Journal of Business & Economic Statistics, April 1992, Vol. 10, No. 2

Inequality Constraints in the Univariate GARCH Model

Daniel B. Nelson and Charles Q. Cao
Graduate School of Business, University of Chicago, Chicago, IL 60637

To keep the conditional variances generated by the GARCH(p, q) model nonnegative, Bollerslev imposed nonnegativity constraints on the parameters of the process. We show that these constraints can be substantially weakened and so should not be imposed in estimation. We also provide empirical examples illustrating the importance of relaxing these constraints.

KEY WORDS: Autoregressive conditional heteroscedasticity (ARCH).

1. INTRODUCTION

Since their introduction by Engle (1982) and Bollerslev (1986), respectively, autoregressive conditional heteroscedastic (ARCH) and generalized autoregressive conditional heteroscedastic (GARCH) models have found extraordinarily wide use. The survey article by Bollerslev, Chou, and Kroner (1982) cited more than 300 papers applying ARCH, GARCH, and other closely related models. As they showed, ARCH and GARCH models have been very successful at modeling time-varying volatility in financial time series. One nettlesome feature of GARCH models, however, has been the inequality constraints imposed to keep the conditional variance nonnegative. As we shall see, estimated parameters frequently violate these constraints. This article shows that inequality constraints less severe than commonly imposed are sufficient to keep the conditional variance nonnegative.

Is this important in practice? An ARCH model estimated using quasi-maximum likelihood methods will not generate negative conditional variances σ_t^2 *in sample*, since the log quasi-likelihood involves a term in $\ln(\sigma_t^2)$, which explodes to $-\infty$ as σ_t^2 approaches 0 and is ill-defined for $\sigma_t^2 \leq 0$. Nevertheless, an estimated ARCH model may have coefficients that allow σ_t^2 to become negative *out of sample* (or, more precisely, assign positive probability to the event that σ_t^2 eventually becomes negative). Such estimated coefficients must either result from sampling error (in which case it may be best to impose the parameter constraints in estimation) or from specification error. In this article, we show that empirically relevant violations of Bollerslev's inequality constraints may be the result neither of sampling error nor of misspecification.

The GARCH(p, q) model sets

$$\xi_t = \sigma_t z_t, \tag{1}$$

$$z_t \sim \text{iid} \quad \text{with } E(z_t) = 0, \qquad \operatorname{var}(z_t) = 1, \tag{2}$$

and

$$\sigma_t^2 = \omega + \sum_{i=1,p} \beta_i \sigma_{t-i}^2 + \sum_{j=1,q} \alpha_j \xi_{t-j}^2, \tag{3}$$

where σ_t^2 is the conditional variance of ξ_t given ξ_{t-1}, $\xi_{t-2}, \ldots, \sigma_{t-1}^2, \sigma_{t-2}^2, \ldots$. As a conditional variance, σ_t^2 must, of course, remain nonnegative with probability 1. To guarantee this nonnegativity, Bollerslev (1986) imposed the conditions

$$\omega \geq 0, \tag{4}$$

$$\beta_i \geq 0 \quad \text{for all } i = 1 \text{ to } p, \tag{5}$$

and

$$\alpha_j \geq 0 \quad \text{for all } j = 1 \text{ to } q. \tag{6}$$

Recursively substituting for lagged values of σ_t^2 in (3), it is easy to see that Conditions (4)–(6) guarantee that $\sigma_t^2 \geq 0$ whenever the σ_{t-i}^2 and ξ_{t-j}^2 in (3) are nonnegative.

Another useful way to guarantee the nonnegativity of σ_t^2 is to substitute for the lagged σ_t^2 terms in (3), writing σ_t^2 as an infinite distributed lag of ξ_t^2 terms—that is, in the terminology of Engle (1982), we write the GARCH(p, q) model in ARCH(∞) form:

$$\begin{aligned}\sigma_t^2 &= \left(1 - \sum_{i=1,p} \beta_i L^i\right)^{-1} \left[\omega + \sum_{j=1,q} \alpha_j \xi_{t-j}^2\right] \\ &= \omega^* + \sum_{k=0,\infty} \phi_k \xi_{t-k-1}^2,\end{aligned} \tag{7}$$

where L is the usual lag (or backshift) operator (i.e., $L(X_t) \equiv X_{t-1}$.). It is clear from (7) that, if ω^* and all of the ϕ_k are nonnegative, $0 \leq \sigma_t^2$. The nonnegativity of ω^* and all the ϕ_k is also necessary for σ_t^2 to remain nonnegative with probability 1 if σ_t^2 is strictly stationary and if for every positive integer n $\{\xi_t^2, \xi_{t-1}^2, \ldots \xi_{t-n}^2\}_{t=-\infty,\infty}$ is strictly stationary with support on the entire nonnegative orthant of R^n. To make ω^* and $\{\phi_k\}_{k=0,\infty}$ well defined, we assume that

$$\text{the roots of } \left(1 - \sum_{i=1,p} \beta_i Z^i\right) \text{ lie outside the unit circle.} \tag{8}$$

