G@RCH 2.2: AN OX PACKAGE FOR ESTIMATING AND FORECASTING VARIOUS ARCH MODELS

Sébastien Laurent

CREPP, Université de Liège, CORE, Université catholique de Louvain, Belgium and Maastricht University, Netherlands

Jean-Philippe Peters

Université de Liège, Belgium.

Abstract. This paper discusses and documents G@RCH 2.2, an Ox package dedicated to the estimation and forecast of various univariate ARCH-type models including GARCH, EGARCH, GJR, APARCH, IGARCH, FIGARCH, HYGARCH, FIEGARCH and FIAPARCH specifications of the conditional variance and an AR(FI)MA specification of the conditional mean.

These models can be estimated by Approximate (Quasi) Maximum Likelihood under four assumptions: normal, Student-*t*, GED or skewed Student errors. Explanatory variables can enter both the conditional mean and the conditional variance equations. *h*-step-ahead forecasts of both the conditional mean and the conditional variance are available as well as many mispecification tests.

We first propose an overview of the package's features, with the presentation of the different specifications of the conditional mean and conditional variance. Then further explanations are given about the estimation methods. Measures of the accuracy of the procedures are also given and the GARCH features provided by G@RCH are compared with those of nine other econometric softwares. Finally, a concrete application of G@RCH 2.2 is provided.

Keywords. ARCH Modelling; Forecasts; Ox; Econometric Software; Financial Time Series.

1. Introduction

It has been recognized for a long time that the dynamic behavior of economic variables is difficult to understand. And this difficulty certainly increases with the observation frequency of the data.

Most time series of asset returns can be characterized as serially dependent. This is revealed by the presence of positive autocorrelation in the squared returns, and sometimes (to a much smaller extent) by autocorrelation in the returns. To fully account for the characteristics of high-frequency financial returns we need to specify a model in which the conditional mean and the conditional variance may be

0950-0804/02/03 0447–39 JOURNAL OF ECONOMIC SURVEYS Vol. 16, No. 3
© Blackwell Publishers Ltd. 2002, 108 Cowley Rd., Oxford OX4 1JF, UK and 350 Main St., Malden, MA 02148, USA.

time-varying. The most widespread modelling approach to capture these properties is to specify a dynamic model for the conditional mean and the conditional variance, such as an ARMA-ARCH model or one of its various extensions (see the seminal paper of Engle, 1982). Another well established stylized fact of financial returns, at least when they are sampled at high frequencies, is that they exhibit fat tails (which corresponds to a kurtosis coefficient larger than three) and are often skewed (which corresponds to a positive or negative skewness coefficient).

The estimation of univariate GARCH models is commonly undertaken by maximizing a Gaussian likelihood function. Even if this hypothesis is unrealistic in practice, the normality assumption may be justified by the fact that the Gaussian Quasi Maximum Likelihood (QML) estimator is consistent assuming that the conditional mean and the conditional variance are specified correctly (Weiss, 1986; Bollerslev and Wooldridge, 1992). The price to pay for this property is that this method it not efficient, the degree of inefficiency increasing with the degree of departure from normality (Engle and González-Rivera, 1991). Searching for a more suitable distribution may thus be of primary importance to gain efficiency. From a practical point of view, the issue of skewness (asymmetry) and kurtosis (fat tails) is important in many respects for financial applications. Indeed, Peiró (1999) emphasizes the relevance of the modelling of higher-order features in asset pricing models,[1] portfolio selection[2] and option pricing theories[3] while Giot and Laurent (2001) show that modelling skewness and kurtosis is crucial in Value-at-Risk applications.

A researcher is thus facing the problem of the specification choice. Which model to select? And which selection criterion to use? It is not our goal to answer these questions. However, it is almost sure that this researcher is going to estimate several candidate models, with different lag orders and perhaps different log-likelihood functions.

Well known statistical packages such as Eviews, Gauss, Matlab, Microfit, PcGive, Rats, SAS, S-Plus or TSP provide various options to estimate sophisticated econometric models in very different areas such as cointegration, panel data, limited dependent model, etc.

The aim of this paper is to provide an overview of a package dedicated to the estimation and forecasting of various univariate ARCH-type models. Contrary to the software mentioned above, G@RCH 2.2 is only concerned with ARCH-type models (Engle, 1982), including some recent contributions in this field such as the GARCH (Bollerslev, 1986), EGARCH (Nelson, 1991), GJR (Glosten, Jagannathan and Runkle, 1993), APARCH (Ding, Granger and Engle, 1993), IGARCH (Engle and Bollerslev, 1986) but also FIGARCH (Baillie, Bollerslev and Mikkelsen, 1996a; Chung, 1999), HYGARCH (Davidson, 2001), FIEGARCH (Bollerslev and Mikkelsen, 1996) and FIAPARCH (Tse, 1998) specifications of the conditional variance and an AR(FI)MA specification of the conditional mean (Baillie, Chung and Tieslau, 1996; Tschernig, 1995; Teyssière, 1997; Lecourt, 2000; or Beine, Laurent and Lecourt, 2000). This package provides a number of features, including two standard errors estimation methods (Approximate ML and Approximate QML) for four distributions (normal, Student-*t*, GED or skewed Student-*t*).

© Blackwell Publishers Ltd. 2002

Moreover, explanatory variables can enter the mean and/or the variance equations. Finally, h-step-ahead forecasts of both the conditional mean and conditional variance are available as well as many mispecification tests (Nyblom, SBT, Pearson goodness-of-fit, Box-Pierce, ...).

The package has been developed using the Ox 3.0 matrix programming language of Doornik (1999).[4] It can be used on several platforms, including Windows, Unix, Linux and Solaris. For most of the specifications, it is generally very fast and its main characteristic is its ease of use. G@RCH 2.2 may be used freely for *non-commercial* purposes and downloaded from the web site **http://www.egss.ulg.ac.be/garch/**.

Two (complementary) versions of the program are available and called the 'Light Version' and the 'Full Version', respectively. The 'Full Version' offers a friendly dialog-oriented interface similar to PcGive and some graphical features by using OxPack, a GiveWin batch client module. This version requires a professional version of Ox and GiveWin.

The 'Light Version' is launched from a simple Ox file. It does not take advantage of the OxPack extension (no dialog-oriented interface and no graphs) and can therefore be used with a free version of Ox. This version thus simply requires any Ox executable and a text editor.

This paper is structured as follows: in Section 2, we present an overview of the package's features, with the presentation of the different specifications of the conditional mean and conditional variance. Comments on estimation procedures (parameters constraints, distributions, tests, forecasts, accuracy of the package and a comparison of its features with those of nine well-known econometric packages) are introduced in Section 3. Then a user guide is provided for both versions of G@RCH 2.2 in Section 4 with an application using the CAC40 stock index. Finally, Section 5 concludes.

2. Features of the package

This section proceeds to describe the models implemented in G@RCH 2.2 and gives some technical details. Our attention will be first devoted to review the specifications of the conditional mean equation. Then, some recent contributions in the ARCH modelling framework will be presented.

2.1. *Mean equation*

Let us consider an univariate time series y_t. If Ω_{t-1} is the information set at time $t-1$, we can define its functional form as:

$$y_t = E(y_t \mid \Omega_{t-1}) + \varepsilon_t, \tag{1}$$

where $E(.\mid.)$ denotes the conditional expectation operator and ε_t is the disturbance term (or unpredictable part), with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_s) = 0, \forall\, t \neq s$.

This is the mean equation which has been studied and modelled in many ways. Two of the most famous specifications are the Autoregressive (AR) and Moving

© Blackwell Publishers Ltd. 2002

Average (MA) models. Mixing these two processes and introducing n_1 explanatory variables in the equation, we obtain this ARMAX(n, s) process,

$$\Psi(L)(y_t - \mu_t) = \Theta(L)\varepsilon_t$$
$$\mu_t = \mu + \sum_{i=1}^{n_1} \delta_i x_{i,t}, \tag{2}$$

where L is the lag operator,[5] $\Psi(L) = 1 - \sum_{i=1}^{n} \psi_i L^i$ and $\Theta(L) = 1 + \sum_{j=1}^{s} \theta_j L^j$. To start the recursion, it is convenient to set the initial conditions at $\varepsilon_t = 0$ for all $t \leqslant \max\{p, q\}$.

Several studies have shown that the dependent variable (interest rate returns, exchange rate returns, etc.) may exhibit significant autocorrelation between observations widely separated in time. In such a case, we can say that y_t displays long memory, or long-term dependence and is best modelled by a fractionally integrated ARMA process (so called ARFIMA process) initially developed in Granger (1980) and Granger and Joyeux (1980) among others.[6] The ARFIMA(n, ζ, s) is given by:

$$\Psi(L)(1 - L)^{\zeta}(y_t - \mu_t) = \Theta(L)\varepsilon_t, \tag{3}$$

where the operator $(1 - L)^{\zeta}$ accounts for the long memory of the process and is defined as:

$$\begin{aligned}(1 - L)^{\zeta} &= \sum_{k=0}^{\infty} \frac{\Gamma(\zeta + 1)}{\Gamma(k + 1)\Gamma(\zeta - k + 1)} L^k \\ &= 1 - \zeta L - \tfrac{1}{2}\zeta(1 - \zeta)L^2 - \tfrac{1}{6}\zeta(1 - \zeta)(2 - \zeta)L^3 - \cdots \\ &= 1 - \sum_{k=1}^{\infty} c_k(\zeta) L^k, \end{aligned} \tag{4}$$

with $0 < \zeta < 1$, $c_1(\zeta) = \zeta$, $c_2(\zeta) = \frac{1}{2}\zeta(1 - \zeta)$, ... and $\Gamma(.)$ denoting the Gamma function (see Baillie, 1996, for a survey on this topic). The truncation order of the infinite summation is set to $t - 1$.

It is worth noting that Doornik and Ooms (1999) recently provided an Ox package for estimating, forecasting and simulating ARFIMA models. However, in contrast to our package, they assume that the conditional variance is constant over time.

2.2. *Variance equation*

The ε_t term in Eq. (1)–(3) is the innovation of the process. Two decades ago, Engle (1982) defined as an Autoregressive Conditional Heteroscedastic (ARCH) process, all ε_t of the form:

$$\varepsilon_t = z_t \sigma_t, \tag{5}$$

© Blackwell Publishers Ltd. 2002

where z_t is an independently and identically distributed (*i.i.d.*) process with $E(z_t) = 0$ and $Var(z_t) = 1$. By definition, ε_t is serially uncorrelated with a mean equal to zero, but its conditional variance equals σ_t^2 and, therefore, may change over time, contrary to what is assumed in the standard regression model.

The models provided by our program are all ARCH-type.[7] They differ based on the functional form of σ_t^2 but the basic principles are the same. Besides the traditional ARCH and GARCH models, we focus mainly on two kinds of models: the asymmetric models and the fractionally integrated models. The former are defined to take account of the so-called 'leverage effect' observed in many stock returns, while the latter allows for long-memory in the variance. Early evidence of the 'leverage effect' can be found in Black (1976), while persistence in volatility is a common finding of many empirical studies; see for instance Bera and Higgins (1993) and Palm (1996) for excellent surveys on ARCH models.

2.2.1. *ARCH Model*

The ARCH (q) model can be expressed as:

$$\begin{aligned} &\varepsilon_t = z_t \sigma_t \\ &z_t \sim i.i.d.\ D(0, 1) \\ &\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \end{aligned} \tag{6}$$

where $D(.)$ is a probability density function with mean 0 and unit variance (it will be defined in Section 3.2).

The ARCH model can describe volatility clustering. The conditional variance of ε_t is indeed an increasing function of the square of the shock that occurred in $t-1$. Consequently, if ε_{t-1} was large in absolute value, σ_t^2 and thus ε_t is expected to be large (in absolute value) as well. Notice that even if the conditional variance of an ARCH model is time-varying ($\sigma_t^2 = E(\varepsilon_t^2 \mid \psi_{t-1})$), the unconditional variance of ε_t is constant and, provided that $\omega > 0$ and $\sum_{i=1}^{q} \alpha_i < 1$, we have:

$$\sigma^2 \equiv E(\varepsilon_t^2) = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i}. \tag{7}$$

Note also that the ARCH model can explain part of the excess kurtosis that we observe in financial time series. As shown by Engle (1982) for the ARCH(1) case under the normality assumption, the kurtosis of ε_t is equal to $3(1-\alpha_1^2)/(1-3\alpha_1^2)$. The kurtosis is thus finite if $\alpha_1 < \frac{1}{3}$ and larger than 3 (the kurtosis of a standard normal distribution) if $\alpha_1 > 0$.

The computation of σ_t^2 in Eq. (6) depends on past (squared) residuals (ε_t^2), that are not observed for $t = 0, -1, \ldots, -q+1$. To initialize the process, the unobserved squared residuals have been set to their sample mean.

© Blackwell Publishers Ltd. 2002

In the rest of the paper, ω is assumed fixed. If n_2 explanatory variables are introduced into the model, $\omega_t = \omega + \sum_{i=1}^{n_2} \omega_i x_{i,t}$ with an exception for the exponential models (EGARCH and FIEGARCH) where $\omega_t = \omega + \ln(1 + \sum_{i=1}^{n_2} \omega_i x_{i,t})$.

Finally, σ_t^2 has obviously to be positive for all t. Sufficient conditions to ensure that the conditional variance in Eq. (6) is positive are given by $\omega > 0$ and $\alpha_i \geqslant 0$. Furthermore, when explanatory variables enter the ARCH equation, these positivity constraints are not valid anymore (even if the conditional variance still has to be non-negative).

2.2.2. *GARCH Model*

Early empirical evidence has shown that a high ARCH order has to be selected to capture the dynamics of the conditional variance (thus involving the estimation of numerous parameters). The Generalized ARCH (GARCH) model of Bollerslev (1986) is an answer to this issue. It is based on an infinite ARCH specification and it allows a reduction in the number of estimated parameters by imposing non-linear restrictions on them. The GARCH (p, q) model can be expressed as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2. \tag{8}$$

Using the lag or backshift operator L, the GARCH (p, q) model is:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \tag{9}$$

with $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_q L^q$ and $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$.

