Econometrica, Vol. 61, No. 4 (July, 1993), 909–927

TEMPORAL AGGREGATION OF GARCH PROCESSES

BY FEIKE C. DROST AND THEO E. NIJMAN[1]

We derive low frequency, say weekly, models implied by high frequency, say daily, ARMA models with symmetric GARCH errors. Both stock and flow variable cases are considered. We show that low frequency models exhibit conditional heteroskedasticity of the GARCH form as well. The parameters in the conditional variance equation of the low frequency model depend upon mean, variance, and kurtosis parameters of the corresponding high frequency model. Moreover, strongly consistent estimators of the parameters in the high frequency model can be derived from low frequency data in many interesting cases. The common assumption in applications that rescaled innovations are independent is disputable, since it depends upon the available data frequency.

KEYWORDS: ARMA-GARCH, GARCH models, heteroskedasticity, temporal aggregation.

1. INTRODUCTION

IT IS WELL KNOWN, nowadays, that many financial time-series such as exchange rates and stock returns exhibit conditional heteroskedasticity, i.e. big shocks are clustered together. GARCH models are often used to parameterize conditional heteroskedasticity. The GARCH model generalizes the ARCH model of Engle (1982) and is proposed by Bollerslev (1986). In applications, GARCH models have been specified for data at different frequencies, typically assuming that the rescaled innovations are i.i.d. and are generated by either normal or t distributions. Implicitly it is assumed that a GARCH process at one frequency, say daily, is consistent with some GARCH process at another frequency, say weekly. The aggregation properties of ARIMA models are well-known: high frequency ARIMA processes aggregate to low frequency ARIMA processes. For an extensive literature we refer to, e.g., Amemiya and Wu (1972), Harvey and Pierse (1984), Palm and Nijman (1984), Lütkepohl (1986), and Nijman and Palm (1990a, b). Little is known about the impact of temporal aggregation upon GARCH processes. Only the limiting cases of an increasing sampling interval and of an increasing sampling frequency have been considered in the literature. Diebold (1988) shows that conditional heteroskedasticity disappears if the sampling time interval increases to infinity. In case of flow variables the implied marginal low frequency distribution converges to the normal distribution. Nelson (1990) considers an increasing sampling frequency. A continuous time model is derived that yields accurate approximations to high frequency data. This model is close to I-GARCH. See also Drost and Nijman (1992b).

[1] Herman Bierens, James Davidson, Svend Hylleberg, the co-editor, and three anonymous referees provided helpful comments. The authors also benefited from discussions with the participants of the International Conference on ARCH Models and Applications to Financial or Monetary Econometrics (Paris, June 25–26, 1990). Both authors gratefully acknowledge financial support of the Royal Netherlands Academy of Arts and Sciences (K.N.A.W.).

It is the purpose of this paper to derive results on temporal aggregation over a finite number of periods. We show that the classical GARCH assumptions are not robust to the specification of the sampling interval. Independent daily rescaled innovations, e.g., imply dependent rescaled innovations at the weekly frequency. In applied work these dependencies are neglected (cf., e.g., Baillie and Bollerslev (1989)). The dependencies also complicate attempts to construct efficient semi-parametric estimators of the variance parameters (cf., e.g., Gallant and Tauchen (1989) and Engle and González-Rivera (1991)).

Three definitions of GARCH are adopted in this paper, which will be referred to as respectively strong, semi-strong, and weak GARCH. The respective definitions are of increasing generality. Strong GARCH requires that rescaled innovations are independent, semi-strong GARCH assumes that rescaled innovations are uncorrelated, while in weak GARCH models only projections of the conditional variance are considered.

In this paper, first of all, we show that the classical (semi-)strong GARCH assumptions on the available data frequency are arbitrary. Generally a (semi-)strong GARCH process aggregates to some weak GARCH process that is not semi-strong GARCH. Second, we show that the assumption of symmetric weak GARCH models at different frequencies is internally consistent. More generally, we show that every ARMA model with symmetric weak GARCH errors aggregates to a model in this class. Third, our results imply that strongly consistent estimation of low frequency parameters is possible with the low frequency data set. A straightforward consistent estimator of these parameters can be derived from the ARMA model that generates the squared observations (cf. Bollerslev (1988)). Simulations suggest that the popular quasi maximum likelihood estimates are also close to the true parameters even if the low frequency model is weak GARCH (see Drost and Nijman (1992a)). Finally we note that high frequency parameters can be identified from the corresponding low frequency ones in many interesting cases.

The paper is organized as follows. Notations and several GARCH concepts are presented in Section 2. Temporal aggregation of the commonly used GARCH(1, 1) model is considered in Section 3. The class of weak GARCH(1, 1) models appears to be closed both in case of stock and flow variables. In both cases (semi-)strong GARCH(1, 1) only aggregates to weak GARCH(1, 1); see Examples 3 and 4. In the stock variable case the low frequency conditional variance parameters depend only upon the high frequency variance parameters. In the case of flow variables the low frequency variance parameters depend also upon the kurtosis of the high frequency observations. In both cases the low frequency parameters are expressed as explicit functions of the high frequency parameters. Our main results are presented in Section 4: the class of ARMA models with weak GARCH errors is closed under temporal aggregation. The proof of this result is deferred to the Appendix. Our results are easily extended to ARIMA models. Various examples illustrate the main theorem. In general the high frequency orders of the ARMA part of ARMA-GARCH models influence the low frequency orders of the GARCH part (see Table I). E.g., temporal aggregation of an ARMA(1,1) model with GARCH(1,1) errors leads to low frequency ARMA(1,1)-GARCH(2,2) models. In Section 5 we illustrate the empirical implications of our results. We compare the daily, weekly, and monthly models of six major exchange rates presented in Baillie and Bollerslev (1989). Finally Section 6 contains some concluding remarks.

2. DEFINITIONS AND NOTATION

In order to define the models properly let $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ be a sequence of stationary errors with finite fourth moments. Define operators $A(L) = 1 + \sum_{i=1}^{q} \alpha_i L^i$ and $B(L) = 1 - \sum_{i=1}^{p} \beta_i L^i$ and let the sequence $\{h_t,\ t \in \mathbb{Z}\}$ be defined as the stationary solution of

$$(1) \qquad B(L) h_t = \psi + \{A(L) - 1\} \varepsilon_t^2.$$

We assume that $B(L)$ and $B(L) + 1 - A(L)$ have roots outside the unit circle and hence are invertible.[2]

Three definitions of GARCH will be adopted in this paper.

DEFINITION 1 (Strong GARCH): The sequence $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ is defined to be *generated by a strong GARCH(p,q) process* if ψ, $A(L)$, and $B(L)$ can be chosen such that

$$(2) \qquad \xi_t = \varepsilon_t \Big/ \sqrt{h_t} \sim \text{i.i.d. } D(0,1),$$

where $D(0,1)$ specifies a distribution with mean zero and unit variance.

DEFINITION 2 (Semi-strong GARCH): The sequence $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ is defined to be *generated by a semi-strong GARCH(p,q) process* if ψ, $A(L)$, and $B(L)$ can be chosen such that

$$(3) \qquad E[\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = 0 \qquad \text{and}$$

$$(4) \qquad E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = h_t.$$

DEFINITION 3 (Weak GARCH): The sequence $\{\varepsilon_t,\ t \in \mathbb{Z}\}$ is defined to be *generated by a weak GARCH(p,q) process* if ψ, $A(L)$, and $B(L)$ can be chosen such that

$$(5) \qquad P[\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = 0 \qquad \text{and}$$

$$(6) \qquad P[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots] = h_t,$$

where $P[x_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots]$ denotes the best linear predictor of x_t in terms of $1, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots,$ i.e.

$$(7) \qquad E(x_t - P[x_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots]) \varepsilon_{t-i}^r = 0 \quad \text{for} \quad i \geqslant 1 \quad \text{and} \quad r = 0, 1, 2.$$

[2] In the literature attention is restricted to parameter values satisfying $\psi > 0$, $\beta_i > 0$, and $\alpha_i > 0$ $(\forall i)$. However, this condition seems to be unnecessarily restrictive. Defining $\pi(L) = \sum_{i=1}^{\infty} \pi_i L^i = B(L)^{-1}\{A(L) - 1\}$ the weaker assumption $\psi > 0$ and $\pi_i > 0$ $(\forall i)$ also guarantees the nonnegativity of h_t. Moreover, one can easily construct examples to show that even the nonnegativity of π_i $(\forall i)$ is not necessary.