ω^* is then finite and nonnegative as long as $\omega \geq 0$. We also assume that

$$\text{the polynomials } \left(1 - \sum_{i=1,p} \beta_i Z^i\right) \text{ and } \sum_{j=1,q} \alpha_j Z^{j-1} \text{ have no common roots.} \quad (9)$$

Under (8)–(9), ω^* and $\{\phi_k\}_{k=0,\infty}$ are well defined and finite. This does not, however, guarantee that $\sigma_t^2 < \infty$ with probability 1 or that $\{\sigma_t^2\}_{t=-\infty,\infty}$ is strictly stationary. The conditions for strict stationarity are stronger; these were worked out for the GARCH(1, 1) case by Nelson (1990) and for the general GARCH(p, q) case by Bougerol and Picard (1992). Under (9), however, (8) is necessary for strict stationarity (Bougerol and Picard 1992).

Define $q^* \equiv \max\{q, p\}$, $\alpha_k \equiv 0$ for $k > q$, and $\beta_k \equiv 0$ for $k > p$. The ϕ_k terms can then be derived as the solution to the difference equation system

$$\phi_0 = \alpha_1$$

$$\phi_1 = \beta_1\phi_0 + \alpha_2$$

$$\phi_2 = \beta_1\phi_1 + \beta_2\phi_0 + \alpha_3$$

$$\cdots$$

$$\phi_{q^*-1} = \beta_1\phi_{q^*-2} + \beta_2\phi_{q^*-3} + \cdots + \beta_{q^*-1}\phi_0 + \alpha_{q^*}, \quad (10)$$

and for integer $k \geq q^*$,

$$\phi_k = \beta_1\phi_{k-1} + \beta_2\phi_{k-2} + \cdots + \beta_{q^*}\phi_{k-q^*}.$$

Requiring that all of the $\{\phi_k\}_{k=0,\infty}$ in (10) be nonnegative imposes an infinite number of inequality constraints on $\{\alpha_j\}_{j=1,q}$ and $\{\beta_i\}_{i=1,p}$. For practical purposes (e.g., in estimation) it is necessary to reduce this to a finite number of inequalities. Fortunately, in certain cases this is straightforward. The GARCH(0, q) [or ARCH(q)] case is trival (i.e., $\omega \geq 0$, $\alpha_j \geq 0$ for all j = 1 to q) and leads to no relaxation of the inequality constraints (4)–(6). In the GARCH(1, q) and GARCH(2, q) cases developed in detail in Section 2, however, we will see that (4)–(6) can be substantially relaxed. The more difficult GARCH(p, q) case for $p \geq 3$ is also briefly considered in Section 2. We give examples of the empirical relevance of the results in Section 2 in Section 3. A brief conclusion is found in Section 4.

2. MAIN RESULTS

2.1 GARCH(1, q)

In this case, $\beta_1 = \beta$, and $\beta_i = 0$ for $i \geq 2$, so our inequality constraints become

$$\omega^* = \omega/(1 - \beta) \geq 0, \quad (11)$$

$$\phi_0 = \alpha_1 \geq 0, \quad (12)$$

$$\phi_1 = \beta\alpha_1 + \alpha_2 \geq 0, \quad (13)$$

$$\phi_2 = \beta^2\alpha_1 + \beta\alpha_2 + \alpha_3 \geq 0, \quad (14)$$

$$\cdots$$

$$\phi_{q-1} = \beta^{q-1}\alpha_1 + \beta^{q-2}\alpha_2 + \cdots + \beta\alpha_{q-1} + \alpha_q \geq 0, \quad (15)$$

and then, for all integer $k \geq q$,

$$\phi_k = \beta^k \cdot \alpha_1 + \beta^{k-1} \cdot \alpha_2 + \cdots + \beta^{k+2-q} \cdot \alpha_{q-1} + \beta^{k+1-q} \cdot \alpha_q = \beta^{k+1-q} \cdot \phi_{q-1} \geq 0. \quad (16)$$

Clearly $\phi_{q-1} = 0$ only if (9) is violated, since in this case the model reduces to an ARCH(q). This, combined with (11)–(16), leads directly to the following result:

Theorem 1. Let (8)–(9) be satisfied. Then ω^* and $\{\phi_k\}_{k=0,\infty}$ are all nonnegative iff (a) $\omega \geq 0$, (b) $\beta \geq 0$, and (c) for all k = 0 to $q - 1$, $\phi_k = \Sigma_{j=0,k}\, \alpha_{j+1}\beta^{k-j} \geq 0$.

The proof is straightforward and is left to the reader.

For the popular GARCH(1, 1) model, Theorem 1 permits no relaxation in (4)–(6). For higher-order models, however, Theorem 1 relaxes (6) by allowing α_i to be negative for $i \geq 2$. In the GARCH(1, 2) model, for example, the conditions of Theorem 1 [along with (8)–(9)] are that (a) $\omega \geq 0$, (b) $0 \leq \beta < 1$, (c) $\beta\alpha_1 + \alpha_2 \geq 0$ and (d) $\alpha_1 \geq 0$. Figure 1 plots this region in (β, α_2/α_1) space when $\omega \geq 0$ and $\alpha_1 > 0$.

