

# GARCH and Irregularly Spaced Data\*

Nour Meddahi<sup>a,†</sup>, Eric Renault<sup>a</sup>, and Bas Werker<sup>b</sup>

<sup>a</sup>Département de sciences économiques, CIRANO, CIREQ, Université de Montréal.

Address: C.P. 6128, succursale Centre-ville, Montréal (Québec), H3C 3J7, Canada.

<sup>b</sup>Finance Group and Econometrics Group, CentER, Tilburg University, Tilburg, The Netherlands.

March 4, 2003

## Abstract

An exact discretization of continuous time stochastic volatility processes observed at irregularly spaced times is used to give insights on how a coherent GARCH model can be specified for such data. The relation of our approach with those in the existing literature is studied.

*Keywords:* Volatility; continuous time model; exact discretization.

*JEL classification:* C22.

---

\*This note succeeds the paper by the same authors that circulated under the title “Modelling High-Frequency Data in Continuous Time.” We thank Torben Andersen, Luc Bauwens, Christian Gouriéroux, and Neil Shephard for helpful comments. The first author thanks CRSNG, IFM2, and MITACS for financial support. Part of this research was conducted when the third author was visiting CIREQ.

<sup>†</sup>Corresponding author. Tel: 1-514-343-2399; fax: 1-514-343-7221. E-mail: [nour.meddahi@umontreal.ca](mailto:nour.meddahi@umontreal.ca); Web-page: <http://www.mapageweb.umontreal.ca/meddahin/>.

Two recent papers, Engle (2000) and Ghysels and Jasiak (1998), proposed different formulations of GARCH models for irregularly spaced data. In this note, we exploit results of Meddahi and Renault (2002) to clarify the advantage of each approach and propose a model that combines these advantages.

In the sequel, we assume that a financial price  $S_t$  is observed at irregularly spaced dates  $t_0, t_1, \dots, t_n$ , with  $0 = t_0 < t_1 < \dots < t_n$ . We denote by  $x_i$ ,  $i=1, \dots, n$ , the  $i$ -th duration ( $x_i \equiv t_i - t_{i-1}$ ), and by  $\varepsilon_i$  the continuously compounded return of  $S_t$  over the period  $(t_{i-1}, t_i]$  ( $\varepsilon_i \equiv \log(S_{t_i}) - \log(S_{t_{i-1}})$ ).

In his simplest volatility model, Engle (2000) assumes that the variable  $\sigma_{i-1}^2$  defined as

$$\sigma_{i-1}^2 = \frac{h_{i-1}}{x_i}, \quad \text{where } h_{i-1} = \text{Var}[\varepsilon_i \mid \varepsilon_j, x_j, j \leq i-1; x_i] \quad (1)$$

follows a GARCH(1,1)-type equation (Bollerslev, 1986). More precisely, under the assumption

$$E[\varepsilon_i \mid \varepsilon_j, x_j, j \leq i-1; x_i] = 0, \quad (2)$$

Engle (2000) specifies

$$\sigma_{i-1}^2 = \omega + \alpha(\varepsilon_{i-1}/\sqrt{x_{i-1}})^2 + \beta\sigma_{i-2}^2. \quad (3)$$

In other words, in order to take into account the unequally spaced feature of the returns, Engle (2000) assumes that the *variance per time-unit*,  $\sigma_{i-1}^2$ , follows a GARCH(1,1) equation.

In contrast, Ghysels and Jasiak (1998) specify a GARCH equation for the *total variance* process  $\tilde{h}_{i-1}$  defined by

$$\tilde{h}_{i-1} = \text{Var}[\varepsilon_i \mid \varepsilon_j, x_j, j \leq i-1]. \quad (4)$$