Observe that all these GARCH definitions require $\Sigma_{i=1}^{p}\beta_i + \Sigma_{i=1}^{q}\alpha_i < 1$. The strong GARCH definition has been adopted by, e.g., Engle (1982) and Bollerslev (1986). The most popular distributions are normal and t distributions. The second definition has been adopted by, e.g., Weiss (1986). Evidently a strong GARCH process will also be semi-strong GARCH. On the other hand a semi-strong GARCH process with time-varying higher order conditional moments of the rescaled innovations $\xi_t = \varepsilon_t/\sqrt{h_t}$ is not strong GARCH (see, e.g., Example 3). Finally the requirements for weak GARCH are met both by strong and semi-strong GARCH processes. Observe that the weak GARCH definition is quite general and captures the characterizing features of the other GARCH formulations. As we will prove below, it is still possible to obtain strongly consistent estimators of the GARCH parameters in this general formulation.

The most general model considered in this paper is the ARMA model with GARCH errors:

$$\Gamma(L)y_t = \Theta(L)\varepsilon_t, \tag{8}$$

where $\Gamma(L) = \Pi_{i=1}^{P}(1-\gamma_i L)$ and $\Theta(L) = \Pi_{i=1}^{Q}(1-\theta_i L)$. Throughout we assume that all standard regularity conditions are fulfilled. This implies that the roots of $\Gamma(L)$ and $\Theta(L)$ are all outside the unit circle and that no roots of $\Gamma(L)$ coincide with roots of $\Theta(L)$.

High frequency observations are assumed to be on y_t $(t = 1, \dots, T)$. If y_t is a stock variable, low frequency observations are assumed to be on y_t $(t = m, 2m, \dots, T)$, where m is some known integer (for simplicity we suppose that T is a multiple of m). If y_t is a flow variable, low frequency observations are assumed to be on $\bar{y}_{(m)t} = \Sigma_{i=0}^{m-1} y_{t-i}$ $(t = m, 2m, \dots, T)$ with m and T as before. Note that this case applies if, e.g., $y_t = \Delta x_t$, where x_t is a stock variable. Extensions to ARIMA models are easily added by constructing slightly more general observation schemes (compare, e.g., Palm and Nijman (1984, Section 2)).

3. AGGREGATION OF GARCH(1,1)

In this section we consider temporal aggregation of GARCH(1,1) models with either stock or flow variables. Example 1 considers stock variables and Example 2 presents aggregation results in case of flow variables. Higher order models are treated in a similar manner. The discussion of the general case is deferred to Section 4. In both cases the class of symmetric weak GARCH(1,1) models is shown to be closed against temporal aggregation. The proofs of these results rely upon aggregation techniques in ARMA models. The high frequency parameters ψ, β, α, and κ_y determine the corresponding low frequency parameters. Explicit formulas for the aggregated GARCH(1,1) model are available. The unconditional kurtosis $\kappa_y = Ey_t^4/(Ey_t^2)^2$ of the observations influences the low frequency GARCH parameters only in case of flow variables. Each example is followed by a discussion of the implications of aggregation. In both cases it is easy to construct (semi-)strong GARCH(1,1) processes showing that the class of (semi)-strong GARCH(1,1) models is not closed under temporal aggregation; see Examples 3 and 4.

EXAMPLE 1 (GARCH(1, 1), Stocks): The class of symmetric weak GARCH(1, 1) processes with stock variables is closed under temporal aggregation. More precisely stated, if $\{y_t,\ t \in \mathbb{Z}\}$ is weak GARCH(1, 1) with symmetric marginal distributions and $h_t = \psi + \beta h_{t-1} + \alpha y_{t-1}^2$ then $\{y_{tm},\ t \in \mathbb{Z}\}$ is symmetric weak GARCH(1, 1) with $h_{(m)tm} = \psi_{(m)} + \beta_{(m)} h_{(m)tm-m} + \alpha_{(m)} y_{tm-m}^2$, where

$$\text{(9)} \qquad \psi_{(m)} = \psi \frac{1-(\beta+\alpha)^m}{1-(\beta+\alpha)}, \qquad \alpha_{(m)} = (\beta+\alpha)^m - \beta_{(m)},$$

and $\beta_{(m)} \in (0,1)$ is the solution of the quadratic equation

$$\text{(10)} \qquad \frac{\beta_{(m)}}{1+\beta_{(m)}^2} = \frac{\beta(\beta+\alpha)^{m-1}}{1+\alpha^2 \dfrac{1-(\beta+\alpha)^{2m-2}}{1-(\beta+\alpha)^2} + \beta^2(\beta+\alpha)^{2m-2}}.$$

PROOF: We restrict attention to $m = 2$. Along the same lines one obtains results for general m. Obviously relation (5) is satisfied for $\{y_{2t},\ t \in \mathbb{Z}\}$. To derive the projection of y_t^2 consider the ARMA(1, 1) model that is known to generate the squared observations (cf. Bollerslev (1988)). Put $\eta_t = y_t^2 - h_t$, observe that the η_t are uncorrelated (use equation (7) with $x_t = y_t^2$) and rewrite the equation determining h_t:

$$\text{(11)} \qquad y_t^2 = \psi + (\beta+\alpha) y_{t-1}^2 + \eta_t - \beta \eta_{t-1}.$$

To obtain the low frequency GARCH model we derive the low frequency model corresponding to the ARMA(1, 1) model in squared observations. We proceed as in Palm and Nijman (1984). Substituting (11) into itself yields

$$\text{(12)} \qquad y_t^2 = \psi(1+\beta+\alpha) + (\beta+\alpha)^2 y_{t-2}^2 + v_t$$

with $v_t = \eta_t + \alpha\eta_{t-1} - \beta(\beta+\alpha)\eta_{t-2}$. Equation (12) determines the autoregressive part of the low frequency model. It remains to determine the moving average structure. It is easily checked that $(\forall k > 1)\ Ev_t v_{t-2k} = 0$. Put $\omega_t = (1 - \lambda L^2)^{-1} v_t$, where λ is such that the ω_t with even indices are uncorrelated, i.e. $-\lambda/(1+\lambda^2) = Ev_t v_{t-2}/Ev_t^2$. Since ω_t is a linear combination of $y_t^2, y_{t-2}^2, \ldots$, it follows that

$$\begin{aligned} &P\left[y_t^2 | y_{t-2}, y_{t-4}, \ldots\right] - \lambda P\left[y_{t-2}^2 | y_{t-4}, y_{t-6}, \ldots\right] \\ &\quad = \psi(1+\beta+\alpha) + \left\{(\beta+\alpha)^2 - \lambda\right\} y_{t-2}^2, \end{aligned}$$

proving that $\{y_{2t},\ t \in \mathbb{Z}\}$ is weak GARCH(1, 1) with parameters $\psi_{(2)} = \psi(1+\beta+\alpha)$, $\beta_{(2)} = \lambda$, and $\alpha_{(2)} = (\beta+\alpha)^2 - \lambda$ (replacing the high frequency parameters ψ, β, and α). *Q.E.D.*

DISCUSSION: The class of ARCH(1) models is also closed under temporal aggregation in case of stock variables ($\beta_{(m)} = 0$ if $\beta = 0$). Note that (10) is a simple quadratic problem admitting a real solution in closed form. Observe that