2.2 GARCH(2, q)

In this case it is convenient to reparameterize the model. Define Δ_1 and Δ_2 to be the roots of $(1 - \beta_1 Z^{-1} - \beta_2 Z^{-2})$ so that

$$1 - \beta_1 L - \beta_2 L^2 = (1 - \Delta_1 L)(1 - \Delta_2 L). \quad (17)$$

Without loss of generality, we assume that $|\Delta_1| \geq |\Delta_2|$, and that $\Delta_1 \neq 0$. If $\Delta_1 = -\Delta_2$, we take $\Delta_1 > 0$. We then have the following result:

Theorem 2. Let (8)–(9) be satisfied. Then Conditions (18)–(22) are necessary and sufficient for $\omega^* \geq 0$ and $\phi_k \geq 0$ for all nonnegative integer k:

$$\omega^* = \omega/[1 - \Delta_1 - \Delta_2 + \Delta_1\Delta_2] \geq 0, \quad (18)$$

$$\Delta_1 \text{ and } \Delta_2 \text{ are real numbers,} \quad (19)$$

$$\Delta_1 > 0, \quad (20)$$

$$\sum_{j=0,q-1} \Delta_1^{-j}\alpha_{j+1} > 0, \quad (21)$$

and

$$\phi_k \geq 0 \quad \text{for } k = 0 \text{ to } q. \quad (22)$$

The proof of Theorem 2 is given in the Appendix. One interesting special case of the GARCH(2, q) model is GARCH(2, 1), which we consider in the following corollary:

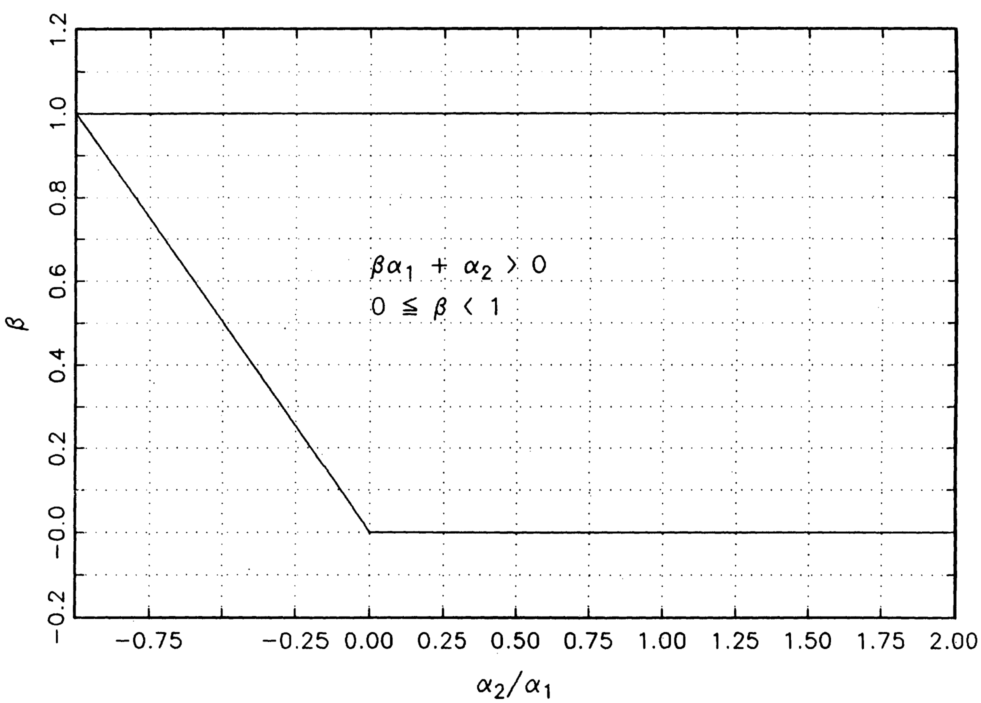

Figure 1. Admissible Parameter Values: The GARCH(1, 2) Case.

Corollary. In the GARCH(2, 1) model, Conditions (8) and (18)–(22) reduce to:

$$\omega \geq 0, \tag{23}$$

$$\alpha_1 \geq 0, \tag{24}$$

$$0 \leq \beta_1, \tag{25}$$

$$\beta_1 + \beta_2 < 1, \tag{26}$$

and

$$\beta_1^2 + 4\beta_2 \geq 0. \tag{27}$$

Figure 2 illustrates the region of (β_1, β_2) space allowed by the corollary to Theorem 2, assuming that α_1 and ω are positive. Again, Conditions (4)–(6) [with the unit circle condition (8) imposed as well] are substantially relaxed.

