GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics

Robert Engle

GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics

Robert Engle

<p>
<big>**T**</big>he great workhorse of applied econometrics is the least squares model. This is a natural choice, because applied econometricians are typically called upon to determine how much one variable will change in response to a change in some other variable. Increasingly however, econometricians are being asked to forecast and analyze the size of the errors of the model. In this case, the questions are about volatility, and the standard tools have become the ARCH/GARCH models.
</p>

The basic version of the least squares model assumes that the expected value of all error terms, when squared, is the same at any given point. This assumption is called homoskedasticity, and it is this assumption that is the focus of ARCH/GARCH models. Data in which the variances of the error terms are not equal, in which the error terms may reasonably be expected to be larger for some points or ranges of the data than for others, are said to suffer from heteroskedasticity. 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. Instead of considering this as a problem to be corrected, ARCH and GARCH models treat heteroskedasticity as a variance to be modeled. As a result, not only are the deficiencies of least squares corrected, but a prediction is computed for the variance of each error term. This prediction turns out often to be of interest, particularly in applications in finance.

The warnings about heteroskedasticity have usually been applied only to cross-section models, not to time series models. For example, if one looked at the

■ *Robert Engle is the Michael Armellino Professor of Finance, Stern School of Business, New York University, New York, New York, and Chancellor's Associates Professor of Economics, University of California at San Diego, La Jolla, California.*

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 the dollars of savings or deficit by poor households are likely to be much smaller in absolute value than high income households. In a cross-section regression of household consumption on income, the error terms seem likely to be systematically larger in absolute value 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 change much over time.

A recent development 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. If the sample is small, the need for a heteroskedasticity correction that does not affect the coefficients, and only asymptotically corrects the standard errors, can be debated.

However, sometimes the natural question facing the applied econometrician is the accuracy of the predictions of the model. In this case, the key issue is the variance of the error terms and what makes them large. 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 the magnitude 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. Financial analysts, looking at plots of daily returns such as in Figure 1, notice that the amplitude of the returns varies over time and describe this as "volatility clustering." The ARCH and GARCH models, which stand for autoregressive conditional heteroskedasticity and *generalized* autoregressive conditional heteroskedasticity, are designed to deal with just this set of issues. They have become widespread tools for dealing with time series heteroskedastic models. 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

Figure 1
Nasdaq, Dow Jones and Bond Returns
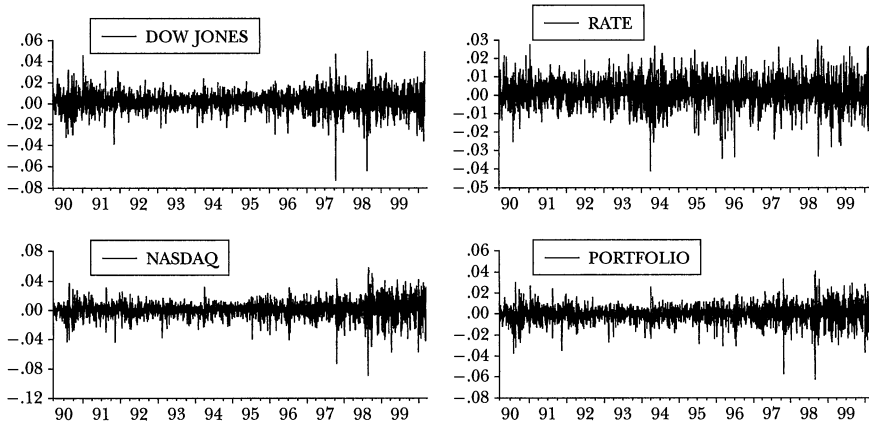
value of r (that is, the expected value of r based on past information) plus the standard deviation of r (that is, the square root of the variance) times the error term for the present period.