If all the roots of the polynomial $|\, 1 - \beta(L) \,| = 0$ lie outside the unit circle, we have:

$$\sigma_t^2 = \omega[1 - \beta(L)]^{-1} + \alpha(L)[1 - \beta(L)]^{-1}\varepsilon_t^2, \tag{10}$$

which may be seen as an ARCH(∞) process since the conditional variance linearly depends on all previous squared residuals. In this case, the conditional variance of y_t can become larger than the unconditional variance given by:

$$\sigma^2 \equiv E(\varepsilon_t^2) = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j},$$

if past realizations of ε_t^2 are larger than σ^2 (Palm, 1996).

As in the ARCH case, some restrictions are needed to ensure σ_t^2 is positive for all t. Bollerslev (1986) shows that imposing $\omega > 0$, $\alpha_i \geqslant 0$ (for $i = 1, \ldots, q$) and $\beta_j \geqslant 0$ (for $j = 1, \ldots, p$) is sufficient for the conditional variance to be positive. In practice, the GARCH parameters are often estimated without the positivity

© Blackwell Publishers Ltd. 2002

restrictions. Nelson and Cao (1992) argued that imposing all coefficients to be nonnegative is too restrictive and that some of these coefficients are found to be negative in practice while the conditional variance remains positive (by checking on a case-by-case basis). Consequently, they relaxed this constraint and gave sufficient conditions for the GARCH$(1, q)$ and GARCH$(2, q)$ cases based on the infinite representation given in Eq. (10). Indeed, the conditional variance is strictly positive provided $\omega[1-\beta(1)]^{-1} > 0$ is positive and all the coefficients of the infinite polynomial $\alpha(L)[1-\beta(L)]^{-1}$ in Eq. (10) are nonnegative. The positivity constraints proposed by Bollerslev (1986) can be imposed during the estimation (see 3.1). If not, these constraints, as well as the ones implied by the ARCH(∞) representation, will be tested *a posteriori* and reported in the output.

2.2.3. *EGARCH Model*

The Exponential GARCH (EGARCH) model is introduced by Nelson (1991). Bollerslev and Mikkelsen (1996) propose to re-express the EGARCH model as follows:

$$\ln \sigma_t^2 = \omega + [1-\beta(L)]^{-1}[1+\alpha(L)]g(z_{t-1}). \tag{11}$$

The value of $g(z_t)$ depends on several elements. Nelson (1991) notes that, '*to accommodate the asymmetric relation between stock returns and volatility changes (…) the value of $g(z_t)$ must be a function of both the magnitude and the sign of z_t*'.[8] That is why he suggests to express the function $g(.)$ as

$$g(z_t) \equiv \underbrace{\gamma_1 z_t}_{\textit{sign effect}} + \underbrace{\gamma_2[\mid z_t \mid - E \mid z_t \mid]}_{\textit{magnitude effect}}. \tag{12}$$

$E \mid z_t \mid$ depends on the assumption made about the unconditional density of z_t. For the normal distribution, $E(\mid z_t \mid) = \sqrt{2/\pi}$. For the skewed Student distribution,

$$E(\mid z_t \mid) = \frac{4\xi^2}{\xi + \frac{1}{\xi}} \frac{\Gamma\left(\frac{1+\upsilon}{2}\right)\sqrt{\upsilon - 2}}{\sqrt{\pi}(\upsilon - 1)\Gamma\left(\frac{\upsilon}{2}\right)},$$

where $\xi = 1$ for the symmetric Student. For the GED, we have

$$E(\mid z_t \mid) = \lambda_\upsilon 2^{1/\upsilon} \frac{\Gamma\left(\frac{2}{\upsilon}\right)}{\Gamma\left(\frac{1}{\upsilon}\right)}.$$

© Blackwell Publishers Ltd. 2002

ξ, υ and λ_υ concern the shape of the non-normal densities and will be defined in Section 3.2.

Note that the use of a *ln* transformation of the conditional variance ensures that σ_t^2 is always positive.

2.2.4. *GJR Model*

This popular model is proposed by Glosten, Jagannathan and Runkle (1993). Its generalized version is given by:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} (\alpha_i \varepsilon_{t-i}^2 + \gamma_i S_{t-i}^- \varepsilon_{t-i}^2) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{13}$$

where S_t^- is a dummy variable.

In this model, it is assumed that the impact of ε_t^2 on the conditional variance σ_t^2 is different when ε_t is positive or negative. The TGARCH model of Zakoian (1994) is very similar to the GJR but models the conditional standard deviation instead of the conditional variance. Finally, Ling and McAleer (2002) have proposed, among other stationarity conditions for GARCH models, the conditions of existence of the second and fourth moment of the GJR.

2.2.5. *APARCH model*

This model was introduced by Ding, Granger and Engle (1993). The APARCH (p, q) model can be expressed as:

$$\sigma_t^\delta = \omega + \sum_{i=1}^{q} \alpha_i, (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^\delta, \tag{14}$$

where $\delta > 0$ and $-1 < \gamma_i < 1$ $(i = 1, ..., q)$.

This model combines the flexibility of a varying exponent with the asymmetry coefficient (to take the 'leverage effect' into account). The APARCH includes seven other ARCH extensions as special cases:[9]

- The ARCH of Engle (1982) when $\delta = 2$, $\gamma_i = 0$ $(i = 1, \ldots, p)$ and $\beta_j = 0$ $(j = 1, \ldots, p)$.
- The GARCH of Bollerslev (1986) when $\delta = 2$ and $\gamma_i = 0$ $(i = 1, \ldots, p)$.
- Taylor (1986)/Schwert (1990)'s GARCH when $\delta = 1$, and $\gamma_i = 0$ $(i = 1, \ldots, p)$.
- The GJR of Glosten, Jagannathan and Runkle (1993) when $\delta = 2$.
- The TARCH of Zakoian (1994) when $\delta = 1$.
- The NARCH of Higgins and Bera (1992) when $\gamma_i = 0 (i = 1, \ldots, p)$ and $\beta_j = 0$ $(j = 1, \ldots, p)$.
- The Log-ARCH of Geweke (1986) and Pentula (1986), when $\delta \rightarrow 0$.

© Blackwell Publishers Ltd. 2002

Following Ding, Granger and Engle (1993), if $\omega > 0$ and $\sum_{i=1}^{q} \alpha_i E(|z| - \gamma_i z)^{\delta} + \sum_{j=1}^{p} \beta_j < 1$, a stationary solution for Eq. (14) exists and is:

$$E(\sigma_t^{\delta}) = \frac{\omega}{1 - \sum_{i=1}^{q} \alpha_i E(|z| - \gamma_i z)^{\delta} - \sum_{j=1}^{p} \beta_j}.$$

Notice that if we set $\gamma = 0$, $\delta = 2$ and z_t has zero mean and unit variance, we have the usual stationarity condition of the GARCH(1, 1) model ($\alpha_1 + \beta_1 < 1$). However, if $\gamma \neq 0$ and/or $\delta \neq 2$, this condition depends on the assumption made about the innovation process.

Ding, Granger and Engle (1993) derived a closed form solution to $\kappa_i = E(|z| - \gamma_i z)^{\delta}$ in the Gaussian case. Lambert and Laurent (2001) show that for the standardized skewed Student:[10]

$$\kappa_i = \{\xi^{-(1+\delta)}(1+\gamma_i)^{\delta} + \xi^{1+\delta}(1-\gamma_i)^{\delta}\} \frac{\Gamma\left(\frac{\delta+1}{2}\right)\Gamma\left(\frac{\upsilon-\delta}{2}\right)(\upsilon-2)^{(1+\delta)/2}}{\left(\xi+\frac{1}{\xi}\right)\sqrt{(\upsilon-2)\pi}\,\Gamma\left(\frac{\upsilon}{2}\right)}.$$

For the GED, we can show that:

$$\kappa_i = \frac{[(1+\gamma_i)^{\delta} + (1-\gamma_i)^{\delta}]2^{(\delta-\upsilon)/\upsilon}\Gamma\left(\frac{\delta+1}{\upsilon}\right)\lambda_{\upsilon}^{\delta}}{\Gamma\left(\frac{1}{\upsilon}\right)}.$$

Note that ξ, υ and λ_{υ} concern the shape of the non-normal densities and will be defined in Section 3.2.

2.2.6. *IGARCH model*

In many high-frequency time-series applications, the conditional variance estimated using a GARCH(p, q) process has the following property:

$$\sum_{j=1}^{p} \beta_j + \sum_{i=1}^{q} \alpha_i \approx 1.$$

If $\sum_{j=1}^{p} \beta_j + \sum_{i=1}^{q} \alpha_i < 1$, the process ($\varepsilon_t$) is second order stationary, and a shock to the conditional variance σ_t^2 has a decaying impact on σ_{t+h}^2, when h increases, and is asymptotically negligible. Indeed, let us rewrite the ARCH(∞)

© Blackwell Publishers Ltd. 2002

representation of the GARCH(p, q), given in Eq. (10), as follows:

$$\sigma_t^2 = \omega^* + \lambda(L)\varepsilon_t^2, \tag{15}$$

where $\omega^* = \omega[1 - \beta(L)]^{-1}$, $\lambda(L) = \alpha(L)[1 - \beta(L)]^{-1} = \sum_{i=1}^{\infty} \lambda_i L^i$ and λ_i are lag coefficients depending nonlinearly on α_i and β_i. For a GARCH(1, 1), $\lambda_i = \alpha_1 \beta_1^{i-1}$. Recall that this model is said to be second order stationary provided that $\alpha_1 + \beta_1 < 1$ since it implies that the unconditional variance exists and equals $\omega/(1 - \alpha_1 - \beta_1)$. As shown by Davidson (2001), the amplitude of the GARCH(1, 1) is measured by $S = \sum_{i=1}^{\infty} \lambda_i = \alpha_1/(1 - \beta_1)$, which determines 'how large the variations in the conditional variance can be' (and hence the order of the existing moments). This concept is often confused with the memory of the model that determines 'how large shocks to the volatility take to dissipate'. In this respect, the GARCH(1, 1) model has a geometric memory $\rho = 1/\beta_1$, where $\lambda_i = O(\rho^{-i})$.

In practice, we often find $\alpha_1 + \beta_1 = 1$. In this case, we are confronted to an Integrated GARCH (IGARCH) model.

Recall that the GARCH(p, q) model can be expressed as an ARMA process. Using the lag operator L, we can rearrange Eq. (8) as:

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)](\varepsilon_t^2 - \sigma_t^2).$$

When the $[1 - \alpha(L) - \beta(L)]$ polynomial contains a unit root, i.e. the sum of all the α_i's and the β_j's is one, we have the IGARCH(p, q) model of Engle and Bollerslev (1986). It can then be written as:

$$\phi(L)(1 - L)\varepsilon_t^2 = \omega + [1 - \beta(L)](\varepsilon_t^2 - \sigma_t^2), \tag{16}$$

where $\phi(L) = [1 - \alpha(L) - \beta(L)](1 - L)^{-1}$ is of order $[\max\{p, q\} - 1]$.

We can rearrange Eq. (16) to express the conditional variance as a function of the squared residuals. After some manipulations, we have its ARCH(∞) representation:

$$\sigma_t^2 = \frac{\omega}{[1 - \beta(L)]} + \{1 - \phi(L)(1 - L)[1 - \beta(L)]^{-1}\}\varepsilon_t^2. \tag{17}$$

For this model, $S = 1$ and thus the second moment does not exist. However, this process is still short memory. To show this Davidson (2001) considers an IGARCH(0, 1) model defined as $\varepsilon_t = \sigma_t z_t$ and $\sigma_t^2 = \varepsilon_{t-1}^2$. This process is often wrongly compared to a random walk since the long-range forecast $\sigma_{t+h}^2 = \varepsilon_t^2$, for any h. However, $\varepsilon_t = z_t \mid \varepsilon_{t-1} \mid$ meaning that the memory of a large deviation persists for only one period.

2.2.7. *Fractionally integrated models*

Volatility tends to change quite slowly over time, and, as shown in Ding, Granger and Engle (1993) among others, the effects of a shock can take a considerable time to decay.[11] Therefore, the distinction between I(0) and I(1) processes seems to be far too restrictive. Indeed, the propagation of shocks in an I(0) process occurs at an exponential rate of decay (so that it only captures the short-memory), while for

© Blackwell Publishers Ltd. 2002

an I(1) process the persistence of shocks is infinite. In the conditional mean, the ARFIMA specification has been proposed to fill the gap between short and complete persistence, so that the short-run behavior of the time-series is captured by the ARMA parameters, while the fractional differencing parameter allows for modelling the long-run dependence.[12]

To mimic the behavior of the correlogram of the observed volatility, Baillie, Bollerslev and Mikkelsen (1996) (hereafter denoted BBM) introduce the Fractionally Integrated GARCH (FIGARCH) model by replacing the first difference operator of Eq. (17) by $(1-L)^d$.

The conditional variance of the FIGARCH (p, d, q) is given by:

$$\sigma_t^2 = \underbrace{\omega[1-\beta(L)]^{-1}}_{\omega^*} + \underbrace{\{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}}_{\lambda(L)}\varepsilon_t^2, \tag{18}$$

or $\sigma_t^2 = \omega^* + \sum_{i=1}^{\infty} \lambda_i L^i \varepsilon_t^2 = \omega^* + \lambda(L)\varepsilon_t^2$, with $0 \leqslant d \leqslant 1$. It is possible to show that $\omega > 0$, $\beta_1 - d \leqslant \phi_1 \leqslant (2-d)/2$ and $d[\phi_1 - (1-d)/2] \leqslant \beta_1(\phi_1 - \beta_1 + d)$ are sufficient to ensure that the conditional variance of the FIGARCH $(1, d, 1)$ is positive almost surely for all t. Setting $\phi_1 = 0$ gives the condition for the FIGARCH $(1, d, 0)$.

