

# An Introduction to Univariate GARCH Models

Timo Teräsvirta  
School of Economics and Management  
University of Aarhus  
Building 1322, DK-8000 Aarhus C  
and  
Department of Economic Statistics  
Stockholm School of Economics  
Box 6501, SE-113 83 Stockholm

*SSE/EFI Working Papers in Economics and Finance, No. 646*

December 7, 2006

## Abstract

This paper contains a survey of univariate models of conditional heteroskedasticity. The classical ARCH model is mentioned, and various extensions of the standard Generalized ARCH model are highlighted. This includes the Exponential GARCH model. Stochastic volatility models remain outside this review.

**Keywords:** ARCH; conditional heteroskedasticity; GARCH; non-linear GARCH; volatility modelling

**JEL Code:** C22

**Acknowledgements.** This research has been supported by Jan Wallander's and Tom Hedelius's Foundation, Grant No. P2005-0033:1. A part of the work for the chapter was done when the author was visiting Sonderforschungsbereich 649 at the Humboldt University Berlin. Comments from Changli He, Marcelo Medeiros and Thomas Mikosch (editor) are gratefully acknowledged. Any shortcomings and errors are the author's sole responsibility.

# 1 Introduction

Financial economists are concerned with modelling volatility in asset returns. This is important as volatility is considered a measure of risk, and investors want a premium for investing in risky assets. Banks and other financial institutions apply so-called value-at-risk models to assess their risks. Modelling and forecasting volatility or, in other words, the covariance structure of asset returns, is therefore important.

The fact that volatility in returns fluctuates over time has been known for a long time. Originally, the emphasis was on another aspect of return series: their marginal distributions were found to be leptokurtic. Returns were modelled as independent and identically distributed over time. In a classic work, Mandelbrot (1963) and Mandelbrot and Taylor (1967) applied so-called stable Paretian distributions to characterize the distribution of returns. Rachev and Mitnik (2000) contains an informative discussion of stable Paretian distributions and their use in finance and econometrics.

Observations in return series of financial assets observed at weekly and higher frequencies are in fact not independent. While observations in these series are uncorrelated or nearly uncorrelated, the series contain higher-order dependence. Models of Autoregressive Conditional Heteroskedasticity (ARCH) form the most popular way of parameterizing this dependence. There are several articles in this *Handbook* devoted to different aspects of ARCH models. This article provides an overview of different parameterizations of these models and thus serves as an introduction to autoregressive conditional heteroskedasticity. The article is organized as follows. Section 2 introduces the classic ARCH model. Its generalization, the Generalized ARCH (GARCH) model is presented in Section 3. This section also describes a number of extensions to the standard GARCH models. Section 4 considers the Exponential GARCH model whose structure is rather different from that of the standard GARCH model, and Section 5 discusses ways of comparing EGARCH models with GARCH ones. Suggestions for further reading can be found at the end.

## 2 The ARCH model

The autoregressive conditional heteroskedasticity (ARCH) model is the first model of conditional heteroskedasticity. According to Engle (2004), the original idea was to find a model that could assess the validity of the conjecture of Friedman (1977) that the unpredictability of inflation was a primary cause of business cycles. Uncertainty due to this unpredictability would affect the

investors' behaviour. Pursuing this idea required a model in which this uncertainty could change over time. Engle (1982) applied his resulting ARCH model to parameterizing conditional heteroskedasticity in a wage-price equation for the United Kingdom. Let  $\varepsilon_t$  be a random variable that has a mean and a variance conditionally on the information set  $\mathcal{F}_{t-1}$  (the  $\sigma$ -field generated by  $\varepsilon_{t-j}$ ,  $j \geq 1$ ). The ARCH model of  $\varepsilon_t$  has the following properties. First,  $\mathbf{E}\{\varepsilon_t|\mathcal{F}_{t-1}\} = 0$  and, second, the conditional variance  $h_t = \mathbf{E}\{\varepsilon_t^2|\mathcal{F}_{t-1}\}$  is a nontrivial positive-valued parametric function of  $\mathcal{F}_{t-1}$ . The sequence  $\{\varepsilon_t\}$  may be observed directly, or it may be an error or innovation sequence of an econometric model. In the latter case,

$$\varepsilon_t = y_t - \mu_t(y_t) \quad (1)$$

where  $y_t$  is an observable random variable and  $\mu_t(y_t) = \mathbf{E}\{y_t|\mathcal{F}_{t-1}\}$ , the conditional mean of  $y_t$  given  $\mathcal{F}_{t-1}$ . Engle's (1982) application was of this type. In what follows, the focus will be on parametric forms of  $h_t$ , and  $\mu_t(y_t)$  will be ignored.