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, virtually no methods were available for the variance before the introduction of ARCH models. The primary descriptive tool was the rolling standard deviation. This is the standard deviation calculated using a fixed number of the most recent observations. For example, this could be calculated every day using the most recent month (22 business days) of data. It is convenient to think of this formulation as the first ARCH model; it assumes that the variance of tomorrow's return is an equally weighted average of the squared residuals from the last 22 days. The assumption of equal weights seems unattractive, as one would think that the more recent events would be more relevant and therefore should have higher weights. Furthermore the assumption of zero weights for observations more than one month old is also unattractive. The ARCH model proposed by Engle (1982) let these weights be parameters to be estimated. Thus, the model allowed the data to determine the best weights to use in forecasting the variance.

A useful generalization of this model is the GARCH parameterization introduced by Bollerslev (1986). This model is also a weighted average of past squared residuals, but it has declining weights that never go completely to zero. It gives parsimonious models that are easy to estimate and, even in its simplest form, has proven surprisingly successful in predicting conditional variances. The most widely used GARCH specification asserts that the best predictor of the variance in the next period is a weighted average of the long-run average variance, the variance

predicted for this period, and the new information in this period that is captured by the most recent squared residual. Such an updating rule is a simple description of adaptive or learning behavior and can be thought of as Bayesian updating.

Consider the trader who knows that the long-run average daily standard deviation of the Standard and Poor's 500 is 1 percent, that the forecast he made yesterday was 2 percent and the unexpected return observed today is 3 percent. Obviously, this is a high volatility period, and today is especially volatile, which suggests that the forecast for tomorrow could be even higher. However, the fact that the long-term average is only 1 percent might lead the forecaster to lower the forecast. The best strategy depends upon the dependence between days. If these three numbers are each squared and weighted equally, then the new forecast would be $2.16 = \sqrt{(1 + 4 + 9)/3}$. However, rather than weighting these equally, it is generally found for daily data that weights such as those in the empirical example of $(.02, .9, .08)$ are much more accurate. Hence the forecast is $2.08 = \sqrt{.02*1 + .9*4 + .08*9}$.

To be precise, we can use h_t to define the variance of the residuals of a regression $r_t = m_t + \sqrt{h_t}\varepsilon_t$. In this definition, the variance of ε is one. The GARCH model for variance looks like this:

$$h_{t+1} = \omega + \alpha(r_t - m_t)^2 + \beta h_t = \omega + \alpha h_t \varepsilon_t^2 + \beta h_t.$$

The econometrician must estimate the constants ω, α, β; updating simply requires knowing the previous forecast h and residual. The weights are $(1 - \alpha - \beta, \beta, \alpha)$, and the long-run average variance is $\sqrt{\omega/(1 - \alpha - \beta)}$. It should be noted that this only works if $\alpha + \beta < 1$, and it only really makes sense if the weights are positive, requiring $\alpha > 0$, $\beta > 0$, $\omega > 0$.

The GARCH model that has been described is typically called the GARCH(1,1) model. The (1,1) in parentheses is a standard notation in which the first number refers to how many autoregressive lags, or ARCH terms, appear in the equation, while the second number refers to how many moving average lags are specified, which here is often called the number of GARCH terms. Sometimes models with more than one lag are needed to find good variance forecasts.

Although this model is directly set up to forecast for just one period, it turns out that based on the one-period forecast, a two-period forecast can be made. Ultimately, by repeating this step, long-horizon forecasts can be constructed. For the GARCH(1,1), the two-step forecast is a little closer to the long-run average variance than is the one-step forecast, and, ultimately, the distant-horizon forecast is the same for all time periods as long as $\alpha + \beta < 1$. This is just the unconditional variance. Thus, the GARCH models are mean reverting and conditionally heteroskedastic, but have a constant unconditional variance.

I turn now to the question of how the econometrician can possibly estimate an equation like the GARCH(1,1) when the only variable on which there are data is r_t. The simple answer is to use maximum likelihood by substituting h_t for σ^2 in the normal likelihood and then maximizing with respect to the parameters. An even

simpler answer is to use software such as EViews, SAS, GAUSS, TSP, Matlab, RATS and many others where there exist already packaged programs to do this.