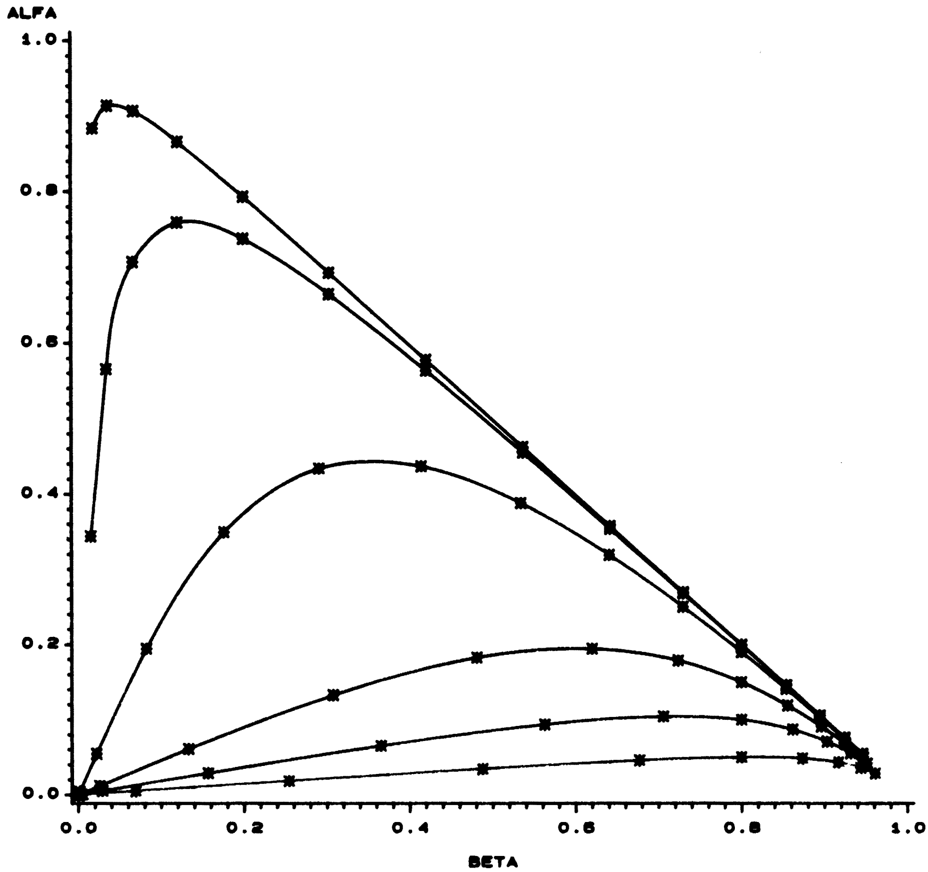

FIGURE 1.—Aggregation of GARCH(1, 1): Stocks. Subsequent marks indicate the variance parameters of weak GARCH(1, 1) models generated by repeated doubling or halving the sampling interval, starting with models where $\beta = .8$ and $\alpha = .05, .1, .15, .19, .199$, and $.1999$ respectively.

$\beta_{(m)} + \alpha_{(m)} = (\beta + \alpha)^m$ tends to zero as m tends to infinity. Hence, conditional heteroskedasticity disappears in the limit if the process is aggregated more and more (compare, e.g., Diebold (1988)). On the other side, assume for the moment that the current observed process is a low frequency process generated by some very high frequency GARCH(1, 1) process. Then the data generating process (DGP) of the original series is close to I-GARCH(1, 1) with $\beta \approx 1$ and $\alpha \approx 0$ (compare, e.g., Nelson (1990) and Drost and Nijman (1992b)).

Next we derive strongly consistent estimators of the high and low frequency parameters only based upon low frequency data. Consider the low frequency ARMA(1, 1) model that is known to generate the squared observations (compare Bollerslev (1988)):

$$y_{tm}^2 = \psi_{(m)} + \left(\beta_{(m)} + \alpha_{(m)}\right) y_{tm-m}^2 + \eta_{tm} - \beta_{(m)} \eta_{tm-m},$$

where $\eta_{tm} = y_{tm}^2 - h_{(m)tm}$ and observe that the η_{tm} are uncorrelated (use equation (7)). Assume that $\{y_{tm}^2\}$ is ergodic. Then one easily derives strongly consistent estimators of the GARCH parameters because the vector process $\{(y_{tm}^2, y_{tm-m}^2, y_{tm-2m}^2)\}$ is also ergodic. This implies that the sample mean and the first two sample autocorrelations converge almost surely and a simple one-one relation between these limits and the GARCH parameters determines the required estimators. The assumption of ergodicity of the low frequency process is, e.g., trivially satisfied if there exists an underlying high frequency strong GARCH process. Then the low frequency GARCH parameters can be consistently estimated using low frequency data. This implies that consistent estimation of the high frequency parameters, based on low frequency data only, is possible (the high frequency parameters are uniquely determined by the corresponding low frequency ones). The problem of multiple high frequency models that are consistent with the low frequency evidence, that can arise in ARMA(1,1) models (cf., e.g., Palm and Nijman (1984)), is absent in the GARCH(1,1) model because of the restriction that all parameters are nonnegative.

The relation between high and low frequency GARCH models is displayed in Figure 1. Subsequent marks to the left on the lines in the figure indicate the effect of doubling the sampling period, while subsequent marks to the right indicate the effect of doubling the sampling frequency. The six lines are generated by doubling and halving the sampling interval starting with models where $\beta = .8$ and $\alpha = .05$, .1, .15, .19, .199, and .1999 respectively. E.g., if the observed sample is weak GARCH with $\beta = .800$ and $\alpha = .050$, then the corresponding parameters for the weak GARCH model where observations with odd indexes are skipped ($m = 2$) are $\beta = .677$ and $\alpha = .046$. Similarly it follows from the figure that the parameter pairs for $m = 4$ and $m = 8$ are equal to (.488, .034) and (.254, .018). If the observations are generated by a very high frequency DGP, then the parameters corresponding to the model where the sampling frequency is doubled are $\beta = .873$ and $\alpha = .048$. Doubling once more and redoubling yields respectively (.917, .043) and (.944, .036). Of course the latter values are only valid if the true underlying DGP contains the calculated frequency. Observe that GARCH(1,1) models are close to ARCH(1) models if the sampling interval is large and that conditional heteroskedasticity disappears when the sampling period is very large. On the other side the figure shows that highly aggregated models with nontrivial variance parameters are generated by a DGP close to the integration in variance model.

Example 2 (GARCH(1,1), Flows): The class of symmetric weak GARCH(1,1) processes with flow variables is closed under temporal aggregation. More precisely stated, if $\{y_t,\ t \in \mathbb{Z}\}$ is weak GARCH(1,1) with symmetric marginal distributions, $h_t = \psi + \beta h_{t-1} + \alpha y_{t-1}^2$, and unconditional kurtosis κ_y, then $\{\bar{y}_{(m)tm},\ t \in \mathbb{Z}\}$ is symmetric weak GARCH(1,1) with

$$\bar{h}_{(m)tm} = \bar{\psi}_{(m)} + \bar{\beta}_{(m)} \bar{h}_{(m)tm-m} + \bar{\alpha}_{(m)} \bar{y}_{(m)tm-m}^2,$$

and kurtosis $\bar{\kappa}_{(m)y}$ where

$$\bar{\psi}_{(m)} = m\psi\frac{1-(\beta+\alpha)^m}{1-(\beta+\alpha)}, \qquad \bar{\alpha}_{(m)} = (\beta+\alpha)^m - \bar{\beta}_{(m)}, \tag{13}$$

$$\begin{aligned}\bar{\kappa}_{(m)y} = 3 + (\kappa_y - 3)/m + 6(\kappa_y - 1) \\ \times \frac{\{m-1-m(\beta+\alpha)+(\beta+\alpha)^m\}\{\alpha-\beta\alpha(\beta+\alpha)\}}{m^2(1-\beta-\alpha)^2(1-\beta^2-2\beta\alpha)},\end{aligned} \tag{14}$$

and $|\bar{\beta}_{(m)}| < 1$ is the solution of the quadratic equation

$$\frac{\bar{\beta}_{(m)}}{1+\bar{\beta}_{(m)}^2} = \frac{a(\beta,\alpha,\kappa_y,m)(\beta+\alpha)^m - b(\beta,\alpha,m)}{a(\beta,\alpha,\kappa_y,m)\{1+(\beta+\alpha)^{2m}\} - 2b(\beta,\alpha,m)}, \tag{15}$$

with

$$\begin{aligned} a(\beta,\alpha,\kappa_y,m) \\ = m(1-\beta)^2 + 2m(m-1)\frac{(1-\beta-\alpha)^2(1-\beta^2-2\beta\alpha)}{(\kappa_y-1)\{1-(\beta+\alpha)^2\}} \\ + 4\frac{\{m-1-m(\beta+\alpha)+(\beta+\alpha)^m\}\{\alpha-\beta\alpha(\beta+\alpha)\}}{1-(\beta+\alpha)^2},\end{aligned} \tag{16}$$

$$b(\beta,\alpha,m) = \{\alpha - \beta\alpha(\beta+\alpha)\}\frac{1-(\beta+\alpha)^{2m}}{1-(\beta+\alpha)^2}. \tag{17}$$