Engle assumed that  $\varepsilon_t$  can be decomposed as follows:

$$\varepsilon_t = z_t h_t^{1/2} \quad (2)$$

where  $\{z_t\}$  is a sequence of independent, identically distributed (iid) random variables with zero mean and unit variance. This implies  $\varepsilon_t|\mathcal{F}_{t-1} \sim D(0, h_t)$  where  $D$  stands for the distribution (typically assumed to be a normal or a leptokurtic one). The following conditional variance defines an ARCH model of order  $q$ :

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 \quad (3)$$

where  $\alpha_0 > 0$ ,  $\alpha_j \geq 0$ ,  $j = 1, \dots, q-1$ , and  $\alpha_q > 0$ . The parameter restrictions in (3) form a necessary and sufficient condition for positivity of the conditional variance. Suppose the unconditional variance  $\mathbf{E}\varepsilon_t^2 = \sigma^2 < \infty$ . The definition of  $\varepsilon_t$  through the decomposition (2) involving  $z_t$  then guarantees the white noise property of the sequence  $\{\varepsilon_t\}$ , since  $\{z_t\}$  is a sequence of iid variables. Although the application in Engle (1982) was not a financial one, Engle and others soon realized the potential of the ARCH model in financial applications that required forecasting volatility.

The ARCH model and its generalizations are thus applied to modelling, among other things, interest rates, exchange rates and stock and stock index returns. Bollerslev, Chou and Kroner (1992) already listed a variety of applications in their survey of these models. Forecasting volatility of these series is different from forecasting the conditional mean of a process because

volatility, the object to be forecast, is not observed. The question then is how volatility should be measured. Using  $\varepsilon_t^2$  is an obvious but not necessarily a very good solution if data of higher frequency are available; see Andersen and Bollerslev (1998) and 2007ANDERSEN for discussion.

### 3 The Generalized ARCH model

#### 3.1 Why Generalized ARCH?

In applications, the ARCH model has been replaced by the so-called generalized ARCH (GARCH) model that Bollerslev (1986) and Taylor (1986) proposed independently of each other. In this model, the conditional variance is also a linear function of its own lags and has the form

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j}. \quad (4)$$

The conditional variance defined by (4) has the property that the unconditional autocorrelation function of  $\varepsilon_t^2$ , if it exists, can decay slowly, albeit still exponentially. For the ARCH family, the decay rate is too rapid compared to what is typically observed in financial time series, unless the maximum lag  $q$  in (3) is long. As (4) is a more parsimonious model of the conditional variance than a high-order ARCH model, most users prefer it to the simpler ARCH alternative.

The overwhelmingly most popular GARCH model in applications has been the GARCH(1,1) model, that is,  $p = q = 1$  in (4). A sufficient condition for the conditional variance to be positive with probability one is  $\alpha_0 > 0$ ,  $\alpha_j \geq 0$ ,  $j = 1, \dots, q$ ;  $\beta_j \geq 0$ ,  $j = 1, \dots, p$ . The necessary and sufficient conditions for positivity of the conditional variance in higher-order GARCH models are more complicated than the sufficient conditions just mentioned and have been given in Nelson and Cao (1992). The GARCH(2,2) case has been studied in detail by He and Teräsvirta (1999b). Note that for the GARCH model to be identified if at least one  $\beta_j > 0$  (the model is a genuine GARCH model) one has to require that also at least one  $\alpha_j > 0$ . If  $\alpha_1 = \dots = \alpha_q = 0$ , the conditional and unconditional variances of  $\varepsilon_t$  are equal and  $\beta_1, \dots, \beta_p$  are unidentified nuisance parameters. The GARCH( $p, q$ ) process is weakly stationary if and only if  $\sum_{j=1}^q \alpha_j + \sum_{j=1}^p \beta_j < 1$ .