2.3 Higher-Order Systems

Deriving necessary and sufficient conditions for $\phi_k \geq 0$ for all k is substantially more difficult when $p \geq 3$. Sufficient conditions alone are a bit easier. For example, if the roots $\{\Delta_1, \ldots \Delta_p\}$ of $(1 - \beta_1 Z^{-1} - \beta_2 Z^{-2} - \cdots - \beta_p Z^{-p})$ are unique, we can write, for $k \geq \max\{q, p\}$,

$$\phi_k = \eta_1 \Delta_1^k + \eta_2 \Delta_2^k + \cdots + \eta_p \Delta_p^k, \tag{28}$$

where $\eta_1, \ldots, \eta_p$ are constants depending on $\phi_1, \ldots, \phi_{\max\{q-1,p\}}$. [For example, see Goldberg (1958) or Sargent (1987). Sargent (1987, pp. 192–194) provided closed-form expressions for the $\{\phi_k\}$.] If Δ_1 is real and positive and if we define $\eta^* \equiv \max_{\{j=2,p\}}|\eta_j|$ and $\Delta^* \equiv \max_{\{j=2,p\}}|\Delta_j|$, then clearly

$$\Delta_1^{-k}\phi_k \geq \eta_1 - (p-1) \cdot \eta^* \cdot (\Delta^*/\Delta_1)^k. \tag{29}$$

If, in addition, $\Delta_1 > |\Delta_j|$ for $j = 2$ to p and $\eta_1 > 0$, the (positive) first term on the right side of (29) dominates the (negative) second term as $k \to \infty$. If the right side of (29) is nonnegative for *some* k^*, it remains nonnegative for *all* $k \geq k^*$. Rearranging (29), it is clear that the right side of (29) must be positive for any k^* greater than $[\ln(\eta_1) - \ln(\eta^* \cdot (p-1))]/\ln(\Delta^*/\Delta_1)$. In this case, therefore, if $\{\phi_k\}_{k=0,k^*}$ is nonnegative, so is $\{\phi_k\}_{k=0,\infty}$.

Presumably, such sufficient (but not necessary) conditions should not be imposed in estimation. In practice, however, it is usually necessary to impose positivity on *in-sample* fitted values of σ_t^2 to keep nonlinear maximization routines from encountering overflows. For the ARCH(p), GARCH(1, q), and GARCH(2, q), the inequality constraints of Sections 2.1 and 2.2 should suffice. For higher-order GARCH models and multivariate GARCH, some other tactic is required. Probably the best method is a version of the penalty function method familiar in nonlinear programming (for example, Luenberger 1973): Insert an appropriate "IF" statement into the subroutine that evaluates the likelihood function. For each t, the "IF" statement should check that $\eta^{-1} \leq \sigma_t^2 \leq \eta$ for some large positive η before $\ln(\sigma_t^2)$ is evaluated. If σ_t^2 is too large or too small, the function evaluation is terminated and some very unfavorable function value returned. This approach can easily be adapted to the multivariate case.

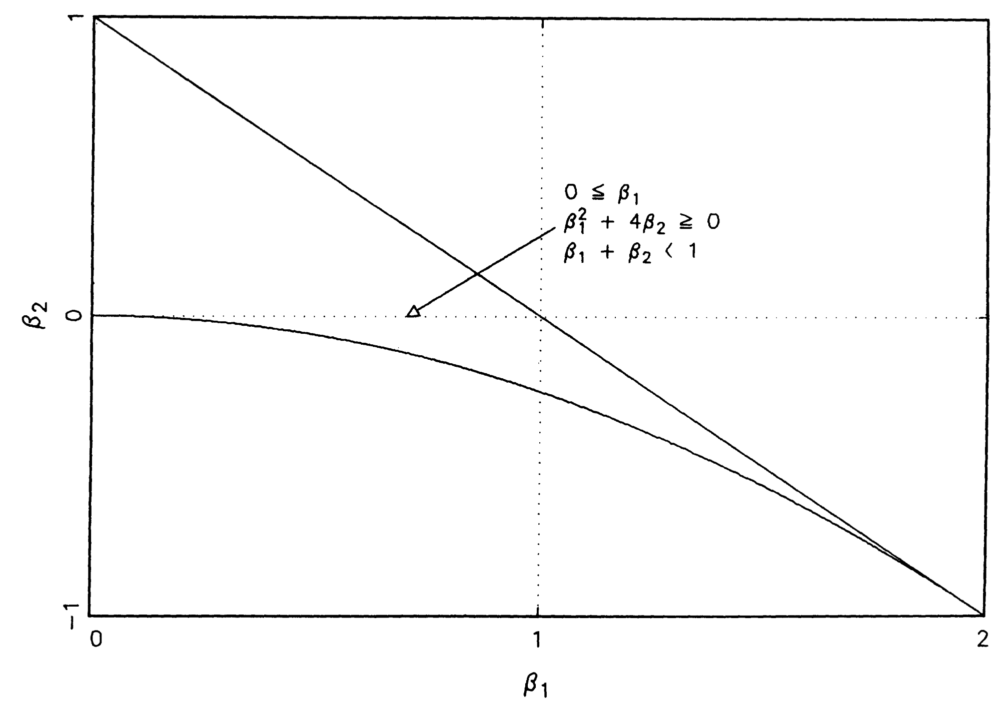

Figure 2. Admissible Parameter Values: The GARCH(2, 1) Case.