But the process is not really mysterious. For any set of parameters ω, α, β and a starting estimate for the variance of the first observation, which is often taken to be the observed variance of the residuals, it is easy to calculate the variance forecast for the second observation. The GARCH updating formula takes the weighted average of the unconditional variance, the squared residual for the first observation and the starting variance and estimates the variance of the second observation. This is input into the forecast of the third variance, and so forth. Eventually, an entire time series of variance forecasts is constructed. Ideally, this series is large when the residuals are large and small when they are small. The likelihood function provides a systematic way to adjust the parameters ω, α, β to give the best fit.

Of course, it is entirely possible that the true variance process is different from the one specified by the econometrician. In order to detect this, a variety of diagnostic tests are available. The simplest is to construct the series of $\{\varepsilon_t\}$, which are supposed to have constant mean and variance if the model is correctly specified. Various tests such as tests for autocorrelation in the squares are able to detect model failures. Often a "Ljung box test" with 15 lagged autocorrelations is used.

A Value-at-Risk Example

Applications of the ARCH/GARCH approach are widespread in situations where the volatility of returns is a central issue. Many banks and other financial institutions use the concept of "value at risk" as a way to measure the risks faced by their portfolios. The 1 percent value at risk is defined as the number of dollars that one can be 99 percent certain exceeds any losses for the next day. Statisticians call this a 1 percent quantile, because 1 percent of the outcomes are worse and 99 percent are better. 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. The long bond is a ten-year constant maturity Treasury bond.[1] 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 returns of the Nasdaq, Dow Jones, bonds and the composite portfolio leading up to the terminal date. Each of these series appears to show the signs of ARCH effects in that the amplitude of the returns varies over time. In the case of the equities, it is clear that this has increased substantially in the latter part of the sample period. Visually, Nasdaq is even more extreme. In Table 1, we present some illustrative

---

[1] The portfolio has constant proportions of wealth in each asset that would entail some rebalancing over time.

Table 1

Portfolio Data

	NASDAQ	Dow Jones	Rate	Portfolio
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

Sample: March 23, 1990 to March 23, 2000.

statistics for each of these three investments separately and for the portfolio as a whole in the final column. From the daily standard deviation, we see that the Nasdaq is the most volatile and interest rates the least volatile of the assets. The portfolio is less volatile than either of the equity series even though it is 80 percent equity—yet another illustration of the benefits of diversification. All the assets show evidence of fat tails, since the kurtosis exceeds 3, which is the normal value, and evidence of negative skewness, which means that the left tail is particularly extreme.

The portfolio shows substantial evidence of ARCH effects as judged by the autocorrelations of the squared residuals in Table 2. The first order autocorrelation is .210, and they gradually decline to .083 after 15 lags. These autocorrelations are not large, but they are very significant. They are also all positive, which is uncommon in most economic time series and yet is an implication of the GARCH(1,1) model. Standard software allows a test of the hypothesis that there is no autocorrelation (and hence no ARCH). The test p-values shown in the last column are all zero to four places, resoundingly rejecting the "no ARCH" hypothesis.

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 ten 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. To do this calculation, one simply sorts the returns and finds the 1 percent worst case. 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 1 percent 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 apparently has increased dramatically over the last quarter of the sample. Each of these numbers is the appropriate value at risk if the next day is equally likely to be the same as the days in the given sample period. This assumption is more likely to be true for the shorter period than for the long one.