The stationary GARCH model has been slightly simplified by 'variance targeting', see Engle and Mezrich (1996). This implies replacing the intercept  $\alpha_0$  in (4) by  $(1 - \sum_{j=1}^q \alpha_j - \sum_{j=1}^p \beta_j)\sigma^2$  where  $\sigma^2 = \mathbb{E}\varepsilon_t^2$ . The estimate

$\hat{\sigma}^2 = T^{-1} \sum_{t=1}^T \varepsilon_t^2$  is substituted for  $\sigma^2$  before estimating the other parameters. As a result, the conditional variance converges towards the 'long-run' unconditional variance, and the model contains one parameter less than the standard GARCH( $p, q$ ) model.

It may be pointed out that the GARCH model is a special case of an infinite-order (ARCH( $\infty$ )) model (2) with

$$h_t = \alpha_0 + \sum_{j=1}^{\infty} \alpha_j \varepsilon_{t-j}^2. \quad (5)$$

The ARCH( $\infty$ ) representation is useful in considering properties of ARCH and GARCH models such as the existence of moments and long memory; see Giraitis, Kokoszka and Leipus (2000). The moment structure of GARCH models is considered in detail in 2007LINDNER.

### 3.2 Families of univariate GARCH models

Since its introduction the GARCH model has been generalized and extended in various directions. This has been done to increase the flexibility of the original model. For example, the original GARCH specification assumes the response of the variance to a shock to be independent of the sign of the shock and just be a function of the size of the shock. Several extensions of the GARCH model aim at accommodating the asymmetry in the response. These include the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993), the asymmetric GARCH models of Engle and Ng (1993) and the quadratic GARCH of Sentana (1995). The GJR-GARCH model has the form (2), where

$$h_t = \alpha_0 + \sum_{j=1}^q \{\alpha_j + \delta_j I(\varepsilon_{t-j} > 0)\} \varepsilon_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j}. \quad (6)$$

In (6),  $I(\varepsilon_{t-j} > 0)$  is an indicator function obtaining value one when the argument is true and zero otherwise.

In the asymmetric models of both Engle and Ng, and Sentana, the centre of symmetry of the response to shocks is shifted away from zero. For example,

$$h_t = \alpha_0 + \alpha_1 (\varepsilon_{t-1} - \gamma)^2 + \beta_1 h_{t-1} \quad (7)$$

with  $\gamma \neq 0$  in Engle and Ng (1993). The conditional variance in Sentana's Quadratic ARCH (QARCH) model (the model is presented in the ARCH form) is defined as follows:

$$h_t = \alpha_0 + \boldsymbol{\alpha}' \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_{t-1}' \mathbf{A} \boldsymbol{\varepsilon}_{t-1} \quad (8)$$

where  $\boldsymbol{\varepsilon}_t = (\varepsilon_t, \dots, \varepsilon_{t-q+1})'$  is a  $q \times 1$  vector,  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)'$  is a  $q \times 1$  parameter vector and  $\mathbf{A}$  a  $q \times q$  symmetric parameter matrix. In (8), not only squares of  $\varepsilon_{t-i}$  but also cross-products  $\varepsilon_{t-i}\varepsilon_{t-j}$ ,  $i \neq j$ , contribute to the conditional variance. When  $\boldsymbol{\alpha} \neq \mathbf{0}$ , the QARCH generates asymmetric responses. The ARCH equivalent of (7) is a special case of Sentana's model. Constraints on parameters required for positivity of  $h_t$  in (8) become clear by rewriting (8) as follows:

$$h_t = \begin{bmatrix} \boldsymbol{\varepsilon}_{t-1} & 1 \end{bmatrix}' \begin{bmatrix} \mathbf{A} & \boldsymbol{\alpha}/2 \\ \boldsymbol{\alpha}'/2 & \alpha_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_{t-1} \\ 1 \end{bmatrix}. \quad (9)$$

The conditional variance  $h_t$  is positive if and only if the matrix in the quadratic form on the right-hand side of (9) is positive definite.

Some authors have suggested modelling the conditional standard deviation instead of the conditional variance: see Taylor (1986), Schwert (1990), and for an asymmetric version, Zakoïan (1994). A further generalization of this idea appeared in Ding, Granger and Engle (1993). These authors proposed a GARCH model for  $h_t^k$  where  $k > 0$  is also a parameter to be estimated. Their power GARCH model is (2) with

$$h_t^k = \alpha_0 + \sum_{j=1}^q \alpha_j |\varepsilon_{t-j}|^{2k} + \sum_{j=1}^p \beta_j h_{t-j}^k, \quad k > 0. \quad (10)$$

The authors argued that this extension provides flexibility lacking in the original GARCH specification of Bollerslev (1986) and Taylor (1986).