Davidson (2001) notes the interesting and counterintuitive fact that the memory parameter of this process is $-d$, and is increasing as d approaches zero, while in the ARFIMA model the memory increases when ζ increases. According to Davidson (2001), the unexpected behavior of the FIGARCH model may be due less to any inherent paradox than to the fact that, embodying restrictions appropriate to a model in levels, it has been transplanted into a model of volatility. The main characteristic of this model is that it is not stationary when $d > 0$. Indeed,

$$\begin{aligned}(1-L)^d &= \sum_{k=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} L^k \\ &= 1 - dL - \tfrac{1}{2}d(1-d)L^2 - \tfrac{1}{6}d(1-d)(2-d)L^3 - \cdots \\ &= 1 - \sum_{k=1}^{\infty} c_k(d)L^k,\end{aligned} \tag{19}$$

where $c_1(d) = d$, $c_2(d) = \frac{1}{2}d(1-d)$, etc. By construction, $\sum_{k=1}^{\infty} c_k(d) = 1$ for any value of d, and consequently, the FIGARCH belongs to the same 'knife-edge-nonstationary' class represented by the IGARCH. To test whether this nonstationarity feature holds, Davidson (2001) proposes a generalized version of the FIGARCH and calls it the HYperbolic GARCH. The HYGARCH is given by Eq. (18), when $\lambda(L)$ is replaced by $1-[1-\beta(L)]^{-1}\phi(L)\{1+\alpha[(1-L)^d - 1]\}$. Note that we report $\ln(\alpha)$ and not α. The $c_k(d)$ coefficients are thus weighted by α. Interestingly, the HYGARCH nests the FIGARCH when $\alpha = 1$ (or equivalently when $\ln(\alpha) = 0$) and if the GARCH component satisfies the usual covariance stationarity restrictions, then this process is stationary with $\alpha < 1$ (or equivalently when $\ln(\alpha) < 0$) (see Davidson, 2001 for more details).

© Blackwell Publishers Ltd. 2002

Chung (1999) underscores some drawbacks of the BBM model: there is a structural problem in the BBM specification since the parallel with the ARFIMA framework of the conditional mean equation is not perfect, leading to difficult interpretations of the estimated parameters. Indeed the fractional differencing operator applies to the constant term in the mean equation (ARFIMA) while it does not in the variance equation (FIGARCH). Chung (1999) proposes a slightly different process:

$$\phi(L)(1-L)^d(\varepsilon_t^2-\sigma^2)=[1-\beta(L)](\varepsilon_t^2-\sigma_t^2), \tag{20}$$

where σ^2 is the unconditional variance of ε_t.

If we keep the same definition of $\lambda(L)$ as in Eq. (18), we can formulate the conditional variance as:

$$\sigma_t^2=\sigma^2+\{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}(\varepsilon_t^2-\sigma^2)$$

or

$$\sigma_t^2=\sigma^2+\lambda(L)(\varepsilon_t^2-\sigma^2). \tag{21}$$

$\lambda(L)$ is an infinite summation which, in practice, has to be truncated. BBM propose to truncate $\lambda(L)$ at 1000 lags (this truncation order has been implemented as the default value in our package, but it may be changed by the user) and initialize the unobserved ε_t^2 at their unconditional moment. Contrary to BBM, Chung (1999) proposes to truncate $\lambda(L)$ at the size of the information set $(t-1)$ and to initialize the unobserved $(\varepsilon_t^2-\sigma^2)$ at 0 (this quantity is small in absolute values and has a zero mean).[13]

The idea of fractional integration has been extended to other GARCH types of models, including the Fractionally Integrated EGARCH (FIEGARCH) of Bollerslev and Mikkelsen (1996) and the Fractionally Integrated APARCH (FIAPARCH) of Tse (1998).[14]

Similarly to the GARCH(p,q) process, the EGARCH(p,q) of Eq. (11) can be extended to account for long memory by factorizing the autoregressive polynomial $[1-\beta(L)]=\phi(L)(1-L)^d$ where all the roots of $\phi(z)=0$ lie outside the unit circle. The FIEGARCH (p,d,q) is specified as follows:

$$\ln(\sigma_t^2)=\omega+\phi(L)^{-1}(1-L)^{-d}[1+\alpha(L)]g(z_{t-1}). \tag{22}$$

Finally, the FIAPARCH (p,d,q) model can be written as:[15]

$$\sigma_t^\delta=\omega+\{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}(|\varepsilon_t|-\gamma\varepsilon_t)^\delta. \tag{23}$$

3. Estimation Methods

3.1. *Parameters Constraints*

When numerical optimization is used to maximize the log-likelihood function with respect to the vector of parameters Ψ, the inspected range of the parameter space

© Blackwell Publishers Ltd. 2002

is $]-\infty;\ \infty[$. The problem is that some parameters might have to be constrained in a smaller interval. For instance, the leverage effect parameter γ of the APARCH model must lie between -1 and 1. To impose these constraints one could estimate Ψ^* (which ranges from $-\infty$ to $+\infty$) instead of Ψ where Ψ is recovered using the non-linear function: $\Psi = x(\Psi^*)$. In our package, $x(.)$ is defined as:

$$x(\Psi^*) = Low + \frac{Up - Low}{1 + e^{-\Psi^*}}, \tag{24}$$

where *Low* is the lower bound and *Up* the upper bound (i.e. in our example, $Low = -1$ and $Up = 1$).

Applying unconstrained optimization of the log-likelihood function with respect to Ψ is equivalent to applying constrained optimization with respect to Ψ^*. Therefore, the optimization process of the program results in $\hat{\Psi}^*$ with the covariance matrix being noted $Cov(\hat{\Psi}^*)$. The estimated covariance of the parameters of interest $\hat{\Psi}$ is:

$$Cov(\hat{\Psi}) = \left(\frac{\partial x(\hat{\Psi}^*)}{\partial \Psi^*}\right) Cov(\hat{\Psi}^*) \left(\frac{\partial x(\hat{\Psi}^*)}{\partial \Psi^*}\right)'. \tag{25}$$

In our case, we have

$$Cov(\hat{\Psi}) = Cov(\hat{\Psi}^*) \frac{\exp(-\hat{\Psi}^*)(Up - Low)}{[1 + \exp(-\hat{\Psi}^*)]^2}.$$

Note that, in G@RCH 2.2, lower and upper bounds of the parameters can be easily modified by the user in the file *startingvalues.txt*.

3.2. *Distributions*

Four distributions are available in our program: the usual Gaussian, the Student-*t*, the Generalized Error Distribution (GED) and the skewed Student distribution.

The GARCH models are estimated using an approximate Maximum Likelihood (ML) approach. It is evident from Eq. (6) (and all the following equations of Section 2) that the recursive evaluation of this function is conditional on unobserved values. The ML estimation is therefore not perfectly exact. To solve the problem of unobserved values, we have set these quantities to their unconditional expected values.

If we express the mean equation as in Eq. (1) and $\varepsilon_t = z_t \sigma_t$, the log-likelihood function of the standard normal distribution is given by:

$$L_{norm} = -\frac{1}{2} \sum_{t=1}^{T} [\ln(2\pi) + \ln(\sigma_t^2) + z_t^2], \tag{26}$$

where T is the number of observations.

© Blackwell Publishers Ltd. 2002

For a Student-t distribution, the log-likelihood is:

$$L_{Stud} = T\left\{ \ln \Gamma\left(\frac{\upsilon+1}{2}\right) - \ln \Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}\ln[\pi(\upsilon-2)] \right\}$$
$$-\frac{1}{2}\sum_{t=1}^{T}\left[\ln(\sigma_t^2) + (1+\upsilon)\ln\left(1+\frac{z_t^2}{\upsilon-2}\right)\right], \tag{27}$$

where υ is the degrees of freedom, $2 < \upsilon \leqslant \infty$ and $\Gamma(.)$ is the gamma function.

The GED log-likelihood function of a normalized random variable is given by:

$$L_{GED} = \sum_{t=1}^{T}\left[\ln(\upsilon/\lambda_\upsilon) - 0.5\left|\frac{z_t}{\lambda_\upsilon}\right|^{\upsilon} - (1+\upsilon^{-1})\ln(2) - \ln \Gamma(1/\upsilon) - 0.5\ \ln(\sigma_t^2)\right], \tag{28}$$

where $0 < \upsilon < \infty$ and

$$\lambda_\upsilon \equiv \sqrt{\frac{\Gamma\left(\frac{1}{\upsilon}\right)2^{-2/\upsilon}}{\Gamma\left(\frac{3}{\upsilon}\right)}}. \tag{29}$$

The main drawback of the last two densities is that despite accounting for fat tails, they are symmetric. Skewness and kurtosis are important in many financial applications (in asset pricing models, portfolio selection, option pricing theory or Value-at-Risk applications among others). To overcome this problem, we can rely on the skewed Student density proposed by Lambert and Laurent (2001) whose log-likelihood is:

$$L_{SkSt} = T\left\{ \ln \Gamma\left(\frac{\upsilon+1}{2}\right) - \ln \Gamma\left(\frac{\upsilon}{2}\right) - 0.5\ \ln[\pi(\upsilon-2)] + \ln\left(\frac{2}{\xi+\frac{1}{\xi}}\right) + \ln(s) \right\}$$
$$-0.5\sum_{t=1}^{T}\left\{\ln \sigma_t^2 + (1+\upsilon)\ln\left[1+\frac{(sz_t+m)^2}{\upsilon-2}\xi^{-2I_t}\right]\right\}. \tag{30}$$

© Blackwell Publishers Ltd. 2002

where

$$I_t = \begin{cases} 1 & \text{if } z_t \geqslant -\dfrac{m}{s} \\ -1 & \text{if } z_t < -\dfrac{m}{s} \end{cases},$$

ξ is the asymmetry parameter, υ is the degree of freedom of the distribution,

$$m = \frac{\Gamma\left(\dfrac{\upsilon+1}{2}\right)\sqrt{\upsilon-2}}{\sqrt{\pi}\Gamma\left(\dfrac{\upsilon}{2}\right)}\left(\xi - \frac{1}{\xi}\right)$$

and

$$s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2}.$$

In principle, the gradient vector and the hessian matrix can be obtained numerically or by evaluating its analytic expressions. Due to the high number of possible models and distributions, we use numerical techniques to approximate the derivatives of the log-likelihood function with respect to the parameter vector.

3.3. *Tests*

In addition to the possibilities offered by GiveWin (ACF, PACF, QQ-plots…), several tests are provided:

- Four Information Criteria (divided by the number of observations):[16]
  - — Akaike $= -2\dfrac{LogL}{n} + 2\dfrac{k}{n}$;
  - — Hannan-Quinn $= -2\dfrac{LogL}{n} + 2\dfrac{k\ \ln[\ln(n)]}{n}$;
  - — Schwartz $= -2\dfrac{LogL}{n} + 2\dfrac{\ln(k)}{n}$;
  - — Shibata $= -2\dfrac{LogL}{n} + \ln\left(\dfrac{n+2k}{n}\right)$.

© Blackwell Publishers Ltd. 2002

- The value of the skewness and the kurtosis of the standardized residuals ($\hat{z}_t$) of the estimated model, their t-tests and p-values. The Jarque-Bera normality test (Jarque and Bera, 1987) is also reported.
- The Box-Pierce statistics at lag l^* for both standardized, i.e. $BP(l^*)$, and squared standardized, i.e. $BP^2(l^*)$, residuals. Under the null hypothesis of no autocorrelation, the statistics $BP(l^*)$ and $BP^2(l^*)$ are respectively $\chi^2(l^* - m - l)$ and $\chi^2(l^* - p - q)$ distributed (see McLeod and Li, 1983).
- The Engle LM ARCH test (Engle, 1982) to test for the presence of ARCH effects in a series.
- The diagnostic test of Engle and Ng (1993) to investigate possible mispecification of the conditional variance equation. The Sign Bias Test (SBT) examines the impact of positive and negative return shocks on volatility not predicted by the model under construction. The negative Size Bias Test (resp. positive Size Bias Test) focuses on the different effects that large and small negative (resp. positive) return shocks have on volatility, which is not predicted by the volatility model. Finally, a joint test for these three tests is also provided.
- The adjusted Pearson goodness-of-fit test. The Pearson goodness-of-fit test compares the empirical distribution of the innovations with the theoretical one. In order to carry out this testing procedure, it is necessary to first classify the residuals in cells according to their magnitude.[17] For a given number of cells denoted g, the Pearson goodness-of-fit statistics is:

$$P(g) = \sum_{i=1}^{g} \frac{(n_i - En_i)^2}{En_i}, \tag{31}$$

  where n_i is the number of observations in cell i and En_i is the expected number of observations (based on the ML estimates). For *i.i.d.* observations, Palm and Vlaar (1997) show that under the null of a correct distribution the asymptotic distribution of $P(g)$ is bounded between a $\chi^2(g-1)$ and a $\chi^2(g-k-1)$ where k is the number of estimated parameters. As explained by Palm and Vlaar (1997), the choice of g is far from being obvious. For $T = 2252$, these authors set g equal to 50. According to König and Gaab (1982), the number of cells must increase at a rate equal to $T^{0.4}$.
- The Nyblom test (Nyblom, 1989; and Lee and Hansen, 1994) to check the constancy of parameters over time. See Hansen (1994) for an overview of this test.

3.4. *Forecasts*

When estimating a model it can be useful to try to understand the mechanism that produces the series of interest. It can also suggest a solution to an economic problem. Is it the only game in town? Certainly not. Indeed, the main purpose of building and estimating a model with financial data is to produce a forecast.

© Blackwell Publishers Ltd. 2002

G@RCH 2.2 also provides forecasting tools. In particular, forecasts of both the conditional mean and the conditional variance are available as well as several forecast error measures.

3.4.1. *Forecasting the conditional mean*

Our first goal is to give the optimal h-step-ahead predictor of y_{t+h} given the information we have up to time t.