However, in order to take into account the unequally spaced feature of the returns, Ghysels and Jasiak (1998) assume a time-varying parameter GARCH equation with

$$\tilde{h}_{i-1} = \omega_{i-1} + \alpha_{i-1}\varepsilon_{i-1}^2 + \beta_{i-1}\tilde{h}_{i-2}, \quad (5)$$

where the parameters  $\omega_{i-1}$ ,  $\alpha_{i-1}$ , and  $\beta_{i-1}$  are functions of the expected duration  $\psi_{i-1}$  defined as  $\psi_{i-1} = E[x_i \mid \varepsilon_j, x_j, j \leq i-1]$  and a structural unknown parameter. The functional forms adopted by Ghysels and Jasiak (1998) are those derived by Drost and Werker (1996) for a weak GARCH representation (Drost and Nijman, 1993) of a GARCH diffusion model (Nelson, 1990) when observations are equally spaced by a length, say,  $\Delta$ . For instance, Drost and Werker (1996) show that  $\alpha_\Delta + \beta_\Delta = \exp(-\kappa\Delta)$  where  $\kappa$  is the mean reverting parameter of the continuous time spot variance process. Therefore Ghysels and Jasiak (1998) assume

$$\alpha_{i-1} + \beta_{i-1} = \exp(-\kappa\psi_{i-1}).$$

It is clear that there are several differences between Engle (2000) and Ghysels and Jasiak (1998) approaches. The first one is in the conditioning informations: Engle (2000) considered the variance of the return  $\varepsilon_i$  given the information  $\mathcal{F}_{i-1}^d = \sigma(\varepsilon_j, x_j, j \leq i-1; x_i)$  while Ghysels and Jasiak (1998) considered the information  $\mathcal{G}_{i-1}^d = \sigma(\varepsilon_j, x_j, j \leq i-1)$ . Clearly, under the assumption (2), one has

$$\tilde{h}_{i-1} = E[h_{i-1} \mid \mathcal{G}_{i-1}^d]. \quad (6)$$

The second difference is in the GARCH formulation: Equation (3), for the variance per unit of time, implies a time-varying parameter GARCH equation for the total variance process  $h_{i-1}$ :

$$h_{i-1} = \omega x_i + \alpha \frac{x_i}{x_{i-1}} \varepsilon_{i-1}^2 + \beta \frac{x_i}{x_{i-1}} h_{i-2}. \quad (7)$$

Therefore, by using (6) and the definition of  $\psi_{i-1}$ , one gets

$$\tilde{h}_{i-1} = \omega \psi_{i-1} + \alpha \frac{\psi_{i-1}}{x_{i-1}} \varepsilon_{i-1}^2 + \beta \frac{\psi_{i-1}}{x_{i-1}} \tilde{h}_{i-2} + \beta \frac{\psi_{i-1}}{x_{i-1}} (h_{i-2} - \tilde{h}_{i-2}), \quad (8)$$

which differs from (5) because it is not a GARCH equation (due to the presence of the last term) and the time-varying coefficients involve not only the expected value of  $x_i$ ,  $\psi_{i-1}$ , but also the duration  $x_{i-1}$ .

A third difference, is that Ghysels and Jasiak (1998) considered a weak GARCH representation of the returns while Engle (2000) considered a semi-strong one (Drost and Nijman, 1993). This difference is not important for our purpose and we therefore ignore it; see Drost and Werker (1996) and Meddahi and Renault (2002) for a discussion.

For future reference, note that one gets from (7) the autoregressive representation for  $\sigma_{i-1}^2$ :

$$\sigma_{i-1}^2 = \omega + (\alpha + \beta) \sigma_{i-2}^2 + s_{i-1}, \quad s_{i-1} \equiv \alpha ((\varepsilon_{i-1} / \sqrt{x_{i-1}})^2 - \sigma_{i-2}^2), \quad E[s_{i-1} | \mathcal{F}_{i-2}^d] = 0. \quad (9)$$

We now consider an exact discretization of a continuous time stochastic volatility model observed at irregularly spaced times. We assume that the price  $S_t$  is given by

$$d \log S_t = \sqrt{v_t} dW_t, \quad (10)$$

where  $W_t$  is a Brownian process,  $v_t$  is a stationary square-integrable and positive process, independent with  $W_t$ , such that

$$\forall \Delta > 0, \quad E[v_{t+\Delta} | v_\tau, \tau \leq t] = \theta + \exp(-\kappa \Delta)(v_t - \theta). \quad (11)$$

Note that we rule out a drift in Eq. (10) in order to get the same assumption (2) which simplifies the exposition. For the same reason, we assume that the processes  $\{W_t\}$  and  $\{v_t\}$  are mutually independent, which excludes the so-called leverage effect. More importantly, we also assume in the rest of the note that the continuous time process driving the durations and the bivariate process  $\{(W_t, v_t)\}$  are mutually independent; see the end of the note for a comment. Finally, Eq. (11) includes the popular Nelson (1990) GARCH diffusion model, the Heston (1993) affine model, the CEV model, as well as the positive Lévy Ornstein-Uhlenbeck model of Barndorff-Nielsen and Shephard (2001).