The proliferation of GARCH models has inspired some authors to define families of GARCH models that would accommodate as many individual models as possible. Hentschel (1995) defined one such family. The first-order GARCH model has the general form

$$\frac{h_t^{\lambda/2} - 1}{\lambda} = \omega + \alpha h_{t-1}^{\lambda/2} f^\nu(z_{t-1}) + \beta \frac{h_{t-1}^{\lambda/2} - 1}{\lambda} \quad (11)$$

where

$$f^\nu(z_t) = |z_t - b| - c(z_t - b).$$

Family (11) contains a large number of well-known GARCH models. The Box-Cox type transformation of the conditional standard deviation  $h_t^{1/2}$  makes it possible, by allowing  $\lambda \rightarrow 0$ , to accommodate models in which the logarithm of the conditional variance is parameterized, such as the exponential GARCH model to be considered in Section 4. Parameters  $b$  and  $c$  in  $f^\nu(z_t)$  allow the inclusion of different asymmetric GARCH models such as the GJR-GARCH or threshold GARCH models in (11).

Another family of GARCH models that is of interest is the one He and Teräsvirta (1999a) defined as follows:

$$h_t^k = \sum_{j=1}^q g(z_{t-j}) + \sum_{j=1}^q c_j(z_{t-j}) h_{t-j}^k, \quad k > 0 \quad (12)$$

where  $\{g(z_t)\}$  and  $\{c(z_t)\}$  are sequences of independent and identically distributed random variables. In fact, the family was originally defined for  $q = 1$ , but the definition can be generalized to higher-order models. For example, the standard GARCH( $q, q$ ) model is obtained by setting  $g(z_t) = \alpha_0/q$  and  $c_j(z_{t-j}) = \alpha_j z_{t-j}^2 + \beta_j, j = 1, \dots, q$ , in (12). Many other GARCH models such as the GJR-GARCH, the absolute-value GARCH, the Quadratic GARCH and the power GARCH model belong to this family.

Note that the power GARCH model itself nests several well-known GARCH models; see Ding et al. (1993) for details. Definition (12) has been used for deriving expressions of fourth moments, kurtosis and the autocorrelation function of  $\varepsilon_t^2$  for a number of first-order GARCH models and the standard GARCH( $p, q$ ) model.

The family of augmented GARCH models, defined by Duan (1997), is a rather general family. The first-order augmented GARCH model is defined as follows. Consider (2) and assume that

$$h_t = \begin{cases} |\lambda \phi_t - \lambda - 1| & \text{if } \lambda \neq 0 \\ \exp\{\phi_t - 1\} & \text{if } \lambda = 0 \end{cases} \quad (13)$$

where

$$\phi_t = \alpha_0 + \zeta_{1,t-1} \phi_{t-1} + \zeta_{2,t-1}. \quad (14)$$

In (14),  $(\zeta_{1t}, \zeta_{2t})$  is a strictly stationary sequence of random vectors with a continuous distribution, measurable with respect to the available information until  $t$ . Duan defined an augmented GARCH(1,1) process as (2) with (13) and (14), such that

$$\begin{aligned} \zeta_{1t} &= \alpha_1 + \alpha_2 |\varepsilon_t - c|^\delta + \alpha_3 \max(0, c - \varepsilon_t)^\delta \\ \zeta_{2t} &= \alpha_4 \frac{|\varepsilon_t - c|^\delta - 1}{\delta} + \alpha_5 \frac{\max(0, c - \varepsilon_t)^\delta - 1}{\delta}. \end{aligned}$$

This process contains as special cases all the GARCH models previously mentioned, as well as the Exponential GARCH model to be considered in Section 4. Duan (1997) generalized this family to the GARCH( $p, q$ ) case and derived sufficient conditions for strict stationarity for this general family as well as conditions for the existence of the unconditional variance of  $\varepsilon_t$ . Furthermore, he suggested misspecification tests for the augmented GARCH model.