For instance, for the following AR(1) process,

$$y_t = \mu + \psi_1(y_{t-1} - \mu) + \varepsilon_t. \tag{32}$$

The optimal[18] h-step-ahead predictor of y_{t+h}, i.e. $\hat{y}_{t+h|t}$, is its conditional expectation at time t (given the estimated parameters $\hat{\mu}$ and $\hat{\psi}_1$):

$$\hat{y}_{t+h|t} = \hat{\mu} + \hat{\psi}_1(\hat{y}_{t+h-1|t} - \hat{\mu}), \tag{33}$$

where $\hat{y}_{t+i|t} = y_{t+i}$ for $i \leqslant 0$.

For the AR(1), the optimal 1-step-ahead forecast equals $\hat{\mu} + \hat{\psi}_1(\hat{y}_t - \hat{\mu})$. For $h > 1$, the optimal forecast can be obtained recursively or directly as $\hat{y}_{t+h|t} = \hat{\mu} + \hat{\psi}_1^h(\hat{y}_t - \hat{\mu})$.

In the general case of an ARFIMA(n, ζ, s) as given in Eq. (3), the optimal h-step-ahead predictor of y_{t+h} is:

$$\begin{aligned}\hat{y}_{t+h|t} = {} & \left[\hat{\mu}_{t+h|t} + \sum_{k=1}^{\infty} \hat{c}_k(\hat{y}_{t+h-k} - \hat{\mu}_{t+h|t})\right] \\ & + \sum_{i=1}^{n} \hat{\psi}_i \left\{ \hat{y}_{t+h-i} - \left[\hat{\mu}_{t+h|t} + \sum_{k=1}^{\infty} \hat{c}_k(\hat{y}_{t+h-i-k} - \hat{\mu}_{t+h|t})\right]\right\} \\ & + \sum_{j=1}^{s} \hat{\theta}_j(\hat{y}_{t+h-j} - \hat{y}_{t+h-j|t}). \end{aligned} \tag{34}$$

Recall that when exogenous variables enter the conditional mean equation, μ becomes $\mu_t = \mu + \sum_{i=1}^{n_1} \delta_i x_{i,t}$ and consequently, provided that the information $x_{i,t+h}$ is available at time t (which is the case for instance if $x_{i,t}$ is a 'day-of-the-week' dummy variable), $\hat{\mu}_{t+h|t}$ is also available at time t. When there is no exogenous variable in the ARFIMA model and $n = 1$, $s = 0$ and $\zeta = 0$ ($c_k = 0$), the forecast of the AR(1) process given in Eq. (33) can be recovered.

3.4.2. *Forecasting the conditional variance*

Independently from the conditional mean, one can forecast the conditional variance. In the simple GARCH(p, q) case, the optimal h-step-ahead forecast of

© Blackwell Publishers Ltd. 2002

the conditional variance, i.e. $\hat{\sigma}^2_{t+h|t}$ is given by:

$$\sigma^2_{t+h|t} = \hat{\omega} + \sum_{i=1}^{q} \hat{\alpha}_i \varepsilon^2_{t+h-i|t} + \sum_{j=1}^{p} \hat{\beta}_j \sigma^2_{t+h-j|t}, \tag{35}$$

where $\varepsilon^2_{t+i|t} = \sigma^2_{t+i|t}$ for $i > 0$ while $\varepsilon^2_{t+i|t} = \varepsilon^2_{t+i}$ and $\sigma^2_{t+i|t} = \sigma^2_{t+i}$ for $i \leqslant 0$. Eq. (35) is usually computed recursively, even if a closed form solution of $\sigma^2_{t+h|t}$ can be obtained by recursive substitution in Eq. (35).

Similarly, one can easily obtain the h-step-ahead forecast of the conditional variance of an ARCH, IGARCH and FIGARCH model. By contrast, for threshold models, the computation of the out-of-sample forecasts is more complicated. Indeed, for the EGARCH, GJR and APARCH models (as well as for their long-memory counterparts), the assumption made on the innovation process may have an effect on the forecast (especially for $h > 1$).

For instance, for the GJR (p, q) model,

$$\hat{\sigma}^2_{t+h|t} = \hat{\omega} + \sum_{i=1}^{q} (\hat{\alpha}_i \varepsilon^2_{t-i+h|t} + \hat{\gamma}_i S^-_{t-i+h|t} \varepsilon^2_{t-i+h|t}) + \sum_{j=1}^{p} \hat{\beta}_j \sigma^2_{t-j+h|t}. \tag{36}$$

When all the γ_i parameters equal 0, one recovers the forecast of the GARCH model. Otherwise, one has to compute $S^-_{t-i+h|t}$. Note first that $S^-_{t+i|t} = S^-_{t+i}$ for $i \leqslant 0$. However, when $i > 1$, $S^-_{t+i|t}$ depends on the distribution choice of z_t. When the distribution of z_t is symmetric around 0 (for the Gaussian, Student and GED density), the probability that ε_{t+i} will be negative is $S^-_{t+i|t} = 0.5$. If z_t is (standardized) skewed Student distributed with asymmetry parameter ξ and degree of freedom υ, $S^-_{t+i|t} = 1/(1+\xi^2)$ since ξ^2 is the ratio of probability masses above and below the mode.

For the APARCH (p, q) model,

$$\begin{aligned}
\hat{\sigma}^\delta_{t+h|t} &= E(\sigma^\delta_{t+h} \mid \Omega_t) \\
&= E\left(\hat{\omega} + \sum_{i=1}^{q} \hat{\alpha}_i (|\varepsilon_{t+h-i}| - \hat{\gamma}_i \varepsilon_{t+h-i})^{\hat{\delta}} + \sum_{j=1}^{p} \hat{\beta}_j \sigma^{\hat{\delta}}_{t+h-j} \mid \Omega_t \right) \\
&= \hat{\omega} + \sum_{i=1}^{q} \hat{\alpha}_i E[(\varepsilon_{t+h-i} - \hat{\gamma}_i \varepsilon_{t+h-i})^{\hat{\delta}} \mid \Omega_t] + \sum_{j=1}^{p} \hat{\beta}_j \sigma^{\hat{\delta}}_{t+h-j|t},
\end{aligned} \tag{37}$$

where $E[(\varepsilon_{t+k} - \hat{\gamma}_i \varepsilon_{t+k})^{\hat{\delta}} \mid \Omega_t] = \kappa_i \sigma^{\hat{\delta}}_{t+k|t}$, for $k > 1$ and $\kappa_i = E(|z| - \gamma_i z)^{\hat{\delta}}$ (see Section 3.2).

For the EGARCH (p, q) model,

$$\begin{aligned}
\ln \hat{\sigma}^2_{t+h|t} &= E(\ln \sigma^2_{t+h} \mid \Omega_t) \\
&= E\{\hat{\omega} + [1 - \hat{\beta}(L)]^{-1}[1 + \hat{\alpha}(L)]\hat{g}(z_{t+h-1}) \mid \Omega_t\} \\
&= [1 - \hat{\beta}(L)]\hat{\omega} + \hat{\beta}(L) \ln \hat{\sigma}^2_{t+h|t} + [1 + \hat{\alpha}(L)]\hat{g}(z_{t+h-1|t}),
\end{aligned} \tag{38}$$

where $\hat{g}(z_{t+k|t}) = \hat{g}(z_{t+k})$ for $k \leqslant 0$ and 0 for $k > 0$.

© Blackwell Publishers Ltd. 2002

Finally, the h-step-ahead forecast of the FIAPARCH and FIEGARCH models are obtained in a similar way.

One of the most popular measures to check the forecasting performance of the ARCH-type models is the Mincer-Zarnowitz regression, i.e. on ex-post volatility regression:

$$\check{\sigma}_t^2 = a_0 + a_1\hat{\sigma}_t^2 + u_t, \tag{39}$$

where $\check{\sigma}_t^2$ is the ex-post volatility, $\hat{\sigma}_t^2$ is the forecasted volatility and a_0, a_1 are parameters to be estimated. If the model for the conditional variance is correctly specified (and the parameters are known) and $E(\check{\sigma}_t^2) = \hat{\sigma}_t^2$, it follows that $a_0 = 0$ and $a_1 = 1$. The R^2 of this regression is often used as a simple measure of the degree of predictability of the ARCH-type model.

However, $\check{\sigma}_t^2$ is never observed. By default, G@RCH 2.2 uses $\check{\sigma}_t^2 = (y_t - \bar{y})^2$, where $\bar{y}$ is the sample mean of y_t. The R^2 of this regression is often lower than 5% and this could lead to the conclusion that GARCH models produce poor forecasts of the volatility (see, among others, Schwert, 1990; or Jorion, 1996). But, as described in Andersen and Bollerslev (1998), the reason of these poor results is the choice of what is considered as the 'true' volatility. G@RCH 2.2 allows selection of any series as the 'observed' volatility (Obs.-Var., see Figure 1). The user may then compute the daily realized volatility as the sum of squared intraday returns and use it as the 'true' volatility. Actually, Andersen and Bollerslev (1998) show that this measure is a more useful one than squared daily returns. Therefore, using 5-minute returns for instance, the realized volatility can be expressed as:

$$\check{\sigma}_t^2 = \sum_{k=1}^{K} y_{k,t}^2, \tag{40}$$

where $y_{k,t}$ is the return of the k^{th} 5-minutes interval of the t^{th} day and K is the number of 5-minutes intervals per day.

Finally, to compare the adequacy of the different distributions, G@RCH 2.2 also allows the computation of density forecasts tests developed in Diebold, Gunther and Tay (1998). The idea of density forecasts is quite simple.[19] Let $f_i(y_i \mid \Omega_i)_{i=1}^m$ be a sequence of m one-step-ahead density forecasts produced by a given model, where Ω_i is the conditioning information set, and $p_i(y_i \mid \Omega_i)_{i=1}^m$ the sequence of densities defining the Data Generating Process y_i (which is never observed). Testing whether this density is a good approximation of the true density $p(.)$ is equivalent to testing:

$$H_0 : f_i(y_i \mid \Omega_i)_{i=1}^m = p_i(y_i \mid \Omega_i)_{i=1}^m \tag{41}$$

Diebold, Gunther and Tay (1998) use the fact that, under Eq. (41), the probability integral transform $\hat{\zeta}_i = \int_{-\infty}^{y_i} f_i(t)dt$ is *i.i.d.* $U(0, 1)$, i.e. independent and identically distributed uniform. To check H_0, they propose to use a goodness-of-fit test and independence test for *i.i.d.* $U(0, 1)$. The *i.i.d.*-ness property of $\hat{\zeta}_i$ can be evaluated by plotting the correlograms of $(\zeta - \bar{\hat{\zeta}})^j$, for $j = 1, 2, 3, 4, \ldots$, to detect potential

© Blackwell Publishers Ltd. 2002

dependence in the conditional mean, variance, skewness, kurtosis, etc. Departure from uniformity can also be evaluated by plotting an histogram of $\hat{\zeta}_i$. According to Bauwens, Giot, Grammig and Veredas (2000), *a humped shape of the $\hat{\zeta}$-histogram would indicate that the issued forecasts are too narrow and that the tails of the true density are not accounted for. On the other hand, a U-shape of the histogram would suggest that the model issues forecasts that either under- or overestimate too frequently*. Moreover, Lambert and Laurent (2001) show that an *inverted S* shape of the histogram would indicate that the errors are skewed, i.e. the true density is probably not symmetric.[20] An illustration is provided in Section 4 with some formal tests and graphical tools.

3.5. *Accuracy*

McCullough and Vinod (1999) and Brooks, Burke and Persand (2001) use the daily German mark/British pound exchange rate data of Bollerslev and Ghysels (1996) to compare the accuracy of GARCH model estimation among several econometric software packages. They choose the GARCH(1,1) model described in Fiorentini, Calzolani and Pamattani (1996) (hereafter denoted FCP) as the benchmark. In this section, we use the same methodology with the same dataset to check the accuracy of our procedures. Coefficients and standard error estimates of G@RCH 2.2 are reported in Table 1 together with the results of McCullough and Vinod (1999) (based on the FORTRAN procedure of FCP and thus entitled 'FCP' in the table).

G@RCH 2.2 gives very satisfactory results since the first four digits (at least) are the same as those of the benchmark for all but two estimations. In addition, it competes well compared to other well known econometric softwares. Table 2 presents the coefficient estimates and the error percentage associated for with the 5 pieces of software. G@RCH, PcGive and TSP (these last two software packages use analytical second-order derivatives for the standard GARCH model) clearly outperform Eviews and S-Plus on this specification.

Moreover, to investigate the accuracy of our forecasting procedures, we have run an 8-step ahead forecast of the model, similar to Brooks, Burk and Persand (2001). Table 4 in Brooks, Burke, and Persand (2001) reports the conditional

Table 1. Accuracy of the GARCH procedure

	Coefficient		Standard Errors		Robust Standard Errors	
	G@RCH	FCP	G@RCH	FCP	G@RCH	FCP
μ	−0.006184	−0.006190	0.008462	0.008462	0.009187	0.009189
ω	0.010760	0.010761	0.002851	0.002852	0.006484	0.006493
α_1	0.153407	0.153134	0.026569	0.026523	0.053595	0.053532
β_1	0.805879	0.805974	0.033542	0.033553	0.072386	0.072461

© Blackwell Publishers Ltd. 2002

Table 2. GARCH Accuracy Comparison

	FCP	G@RCH	Eviews	PcGive	TSP	S-Plus
μ	−0.00619	−0.00618	−0.00541	−0.00625	−0.00619	−0.00919
ω	0.010761	0.010760	0.009581	0.010760	0.010761	0.011696
α_1	0.153134	0.153407	0.142284	0.153397	0.153134	0.154295
β_1	0.805974	0.805879	0.821336	0.805886	0.805974	0.800276
μ	—	0.10%	12.58%	0.91%	0.00%	48.41%
ω	—	0.01%	10.96%	0.01%	0.00%	8.69%
α_1	—	0.18%	7.08%	0.17%	0.00%	0.76%
β_1	—	0.01%	1.91%	0.01%	0.00%	0.71%

variance forecasts given by six well-known software packages and the correct values. Contrary to E-Views, Matlab and SAS, G@RCH 2.2 hits the benchmarks for all steps to the third decimal (note that GAUSS, Microfit and Rats also do).