The basic GARCH(1,1) results are given in Table 3. Under this table it lists the dependent variable, PORT, and the sample period, indicates that it took the algorithm 16 iterations to maximize the likelihood function and computed stan-

Table 2

Autocorrelations of Squared Portfolio Returns

	AC	Q-Stat	Prob
1	0.210	115.07	0.000
2	0.183	202.64	0.000
3	0.116	237.59	0.000
4	0.082	255.13	0.000
5	0.122	294.11	0.000
6	0.163	363.85	0.000
7	0.090	384.95	0.000
8	0.099	410.77	0.000
9	0.081	427.88	0.000
10	0.081	445.03	0.000
11	0.069	457.68	0.000
12	0.080	474.29	0.000
13	0.076	489.42	0.000
14	0.074	503.99	0.000
15	0.083	521.98	0.000

Sample: March 23, 1990 to March 23, 2000.

Table 3

GARCH(1,1)

Variable	Variance Equation		Z-Stat	P-Value
	Coef	St. Err		
C	1.40E-06	4.48E-07	3.1210	0.0018
ARCH(1)	0.0772	0.0179	4.3046	0.0000
GARCH(1)	0.9046	0.0196	46.1474	0.0000

Notes: Dependent Variable: PORT.
Sample (adjusted): March 23, 1990 to March 23, 2000.
Convergence achieved after 16 iterations.
Bollerslev-Woodridge robust standard errors and covariance.

dard errors using the robust method of Bollerslev-Wooldridge. The three coefficients in the variance equation are listed as C, the intercept; ARCH(1), the first lag of the squared return; and GARCH(1), the first lag of the conditional variance. Notice that the coefficients sum up to a number less than one, which is required to have a mean reverting variance process. Since the sum is very close to one, this process only mean reverts slowly. Standard errors, Z-statistics (which are the ratio of coefficients and standard errors) and *p*-values complete the table.

The standardized residuals are examined for autocorrelation in Table 4. Clearly, the autocorrelation is dramatically reduced from that observed in the portfolio returns themselves. Applying the same test for autocorrelation, we now

Table 4
Autocorrelations of Squared Standardized Residuals

	AC	Q-Stat	Prob
1	0.005	0.0589	0.808
2	0.039	4.0240	0.134
3	−0.011	4.3367	0.227
4	−0.017	5.0981	0.277
5	0.002	5.1046	0.403
6	0.009	5.3228	0.503
7	−0.015	5.8836	0.553
8	−0.013	6.3272	0.611
9	−0.024	7.8169	0.553
10	−0.006	7.9043	0.638
11	−0.023	9.3163	0.593
12	−0.013	9.7897	0.634
13	−0.003	9.8110	0.709
14	0.009	10.038	0.759
15	−0.012	10.444	0.791

find the p-values are about 0.5 or more, indicating that we can accept the hypothesis of "no residual ARCH."

The forecast standard deviation for the next day is 0.0146, which is almost double the average standard deviation of 0.0083 presented in the last column of Table 1. If the residuals were normally distributed, then this would be multiplied by 2.327, because 1 percent of a normal random variable lies 2.327 standard deviations below the mean. The estimated normal value at risk = $33,977. As it turns out, the standardized residuals, which are the estimated values of $\{\varepsilon_t\}$, are not very close to a normal distribution. They have a 1 percent quantile of 2.844, which reflects the fat tails of the asset price distribution. Based on the actual distribution, the estimated 1 percent value at risk is $39,996. Notice how much this value at risk has risen to reflect the increased risk in 2000.

Finally, the value at risk can be computed based solely on estimation of the quantile of the forecast distribution. This has recently been proposed by Engle and Manganelli (2001), adapting the quantile regression methods of Koenker and Basset (1978) and Koenker and Hallock in this symposium. Application of their method to this data set delivers a value at risk = $38,228.