REMARK 1: For strong GARCH models the distribution of the rescaled innovations is usually given. To make the aggregation results directly applicable to this situation the relation between the kurtosis of the rescaled innovations $\kappa_\xi = E\xi^4$ and κ_y is given:

$$\kappa_y = \kappa_\xi \frac{1-(\beta+\alpha)^2}{1-(\beta+\alpha)^2-(\kappa_\xi-1)\alpha^2}. \tag{18}$$

PROOF: Once more we restrict derivations to the case with $m = 2$. Similar but tedious calculations prove the general case. Equation (5) is easily checked. In order to derive the projection $P[\bar{y}_{(2)t}^2 | \bar{y}_{(2)t-2}, \bar{y}_{(2)t-4}, \ldots]$ let $\eta_t = y_t^2 - h_t$ be defined as in the preceding example. Then equation (12) implies

$$\bar{y}_{(2)t}^2 = 2\psi(1+\beta+\alpha) + (\beta+\alpha)^2 \bar{y}_{(2)t-2}^2 + \bar{v}_t,$$

where

$$\begin{aligned}\bar{v}_t = \eta_t + (1+\alpha)\eta_{t-1} + \{\alpha - \beta(\beta+\alpha)\}\eta_{t-2} - \beta(\beta+\alpha)\eta_{t-3} \\ + 2y_t y_{t-1} - 2(\beta+\alpha)^2 y_{t-2}y_{t-3}.\end{aligned}$$

As before we derive the low frequency ARMA(1,1) equation corresponding to this model in squared observations. The components of $\bar{v}_t$ are uncorrelated and hence $(\forall k > 1)$ $E\bar{v}_t\bar{v}_{t-2k} = 0$. Put $\bar{\omega}_t = (1 - \bar{\lambda}L^2)^{-1}\bar{v}_t$, where $\bar{\lambda}$ is such that the $\bar{\omega}_{2t}$ are uncorrelated. Thus the low frequency ARMA(1,1) equation is given by

$$\bar{y}_{(2)t}^2 = 2\psi(1+\beta+\alpha) + (\beta+\alpha)^2\bar{y}_{(2)t-2}^2 + \bar{\omega}_t - \bar{\lambda}\bar{\omega}_{t-2}.$$

Finally a simple projection argument implies that the low frequency model is weak GARCH(1,1) with parameters $\bar{\psi}_{(2)} = 2\psi(1+\beta+\alpha)$, $\bar{\beta}_{(2)} = \bar{\lambda}$, and $\bar{\alpha}_{(2)} = (\beta+\alpha)^2 - \bar{\lambda}$ (replacing the parameters ψ, β, and α). The other conclusions of the proposition are easy.

The main difference with the case of stock variables is the presence of the cross-products $y_t y_{t-1}$ and $y_{t-2}y_{t-3}$ in $\bar{v}_t$. This complicates the computation of $\bar{\lambda}$ since $Ey_t^2y_{t-1}^2/E\eta_t^2$ has to be expressed in κ_y, α, and β. More details are provided in Drost and Nijman (1992a). *Q.E.D.*

Discussion: Observe that the class of weak ARCH(1) models with flow variables is not closed under temporal aggregation. Generally high frequency ARCH(1) models aggregate to low frequency GARCH(1,1) with $\bar{\beta}_{(m)} \neq 0$. The relations (13)–(17) imply that the variance parameters of the low frequency model depend upon the high frequency variance parameters but also on the kurtosis of the observations. This contrasts the case of flow variables with the case of stock variables. As in the preceding example $\bar{\beta}_{(m)} + \bar{\alpha}_{(m)} = (\beta+\alpha)^m$, implying that conditional heteroskedasticity disappears when m is large and that very high frequency processes are close to I-GARCH. Moreover $\bar{\kappa}_{(m)y} \to 3$ as $m \to \infty$, suggesting asymptotic normality of $\bar{y}_{(m)t}/\sqrt{m}$. This result has been established by Diebold (1988). As before the one-one correspondence between high and low frequency parameters and the ARMA equation in squared observations implied by the GARCH model permits consistent estimation of high frequency parameters from low frequency data.

The results of temporal aggregation of GARCH(1,1) models with flow variables are illustrated in Figure 2. Six different GARCH DGP's are considered. The parameter vectors of β, α, and κ_y are given by (.871, .051, 3.11), (.887, .035, 3.05), (.800, .050, 3.06), (.871, .051, 10.47), (.887, .035, 9.62), and (.800, .050, 9.70). For strong GARCH models these kurtosis values correspond to normal rescaled innovations for the first three parameter vectors and t_5 distributed rescaled innovations for the last ones (see Remark 1, equation (18)). Similar to Figure 1 subsequent marks to the left indicate the effect of doubling the sampling period. Stars "$*$" correspond to low frequency models obtained with the first three parameter vectors and diamonds "$\diamond$" to low frequency models obtained with the last ones. From the figure it is evident that differences in the high frequency kurtosis produce different low frequency variance parameters. E.g., aggregation of the first, respectively fourth, model with $m = 2$ yields (.800, .050, 3.26), respectively (.777, .073, 7.68). The first and fifth parameter vectors are chosen such that the parameters α and β of the third pair are