### 3.3 Nonlinear GARCH

#### 3.3.1 Smooth transition GARCH

As mentioned above, the GARCH model has been extended to characterize asymmetric responses to shocks. The GJR-GARCH model, obtained as setting  $g(z_t) = \alpha_0$  and  $c_j(z_{t-j}) = (\alpha_j + \omega_j I(z_{t-j} > 0))z_{t-j}^2 + \beta_j, j = 1, \dots, q$ , in (12), is an example of that. A nonlinear version of the GJR-GARCH model is obtained by making the transition between regimes smooth. Hagerud (1997), Gonzalez-Rivera (1998) and Anderson, Nam and Vahid (1999) proposed this extension. A smooth transition GARCH (STGARCH) model may be defined as equation (2) with

$$h_t = \alpha_{10} + \sum_{j=1}^q \alpha_{1j} \varepsilon_{t-j}^2 + (\alpha_{20} + \sum_{j=1}^q \alpha_{2j} \varepsilon_{t-j}^2) G(\gamma, c; \varepsilon_{t-j}) + \sum_{j=1}^p \beta_j h_{t-j} \quad (15)$$

where the transition function

$$G(\gamma, c; \varepsilon_{t-j}) = (1 + \exp\{-\gamma \prod_{k=1}^K (\varepsilon_{t-j} - c_k)\})^{-1}, \quad \gamma > 0. \quad (16)$$

When  $K = 1$ , (16) is a simple logistic function that controls the change of the coefficient of  $\varepsilon_{t-j}^2$  from  $\alpha_{1j}$  to  $\alpha_{1j} + \alpha_{2j}$  as a function of  $\varepsilon_{t-j}$ , and similarly for the intercept. In that case, as  $\gamma \rightarrow \infty$ , the transition function becomes a step function and represents an abrupt switch from one regime to the other. Furthermore, at the same time setting  $c_1 = 0$  yields the GJR-GARCH model because  $\varepsilon_t$  and  $z_t$  have the same sign. When  $K = 2$  and, in addition,  $c_1 = -c_2$  in (16), the resulting smooth transition GARCH model is still symmetric about zero, but the response of the conditional variance to a shock is a nonlinear function of lags of  $\varepsilon_t^2$ . Smooth transition GARCH models are useful in situations where the assumption of two distinct regimes is too rough an approximation to the asymmetric behaviour of the conditional variance. Hagerud (1997) also discussed a specification strategy that allows the investigator to choose between  $K = 1$  and  $K = 2$  in (16). Values of  $K > 2$  may also be considered, but they are likely to be less common in applications than the two simplest cases.

The smooth transition GARCH model (15) with  $K = 1$  in (16) is designed for modelling asymmetric responses to shocks. On the other hand, the standard GARCH model has the undesirable property that the estimated model often exaggerates the persistence in volatility (the estimated sum of the  $\alpha$ -



and  $\beta$ -coefficients is close to one). This in turn results in poor volatility forecasts. In order to remedy this problem, Lanne and Saikkonen (2005) proposed a smooth transition GARCH model whose first-order version has the form

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \delta_1 G_1(\boldsymbol{\theta}; h_{t-1}) + \beta_1 h_{t-1}. \quad (17)$$

In (17),  $G_1(\boldsymbol{\theta}; h_{t-1})$  is a continuous bounded function such as (16): Lanne and Saikkonen use the cumulative distribution function of the gamma-distribution. A major difference between (15) and (17) is that in the latter model the transition variable is a lagged conditional variance. In empirical examples given in the paper, this parameterization clearly alleviates the problem of exaggerated persistence. The model may also be generalized to include a term of the form  $G_1(\boldsymbol{\theta}; h_{t-1})h_{t-1}$ , but according to the authors, such an extension appeared unnecessary in practice.

### 3.3.2 Threshold GARCH and extensions

If (15) is defined as a model for the conditional standard deviation such that  $h_t$  is replaced by  $h_t^{1/2}$ ,  $h_{t-j}$  by  $h_{t-j}^{1/2}$ ,  $j = 1, \dots, p$ , and  $\varepsilon_{t-j}^2$  by  $|\varepsilon_{t-j}|$ ,  $j = 1, \dots, q$ , then choosing  $K = 1$ ,  $c_1 = 0$  and letting  $\gamma \rightarrow \infty$  in (16) yields the threshold GARCH (TGARCH) model that Zakořian (1994) considered. The TGARCH( $p, q$ ) model is thus the counterpart of the GJR-GARCH model in the case where the entity to be modelled is the conditional standard deviation instead of the conditional variance. Note that in both of these models, the threshold parameter has a known value (zero). In Zakořian (1994), the conditional standard deviation is defined as follows:

$$h_t^{1/2} = \alpha_0 + \sum_{j=1}^q (\alpha_j^+ \varepsilon_{t-j}^+ - \alpha_j^- \varepsilon_{t-j}^-) + \sum_{j=1}^q \beta_j h_{t-j}^{1/2} \quad (18)$$

where  $\varepsilon_{t-j}^+ = \max(\varepsilon_{t-j}, 0)$  and  $\varepsilon_{t-j}^- = \min(\varepsilon_{t-j}, 0)$ . Rabemananjara and Zakořian (1993) introduced an even more general model in which  $h_t^{1/2}$  can obtain negative values, but it has not gained wide acceptance. Nevertheless, these authors provide evidence of asymmetry in the French stock market by fitting the TGARCH model (18) to the daily return series of stocks included in the CAC 40 index of the Paris Bourse.

The TGARCH model is linear in parameters because the threshold parameter is assumed to equal zero. A genuine nonlinear threshold model is the Double Threshold ARCH (DTARCH) model of Li and Li (1996). It is called a double threshold model because both the autoregressive conditional mean and the conditional variance have a threshold-type structure as defined in

Tong (1990). The conditional mean is defined as follows:

$$y_t = \sum_{k=1}^K (\phi_{0k} + \sum_{j=1}^{p_k} \phi_{jk} y_{t-j}) I(c_{k-1}^{(m)} < y_{t-b} \leq c_k^{(m)}) + \varepsilon_t \quad (19)$$

and the conditional variance has the form

$$h_t = \sum_{\ell=1}^L (\alpha_{0\ell} + \sum_{j=1}^{p_\ell} \alpha_{j\ell} \varepsilon_{t-j}^2) I(c_{\ell-1}^{(v)} < y_{t-d} \leq c_\ell^{(v)}). \quad (20)$$

Furthermore,  $k = 1, \dots, K$ ,  $\ell = 1, \dots, L$ , and  $b$  and  $d$  are delay parameters,  $b, d \geq 1$ . The number of regimes in (19) and (20),  $K$  and  $L$ , respectively, need not be the same, nor do the two threshold variables have to be equal. Other threshold variables than lags of  $y_t$  are possible. For example, replacing  $y_{t-d}$  in (20) by  $\varepsilon_{t-d}$  or  $\varepsilon_{t-d}^2$  may sometimes be an interesting possibility.

Another variant of the DTARCH model is the model that Audrino and Bühlmann (2001) who introduced it called the Tree-Structured GARCH model. It has an autoregressive conditional mean:

$$y_t = \phi y_{t-1} + \varepsilon_t \quad (21)$$

where  $\varepsilon_t$  is decomposed as in (2), and the first-order conditional variance

$$h_t = \sum_{k=1}^K (\alpha_{0k} + \alpha_{1k} y_{t-1}^2 + \beta_k h_{t-1}) I\{(y_{t-1}, h_{t-1}) \in \mathcal{R}_k\}. \quad (22)$$

In (22),  $\mathcal{R}_k$  is a subset in a partition  $\mathcal{P} = \{\mathcal{R}_1, \dots, \mathcal{R}_K\}$  of the sample space of  $(y_{t-1}, h_{t-1})$ . For example, if  $K = 2$ , either  $\mathcal{R}_1 = \{y_{t-1} > c_y, h_{t-1} > 0\}$  or  $\mathcal{R}_1 = \{-\infty < y_{t-1} < \infty, h_{t-1} > c_h\}$ ,  $c_h > 0$ , and  $\mathcal{R}_2$  is the complement of  $\mathcal{R}_1$ . Note that, strictly speaking, equation (22) does not define a GARCH model unless  $\phi = 0$  in (21), because the squared variable in the equation is  $y_{t-1}^2$ , not  $\varepsilon_{t-1}^2$ . A practical problem is that the tree-growing strategy of Audrino and Bühlmann (2001) does not seem to prevent underidentification: if  $K$  is chosen too large, (22) is not identified. A similar problem is present in the DTARCH model as well as in the STGARCH one. Hagerud (1997) and Gonzalez-Rivera (1998), however, do provide linearity tests in order to avoid this problem in the STGARCH framework.