Finally, Lombardi and Gallo (2001) extend the work of Fiorentini, Calzolani and Pamattani (1996) to the FIGARCH model of Baillie, Bollerslev and Mikkelsen (1996) and derive analytic expressions for the second-order derivatives of this model in the Gaussian case. For the same DEM/UKP database as in the previous example, Table 3 reports the coefficient estimates and their standard errors for our package (using numerical gradients and the BFGS optimization method) and for Lombardi and Gallo (2001) (using analytical gradients and the Newton-Raphson algorithm; results correspond to the columns entitled 'LG').

Results show that G@RCH 2.2 provides accurate estimates, even for an advanced model such as the FIGARCH. As expected, it is however more time-consuming than the C code of Lombardi and Gallo (2001)[21] (163 sec. vs 43 sec. using a PIII processor with 450 Mhz).

3.6. *Features comparison*

The goal of this section is to compare more objectively the features offered by G@RCH 2.2 with respect to nine other well known econometric software packages, namely PcGive 10 (also programmed in Ox), GAUSS and its Fanpac package, Eviews 4, S-Plus 6 and its GARCH module, Rats 5.0 and its *garch.src*

Table 3. Accuracy of the FIGARCH procedure

	Coefficient		Standard Errors	
	G@RCH	LG	G@RCH	LG
μ	0.003606	0.003621	0.009985	0.009985
ω	0.015772	0.015764	0.003578	0.003581
α_1	0.198134	0.198448	0.042508	0.042444
β_1	0.675652	0.675251	0.051800	0.051693
d	0.570702	0.569951	0.075039	0.074762

© Blackwell Publishers Ltd. 2002

Table 4. GARCH Features Comparison

	G@RCH	PcGive	Fanpac	Eviews	S-Plus	Rats	TSP	Microfit	SAS	Stata
Version	2.0	10	–	4.0	6	5.0	4.5	4	8.2	7
Conditional mean										
Explanatory variables	+	+	+	+	+	+	+	+	+	+
ARMA	+	+	+	+	+	+	+	+	+	+
ARFIMA	+	−	−	−	−	−	−	−	−	−
ARCH-in-Mean	−	+	+	+	+	+	+	−	+	+
Conditional variance										
Explanatory variables	+	+	+	+	+	+	+	+	+	+
GARCH	+	+	+	+	+	+	+	+	+	+
IGARCH	+	−	+	−	−	+	−	−	+	−
EGARCH	+	+	+	+	+	+	−	+	+	+
GJR	+	+	−	+	+	+	−	−	−	+
APARCH	+	−	−	−	+	−	−	−	−	+
C-GARCH	−	−	−	+	+	−	−	−	−	−
FIGARCH	+	−	+	−	+	−	−	−	−	−
FIEGARCH	+	−	−	−	+	−	−	−	−	−
FIAPARCH	+	−	−	−	−	−	−	−	−	−
HYGARCH	+	−	−	−	−	−	−	−	−	−
Distributions										
Normal	+	+	+	+	+	+	+	+	+	+
Student-t	+	+	+	−	+	+	−	+	+	−
GED	+	+	−	−	+	+	−	−	−	−
Skewed-t	+	−	−	−	−	−	−	−	−	−
Double Exponential	−	−	−	−	+	−	−	−	−	−
Estimation										
MLE	+	+	+	+	+	+	+	+	+	+
QMLE	+	+	+	+	−	−	−	−	−	+

A '+' (resp. '−') means that the corresponding option is (resp. is not) available for this software. C-GARCH corresponds to the Component GARCH of Engle and Lee (1999).

© Blackwell Publishers Ltd. 2002

procedure,[22] TSP 4.5, Microfit 4, SAS 8.2 and Stata 7. It is thus not our intention to evaluate a program against another, but we will rather present an overview of what can or cannot be done with these packages.

The proposed models and options differ widely from one program to another as can be seen in Table 4. Regarding the range of different univariate models, many programs propose asymmetric models, very few (G@RCH, S-Plus with the FIGARCH and the FIEGARCH and Fanpac with the FIGARCH) offer long memory models in the variance equation and none (except G@RCH) offers a fractionally integrated specification in the mean. As for the distribution, the choice is often limited to symmetric densities (except G@RCH which provides a skewed Student likelihood). Finally, robust standard errors are proposed in 5 of the 10 packages we have compared (G@RCH, PcGive, GAUSS Fanpac, Eviews and Stata).

4. Application

4.1. *Data and methodology*

To illustrate the G@RCH 2.2 package with a concrete application, we analyze the French CAC40 stock index for the years 1995–1999 (1249 daily observations). It is computed by the exchange as a weighted measure of the prices of its components and is available in the database on an intraday basis with the price index being computed every 15 minutes. For the time period under review, the opening hours of the French stock market were 10.00 am to 5.00 pm, thus 7 hours of trading per day. This translates into 28 intraday returns used to compute the daily realized volatility. Intraday prices are the outcomes of a linear interpolation between the closest recorded prices below and above the time set in the grid. Correspondingly, all returns are computed as the first difference in the regularly time-spaced log prices of the index. Because the exchange is closed from 5.00 pm to 10.00 am the next day, the first intraday return is the first difference between the log price at 10.15 am and the log price at 5.00 pm the day before. Then, the intraday data are used to compute the daily realized volatility using Eq. (40). Finally, daily returns in percentage are defined as 100 times the first difference of the log of the closing prices.[23]

The estimation of the parameters is carried out for the 800 observations while forecasting is computed for the last observations.

4.2. *Using the 'full version'*

Once the installation process is correctly completed following the instructions of the *readme.txt* file, the user may open the database in GiveWin (in the example 'CAC15.xls'), and then select the OxPack module.

Once the package has been selected, one can launch the **Model/Formulate** menu. The list of all the variables of the database appears in the *Database* section (see Figure 1). There are four possible statuses for each variable: dependent

© Blackwell Publishers Ltd. 2002

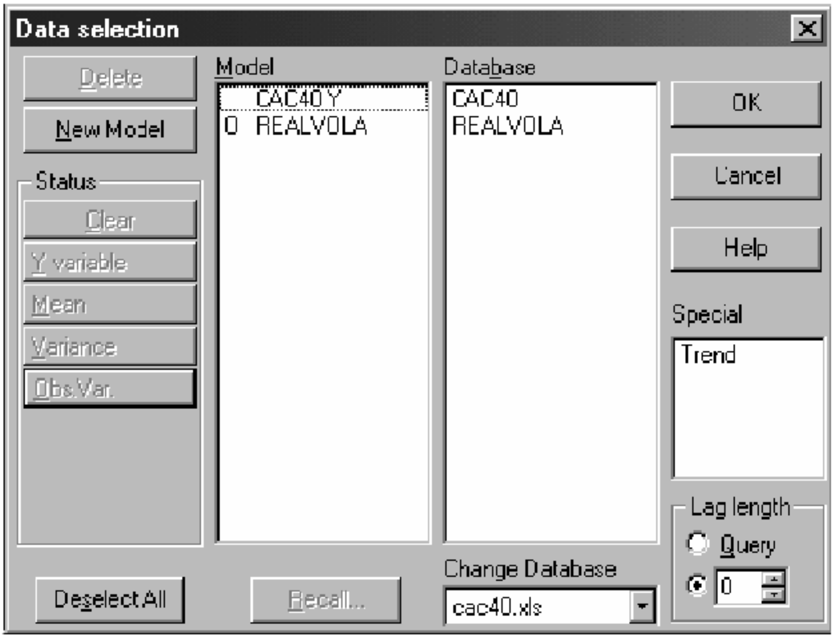

Figure 1. Selecting the variables.

variable (Y variable), regressor in the mean (Mean), regressor in the variance (Variance) or observed volatility (Obs. Var.). The program provides estimates for univariate models,[24] so only one Y variable per model is accepted. However one can include several regressors in the mean and the variance equations and the same variable can be a regressor in both equations.

Once the OK button is pressed, the **Model/Model Settings** box automatically appears. This box allows selection and specification of the model: AR(FI)MA orders for the mean equation, GARCH orders, type of GARCH model for the variance equation and the distribution (Figure 2). The default specification is an ARMA(0,0)-GARCH(1,1) with normal errors. In our application, we select an ARMA(1,0)-APARCH(1,1) specification with a skewed Student likelihood.

As explained in Section 3.1, it is possible to constrain the parameters to range between a lower and an upper bound by selecting the **Bounded Parameters** option. The defaults bounds can be changed in the *startingvalues.txt* file.

In the next window, the user is asked to make a choice regarding the starting values (Figure 3): they might (1) let the program use the predefined starting values,[25] (2) enter them manually, element by element, or (3) enter the starting values in a vector form (the required form is ‘value1;value2;value3’).

Then, the estimation method for standard deviations is selected: ML or QML (with a specified pseudo-likelihood) or both. In this window (see Figure 4), one may also select the sample and some maximization options (such has the number

© Blackwell Publishers Ltd. 2002

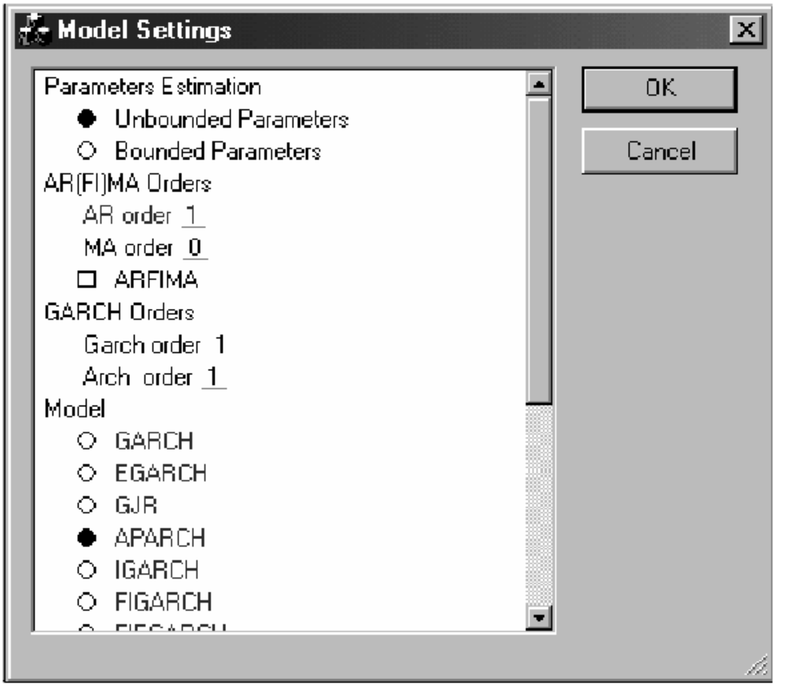

Figure 2. Model Settings.

of iterations between intermediary results printings) when clicking on the *Options* button.

The estimation procedure is then launched and the program comes back to GiveWin. Let us assume that the element-by-element method has been selected. A new window appears (see Figure 5) with all the possible parameters to be estimated. Depending on the specification, some parameters have a value, others do not. The user should replace only the former, since they correspond to the parameters to be estimated for the specified model.

Once this step is completed, the program starts the iteration process. The final output is divided by default into two main parts: first, the model specification reminder; second, the estimated values and other useful statistics of the parameters.[26] The output is given in the box 'Output 1'.

After the estimation of the model, new options are available in **OxPack: Menu/Tests, Menu/Graphic Analysis, Menu/Forecasts, Menu/Exclusion Restrictions, Menu/Linear Restrictions** and **Menu/Store**.

The **Menu/Graphic Analysis** option allows ploting using different graphics (see Figure 6 for details). Just as any other graphs in the GiveWin environment, they can be easily edited (color, size, ...) and exported in many formats (.eps, .ps, .wmf,

© Blackwell Publishers Ltd. 2002

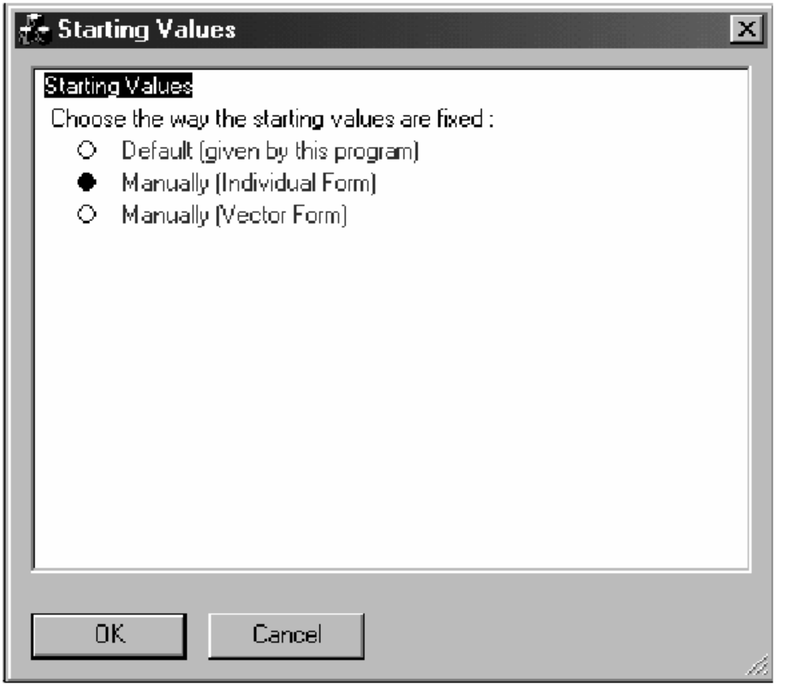

Figure 3. Selecting the Starting Values Method.