2.4 Start-up Values

Although the GARCH process σ_t^2 as written in (7) extends into the infinite past, in practice (e.g., in estimation) it is necessary to compute $\{\sigma_t^2\}$ recursively beginning at time 0, using (3) and assuming arbitrary fixed values for $\{\sigma_{-1}^2, \ldots, \sigma_{-p}^2, \xi_{-1}^2, \ldots, \xi_{-q}^2\}$. Although (4)–(6) guarantee that $\{\sigma_t^2\}_{t=0,\infty}$ remains nonnegative given arbitrary nonnegative $\{\sigma_{-1}^2, \ldots, \sigma_{-p}^2, \xi_{-1}^2, \ldots, \xi_{-q}^2\}$, the weaker condition that ω^* and $\{\phi_k\}_{k=0,\infty}$ are nonnegative does not guarantee this. Fortunately, it is not difficult to select start-up values that keep $\{\sigma_t^2\}_{t=0,\infty}$ nonnegative with probability 1 given nonnegative ω^* and $\{\phi_k\}_{k=0,\infty}$. One way to do this is to arbitrarily choose any $\xi^2 \geq 0$ and set $\xi_t^2 \equiv \xi^2$ for all $t = -1$ to $-\infty$. Then set $\sigma_t^2 \equiv \sigma^2$ for $1 - p \leq t \leq 0$, where

$$\sigma^2 \equiv \left(1 - \sum_{i=1,p} \beta_i\right)^{-1} \left[\omega + \xi^2 \sum_{j=1,q} \alpha_j\right]$$
$$= \omega^* + \xi^2 \sum_{k=0,\infty} \phi_k. \quad (30)$$

This keeps σ_t^2 nonnegative for all $t \geq 0$ with probability 1, since

$$\sigma_t^2 = \omega^* + \sum_{k=0,t-1} \phi_k \xi_{t-k-1}^2 + \sum_{k=t,\infty} \phi_k \xi^2 \geq 0. \quad (31)$$

If $\Sigma_{i=1,p} \beta_i + \Sigma_{j=1,q} \alpha_j < 1$, (which is not required for strict stationarity of $\{\xi_t\}$), one can set σ^2 and ξ^2 equal to their (common) unconditional mean; that is,

$$\sigma^2 \equiv \xi^2 \equiv \omega \Big/ \left(1 - \sum_{i=1,p} \beta_i - \sum_{j=1,q} \alpha_j\right). \quad (32)$$

3. EMPIRICAL EXAMPLES

Several violations of Bollerslev's original inequality constraints have been reported in the ARCH literature, and we suspect that many more would be reported were it not that many researchers see these violations as evidence of misspecification or of sampling error. Unfortunately, there is no easy way to verify this suspicion; unless researchers actually report negative coefficient values, it is usually impossible to tell whether the Bollerslev inequality constraints were imposed or not. In the widely circulated GARCH estimation code of Kroner (1990), the Bollerslev inequality constraints are not automatically imposed but can be imposed at the user's option.

Some of the reported violations of the Bollerslev inequality constraints violate our weaker constraints as well. For example, Engle (1983) and Engle, Lilien, and Robins (1987) found that they had to impose linearly declining weights on the ARCH(p) model to prevent some parameters from becoming negative. Our inequality constraints are no weaker than Bollerslev's for the ARCH(p) case. Hence this violation results from sampling error or misspecification. [These are the only possibilities, since Engle (1983) and Engle et al. (1987) assumed conditional normality, making the support of the errors unbounded.]

Violations of the Bollerslev inequality constraints are rarer for GARCH models. For example, Bollerslev (1986) analyzed the same Consumer Price Index data as Engle (1983) but found that a GARCH(1, 1) model fit the data well with no inequality constraint violations.

Several violations of the Bollerslev inequality constraints in GARCH models have been reported, however, usually satisfying the weaker inequality constraints of Theorems 1 and 2.

One example was provided by French, Schwert, and Stambaugh (1987, table 5), who estimated a GARCH(1, 2)-M model for daily capital gains on the Standard and Poor's (S&P) 500 from 1928 to 1984 and found estimates for the σ_t^2 process of

$$\underset{(6\cdot 10^{-8})}{\sigma_t^2 = 6.3\cdot 10^{-7}} + \underset{(.003)}{.918\cdot\sigma_{t-1}^2} + \underset{(.007)}{.121\cdot\xi_{t-1}^2} - \underset{(.007)}{.043\cdot\xi_{t-2}^2}, \tag{33}$$

where ξ_t is the time t residual in returns (standard errors are in parentheses). The final term on the right side of (33) (i.e., α_2) is several standard deviations below 0. French et al. (1987) estimated the model over various subperiods and with an alternative specification for the conditional mean. In each case, however, the estimated α_2 was negative, usually significantly so. [However, α_2 becomes insignificantly different from 0 when a Student's likelihood is used in place of a normal likelihood—see Baillie and DeGennaro (1990)]. Nevertheless, *all* of the models that French et al. estimated using daily data satisfy the conditions of Theorem 1. On the other hand, French et al. (1987, table 6a) also estimated GARCH(1, 2)-M models using monthly returns data, and in one subperiod (1953–1984) the estimated α_1 coefficient was negative (though insignificant), violating the conditions of Theorem 1.