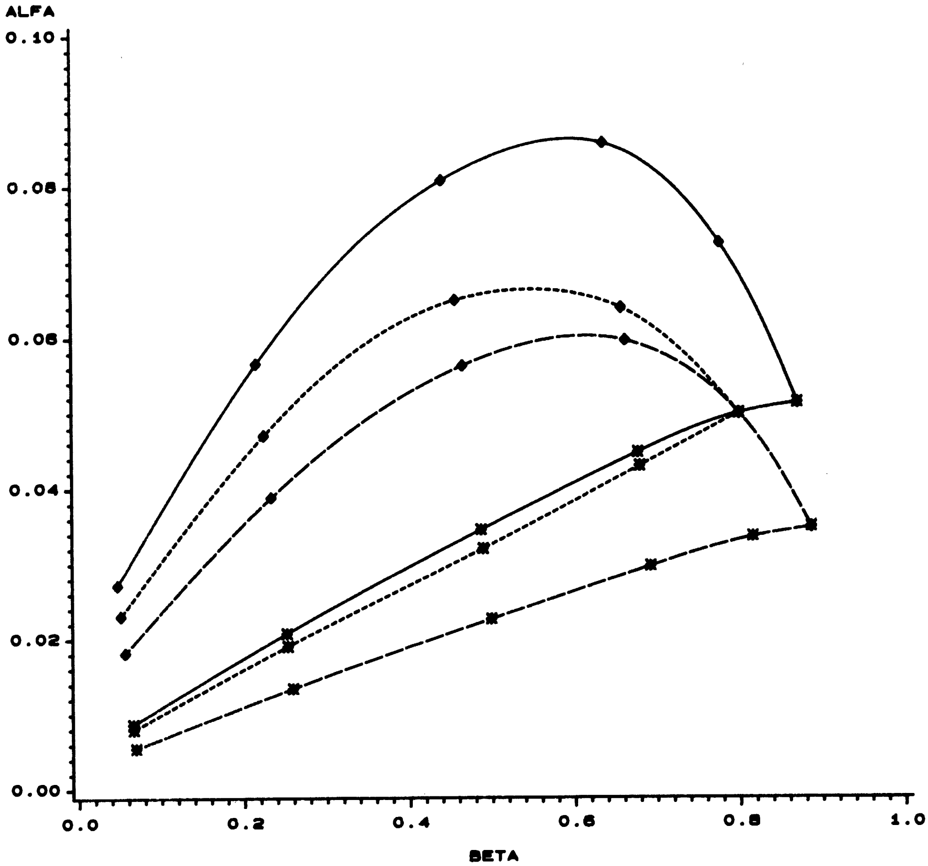

FIGURE 2.—Aggregation of GARCH(1, 1): Flows. Subsequent marks indicate the variance parameters of weak GARCH(1, 1) models generated by repeated doubling the sampling interval, starting with strong GARCH(1, 1) models where $(\beta, \alpha) = (.887, .035)$, $(.871, .051)$, and $(.800, .050)$. The rescaled innovations are either normal ($*$) or student distributed with five degrees of freedom ($\diamond$).

obtained if the sampling period is doubled. This illustrates that knowledge of the low frequency variance parameters is insufficient to identify the high frequency ones unless information about $\bar{\kappa}_{(m)y}$ is available. From the applied point of view, nevertheless, the impact of the kurtosis on the low frequency variance parameters appears to be minor. Numerical results for exchange rates will be presented in Section 5.

Drost and Nijman (1992b) show that the kurtosis $\bar{\kappa}_{(m)y}$ is a function of the variance parameters only, rather than an independent parameter, if it is assumed that the data are generated by a GARCH model at arbitrarily high frequency. If the existence of an underlying continuous time model is assumed, high frequency parameters can be identified from low frequency data only.

EXAMPLE 3 (Strong GARCH not closed): To show that the class of strong GARCH models is not closed under temporal aggregation consider the classical model

$$y_t = \xi_t \sqrt{\psi + \alpha y_{t-1}^2} = \xi_t \sqrt{h_t}$$

with $\xi_t \sim$ i.i.d. $N(0,1)$. In the low frequency stock model with, e.g., $m=2$, relation (2) is violated. Let $\nu_t = y_t/\sqrt{h_{(2)t}} = y_t/\sqrt{\psi(1+\alpha)+\alpha^2 y_{t-2}^2}$ denote the rescaled low frequency innovations and note that

$$\nu_t = \xi_t \sqrt{\lambda_t + \xi_{t-1}^2(1-\lambda_t)} \tag{19}$$

with $\lambda_t = \psi/h_{(2)t}$. Therefore rescaled low frequency innovations depend upon past observations. From (19) one can show that the conditional fourth moment of ν_t depends on the information set. Rescaled innovations are not i.i.d. and hence the low frequency model implied by a high frequency strong ARCH(1) model does not satisfy the strong GARCH assumption. Therefore the common assumption that rescaled disturbances are i.i.d. at the available data frequency is disputable. Ideally, economic theory should indicate at which time intervals innovations occur. It is easily checked that in this particular example the low frequency model is semi-strong GARCH. This property does not generalize to the case of flow variables or to higher order models.

EXAMPLE 4 (Semi-strong GARCH not closed): To show that the class of semi-strong GARCH models is not closed under temporal aggregation consider the model

$$y_t = \xi_t \sqrt{\psi + \alpha y_{t-1}^2} = \xi_t \sqrt{h_t}$$

where the ξ_t are i.i.d. with the distribution determined by $P\{\xi_t = 0\} = 1-\alpha$ and $P\{\xi_t = -1/\sqrt{\alpha}\} = P\{\xi_t = 1/\sqrt{\alpha}\} = \alpha/2$. A possible realization of the process is, e.g., $\sqrt{\psi/\alpha}(0,-1,\sqrt{2},\sqrt{3},-\sqrt{4},0,0,1,0,0,0,1,-\sqrt{2},0,\dots)$. In the low frequency flow model with, e.g., $m=2$ the semi-strong GARCH definition is violated. Assume, for instance, that $\{\bar{y}_{(2)2t}\}$ is semi-strong GARCH(p,q) for some p and q. Then it is clear that $p=q=1$ since the conditional expectations are necessarily equal to the projections in Example 2. Thus, with $\bar{\psi}$, $\bar{\alpha}$, and $\bar{\beta}$ determined by (13)–(17),

$$\bar{h}_{(2)t} = E\left[\bar{y}_{(2)t}^2 | \bar{y}_{(2)t-2}, \bar{y}_{(2)t-4}, \dots\right] = \bar{\psi} + \bar{\alpha}\bar{y}_{(2)t-2}^2 + \bar{\beta}\bar{h}_{(2)t-2}.$$

However, the squared process $\{\alpha y_t^2/\psi\}$ consists of sequences of successive integers of different length starting at zero. Hence if $\bar{y}_{(2)t} = \sqrt{\psi n/\alpha}$ for some integer $n \geqslant 2$, then $y_t = 0$ and $y_{t-1} = \sqrt{\psi n/\alpha}$. This remark implies, for $n \geqslant 2$,

$$\begin{aligned} &E\left[\bar{y}_{(2)t}^2 | \bar{y}_{(2)t-2} = \sqrt{\psi n/\alpha}, \bar{y}_{(2)t-4}, \dots\right] \\ &\quad = E\left[y_t^2 + y_{t-1}^2 + 2y_t y_{t-1} | y_{t-2} = 0, y_{t-3} = \sqrt{\psi n/\alpha}, \bar{y}_{(2)t-4}, \dots\right] \\ &\quad = E\left[\psi + (1+\alpha)y_{t-1}^2 | y_{t-2} = 0, y_{t-3} = \sqrt{\psi n/\alpha}, \bar{y}_{(2)t-4}, \dots\right] \\ &\quad = \psi(2+\alpha). \end{aligned}$$

TABLE I

UPPER BOUNDS ON THE ORDERS OF LOW FREQUENCY MODELS IMPLIED BY HIGH FREQUENCY MODELS.

High Frequency Model		Low Frequency Model
ARCH(q)	stock	GARCH($q-1,q$)
ARCH(q)	flow	GARCH(q,q)
GARCH(1,1)	stock	GARCH(1,1)
GARCH($1,q^{a}$)	stock	GARCH($q-1,q$)
GARCH($1,q$)	flow	GARCH(q,q)
MA(1)-ARCH(q)	stock	GARCH(q,q)
MA(1)-ARCH(q)	flow	MA(1)-GARCH($q+1,q+1$)
AR(1)-ARCH(q)	stock	AR(1)-GARCH(q,q)
AR(1)-ARCH(q)	flow	ARMA(1,1)-GARCH($q+1,q+1$)
ARMA(1,1)-GARCH($1,q$)	stock	ARMA(1,1)-GARCH($q+1,q+1$)
ARMA(1,1)-GARCH($1,q$)	flow	ARMA(1,1)-GARCH($q+1,q+1$)

[a] $q \geqslant 2$.

The RHS does not depend on n and $\bar{y}_{(2)t-4}, \bar{y}_{(2)t-6}, \ldots$. Hence $\bar{\alpha} = \bar{\beta} = 0$, contradicting both the values derived above and the assumption of semi-strong GARCH.