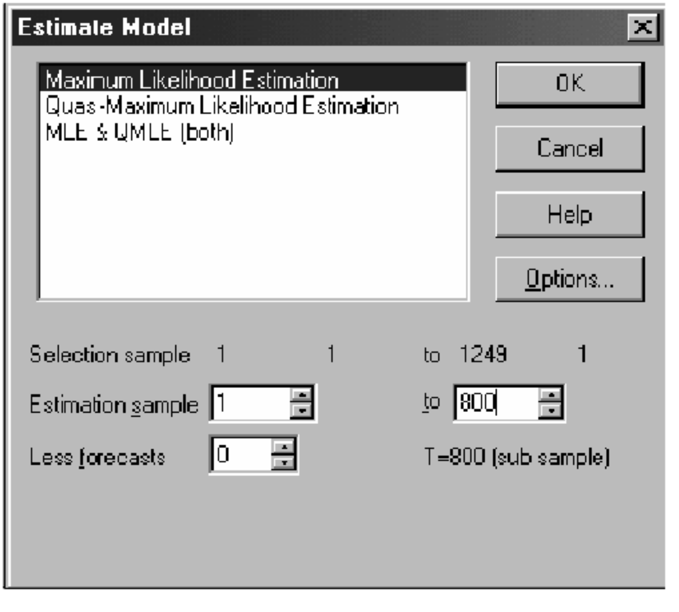

Figure 4. Standard Errors Estimation Methods.

© Blackwell Publishers Ltd. 2002

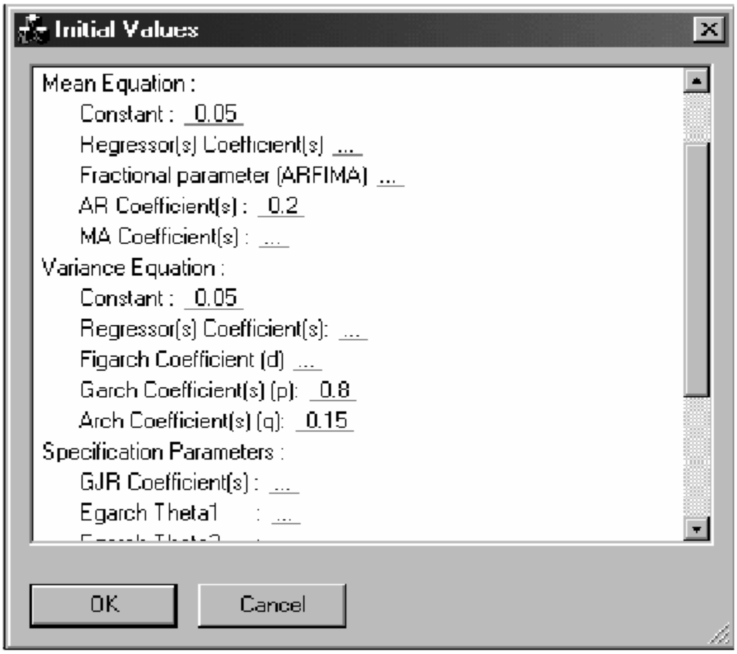

Figure 5. Entering the Starting Values.

Output 1

****
* SPECIFICATIONS **
****

Mean Equation: ARMA (1, 0) model.
No regressor in the mean.
Variance Equation : APARCH (1, 1) model.
No regressor in the variance.
The distribution is a Skewed Student distribution, with a tail coefficient of 15.72 and an asymmetry coefficient of -0.08751.
Strong convergence using numerical derivatives

Maximum Likelihood Estimation

	Coefficient	Std.Error	t-value	t-prob
Cst(M)	0.065337	0.037157	1.758	0.0791
AR(1)	0.004704	0.037117	0.1267	0.8992
Cst(V)	0.017498	0.013488	1.297	0.1949
Beta1	0.947590	0.020193	46.93	0.0000
Alpha1	0.038464	0.017776	2.164	0.0308
Gamma1	0.676364	0.348702	1.940	0.0528
Delta	1.462837	0.533581	2.742	0.0063
Asymmetry	-0.087512	0.054314	-1.611	0.1075
Tail	15.718323	8.087414	1.944	0.0523

No. Observations: 800 — No. Parameters: 9
Mean (Y): 0.08103 — Variance (Y): 1.27405
Log Likelihood: -1190.521 — Alpha[1]+Beta[1]: 0.98605

The sample mean of squared residuals was used to start recursion.
The condition for existence of $E(\sigma^{\delta})$ and $E(|e^{\delta}|)$ is observed.
The constraint equals 0.9926 and should be < 1.
Vector of estimated parameters:
0.065337; 0.004704; 0.017498; 0.947590; 0.038464; 0.676364; 1.462837;-0.087512; 15.718323

© Blackwell Publishers Ltd. 2002

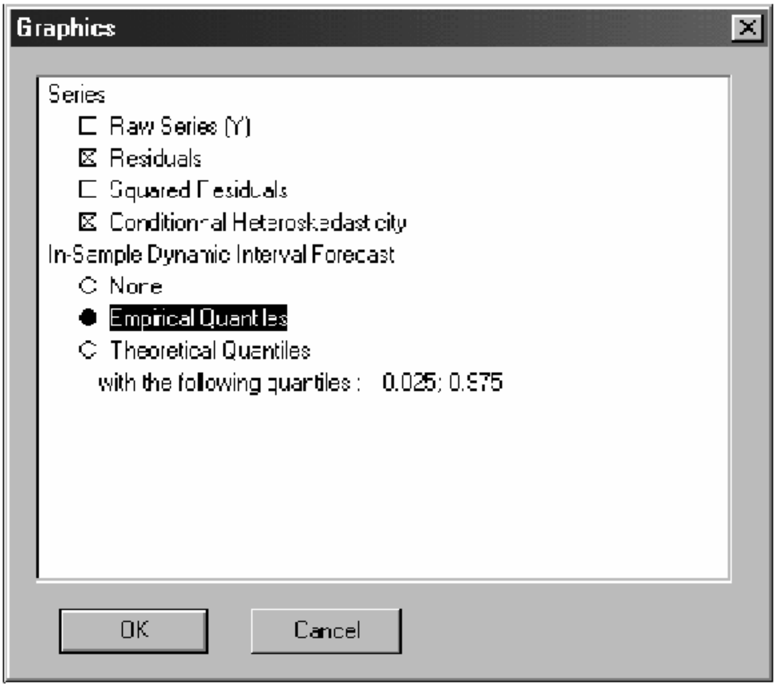

Figure 6. Graphics Menu.

.emf and .gwg). Figure 7 provides the graphs of the squared residuals and the conditional mean with a 95% confidence interval.

The **Menu/Tests** option allows different tests to be run (see Section 3.2 for further explanations). It also allows the printing of the variance-covariance matrix of the estimated parameters (Figure 8). The results of these tests are printed in GiveWin. An example of output is reported in the next box ('Output 2').

We do not intend to comment upon this application in detail. However, looking at these results, one can briefly argue that the model seems to capture the dynamics of the first and second moments of the CAC40 (see the Box-Pierce statistics). Moreover, the Sign Bias tests shows that there is no remaining leverage component in the innovations while the Nyblom stability test suggests that the estimated parameters are quite stable during the investigated period. Finally, our model specification is not rejected by the goodness-of-fit tests for various lag lengths.

To obtain the h-step-ahead forecasts, access the menu **Test/Forecast** and set the number of forecasts, pre-sample observations (to be plotted) as well as some other graphical options.

Figure 9 shows 10 pre-sample observations and the forecasts up to horizon 10 of the conditional mean. The forecasted bands are $\pm 2\hat{\sigma}_{t+h|t}$ (note that the critical value 2 can be changed).

© Blackwell Publishers Ltd. 2002

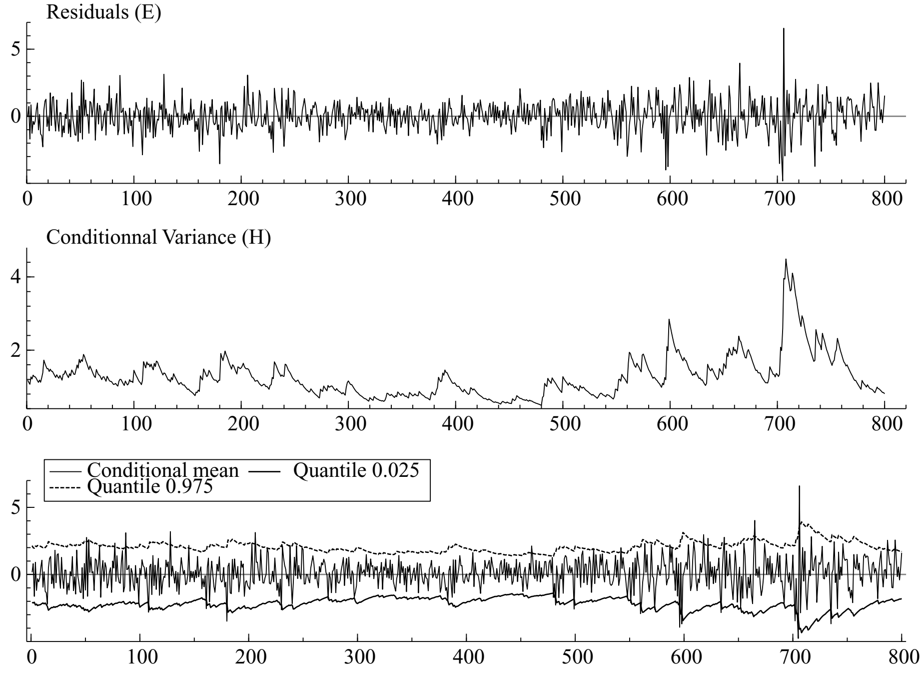

Figure 7. Graphical Analysis.

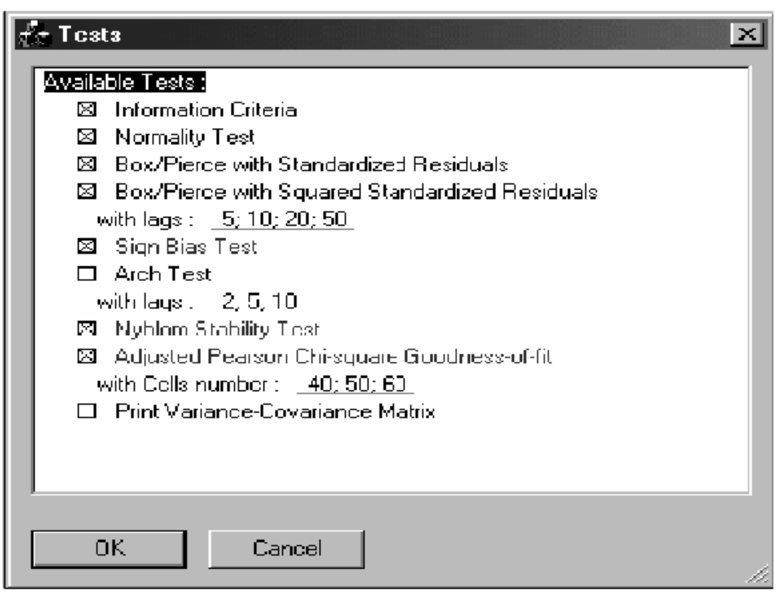

Figure 8. Tests Dialog Box.

© Blackwell Publishers Ltd. 2002

Output 2

TESTS:

Information Criterium (minimize)
Akaike 2.998802 Shibata 2.998553
Schwarz 3.051504 Hannan-Quinn 3.019048

	Statistic	t-value	t-prob
Skewness	-0.2135	2.47	0.0135
Excess Kurtosis	0.4684	2.713	0.006674
Jarque-Bera	13.39	13.39	0.001235

BOX-PIERCE:

H0: No serial correlation $\Rightarrow$ Accept H0 when prob. is High [Q < Chisq(lag)]

Box-Pierce Q-statistics on residuals

$\rightarrow$ P-values adjusted by 1 degree(s) of freedom

Q(10) = 14.47 [0.1064]

Q(20) = 21.67 [0.3012]

Box-Pierce Q-statistics on squared residuals

$\rightarrow$ P-values adjusted by 2 degree(s) of freedom

Q(10) = 9.887 [0.2731]

Q(20) = 16.13 [0.5838]

Diagnostic test based on the news impact curve (EGARCH vs. GARCH)

	Test	Prob
Sign Bias t-Test	0.98838	0.32297
Negative Size Bias t-Test	0.14581	0.88407
Positive Size Bias t-Test	0.62400	0.53263
Joint Test for the Three Effects	5.13914	0.16189

Joint Statistic of the Nyblom test of stability: 2.727
Individual Nyblom Statistics:

Cst(M)	0.72438
AR(1)	0.68524
Cst(V)	0.51505
Beta1	0.42785
Alpha1	0.46229
Gamma1	0.43489
Delta	0.54130
Asymmetry	0.21342
Tail	0.08950

Rem: Asymptotic 1% critical value for individual statistics = 0.75.
Asymptotic 5% critical value for individual statistics = 0.47.