We estimated GARCH models up to the order of GARCH(1, 3) and GARCH(2, 2), using S&P 500 daily returns data from 1928 to 1989, and found the same pattern in all subperiods; significantly negative estimated α_2's are common. We found no examples of negative estimated β_2's in GARCH(2, q) models, however, so for this data the inequality constraint relaxations in Theorem 1 are more empirically relevant than those in Theorem 2. All of our estimated models satisfied the conditions of Theorems 1 and 2.

One of the models estimated by Engle, Ito, and Lin (1990, table IV) provides another example: They estimated a GARCH(1, 4) model for exchange-rate movements, treating the exchange-rate movements in different world markets as separate observations. Their estimated model was

$$\underset{(2\cdot 10^{-4})}{\sigma_t^2 = 6\cdot 10^{-4}} + \underset{(.0087)}{.9581\cdot\sigma_{t-1}^2} + \underset{(.0204)}{.1169\cdot\xi_{t-1}^2} - \underset{(.0292)}{.0627\cdot\xi_{t-2}^2} - \underset{(.0292)}{.0047\cdot\xi_{t-3}^2} - \underset{(.0103)}{.0181\cdot\xi_{t-4}^2}. \tag{34}$$

Since only the α_2 term is significantly negative, one might be tempted to impose the Bollerslev inequality constraints in estimation. There is no need, however, since the estimated model satisfies the conditions of Theorem 1.

To further check the empirical relevance of Theorems 1 and 2, we next reconsider the exchange-rate series analyzed by Baillie and Bollerslev (1989). Tim Bollerslev graciously provided the data. Our estimated model is slightly different from Baillie and Bollerslev's; they included day-of-the-week dummy variables, which we omit, and we use data from June 1, 1973, to January 28, 1985, whereas they used data from March 1, 1980, to January 28, 1985. We estimated GARCH models of order up to GARCH(2, 2) and GARCH(1, 3) for the British pound/dollar, Deutschmark/dollar, yen/dollar, franc/dollar, and lira/dollar exchange rates, respectively. We encountered many instances of negative estimated α's and β's in every case satisfying the requirements of Theorems 1 and 2. In three cases out of the five, the model selected by the Akaike information criterion (AIC) (Akaike 1973) involves a negative α, though never a negative β. The results for the models favored by the AIC are reported in Table 1. Using the model-selection criterion of Schwarz (1978), models involving negative estimated α's are selected in two out of five cases. In this data, as in the S&P 500 data, Theorem 1 appears highly empirically relevant, Theorem 2 less so.

4. CONCLUSION

Although negative coefficients in GARCH models may result from misspecification or sampling error, this is not always the case. Our weaker set of sufficient conditions to guarantee that $\sigma_t^2 \geq 0$ almost surely for all t is empirically relevant, as the examples in Section 3

Table 1. GARCH Models for Daily Exchange Rate (selected by AIC) June 1, 1973–January 28, 1985

Exchange rate	Selected model	b_0	ω	α_1	α_2	α_3	β_1	β_2	LF
British pound	GARCH(1, 2)	−.0219	.0060	.1872	−.0812		.8884		−2,289.94
		(.0093)	(.0004)	(.0141)	(.0146)		(.0058)		
Deutschmark	GARCH(2, 2)	−.0009	.0186	.0573	.2262		.3833	.3100	−2,447.24
		(.0091)	(.0021)	(.0170)	(.0203)		(.0925)	(.0760)	
Japanese yen	GARCH(1, 3)	.0016	.0002	.1888	.0752	−.2344	.9730		−2,086.27
		(.0067)	(.0001)	(.0176)	(.0236)	(.0089)	(.0011)		
French franc	GARCH(1, 2)	−.0002	.0079	.1024	.1444		.7735		−2,356.21
		(.0063)	(.0005)	(.0175)	(.0188)		(.0097)		
Italian lira	GARCH(1, 3)	−.0044	.0003	.3058	−.0485	−.0573	.8451		−1,593.92
		(.0036)	(.0001)	(.0159)	(.0265)	(.0210)	(.0057)		

NOTE: The numbers in the parentheses are the standard deviations. LF is log-likelihood. The best-fitted models are selected by the Akaike information criterion (Akaike 1973). $100\log(s_t/s_{t-1}) = b_0 + \varepsilon_t$, $\varepsilon_t|I_{t-1} \sim N(0, \sigma_t^2)$, and $\sigma_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_p\sigma_{t-p}^2$.

indicate. Practitioners should therefore probably *not* impose the Bollerslev inequality constraints in estimation.