4. AGGREGATION OF ARMA-GARCH

In this section we derive the low frequency ARMA-GARCH model corresponding to a general high frequency ARMA model with weak GARCH errors. As a corollary it follows that the class of ARMA-GARCH models is closed under temporal aggregation both in the cases of stock and flow variables. Furthermore we present two corollaries for pure GARCH models. The order of low frequency models are explicitly given and some important cases are summarized in Table I. Low frequency parameters are determined by high frequency mean, kurtosis, and variance parameters. Numerical techniques are often necessary to derive the low frequency parameters (similar to the ARIMA case; c.f., Palm and Nijman (1984)). The proof of Theorem 1 is deferred to the Appendix. A relaxation of the symmetry condition is also discussed in the Appendix. In this paragraph $[x]$ is the largest integer less than or equal to x.

THEOREM 1 (ARMA-GARCH): *Let $\{y_t, \ t \in \mathbb{Z}\}$ be generated by the ARMA(P,Q) model* (8) *with symmetric weak GARCH(p,q) errors determined by* (1). *Put $W(L) = \sum_{i=0}^{w} w_i L^i$; then $\{W(L)y_{tm}, \ t \in \mathbb{Z}\}$ follows an ARMA($\tilde{P},\tilde{Q}$) process with symmetric weak GARCH($\tilde{p},\tilde{q}$) errors, where $\tilde{P} = P$, $\tilde{Q} = P + [(Q - P + w)/m]$, and $\tilde{p} = \tilde{q} = r + \frac{1}{2}\tilde{Q}(\tilde{Q}+1)$, where $r = \max(p,q)$.*

If $\prod_{i=1}^{P}(1 + \gamma_i L + \cdots + \gamma_i^{m-1}L^{m-1})\Theta(L)W(L) = \tilde{\Theta}(L^m)$ for some polynomial $\tilde{\Theta}$, then the low frequency error process is GARCH($r + [(p-r)/m], r$).

COROLLARY 2: *The classes of ARMA models with symmetric weak GARCH errors and either stock or flow variables are closed under temporal aggregation.*

COROLLARY 3 (GARCH, Stocks): *If* $\{y_t,\ t \in \mathbb{Z}\}$ *is symmetric weak GARCH*(p,q), *then* $\{y_{tm},\ t \in \mathbb{Z}\}$ *is symmetric weak GARCH*$(r + [(p-r)/m], r)$.

COROLLARY 4 (GARCH, Flows): *If* $\{y_t,\ t \in \mathbb{Z}\}$ *is symmetric weak GARCH*(p,q), *then* $\{\bar{y}_{(m)tm},\ t \in \mathbb{Z}\}$ *is symmetric weak GARCH*(r,r).

In Table I we present upper bounds on the orders of some low frequency models. Observe that the orders of the variance equation are influenced by the properties of the mean equation. The calculated true low frequency orders may be smaller than the calculated upper bounds. Consider, e.g., a high frequency ARMA(1, 1)-GARCH(1, 1) model. Then the corresponding low frequency models with $m = 2$ and $m = 4$ are ARMA(1, 1)-GARCH(2, 2). However, if we started at the level $m = 2$ with an ARMA(1, 1)-GARCH(2, 2) model, Theorem 1 suggests an ARMA(1, 1)-GARCH(3, 3) model for $m = 4$. Hence if the model under consideration is already aggregated, there may be intricate connections between the parameters. These connections influence the orders of more highly aggregated models. Furthermore seasonal terms in the polynomials may influence the orders of the low frequency model. In the proof of Theorem 1 we did not exploit possible special structures in the ARMA polynomials determining the mean equation and the GARCH part. If these polynomials contain seasonal factors of type $1 - cL^m$, then other factorizations can be used to show that the orders of the low frequency ARMA-GARCH model are smaller.

Note also that high frequency parameters can be identified from low frequency parameters in many interesting cases. The aliasing problem implying that infinitely many high frequency models are consistent with the low frequency parameters (well-known for ARMA models; see, e.g., Phillips (1973), Hansen and Sargent (1983), and Palm and Nijman (1984)) is generally absent in GARCH models. The condition from Palm and Nijman (1984) that the number of autoregressive parameters is at least equal to the number of moving average parameters is satisfied in the ARMA model for squared observations if $p \neq q$ or $\beta_p \neq \alpha_q$. In higher order GARCH models the existence of a finite number of observationally equivalent GARCH processes cannot be excluded (compare the results on ARMA processes of Nijman and Palm (1990a)).

5. EMPIRICAL EXAMPLE: EXCHANGE RATES

In order to illustrate the empirical implications of the results in the previous sections we compare the estimated daily, weekly, and monthly models of six exchange rates presented in Baillie and Bollerslev (1989) (BB from now on). BB analyzes observations on the exchange rates of the French Franc, Italian Lira, Japanese Yen, Swiss Franc, British Pound, and German Mark with respect to the US Dollar from the New York Exchange Market between March 1, 1980 and January 28, 1985. BB assumes that rescaled innovations in the daily and weekly GARCH(1, 1) model for the returns are t distributed. Their estimates of the parameters for the daily and weekly model are reproduced in the first six

TABLE II

ESTIMATES OF WEEKLY PARAMETERS FOR SIX EXCHANGE RATES. COMPARISON OF DIRECT WEEKLY ESTIMATES WITH ESTIMATES IMPLIED BY DAILY ESTIMATES.

	Daily Estimates			Weekly Estimates			Implied Weekly		
	β	α	κ_ξ	$\bar{\beta}_{(5)}$	$\bar{\alpha}_{(5)}$	$\bar{\kappa}_{(5)\xi}$	$\bar{\beta}_{(5)}$	$\bar{\alpha}_{(5)}$	$\bar{\kappa}_{(5)\xi}$
FF/$	.829	.114	4.92	.655	.144	5.13	.589	.157	5.81
IL/$	.848	.113	3.89	.658	.187	3.00	.663	.157	5.13
JY/$	.941	.049	5.62	.927	.072	6.65	.839	.112	4.71
SF/$	.907	.073	3.41	.784	.121	3.00	.792	.112	4.05
BP/$	.910	.061	4.16	.842	.049	3.00	.768	.096	4.09
GM/$	.881	.085	3.41	.636	.249	3.00	.728	.113	4.11

columns of Table II. Weiss (1986) and Bollerslev and Wooldridge (1992) have shown that the quasi maximum likelihood estimator (QMLE), that is based upon strong GARCH and conditional normal distributions, is consistent if the conditional variance of the semi-strong GARCH process is correctly specified. Similar results are not available for QMLE based upon conditional t distributions (used by BB). Moreover, this paper shows that the assumption of semi-strong exchange returns at all frequencies is internally inconsistent. No analytical results are available for QMLE applied to weak GARCH models. Drost and Nijman (1992a) have carried out a number of simulation experiments where data were generated by strong GARCH models. Subsequently the parameters at lower frequencies were estimated by QMLE. For large samples the true GARCH parameters were close to the estimated ones. These results suggest that the asymptotic bias of the QMLE, if there is any, is small. For ease of reference with BB we present the calculated κ_ξ of the rescaled innovations in Table II instead of κ_y although not all presented processes can be strong GARCH. The κ_y are obtained along the lines of Remark 1. Note that the dummy variables contained in the BB model drop out in the weekly model or are insignificant. We use the results of Example 2 to estimate the variance and kurtosis parameters in the weekly GARCH(1, 1) model. Plugging in the daily estimates in equations (13)–(18) we obtain an alternative estimation procedure to direct weekly estimates. These estimates are given in the last three columns of Table II. Except for the Japanese Yen and given the large standard errors (cf. Baillie and Bollerslev (1989)) the direct estimators are close to the implied weekly estimates. The latter estimates, based on daily data, are probably better. The problems with the Japanese Yen are probably caused by the high kurtosis. BB's estimate of the kurtosis in the weekly JY/$ model equals $\kappa_\xi = 6.65$, implying that the fourth moment of the observations does not exist. This probably invalidates both the estimation procedure and our aggregation results for the JY/$ case.

From the results in Example 2 it is clear that the high frequency GARCH(1, 1) model can be identified from low frequency data. Hourly parameter estimates are obtained by plugging in daily estimates in equations (13)–(18) (we take 8 hours a day and we neglect possible seasonal patterns). We also present the high frequency variance parameter estimates if κ_ξ in the high frequency model

TABLE III

ESTIMATES OF HOURLY PARAMETERS FOR SIX EXCHANGE RATES IMPLIED BY DAILY ESTIMATES.