As pointed out by Meddahi and Renault (2002), when one considers temporal aggregation of stochastic volatility models, one has to incorporate the spot variance  $v_t$  in the relevant conditioning information. This leads us to study the variance of the return  $\varepsilon_i$  given the information  $\mathcal{F}_{i-1}^c = \sigma(\varepsilon_j, v_j, x_j, j \leq i-1; x_i)$ . In the appendix, we show that

$$Var[\varepsilon_i | \mathcal{F}_{i-1}^c] = \theta x_i + c(\kappa x_i)(v_{i-1} - \theta)x_i, \quad (12)$$

where  $c(x) \equiv (1 - \exp(-x))/x$ . Note that  $c(x) \approx 1$  for small  $x$ , which shows that the variance of the innovation is approximately equal to  $v_{i-1}x_i$ . This approximate linearity leads us, following Engle (2000), to study the discrete time behavior of the *variance per time-unit*:

$$f_{i-1} \equiv \frac{\text{Var}[\varepsilon_i | \mathcal{F}_{i-1}^c]}{x_i}. \quad (13)$$

In the appendix, we show that

$$f_{i-1} = \omega_i + \gamma_i f_{i-2} + \nu_{i-1}, \quad \text{where} \quad (14)$$

$$\gamma_i = \exp(-\kappa x_{i-1}) \frac{c(\kappa x_i)}{c(\kappa x_{i-1})}, \quad \omega_i = \theta(1 - \gamma_i), \quad \nu_{i-1} = c(\kappa x_i)(v_{i-1} - E[v_{i-1} | \mathcal{F}_{i-2}^c]). \quad (15)$$

Therefore,  $f_i$  is an autoregressive process of order one and  $\varepsilon_i$  is a SR-SARV process as in Andersen (1994) and Meddahi and Renault (2002). However, an important difference with these two studies is that  $f_{i-1}$  is an AR(1) process with time-varying parameters.

This time-varying feature of the parameters is in contrast with the specification (9) adopted by Engle (2000), who does not take into account the effects of temporal aggregation on the model parameters. From (15) we see that this effect is quite important and plays a role in other specifications as well. For example, for long durations  $x_i$ , the conditional variance per time-unit over the next duration should be close to the unconditional one. As we see from (12),  $f_{i-1} \rightarrow \theta$  for  $x_i \rightarrow \infty$ . Specifying time-invariant parameters in (3) contradicts this natural assumption.

In a different way, Ghysels and Jasiak (1998) highlighted the importance of taking into account the temporal aggregation effect. However, they proposed an ad hoc functional form for the time-varying parameter, which in turn is not supported by an exact discretization of a continuous time process. In addition, our exact discretization clearly shows that it is better to model the variance per time-unit (as in Engle, 2000) instead of the total variance over the next duration (as in Ghysels and Jasiak, 1998) because total variances are foremost influenced by the associated duration. For example, in (5) a high variance for the current duration induces a high volatility for the following duration, even if both durations are of unequal length.

Finally, regarding the conditioning information, our approach clearly favors the one adopted by Engle (2000), i.e., by incorporating the current duration in the information. Of course, one can always use the formula (6) and get a volatility model given the information  $\mathcal{G}_{i-1}^d$  as in Ghysels and Jasiak (1998). However, from (14), one needs to compute for instance  $E[\gamma_i | \mathcal{G}_{i-1}^d]$ , which needs a specification of the dynamics of the durations (as in Engle and Russell, 1998, for instance).