Adjusted Pearson Chi-square Goodness-of-fit test

Lags	Statistic	P-Value(lag-1)	P-Value(lag-k-1)
40	24.9000	0.961261	0.729877
50	26.7500	0.995994	0.946240
60	32.6500	0.997893	0.972622

Rem.: k = # estimated parameters

© Blackwell Publishers Ltd. 2002

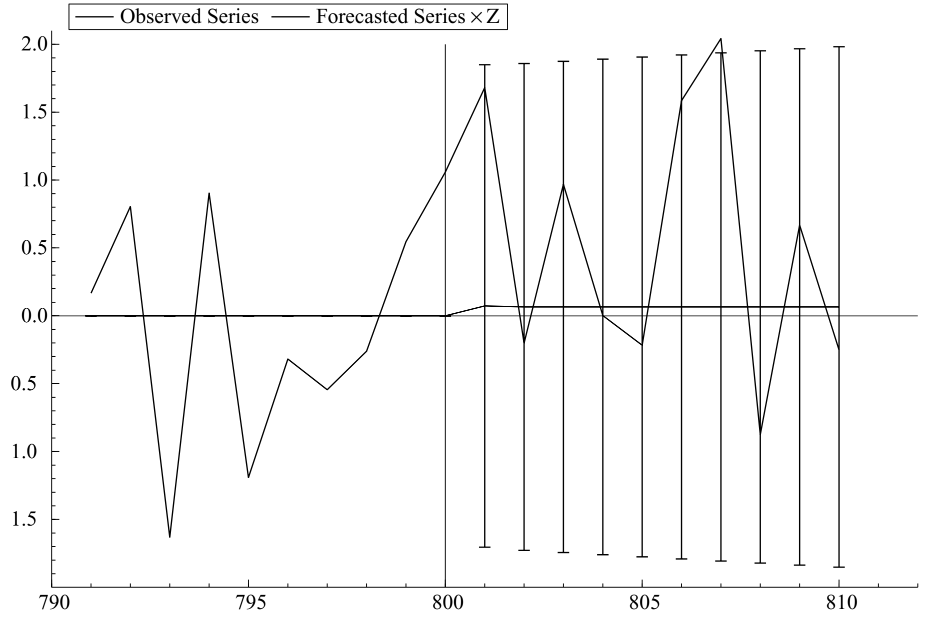

Figure 9. Forecasts from an AR(1)-APARCH(1,1).

4.3. *Using the 'light version'*

First, to specify the model you want to estimate, you have to edit *GarchEstim.ox* with any text editor. We recommend OxEdit. It is shareware that highlights Ox syntax in color (see http://www.oxedit.com for more details). An example of the *GarchEstim.ox* file is displayed below.

The ***GarchEstim*** file consists of five parts:

- the 'Data' part deals with the database, the sample and the variables selection;
- the 'Specification' part is related to the choice of the model, the lag orders and the shape of the distribution;
- the 'Tests & Forecasts' part allows computation of different tests and parameterisation of the forecasts. Note that **BOXPIERCE**, **ARCHLAGS** and **PEARSON** all require a vector of integers corresponding to the lags used in the computation of the statistics;
- the 'Output' part includes several options including **MLE** that refers to the computation method of the standard deviations of the estimated parameters, **TESTONLY**, which is useful when you want to run some tests on the raw series, prior to any estimation and **GRAPHS** and **FOREGRAPHS**, to print graphs for estimation and forecasting, respectively;[27]
- the 'Parameters' part consists of five procedures. **BOUNDS** to constraint several parameters to range between a lower and an upper bound (see

© Blackwell Publishers Ltd. 2002

```
                                                                       GarchEstim.ox
#import <packages/garch/garch>
main()
{
    decl garchobj;
    garchobj = new Garch();
//*** DATA ***//
    garchobj.Load("/data/cac40.xls");
    garchobj.Info();

    garchobj.Select(Y_VAR, {"CAC40",0,0});
//  garchobj.Select(X_VAR, {"NAME",0,0});         // REGRESSOR IN THE MEAN
//  garchobj.Select(Z_VAR, {"NAME",0,0});         // REGRESSOR IN THE VARIANCE
//  garchobj.Select(O_VAR, {"REALVOLA",0,0});     // REALIZED VOLATILITY
    garchobj.SetSelSample(-1, 1, 1000, 1);

//*** SPECIFICATIONS ***//
    garchobj.CSTS(1,1);            // cst in Mean (1 or 0), cst in Variance (1 or 0)
    garchobj.DISTRI(1);            // 0 for Gauss, 1 for Student, 2 for GED, 3 for Skewed-Student
    garchobj.ARMA_ORDERS(1,0);     // AR order (p), MA order (q).
    garchobj.ARFIMA(0);            // 1 if Arfima wanted, 0 otherwise
    garchobj.GARCH_ORDERS(1,1);    // p order, q order
    garchobj.MODEL(1);             //  1:GARCH     2:EGARCH    3:GJR   4:APARCH    5:IGARCH
                                   //  6:FIGARCH(BBM)  7:FIGARCH(Chung)    8:FIEGARCH(BBM only)
                                   //  9:FIAPARCH(BBM) 10: FIAPARCH(Chung) 11: HYGARCH(BBM)

    garchobj.TRUNC(1000);          // Truncation order (only F.I. models with BBM method)

//*** TESTS & FORECASTS ***//
    garchobj.BOXPIERCE(<10;15;20>); // Lags for the Box-Pierce Q-statistics, <> otherwise
    garchobj.ARCHLAGS(<2;5;10>);   // Lags for Engle's LM ARCH test, <> otherwise
    garchobj.NYBLOM(1);            // 1 to compute the Nyblom stability test, 0 otherwise
    garchobj.PEARSON(<40;50;60>);  // Cells for the adjusted Pearson Chi-square Goodness-of-fit test
    garchobj.FORECAST(0,9,1);      // Arg.1 : 1 to launch the forecasting procedure, 0 otherwize
                                   // Arg.2 : Number of forecasts
                                   // Arg.3 : 1 to Print the forecasts, 0 otherwise

//*** OUTPUT ***//
    garchobj.MLE(1);               // 0 : both, 1 : MLE, 2 : QMLE
    garchobj.COVAR(0);             // if 1, prints variance-covariance matrix of the parameters.
    garchobj.ITER(0);              // Interval of iterations between printed intermediary results
    garchobj.TESTSONLY(0,0);       // Arg.1 : if 1, runs tests for the raw Y series, prior to ...
                                   // Arg.2 : if 1, runs tests after the estimation.
    garchobj.GRAPHS(0,0,"");       // Arg.1 : if 1, displays graphics of the estimations.
                                   // Arg.2 : if 1, saves these graphics in a EPS file
                                   // Arg.3 : Name of the saved file.
    garchobj.FOREGRAPHS(1,0,"");   // Same as GRAPHS(p,s,n) but for the graphics of the forecasts.

//*** PARAMETERS ***//
    garchobj.BOUNDS(1);            // 1 if bounded parameters wanted, 0 otherwise
    garchobj.FixParam(1);          // 1 to fix some parameters to their starting values, 0 otherwize
    garchobj.FixedParam(<0;0;0;0;0;0;1>);
                                   // 1 to fix and 0 to estimate the corresponding parameter
    garchobj.DoEstimation(<>);

// m_vPar = m_clevel | m_vbetam | m_dARFI | m_vAR | m_vMA | m_calpha0 | m_vgammav | m_dD | m_vbetav |
//          m_valphav | m_vleverage | m_vtheta1 | m_vtheta2 | m_vpsy | m_ddelta | m_cA | m_cV | m_vHY
//garchobj.DoEstimation(<0.02;0.05;0.45;0.22;0.01;0.025;0.8;0.1;-0.15;0.2;6>);

    garchobj.STORE(0,0,0,0,0,"01",0); //  Arg.1,2,3,4,5 : if 1 -> stored. (Res-SqRes-CondV-...
                                      //  Arg.6 : Suffix. The name of the saved series will be...
                                      //  Arg.7 : if 0, saves as an Excel spreadsheet (.xls)...
    delete garchobj;
}
```

Section 3.1), **FixedParam** to fix some parameters to their starting values, **DoEstimation** that launches the estimation of the model and the **STORE** function allowing storage of some series. The arguments of the **DoEstimation** procedure are a vector containing starting values of the parameters in a specified order (but the user can also let the program choose defaults values).

© Blackwell Publishers Ltd. 2002

Note that the 'Light Version' is more than just a replication of the 'Full Version' without the graphical interface. Indeed, G@RCH uses the object-oriented programming features of Ox and provides a new class called **Garch**. All the functions of this class can thus be used within an Ox program. To illustrate the potentiality of the package, we also provide *Forecast.ox*, an example that computes 448 one-step-ahead forecasts of the conditional mean and conditional variance (using the estimated parameters presented in the previous section), compute the Mincer-Zarnowitz regression and perform some out-of-sample density forecast tests as suggested by Diebold, Gunther and Tay (1998).

The interesting part of *Forecast.ox* is printed in the next box. This code has been used to produce Figure 10 and the outputs associated with this forecasting experiment (see page 480).

In the first four panels of Figure 10, we show the correlograms of $(\hat{\zeta} - \bar{\hat{\zeta}})^j$, for $j = 1, 2, 3, 4$. This graphical tool has been proposed by Diebold, Gunther and Tay (1998) to detect potential remaining dependence in the conditional mean, variance, skewness, kurtosis. In our example, it seems that the probability integral transform is independently distributed.

Panel 5 of Figure 10 also shows the histogram (with 30 cells) of $\hat{\zeta}$ with the 95% confidence bands. From this figure, it is clear that the AR(1)-APARCH(1,1) model

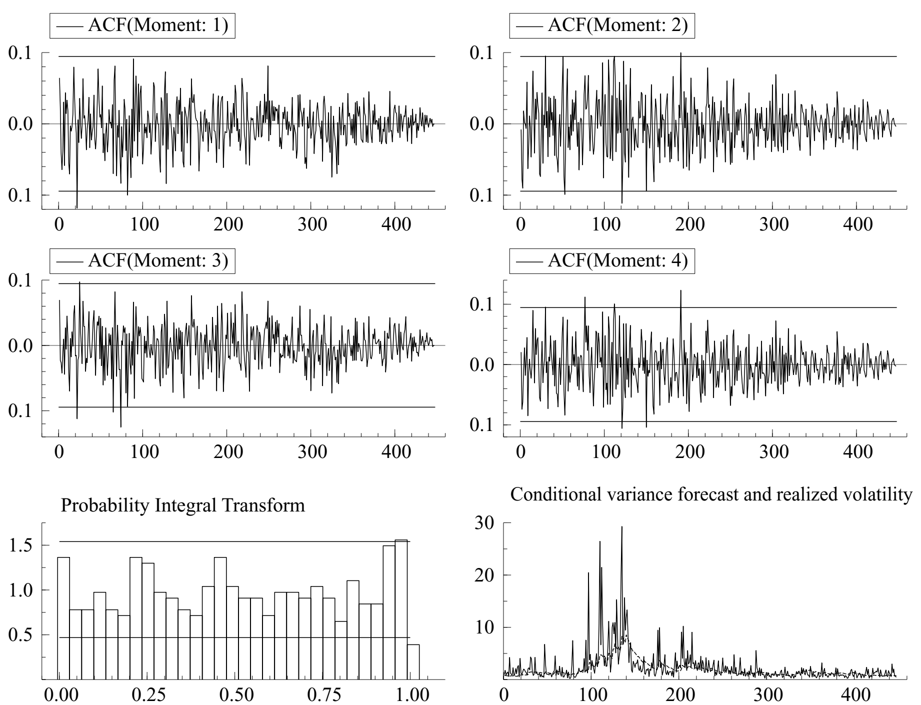

Figure 10. Density Forecast Analysis.

© Blackwell Publishers Ltd. 2002

```
                                                                    Forecast.ox
 #import <packages/garch/garch>
 main()
 {
     decl garchobj;
     garchobj = new Garch();

     ...

    garchobj.DoEstimation(<>);
    decl number_of_forecasts=448; // number of h_step_ahead forecasts
    decl step=1;                  // specify h (h-step-ahead forecasts)
    decl T=garchobj.GetcT();
    decl par=garchobj.PAR()[][0];
    println("!!! Please Wait while computing the forecasts !!!");
    decl forc=<>,h,yfor=<>,Hfor=<>;
    decl RV=columns(garchobj.GetGroup(O_VAR));
    decl shape=<>;
    if (garchobj.GetDistri()==1 || garchobj.GetDistri()==2)  // Except for the HYGARCH
        shape=par[rows(par)-1];
    else if (garchobj.GetDistri()==3)
        shape=par[rows(par)-2:rows(par)-1];
    for (h=0; h<number_of_forecasts; ++h)
    {
        garchobj.FORECAST(1,step,0);
        garchobj.SetSelSample(-1, 1, T+h, 1);
        garchobj.InitData();
        yfor|=garchobj.GetForcData(Y_VAR, step);
        forc|=garchobj.FORECASTING();
        if (RV==1)
            Hfor|=garchobj.GetForcData(O_VAR, step);        // If you use the realized volatility
    }
    decl cd=garchobj.CD(yfor-forc[][0],forc[][1],garchobj.GetDistri(),shape);
    println("Density Forecast Test on Standardized Forecast Errors");
    garchobj.APGT(cd,20|30,rows(par));
    garchobj.AUTO(cd, number_of_forecasts, -0.1, 0.1, 0);
    garchobj.confidence_limits_uniform(cd,30,0.95,1,4);
    if (RV==0)
    {
        DrawTitle(5, "Conditional variance forecast and absolute returns");
        Hfor = (yfor - meanc(yfor)).^2;
    }
    else
        DrawTitle(5, "Conditional variance forecast and realized volatility");
    Draw(5, (Hfor~forc[][1])');
    ShowDrawWindow();
    garchobj.MZ(Hfor, forc, number_of_forecasts);
    garchobj.FEM(forc, yfor~Hfor);

    garchobj.STORE(0,0,0,0,0,"01",0); //  Arg.1,2,3,4,5 ...
                                      //  Arg.6 : Suffix. ...
                                      //  Arg.7 : if 0, ...
    delete garchobj;
}
```

coupled with a skewed Student distribution for the innovations performs very well with the dataset we have investigated. This conclusion is reinforced by the Pearson Chi-square goodness-of-fit test printed hereafter that provides a statistical version of the graphical test presented in Figure 10. Finally, the program performs the Mincer-Zarnowitz regression given in Eq. (39) that regresses the observed volatility (in our case the realized volatility) on a constant and a vector of 448 one-step-ahead forecasts of the conditional variance (produced by the APARCH model).[28] The results (reported in the next box) suggest that the APARCH model gives good forecasts of the conditional variance. Indeed, looking at the estimated parameters of this regression, one can hardly conclude that the APARCH model provides biases

© Blackwell Publishers Ltd. 2002

Density Forecast Test on Standardized Forecast Errors Adjusted Pearson Chi-square Goodness-of-fit test

Lags	Statistic	P-Value(lag-1)	P-Value(lag-k-1)
20	21.0179	0.335815	0.020969
30	26.5089	0.598181	0.149654

Rem.: k = number of estimated parameters

Mincer-Zarnowitz regression on the forecasted volatility

	Coefficient	Std.Error	t-value	t-prob
a_0	-0.225818	0.264837	-0.8527	0.3940
a_1	1.370648	0.176086	7.784	0.0000

R^2: 0.402914

Note: S.E. are Heteroskedastic Consistent (White, 80)

forecasts. Moreover, the R^2 of this regression is higher than 40% (See Andersen and Bollerslev (1998) for more details).