What actually did happen on March 24, 2000, and subsequently? The portfolio lost more than $1000 on March 24 and more than $3000 on March 27. The biggest hit was $67,000 on April 14. We all know that Nasdaq declined substantially over the next year. The Dow Jones average was much less affected, and bond prices increased as the Federal Reserve lowered interest rates. Figure 2 plots the value at risk estimated each day using this methodology within the sample period and the losses that occurred the next day. There are about 1 percent of times the value at risk is exceeded, as is expected, since this is in-sample. Figure 3 plots the same graph for the next year and a quarter, during
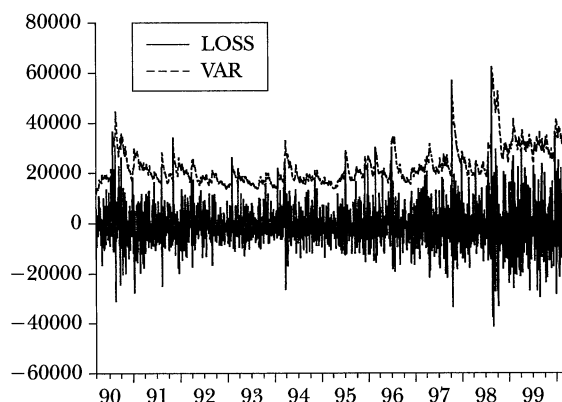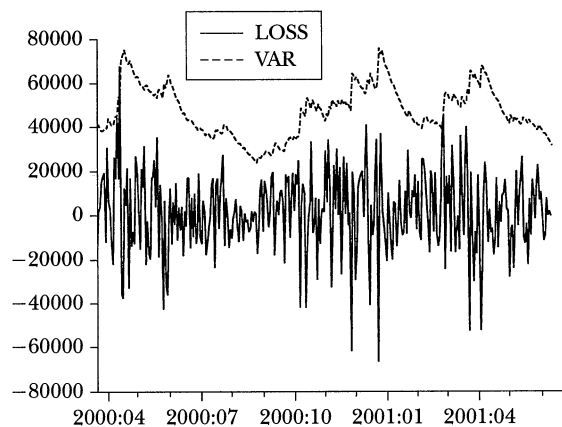
Figure 2

Value at Risk and Portfolio Losses In-Sample

Figure 3

Value at Risk and Portfolio Losses Out of Sample

which the equity market tanks and the bond yields fall. The parameters are not reestimated, but the formula is simply updated each day. The computed value at risk rises substantially from the $40,000 initial figure as the volatility rises in April 2000. Then the losses decline, so that the value at risk is well above the realized losses. Toward the end of the period, the losses approach the value at risk again, but at a lower level. In this year and a quarter, the value at risk is exceeded only once; thus, this is actually a slightly conservative estimate of the risk. It is not easy to determine whether a particular value-at-risk number is correct, although statistical tests can be formulated for this in the same way they are formulated for volatilities. For example, Engle and Manganelli (2001) present a "dynamic quantile test."

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. I will briefly mention three modifications, although the number of volatility models that can be found in the literature is now quite extraordinary.

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.

ARCH/GARCH models thus far have ignored information on the direction of returns; only the magnitude matters. However, there is very convincing evidence that the direction does affect volatility. Particularly for broad-based equity indices and bond market indices, it appears that market declines forecast higher volatility than comparable market increases do. There is now a variety of asymmetric GARCH models, including the EGARCH model of Nelson (1991), the TARCH model— threshold ARCH—attributed to Rabemananjara and Zakoian (1993) and Glosten, Jaganathan and Runkle (1993), and a collection and comparison by Engle and Ng (1993).

The goal of volatility analysis must ultimately be to explain the causes of volatility. While time series structure is valuable for forecasting, it does not satisfy our need to explain volatility. The estimation strategy introduced for ARCH/GARCH models can be directly applied if there are predetermined or exogenous variables. Thus, we can think of the estimation problem for the variance just as we do for the mean. We can carry out specification searches and hypothesis tests to find the best formulation. Thus far, attempts to find the ultimate cause of volatility are not very satisfactory. Obviously, volatility is a response to news, which must be a surprise. However, the timing of the news may not be a surprise and gives rise to predictable components of volatility, such as economic announcements. It is also possible to see how the amplitude of news events is influenced by other news events. For example, the amplitude of return movements on the United States stock market may respond to the volatility observed earlier in the day in Asian markets as well as to the volatility observed in the United States on the previous day. Engle, Ito and Lin (1990) call these "heat wave" and "meteor shower" effects.