ACKNOWLEDGMENTS

We thank Phillip Braun, an associate editor, and two anonymous referees for helpful comments. The Center for Research in Security Prices and the National Science Foundation provided financial support.

APPENDIX: PROOF OF THEOREM 2

The necessity and sufficiency of (18) for $\omega^* \geq 0$ is obvious, so we turn to (19). For $k \geq q+1$, ϕ_k evolves according to the homogeneous difference equation $\phi_k = \beta_1\phi_{k-1} + \beta_2\phi_{k-2}$. When Δ_1 and Δ_2 are complex, we may write (for example, see Fuller 1976, pp. 43–45; Goldberg 1958, chap. 3)

$$\phi_k = b \cdot \Delta_1^k + b^* \cdot \Delta_2^k, \qquad \text{(A.1)}$$

where $b = r \cdot [\cos(\gamma) + i \cdot \sin(\gamma)]$ (with $r \geq 0$) is a constant depending on ϕ_q and ϕ_{q-1} and b^* is its complex conjugate; Δ_1 can be written as $\rho \cdot [\cos(\theta) + i \cdot \sin(\theta)]$ (with $\rho > 0$, since $\Delta_1 \neq 0$) and Δ_2 is its complex conjugate. Substituting for Δ_1, Δ_2, b, and b^* in (A.1) we may write

$$\rho^{-k}\phi_k = 2 \cdot r \cdot \cos[\gamma + \theta k], \qquad \text{(A.2)}$$

which has an oscillating sign as $k \to \infty$ unless θ is either an integer multiple of 2π or $r = 0$. We can rule out θ being an integer multiple of 2π, since this would make Δ_1 and Δ_2 real numbers. We can also rule out $r = 0$, since this would make $\phi_k = 0$ for all sufficiently large k, reducing the model to an ARCH(p) for finite p, thereby violating (9). Therefore, (19) is proved.

For the real roots case, note that under condition (8)

$$[(1 - \Delta_1 L)(1 - \Delta_2 L)]^{-1} = \sum_{i=0,\infty}\sum_{j=0,\infty} \Delta_1^i \Delta_2^j L^{i+j} = \sum_{k=0,\infty} \gamma_k L^k, \qquad \text{(A.3)}$$

where

$$\gamma_k \equiv \Delta_1^k \sum_{j=0,k} (\Delta_2/\Delta_1)^j. \qquad \text{(A.4)}$$

It is then easy to verify that

$$\phi_k = \sum_{j=0,\min\{k,q-1\}} \gamma_{k-j} \cdot \alpha_{j+1}. \qquad \text{(A.5)}$$

First, suppose that $\Delta_1 = \Delta_2 \equiv \Delta$. Equation (A.5) becomes

$$\phi_k = \sum_{j=0,\min\{k,q-1\}} \Delta^{k-j}(k - j + 1) \cdot \alpha_{j+1}. \qquad \text{(A.6)}$$

Obviously, (22) is necessary if $\phi_k \geq 0$ for all k. To see that (20) and (21) are also necessary, take $k \geq q - 1$, and divide (A.6) by $|\Delta|^k$. We obtain

$$|\Delta|^{-k}\phi_k = \operatorname{sign}(\Delta^k) \sum_{j=0,q-1} [\Delta^{-j}(1 - j) \cdot \alpha_{j+1} + k \cdot \Delta^{-j} \cdot \alpha_{j+1}]. \qquad \text{(A.7)}$$

Clearly the second term in the summation dominates as $k \to \infty$ unless $\Sigma_{j=0,q-1} \Delta^{-j}\alpha_{j+1} = 0$, which would violate (9). To keep the sign of this asymptotically dominant term positive, (20) and (21) are clearly necessary.

To see that (19)–(22) are also sufficient for $\phi_k \geq 0$ for all k, take $k \geq q - 1$ and rewrite (A.7) as

$$\Delta^{-k}\phi_k = (k + 1) \cdot \sum_{j=0,q-1} \Delta^{-j} \cdot \alpha_{j+1} - \sum_{j=0,q-1} \Delta^{-j} \cdot j \cdot \alpha_{j+1}. \qquad \text{(A.8)}$$

Under (19)–(22), the first term on the right side of (A.8) is positive and increasing in k, whereas the second term is constant, so it is clear that if $\Delta^{-k}\phi_k$ is nonnegative when $k = q$, it is nonnegative whenever $k > q$. Clearly therefore (19)–(22) imply $\phi_k \geq 0$ for all k.