5. Conclusions

This paper documents the software G@RCH 2.2, an Ox package allowing to estimation and forecasting of numerous univariate ARCH-type processes including GARCH, EGARCH, GJR, APARCH, IGARCH, FIGARCH, HYGARCH, FIEGARCH and FIAPARCH specifications of the conditional variance. Several features of the program are worth noting since they are unavailable in most of the traditional econometric softwares: the asymmetric and fractionally integrated processes, four distributions (normal, Student-*t*, GED and skewed Student-*t*), (editable) parameters bounds, several mispecification tests and *h*-step-ahead forecasts.

G@RCH 2.2 is free of charge when used for educational or research purposes and can be downloaded at **http://www.egss.ulg.ac.be/garch/**.

Acknowledgements

While remaining responsible for any error in this paper, the authors would like to thank F. Palm, J-P. Urbain, J. Davidson and M. McAleer for useful comments and suggestions.

Notes

1. Asset pricing models are indeed incomplete unless the full conditional model is specified.
2. Chunhachinda, Dandapani, Hamid and Prakash (1997) find that the incorporation of skewness into the investor's portfolio decision causes a major change in the construction of the optimal portfolio.

© Blackwell Publishers Ltd. 2002

3. Corrado and Su (1996) and Corrado and Su (1997) show that when skewness and kurtosis adjustment terms are added to the Black and Scholes formula, improved accuracy is obtained for pricing options.
4. For a comprehensive review of this language, see Cribari-Neto and Zarkos (2001).
5. Recall that $L^k y_t = y_{t-k}$.
6. ARFIMA models have been combined with an ARCH-type specification by Baillie, Chung and Tieslau (1996), Tschernig (1995), Teyssière (1997), Lecourt (2000) and Beine, Laurent and Lecourt (2000).
7. For stochastic volatility models, see Koopman, Shepard and Doornik (1998).
8. Note that with the EGARCH parameterization of Bollerslev and Mikkelsen (1996), it is possible to estimate an EGARCH $(p, 0)$ since $\ln \sigma_t^2$ depends on $g(z_{t-1})$, even when $q = 0$.
9. Complete developments leading to these conclusions are available in Ding, Granger and Engle (1993).
10. For the symmetric Student density, $\xi = 1$.
11. In their study of the daily S&P500 index, they find that the squared returns series has positive autocorrelations over more than 2,500 lags (or more than 10 years!).
12. See Bollerslev and Mikkelsen (1996, p. 158) for a discussion on the importance of non-integer values of integration when modelling long-run dependencies in the conditional mean of economic time series.
13. See Chung (1999) for more details.
14. Notice that the GJR has not been extended to the long-memory framework. It is however nested in the FIAPARCH class of models.
15. When using the BBM option in G@RCH for the FIEGARCH and FIAPARCH, $(1 - L)^d$ and $(1 - L)^{-d}$ are truncated at some predefined value (see above). It is also possible to truncate this polynomial at the information size at time t, i.e. $t - 1$.
16. *LogL* = log-likelihood value, n is the number observations and k the number of estimated parameters.
17. See Palm and Vlaar (1997) for more details.
18. By optimal, we mean optimal under expected quadratic loss, or in a mean square error sense.
19. For more details about density forecasts and applications in finance, see the special issue of *Journal of Forecasting* (Timmermann, 2000).
20. Confidence intervals for the $\hat{\zeta}$-histogram can be obtained by using the properties of the histogram under the null hypothesis of uniformity.
21. This C code is available at **http://www.ds.unifi.it/~mjl/** in the ‘software’ section. Note that the only configuration available is a FIGARCH $(1, d, 1)$ with a constant in the mean and variance equations and a Gaussian likelihood.
22. This file is available at **http://www.estima.com/procindx.htm** for download.
23. By definition and using the properties of the log distribution, the sum of the intraday returns is equal to the observed daily return based on the closing prices.
24. The extension of this package to multivariate GARCH models is currently under development.
25. Note that these default values can be modified by the user as they are stored in the *startingvalues.txt* file installed with the package.
26. Recall that the estimations are based on the numerical evaluation of the gradients.
27. Graphics will only be displayed when using GiveWin as front-end.
28. The realized and one-step-ahead forecasts are plotted in the last panel of Figure 10.

© Blackwell Publishers Ltd. 2002

References

Andersen, T. and Bollerslev, T. (1998) Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, 39, 885–905.

Baillie, R. (1996) Long Memory Processes and Fractional Integration in Econometrics. *Journal of Econometrics*, 73, 5–59.

Baillie, R., Bollerslev, T. and Mikkelsen, H. (1996) Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 74, 3–30.

Baillie, R., Chung, C. and Tieslau, M. (1996) Analyzing Inflation by the Fractionally Integrated ARFIMA-GARCH Model. *Journal of Applied Econometrics*, 11, 23–40.

Bauwens, L., Giot, P. Grammig, J. and Veredas, D. (2000) A Comparison of Financial Duration Models Via Density Forecasts. CORE DP 2060.

Beine, M., Laurent, S. and Lecourt, C. (2000) Accounting for Conditional Leptokurtosis and Closing Days Effects in FIGARCH Models of Daily Exchange Rates. *Forthcoming in Applied Financial Economics*.

Bera, A. and Higgins, M. (1993) ARCH Models: Properties, Estimation and Testing. *Journal of Economic Surveys*.

Black, F. (1976) Studies of Stock Market Volatility Changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, pp. 177–181.

Bollerslev, T. (1986) Generalized Autoregressive Condtional Heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

Bollerslev, T. and Ghysels, E. (1996) Periodic Autoregressive Conditional Heteroskedasticity. *Journal of Business and Economics Statistics*, 14, 139–152.

Bollerslev, T. and Mikkelsen, H. O. (1996) Modeling and Pricing Long-Memory in Stock Market Volatility. *Journal of Econometrics*, 73, 151–184.

Bollerslev, T. and Wooldridge, J. (1992) Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances. *Econometric Reviews*, 11, 143–172.

Brooks, C., Burke, S. and Persand, G. (2001) Benchmarks and the Accuracy of GARCH Model Estimation. *International Journal of Forecasting*, 17, 45–56.

Chung, C.-F. (1999) Estimating the Fractionnally Intergrated GARCH Model. National Taïwan University working paper.

Chunhachinda, P., Dandapani, K. Hamid, S. and Prakash, A. (1997) Portfolio Selection and Skewness: Evidence from International Stock Markets. *Journal of Banking and Finance*, 21, 143–167.

Corrado, C. and Su, T. (1996) Skewness and Kurtosis in S&P 500 Index Returns Implied by Option Prices. *Journal of Financial Research*, 19, 175–192.

—— (1997) Implied Volatility Skews and Stock Return Skewness and Kurtosis Implied by Stock Option Prices. *European Journal of Finance*, 3, 73–85.

Cribari-Neto, F. and Zarkos, S. (2001) Econometric and Statistical Computing Using Ox. *Forthcoming* in Computational Economics.

Davidson, J. (2001) Moment and Memory Properties of Linear Conditional Heteroscedasticity Models. Manuscript, Cardiff University.

Diebold, F. X., Gunther, T. A. and Tay, A. S. (1998) Evaluating Density Forecasts, with Applications to Financial Risk Management. *International Economic Review*, 39, 863–883.

Ding, Z., Granger, C. W. J. and Engle, R. F. (1993) A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance* 1, 83–106.

Doornik, J. A. (1999) *An Object Oriented Matrix Programming Language*. Timberlake Consultant Ltd., third edn.

Doornik, J. A. and Ooms, M. (1999) A Package for Estimating, Forecasting and Simulating Arfima Models: Arfima package 1.0 for Ox. Discussion paper, Econometric Intitute, Erasmus University Rotterdam.

© Blackwell Publishers Ltd. 2002

Engle, R. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987–1007.

Engle, R. and Bollerslev, T. (1986) Modeling the Persistence of Conditional Variances. *Econometric Reviews*, 5, 1–50.

Engle, R. and González-Rivera, G. (1991) Semiparametric ARCH Model. *Journal of Business and Economic Statistics*, 9, 345–360.

Engle, R. and Lee, G. (1999) *A Permanent and Transitory Component Model of Stock Return Volatility*. pp. 475–497. In R. Engle and H. White (eds), Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger. Oxford University Press.

Engle, R. and Ng, V. (1993) Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, 48, 1749–1778.

Fiorentini, G., Calzolari, G. and Panattoni, L. (1996) Analytic Derivatives and the Computation of GARCH Estimates. *Journal of Applied Econometrics*, 11, 399–417.

Geweke, J. (1986) Modeling the Persistece of Conditional Variances: A Comment. *Econometric Review*, 5, 57–61.

Giot, P. and Laurent, S., (2001) Valut-at-Risk for Long and Short Positions. CORE DP 2001–22, Maastricht University METEOR RM/01/005.

Glosten, L., Jagannathan, R. and Runkle, D. (1993) On the Relation Between Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48, 1779–1801.

Granger, C. (1980) Long Memory Relationships and the Aggregation of Dynamic Models. *Journal of Econometrics*, 14, 227–238.

Granger, C. and Joyeux, R. (1980) An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis* 1, 15–29.

Hansen, B. (1994) Autoregressive Conditional Density Estimation. *International Economic Review*, 35, 705–730.

Higgins, M. and Bera, A. (1992) A Class of Nonlinear ARCH Models. *International Economic Review*, 33, 137–158.

Jarque, C. and Bera, A., (1987) A Test for Normality of Observations and Regression Resid-uals. *International Statistical Review*, 55, 163–172.

Jorion, P. (1996) *Risk and Turnover in the Foreign Exchange Market*. In J. A. Frankel, G. Galli, and A. Giovanni (eds), The Microstructure of Foreign Exchange Markets. Chicago: The University of Chicago Press.

König, H. and Gaab, W. (1982) *The Advanced Theory of Statistics*, vol. 2 of *Inference and Relationships*. Haffner.

Koopman, S., Shepard, N. and Doornik, J. (1998) Statistical Algorithms for Models in State Space using SsfPack 2.2. *Econometrics Journal* 1, 1–55.

Lambert, P. and Laurent, S. (2001) Modelling Financial Time Series Using GARCH-Type Models and a Skewed Student Density. Mimeo, Université de Liège.

Lecourt, C. (2000) Dépendance de Court et Long Terme des Rendements de Taux de Change. *Economie et Prévision*, 5, 127–137.

Lee, S. and Hansen, B. (1994) Asymptotic Properties of the Maximum Likelihood Estimator and Test of the Stability of Parameters of the GARCH and IGARCH Models. *Econometric Theory*, 10, 29–52.

Ling, S. and McAleer, M. (2002) Stationarity and the Existence of Moments of a Family of GARCH processes. *Journal of Econometrics*, 106, 109–117.

Lombardi, M. and Gallo, G. (2001) Analytic Hessian Matrices and the Computation of FIGARCH Estimates. Manuscript, Università degli studi di Firenze.

McCullough, B. and Vinod, H. (1999) The Numerical Reliability of Econometric Software. *Journal of Economic Literature*, 37, 633–665.

McLeod, A. and Li, W. (1983) Diagnostic Checking ARMA Time Series Models Using Squared Residuals Autocorrelations. *Journal of Time Series Analysis* 4, 269–273.

Nelson, D. (1991) Conditional Heteroskedasticity in Asset Returns: a New Approach. *Econometrica*, 59, 349–370.

© Blackwell Publishers Ltd. 2002

Nelson, D. and Cao, C. (1992) Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics*, 10, 229–235.
Nyblom, J. (1989) Testing for the Constancy of Parameters Over Time. *Journal of the American Statistical Association*, 84, 223–230.
Palm, F. (1996) GARCH Models of Volatility. In G. S. Maddala, C. R. Rao (eds), *Handbook of Statistics* (pp. 209–240).
Palm, F. and Vlaar, P. (1997) Simple Diagnostics Procedures for Modelling Financial Time Series. *Allgemeines Statistisches Archiv*, 81, 85–101.
Peiró, A. (1999) Skewness in Financial Returns. *Journal of Banking and Finance*, 23, 847–862.
Pentula, S. (1986) Modeling the Persistece of Conditional Variances: A Comment. *Econometric Review*, 5, 71–74.
Schwert, W. (1990) Stock Volatility and the Crash of '87. *Review of Financial Studies* 3, 77–102.
Taylor, S. (1986) *Modelling Financial Time Series* Wiley, New York.
Teyssière, G. (1997) Double Long-Memory Financial Time Series. Paper presented at the ESEM, Toulouse.
Timmermann, A. (2000) Density Forecasting in Economics and Finance. *Journal of Forecasting*, 19, 120–123.
Tschernig, R. (1995) Long Memory in Foreign Exchange Rates Revisited. *Journal of International Financial Markets, Institutions and Money*, 5, 53–78.
Tse, Y. (1998) The Conditional Heteroscedasticity of the Yen-Dollar Exchange Rate. *Journal of Applied Econometrics*, 193, 49–55.
Weiss, A. (1986) Asymptotic Theory for ARCH Models: Estimation and Testing. *Econometric Theory*, 2, 107–131.
Zakoian, J.-M. (1994) Threshold Heteroskedasticity Models. *Journal of Economic Dynamics and Control*, 15, 931–955.

© Blackwell Publishers Ltd. 2002