### 3.4 Time-varying GARCH

An argument brought forward in the literature, see for instance Mikosch and Stărică (2004), is that in applications the assumption of the GARCH models

having constant parameters may not be appropriate when the series to be modelled are long. Parameter constancy is a testable proposition, and if it is rejected, the model can be generalized. One possibility is to assume that the parameters change at specific points of time, divide the series into subseries according to the location of the break-points, and fit separate GARCH models to the subseries. The main statistical problem is then finding the number of break-points and their location because they are normally not known in advance. Chu (1995) has developed tests for this purpose.

Another possibility is to modify the smooth transition GARCH model (15) to fit this situation. This is done by defining the transition function (16) as a function of time:

$$G(\gamma, c; t^*) = (1 + \exp\{-\gamma \prod_{k=1}^K (t^* - c_k)\})^{-1}, \quad \gamma > 0$$

where  $t^* = t/T$ . Standardizing the time variable between zero and unity makes interpretation of the parameters  $c_k, k = 1, \dots, K$ , easy as they indicate where in relative terms the changes in the process occur. The resulting time-varying parameter GARCH (TV-GARCH) model has the form

$$h_t = \alpha_0(t) + \sum_{j=1}^q \alpha_j(t) \varepsilon_{t-j}^2 + \sum_{j=1}^p \beta_j(t) h_{t-j} \quad (23)$$

where  $\alpha_0(t) = \alpha_{01} + \alpha_{02}G(\gamma, c; t^*)$ ,  $\alpha_j(t) = \alpha_{j1} + \alpha_{j2}G(\gamma, c; t^*)$ ,  $j = 1, \dots, q$ , and  $\beta_j(t) = \beta_{j1} + \beta_{j2}G(\gamma, c; t^*)$ ,  $j = 1, \dots, p$ . This is the most flexible parameterization. Some of the time-varying parameters in (23) may be restricted to constants *a priori*. For example, it may be assumed that only the intercept  $\alpha_0(t)$  is time-varying. This implies that the 'baseline volatility' or unconditional variance is changing over time. If change is allowed in the other GARCH parameters then the model is capable of accommodating systematic changes in the amplitude of the volatility clusters that cannot be explained by a constant-parameter GARCH model.

This type of time-varying GARCH is mentioned here because it is a special case of the smooth transition GARCH model. Other time-varying parameter models of conditional heteroskedasticity, such as nonstationary ARCH models with locally changing parameters, are discussed in 2007SPOKOINY.

### 3.5 Markov-switching ARCH and GARCH

Markov-switching or hidden Markov models of conditional heteroskedasticity constitute another class of nonlinear models of volatility. These models are

an alternative way of modelling volatility processes that contains breaks. Hamilton and Susmel (1994) argued that very large shocks, such as the one affecting the stocks in October 1987, may have consequences for subsequent volatility so different from consequences of small shocks that a standard ARCH or GARCH model cannot describe them properly. Their Markov-switching ARCH model is defined as follows:

$$h_t = \sum_{i=1}^k (\alpha_0^{(i)} + \sum_{j=1}^q \alpha_j^{(i)} \varepsilon_{t-j}^2) I(s_t = i) \quad (24)$$

where  $s_t$  is a discrete unobservable random variable obtaining values from the set  $S = \{1, \dots, k\}$ . It follows a (usually first-order) homogeneous Markov chain:

$$\Pr\{s_t = j | s_{t-1} = i\} = p_{ij}, \quad i, j = 1, \dots, k.$$

Cai (1994) considered a special case of (24) in which only the intercept  $\alpha_0^{(i)}$  is switching, and  $k = 2$ . But then, his model also contains a switching conditional mean. Furthermore, Rydén, Teräsvirta and Åsbrink (1998) showed that a simplified version of (24) where  $\alpha_j^{(i)} = 0$  for  $j \geq 1$  and all  $i$ , is already capable of generating data that display most of the stylized facts that Granger and Ding (1995) ascribe to high-frequency, daily, say, financial return series. This suggests that a Markov-switching variance alone without any ARCH structure may in many cases explain a large portion of the variation in these series.