Next, suppose that Δ_1 and Δ_2 are real and distinct. Equation (A.5) becomes

$$\phi_k = (\Delta_1 - \Delta_2)^{-1} \sum_{j=0,\min\{k,q-1\}} \times (\Delta_1^{k+1-j} - \Delta_2^{k+1-j}) \cdot \alpha_{j+1}. \qquad \text{(A.9)}$$

First we show that (20)–(22) are necessary for $\phi_k \geq 0$ for all k. The necessity of (22) is obvious. To see that (20)–(21) are also necessary, suppose first that $\Delta_1 = -\Delta_2 > 0$. Equation (20) holds trivially, and for $k \geq q - 1$, the nonnegativity of (A.9) is equivalent to

$$2\Delta_1^{-k}\phi_k = \sum_{j=0,q-1} \Delta_1^{-j} \times [1 - (-1)^{k+1-j}] \cdot \alpha_{j+1} \geq 0. \qquad \text{(A.10)}$$

Since we require $2\Delta_1^{-k}\phi_k \geq 0$ and $2\Delta_1^{-k-1}\phi_{k+1} \geq 0$, clearly the sum $2\Delta_1^{-k}\phi_k + 2\Delta_1^{-k-1}\phi_{k+1} = 2 \cdot \Sigma_{j=0,q-1} \Delta_1^{-j}\alpha^{j+1} \geq 0$, implying (21). Next, suppose that $\Delta_1 \neq -\Delta_2$. Since $|\Delta_1| > |\Delta_2|$, the $\Delta_1^{k+1} \Sigma_{j=0,q-1} \Delta_1^{-j}\alpha_{j+1}$ term dominates (A.9) as $k \to \infty$. If $\Delta_1 < 0$, the sign of this term oscillates as $k \to \infty$ unless $\Sigma_{j=0,q-1} \Delta_1^{-j}\alpha_{j+1} = 0$, which would violate (9). To keep this term positive, (20) and (21) are therefore necessary.

Finally, we show that (19)–(22) are sufficient for $\phi_k \geq 0$ for all k. Again, suppose first that $\Delta_1 = -\Delta_2 > 0$. Clearly, if in (A.10) $\phi_{k^*} \geq 0$ and $\phi_{k^*+1} \geq 0$, then $\phi_k \geq 0$ for all $k \geq k^*$. Therefore if $\phi_k \geq 0$ for $k = 1$ to q, $\phi_k \geq 0$ for all k. Finally, suppose that $\Delta_1 \neq -\Delta_2$. Under (19)–(22) we may write, for $k \geq q - 1$,

$$\Delta_1^{-k-1}(\Delta_1 - \Delta_2)\phi_k = \sum_{j=0,q-1} \Delta_1^{-j}\alpha_{j+1}[1 - (\Delta_2/\Delta_1)^{k+1-j}]. \qquad \text{(A.11)}$$

By (21), the $\Sigma_{j=0,q-1} \Delta_1^{-j}\alpha_{j+1}$ term is positive and asymptotically dominates the right side of (A.11). The $-\Sigma_{j=0,q-1} \Delta_1^{-j}\alpha_{j+1}(\Delta_2/\Delta_1)^{k+1-j}$ term is of declining magnitude (but possiblly oscillating sign) as $k \to \infty$. Once again, if in (A.11) $\phi_{k^*} \geq 0$, then $\phi_k \geq 0$ for all $k \geq k^*$. Therefore if $\phi_k \geq 0$ for $k = 1$ to q, $\phi_k \geq 0$ for all k.

The extension to the corollary is straightforward and is left to the reader.

[Received January 1991. Revised September 1991.]

REFERENCES

Akaike, H. (1973), "Information Theory and an Extension of the Maximum Likelihood Principle," in *Second International Symposium on Information Theory*, eds. B. N. Petrov and F. Csaki, Budapest: Akademiai Kiado, pp. 267–281.

Baillie, R. T., and Bollerslev, T. (1989), "The Message in Daily Exchange Rates: A Conditional Variance Tale," *Journal of Business & Economic Statistics*, 7, 297–305.

Baillie, R. T., and DeGennaro, R. P. (1990), "Stock Returns and Volatility," *Journal of Financial and Quantitative Analysis*, 25, 203–214.

Bollerslav, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Ecnometrics*, 31, 307–327.

Bollerslev, T., Chou, R. Y., and Kroner, K. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 37.

Bougerol, P., and Picard, N. (1992), "Stationarity of GARCH Processes and of Some Non-Negative Time Series," *Journal of Econometrics*, 37.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1008.

——— (1983), "Estimates of the Variance of U.S. Inflation Based Upon the ARCH Model," *Journal of Money, Credit, and Banking*, 15, 286–301.

Engle, R. F., Ito, T., and Lin, W. L. (1990), "Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market," *Econometrica*, 58, 525–542.

Engle, R. F., Lilien, D. M., and Robins, R. P. (1987), "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model," *Econometrica*, 55, 391–407.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987), "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3–29.

Fuller, W. A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.

Goldberg, S. (1958), *Introduction to Difference Equations*, New York: Dover.

Kroner, K. F. (1990), "MGARCH," unpublished FORTRAN code, University of Arizona, Dept. of Economics.

Luenberger, D. G. (1973), *Introduction to Linear and Nonlinear Programming*, Reading, MA: Addison-Wesley.

Nelson, D. B. (1990), "Stationarity and Persistance in the GARCH(1, 1) Model," *Econometric Theory*, 6, 318–334.

Sargent, T. J. (1987), *Macroeconomic Theory* (2nd ed.), Boston: Academic Press.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464.