We conclude this note by making four remarks. 1) From Eq. (14), one may derive the corresponding coefficients  $\alpha_i$  and  $\beta_i$  as in (3) and (5) (with the restriction  $\alpha_i + \beta_i = \gamma_i$ ) if one fully specifies the process  $v_t$ . 2) Interestingly, (14) implies the following multi-period conditional moment restriction

$$E[\varepsilon_i^2/x_i - \omega_i - \gamma_i \varepsilon_{i-1}^2/x_{i-1} | \mathcal{G}_{i-2}^d] = 0,$$

which is appealing because the conditioning information  $\mathcal{G}_{i-2}^d$  does not include the latent variable  $v_i$ ; see Meddahi and Renault (2002) for further discussions. 3) Our approach can

be extended to non-linear stochastic volatility models by using the eigenfunction approach of Meddahi (2001). 4) Renault and Werker (2003) show that an additional term appears in (14) when the independence assumption of the continuous time process driving the durations and the bivariate process  $\{(W_t, v_t)\}$  does not hold.

## Appendix

**Proof of (12).** Observe that  $\varepsilon_i = \int_{t_{i-1}}^{t_i} \sqrt{v_u} dW_u$ . Therefore, by using the independence of the processes  $\{v_t\}$ ,  $\{W_t\}$ , and the process driving the durations, one gets

$$Var[\varepsilon_i | \mathcal{F}_{i-1}^c] = E\left[\int_{t_{i-1}}^{t_i} v_u du \mid \mathcal{F}_{i-1}^c\right] = \int_{t_{i-1}}^{t_i} E[v_u \mid v_{i-1}] du.$$

By using (11), one obtains easily the formula (12). ■

**Proof of (14).** By combining (11), (12), and (13), one gets:

$$\begin{aligned} f_{i-1} &= \theta + c(\kappa x_i)(v_{i-1} - \theta) = \theta + c(\kappa x_i) (\exp(-\kappa x_{i-1})(v_{i-2} - \theta) + v_{i-1} - E[v_{i-1} \mid v_{i-2}]) \\ &= \theta + c(\kappa x_i)(\exp(-\kappa x_{i-1}) \frac{(f_{i-2} - \theta)}{c(\kappa x_{i-1})} + v_{i-1} - E[v_{i-1} \mid v_{i-2}]), \end{aligned}$$

which coincides with (14) given that  $E[v_{i-1} \mid v_{i-2}] = E[v_{i-1} \mid \mathcal{F}_{i-2}^c]$  (under the independence of the duration and volatility processes). ■

## References

- Andersen, T.G., 1994, Stochastic autoregressive volatility: a framework for volatility modeling, *Mathematical Finance* 4, 75-102.
- Barndorff-Nielsen, O.E. and N. Shephard, 2001, Non-Gaussian OU based models and some of their uses in financial economics, *Journal of the Royal Statistical Society, B*, 63, 167-241.
- Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307-327.
- Drost, F.C. and TH.E. Nijman, 1993, Temporal aggregation of GARCH processes, *Econometrica* 61, 909-927.
- Drost, F.C. and B.J.M. Werker, 1996, Closing the GARCH gap: continuous time GARCH modeling, *Journal of Econometrics* 74, 31-58.
- Engle, R.F., 2000, The econometrics of ultra-high frequency data, *Econometrica* 68, 1-22.
- Engle, R.F. and J.R. Russell, 1998, Autoregressive conditional duration: a new model for irregularly spaced transaction data, *Econometrica* 66, 1127-1161.
- Ghysels, E. and J. Jasiak, 1998, GARCH for irregularly spaced financial data: The ACD-GARCH model, *Studies in Nonlinear Dynamics and Econometrics* 2, 133-149.
- Heston, S.L., 1993, A closed form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327-344.
- Meddahi, N., 2001, An eigenfunction approach for volatility modeling, CIRANO working paper 2001s-70.
- Meddahi, N. and E. Renault, 2002, Temporal aggregation of volatility models, *Journal of Econometrics*, forthcoming.
- Nelson, D.B., 1990, ARCH models as diffusion approximations, *Journal of Econometrics* 45, 7-39.
- Renault, E. and B.J.M. Werker, 2003, Stochastic volatility models with transaction time risk, Working paper, Tilburg University.