A similar issue arises when examining several assets in the same market. Does the volatility of one influence the volatility of another? In particular, the volatility of an individual stock is clearly influenced by the volatility of the market as a whole. This is a natural implication of the capital asset pricing model. It also appears that there is time variation in idiosyncratic volatility (for example, Engle, Ng and Rothschild, 1992).

This discussion opens the door to multivariate modeling where not only the volatilities but also the correlations are to be investigated. There are now a large number of multivariate ARCH models to choose from. These turn out often to be difficult to estimate and to have large numbers of parameters. Research is continuing to examine new classes of multivariate models that are more convenient for fitting large covariance matrices. This is relevant for systems of equations such as vector autoregressions and for portfolio problems where possibly thousands of assets are to be analyzed.

Conclusion

ARCH and GARCH models have been applied to a wide range of time series analyses, but applications in finance have been particularly successful and have been the focus of this introduction. Financial decisions are generally based upon the tradeoff between risk and return; the econometric analysis of risk is therefore an integral part of asset pricing, portfolio optimization, option pricing and risk management. This paper has presented an example of risk measurement that could be the input to a variety of economic decisions. The analysis of ARCH and GARCH models and their many extensions provides a statistical stage on which many theories of asset pricing and portfolio analysis can be exhibited and tested.

References

Bollerslev, Tim. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics.* April, 31:3, pp. 307–27.

Bollerslev, Tim and Jeffrey M. Wooldridge. 1992. "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances." *Econometric Reviews.* 11:2, pp. 143–72.

Engle, Robert F. 1982. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica.* 50:4, pp. 987–1007.

Engle, Robert and Gary G. J. Lee. 1999. "A Permanent and Transitory Component Model of Stock Return Volatility," in *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger.* Robert F. Engle and Halbert

White, eds. Oxford: Oxford University Press, pp. 475–97.

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

Engle, Robert F. and Simone Manganelli. 2001. "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles." Manuscript, University of California, San Diego. Revision of NBER Working Paper No. W7341 (1999).

Engle, Robert F. and Joseph Mezrich. 1996. "GARCH for Groups." *RISK.* 9:8, pp. 36–40.

Engle, Robert F. and Victor Ng. 1993. "Measuring and Testing the Impact of News on

Volatility." *Journal of Finance*. December, 48:5, pp. 1749–78.

Engle, Robert, Takatoshi Ito and Wen-Ling Lin. 1990. "Meteor Showers or Heat Waves? Het eroskedastic Intra-Daily Volatility in the Foreign Exchange Market." *Econometrica*. May, 58:3, pp. 525–42.

Engle, Robert, Victor Ng and M. Rothschild. 1992. "A Multi-Dynamic Factor Model for Stock Returns." *Journal of Econometrics*. April/May, 52: 1–2, pp. 245–66.

Glosten, Lawrence R., Ravi Jagannathan and David E. Runkle. 1993. "On the Relation Be-

tween the Expected Value and the Volatility of the Nominal Excess Returns on Stocks." *Journal of Finance*. 48:5, pp. 1779–801.

Koenker, Roger and Gilbert Bassett. 1978. "Regression Quantiles." *Econometrica*. January, 46:1, pp. 33–50.

Nelson, Daniel B. 1991. "Conditional Heteroscedasticity in Asset Returns: A New Approach." *Econometrica*. 59:2, pp. 347–70.

Rabemananjara, R. and J. M. Zakoian. 1993. "Threshold Arch Models and Asymmetries in Volatility." *Journal of Applied Econometrics*. January/March, 8:1, pp. 31–49.