	Daily Data			Hourly Estimates Implied by Daily Estimates with κ_ξ from Superimposed N			Superimposed t_6		
	$\beta_{(1/8)}$	$\alpha_{(1/8)}$	$\kappa_{(1/8)\xi}$	$\beta_{(1/8)}$	$\alpha_{(1/8)}$	$\bar{\kappa}_{(1)\xi}$	$\beta_{(1/8)}$	$\alpha_{(1/8)}$	$\bar{\kappa}_{(1)\xi}$
FF/$	.957	.035	7.36	.934	.058	3.99	.954	.039	4.63
IL/$	.936	.059	2.66	.941	.054	3.96	.959	.036	4.60
JY/$	.990	.009	18.12	.978	.020	3.32	.985	.014	3.78
SF/$	.958	.039	2.31	.965	.032	3.53	.976	.022	4.04
BP/$	.979	.017	7.62	.967	.029	3.42	.977	.019	3.90
GM/$	.928	.067	1.69	.956	.040	3.65	.969	.027	4.19

is assumed to be known (in that case the estimated kurtosis in the low frequency model is neglected). We compare two superimposed values: the low value $\kappa_\xi = 3$ (e.g., normal innovations) and the relatively high value $\kappa_\xi = 6$ (e.g., t_6 innovations). Observe that the estimated daily kurtosis values in Table II are between these two values. For these cases we present the implied daily kurtosis instead of repeating the superimposed hourly kurtosis. Table III illustrates that the parameters do not change dramatically if the kurtosis varies in a reasonable range. A warning applies at this moment. Calculation of high frequency parameters is not always possible by plugging in the low frequency parameters in equations (13)–(18). A solution does not always exist. If the inserted parameters are the true parameters, this implies that an underlying high frequency model does not exist. Disaggregating more and more is not feasible. In applications, however, the fact that no solutions exist can be caused by discrepancies between true and estimated parameters. This problem of embeddability is well-known in the literature of Markov models. Stroock and Varadhan (1979) present general conditions such that a sequence of discrete time models converges to an Ito process. See also De Haan and Karandikar (1989). Nelson (1990) and Drost and Nijman (1992b) discuss the implications of embeddability in the context of GARCH processes.

Finally we present in Table IV the monthly variance and kurtosis estimates implied by the daily and weekly ones. In both cases the monthly model still contains strong conditional heteroskedasticity, although BB cannot reject the homoskedasticity assumption by direct estimation of the monthly parameters. The estimates in Table IV seem to be better than the direct estimates based on 59 monthly observations. Note that the estimated weekly parameters in the JY/$ model imply that the fourth moment of the returns does not exist. Hence we cannot apply the temporal aggregation results in this case.

6. CONCLUDING REMARKS

In this paper we derived the low frequency model that is implied by an assumed high frequency GARCH model. We restricted ourselves primarily to properties of the parameters in the mean and variance equations. Moreover we showed that the i.i.d. assumption on the rescaled innovations at the data frequency which one happens to have available is arbitrary. The low frequency

TABLE IV

ESTIMATES OF MONTHLY PARAMETERS FOR SIX EXCHANGE RATES. COMPARISON OF MONTHLY ESTIMATES IMPLIED BY DAILY AND WEEKLY ESTIMATES.

	Monthly Estimates Implied by					
	Daily Estimates			Weekly Estimates		
	$\bar{\beta}_{(20)}$	$\bar{\alpha}_{(20)}$	$\bar{\kappa}_{(20)\xi}$	$\bar{\beta}_{(20)}$	$\bar{\alpha}_{(20)}$	$\bar{\kappa}_{(20)\xi}$
FF/$	.206	.103	6.27	.299	.109	5.56
IL/$	.325	.126	6.04	.391	.119	4.48
JY/$	.661	.157	5.48	—[a]	—[a]	—[a]
SF/$	.553	.115	4.54	.570	.101	3.90
BP/$	.472	.083	4.17	.593	.037	3.26
GM/$	.411	.090	4.36	.426	.187	6.26

[a] Temporal aggregation results not applicable since fourth moment does not exist.

variance parameters generally depend on mean, variance, and kurtosis parameters of the high frequency model. Furthermore we showed that the orders of the low frequency GARCH process can be affected by properties of the high frequency mean equation. Identification of the parameters in a high frequency strong GARCH model from low frequency data is often possible. In addition we showed how high frequency observations can be used to obtain estimates of the low frequency variance parameters which are likely to be better than direct estimates from low frequency data. A menu-driven computer program yielding aggregation and disaggregation results for GARCH(1, 1) models is available on request from the authors. Estimates of parameters in a variance equation for monthly exchange rates derived in this way suggest strong conditional heteroskedasticity, as opposed to direct estimates. Simulation results show that the classical quasi maximum likelihood estimator yields parameter estimates close to the true parameters of weak GARCH models. Although we do not claim that this estimator is consistent, these simulations suggest that the bias is negligible in applications.

Dept. of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

Manuscript received November, 1990; final revision received December, 1992.

APPENDIX

PROOF OF THEOREM 1: We introduce some more notation. If $P(L)=\prod_{i=1}^{p}(1-p_iL)$ is some polynomial of order p, define $P_m(L^m)$ and $\bar{P}_m(L)$ by

$$P_m(L^m)=\prod_{i=1}^{p}(1-p_i^mL^m),\quad \text{and}$$

$$\bar{P}_m(L)=\prod_{i=1}^{p}\left(1+p_iL+\cdots+p_i^{m-1}L^{m-1}\right).$$

Note that $P_m(L^m)=\bar{P}_m(L)P(L)$.

First we derive the ARMA structure of the low frequency process $\tilde{y}_t = W(L)y_t$ ($t = \dots, m, 2m, 3m, \dots$). Multiplying the ARMA equation (8) by $\bar{\Gamma}_m(L)W(L)$ we obtain

$$\Gamma_m(L^m)\tilde{y}_t = \bar{\Gamma}_m(L)\Theta(L)W(L)\varepsilon_t = \nu_t.$$

The autoregressive part of the low frequency model is evident from this equation. To determine the moving average structure note that $(\forall k > \tilde{Q})\, E\nu_t\nu_{t-mk} = 0$. This determines the order of the MA part. To derive the MA parameters proceed as in Palm and Nijman (1984) and construct a polynomial $\tilde{\Theta}(L^m) = \sum_{i=0}^{\tilde{Q}}\tilde{\theta}_i L^{im} = 1 + \tilde{\theta}_1 L^m + \cdots + \tilde{\theta}_{\tilde{Q}} L^{\tilde{Q}m}$ such that $\zeta_{tm} = \tilde{\Theta}(L^m)^{-1}\nu_{tm}$ is an uncorrelated sequence. Explicit determination of the MA parameters $\tilde{\theta}_i$ ($i = 1, \dots, \tilde{Q}$) usually requires numerical procedures. Combining these results the low frequency ARMA equation is given by

$$\Gamma_m(L^m)\tilde{y}_t = \tilde{\Theta}(L^m)\zeta_t.$$

Next we derive the GARCH structure of this equation. By construction the ζ_{tm} are uncorrelated and hence the proof is complete if the projection $P[\zeta_t^2|\zeta_{t-m}, \zeta_{t-2m}, \dots]$ is of the required form. Put $\eta_t = \varepsilon_t^2 - h_t, C(L) = B(L) + 1 - A(L)$ with order r and observe that

$$C(L)\varepsilon_t^2 = \psi + B(L)\eta_t. \tag{20}$$

Let $\Psi(L) = \sum_{i=0}^{\infty}\psi_i L^i = \tilde{\Theta}(L^m)^{-1}\bar{\Gamma}_m(L)\Theta(L)W(L)$ such that $\zeta_t = \Psi(L)\varepsilon_t$ and put $\Psi^2(L) = \sum_{i=0}^{w}\psi_i^2 L^i$. Multiplying equation (20) by $\bar{C}_m(L)\Psi^2(L)$ and rearranging some terms yields

$$\begin{aligned} C_m(L^m)\zeta_t^2 &= \bar{C}_m(1)\Psi^2(1)\psi + \bar{C}_m(L)B(L)\Psi^2(L)\eta_t \\ &\quad + C_m(L^m)\big(\{\Psi(L)\varepsilon_t\}^2 - \Psi^2(L)\varepsilon_t^2\big). \end{aligned}$$

