle, Robert F., and Ng, Victor, 1993, Measuring and Testing the Impact of News on Volatility, *Journal of Finance*, 48, 1749-1778.

Glosten, Lawrence R., Jagannathan, Ravi and Runkle, David E., 1993, On the Relation between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks, *Journal of Finance*, 48(5), 1779-1801.

Nelson, Daniel B., 1991, Conditional Heteroscedasticity in Asset Returns: A New Approach, *Econometrica*, 59(2), 347-370.