The second and third term of the right-hand side are uncorrelated since η_t and $\varepsilon_{t-i}\varepsilon_{t-j}$ are uncorrelated by assumption ($\forall i \neq j \in \mathbb{Z}$). The moving average structure of the right-hand side is not necessarily finite. To obtain a suitable ARMA structure in ζ_t^2 we have to multiply the latter equation once more by a polynomial. Note that $\tilde{\Theta}(L^m)\Psi(L)$ is a polynomial of finite order $(m-1)P + Q + w$, hence (with $\tilde{\theta}_0 = 1$)

$$\sum_{j=0}^{\tilde{Q}}\tilde{\theta}_j\psi_{i-jm} = 0 \quad \text{for} \quad i > (m-1)P + Q + w.$$

This difference equation in ψ of order $\tilde{Q}$ determines a difference equation in ψ^2 of order $\frac{1}{2}\tilde{Q}(\tilde{Q}+1)$, say

$$\sum_{j=0}^{(1/2)\tilde{Q}(\tilde{Q}+1)}\tilde{\psi}_j\psi_{i-jm}^2 = 0 \quad \text{for} \quad i > (m-1)P + Q + w + \tfrac{1}{2}m\tilde{Q}\big(\tilde{Q}-1\big).$$

Put

$$\tilde{\Psi}(L^m) = \sum_{i=0}^{(1/2)\tilde{Q}(\tilde{Q}+1)}\tilde{\psi}_i L^{im}$$

and observe that $\tilde{\Psi}(L^m)\Psi^2(L)$ is a polynomial of order $(m-1)P + Q + w + \frac{1}{2}m\tilde{Q}(\tilde{Q}-1)$. Pre-multiplying with $\tilde{\Psi}(L^m)$ yields

$$\begin{aligned} &\tilde{\Psi}(L^m)C_m(L^m)\zeta_t^2 \\ &\quad = \tilde{\Psi}(1)\bar{C}_m(1)\Psi^2(1)\psi + \bar{C}_m(L)B(L)\tilde{\Psi}(L^m)\Psi^2(L)\eta_t \\ &\qquad + C_m(L^m)\tilde{\Psi}(L^m)\big(\{\Psi(L)\varepsilon_t\}^2 - \Psi^2(L)\varepsilon_t^2\big). \end{aligned}$$

By construction of $\tilde{\Psi}(L^m)$ the variables in the right-hand side are: η_{t-i}, ($i \in \{0, \dots, (m-1)r + q + (m-1)P + Q + w + \frac{1}{2}m\tilde{Q}(\tilde{Q}-1)\}$) and $\varepsilon_{t-i}\varepsilon_{t-j}$ ($i \neq j$, i or $j \in \{0, \dots, mr + (m-1)P + Q + w + \frac{1}{2}m\tilde{Q}(\tilde{Q}-1)\}$). Hence the order of the MA term in the low frequency model is equal to $r+\frac{1}{2}\tilde{Q}(\tilde{Q}+1)$. Calculation of the autocorrelations of the MA term requires knowledge of $E\varepsilon_t^2\varepsilon_{t-i}^2/E\eta_t^2$. Tedious calculations imply that these quantities depend only upon the high frequency kurtosis κ_y and the high frequency variance parameters. This determines the autocorrelations and similar to the ARMA mean equation one has to determine a low frequency MA term with the same autocorrelation structure. This low frequency MA term determines the β's of the low frequency GARCH equation. The α's are obtained by subtracting the low frequency MA polynomial from $\tilde{\Psi}(L^m)C_m(L^m)$.

Under the conditions of the second part of the theorem $\zeta_t=\varepsilon_t$. In this case multiply relation (20) by $\overline{C}_m(L)$ and obtain

$$C_m(L^m)\zeta_t^2=\overline{C}_m(1)\psi+\overline{C}_m(L)B(L)\eta_t.$$

Proceed as before and replace the right-most term by the corresponding low frequency term. The proof is complete. *Q.E.D.*

REMARK 2: The symmetry condition can be relaxed in Theorem 1. The proof needs ($\star$) $(\forall 0\leqslant i\leqslant j)$ $Ey_ty_{t-i}y_{t-j}=0$ and $(*)$ $(\forall 0\leqslant i\leqslant j\leqslant k,\ i\neq 0$ or $j\neq k)$ $Ey_ty_{t-i}y_{t-j}y_{t-k}=0$. Recall that equation (3) implies such results for $i\neq 0$. Conditions ($\star$) and $(*)$ are somewhat technical and, therefore, they are replaced by the more appealing symmetry condition.

REFERENCES

AMEMIYA, T., AND R. Y. WU (1972): "The Effect of Aggregation on Prediction in the Autoregressive Model," *Journal of the American Statistical Association*, 67, 628–632.

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

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

——— (1988): "On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process," *Journal of Time Series Analysis*, 9, 121–131.

BOLLERSLEV, T., AND J. M. WOOLDRIDGE (1992): "Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Reviews*, 11, 143–172.

DIEBOLD, F. X. (1988): *Empirical Modeling of Exchange Rates*. New York: Springer Verlag.

DE HAAN, L., AND R. L. KARANDIKAR (1989): "Embedding a Stochastic Difference Equation into a Continuous-Time Process," *Stochastic Processes and Their Applications*, 32, 225–235.

DROST, F. C., AND T. E. NIJMAN (1992a): "Temporal Aggregation of GARCH Processes," CenTER Discussion Papers 9066 and 9240, Tilburg University.

——— (1992b): "A Test of the Validity of a GARCH Diffusion Limit," Mimeographed, Tilburg University.

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987–1008.

ENGLE, R. F., AND T. BOLLERSLEV (1986): "Modeling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1–50.

ENGLE, R. F., AND G. GONZÁLEZ-RIVERA (1991): "Semiparametric ARCH Models," *Journal of Business and Economic Statistics*, 9, 345–359.

GALLANT, A. R., AND G. TAUCHEN (1989): "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica*, 57, 1091–1120.

HANSEN, L. P., AND T. J. SARGENT (1983): "The Dimensionality of the Aliasing Problem in Models with Rational Spectral Densities," *Econometrica*, 51, 377–388.

HARVEY, A. C., AND R. G. PIERSE (1984): "Estimating Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, 79, 125–131.

LÜTKEPOHL, H. (1986): "Forecasting Vector ARMA Processes with Systematically Missing Observations," *Journal of Business and Economic Statistics*, 4, 375–390.

NELSON, D. B. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7–38.

NIJMAN, T. E., AND F. C. PALM (1990a): "Parameter Identification in ARMA-Processes in the Presence of Regular but Incomplete Sampling," *Journal of Time Series Analysis*, 11, 239–248.

——— (1990b): "Predictive Accuracy Gain from Disaggregate Sampling in ARIMA Models," *Journal of Business and Economic Statistics*, 8, 405–415.

Palm, F. C., and T. E. Nijman (1984): "Missing Observations in the Dynamic Regression Model," *Econometrica*, 52, 1415–1435.

Phillips, P. C. B. (1973): "The Problem of Identification in Finite Parameter Continuous Time Models," *Journal of Econometrics*, 1, 351–362.

Stroock, D. W., and S. R. S. Varadhan (1979): *Multidimensional Diffusion Processes*. Berlin: Springer Verlag.

Weiss, A. A. (1986): "Asymptotic Theory for ARCH Models: Estimation and Testing," *Econometric Theory*, 2, 107–131.