Nevertheless, it can be argued that shocks drive economic processes, and this motivates the ARCH structure. If the shocks have a persistent effect on volatility, however, a parsimonious GARCH representation may be preferred to (24). Generalizing (24) into a GARCH model involves one major difficulty. A straightforward (first-order) generalization would have the following form:

$$h_t = (\alpha_0^{(i)} + \alpha_1^{(i)} \varepsilon_{t-1}^2 + \beta_1^{(i)} h_{t-1}) I(s_t = i). \quad (25)$$

From the autoregressive structure of (25) it follows that  $h_t$  is completely path-dependent: its value depends on the unobservable  $s_{t-j}$ ,  $j = 0, 1, 2, \dots, t$ . This makes the model practically impossible to estimate because in order to evaluate the log-likelihood, these unobservables have to be integrated out of this function. Simplifications of the model that circumvent this problem can be found in Gray (1996) and Klaassen (2002). A good discussion about their models can be found in Haas, Mittnik and Paolella (2004). These authors present another Markov-switching (MS-) GARCH model whose fourth-moment structure they are able to work out. That does not seem possible

for the other models. The MS-GARCH model of Haas et al. (2004) is defined as follows:

$$\varepsilon_t = z_t \sum_{i=1}^k h_{it}^{1/2} I(s_t = i)$$

where  $s_t$  is defined as in (24). Furthermore,

$$\mathbf{h}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \varepsilon_{t-1}^2 + \mathbf{B} \mathbf{h}_{t-1}$$

where  $\mathbf{h}_t = (h_{1t}, \dots, h_{kt})'$ ,  $\boldsymbol{\alpha}_i = (\alpha_{i1}, \dots, \alpha_{ik})'$ ,  $i = 0, 1$ , and  $\mathbf{B} = \text{diag}(\beta_{11}, \dots, \beta_{k1})'$ . Thus, each volatility regime has its own GARCH equation. The conditional variance in a given regime is only a function of the lagged conditional variance in the same regime, which is not the case in the other models. The identification problem mentioned in Section 3.3.2 is present here as well. If the true model has fewer regimes than the specified one, the latter contains unidentified nuisance parameters. Liu (2006) provides a number of results, including conditions for strict stationarity and the existence of higher-order moments, for this MS-GARCH model.

More information about Markov-switching ARCH and GARCH models can be found in Lange and Rahbek (2007).

### 3.6 Integrated and fractionally integrated GARCH

In applications it often occurs that the estimated sum of the parameters  $\alpha_1$  and  $\beta_1$  in the standard first-order GARCH model (4) with  $p = q = 1$  is close to unity. Engle and Bollerslev (1986), who first paid attention to this phenomenon, suggested imposing the restriction  $\alpha_1 + \beta_1 = 1$  and called the ensuing model an integrated GARCH (IGARCH) model. The IGARCH process is not weakly stationary as  $E\varepsilon_t^2$  is not finite. Nevertheless, the term "integrated GARCH" may be somewhat misleading as the IGARCH process is strongly stationary. Nelson (1991) showed that under mild conditions for  $\{z_t\}$  and assuming  $\alpha_0 > 0$ , the GARCH(1,1) process is strongly stationary if

$$E \ln(\alpha_1 + \beta_1 z_t^2) < 0. \quad (26)$$

The IGARCH process satisfies (26). The analogy with integrated processes, that is, ones with a unit root, is therefore not as straightforward as one might think. For a general discussion of stationarity conditions in GARCH models, see 2007LINDNER.

Nelson (1991) also showed that when an IGARCH process is started at some finite time point, its behaviour depends on the intercept  $\alpha_0$ . On the one hand, if the intercept is positive then the unconditional variance of the

process grows linearly with time. In practice this means that the amplitude of the clusters of volatility to be parameterized by the model on the average increases over time. The rate of increase need not, however, be particularly rapid. One may thus think that in applications with, say, a few thousand observations, IGARCH processes nevertheless provide a reasonable approximation to the true data-generating volatility process. On the other hand, if  $\alpha_0 = 0$  in the IGARCH model, the realizations from the process collapse to zero almost surely. How rapidly this happens, depends on the parameter  $\beta_1$ .

Although the investigator may be prepared to accept an IGARCH model as an approximation, a potentially disturbing fact is that this means assuming that the unconditional variance of the process to be modelled does not exist. It is not clear that this is what one always wants to do. There exist other explanations to the fact that the sum  $\alpha_1 + \beta_1$  estimates to one or very close to one. First Diebold (1986) and later Lamoureux and Lastrapes (1990) suggested that this often happens if there is a switch in the intercept of a GARCH model during the estimation period. This may not be surprising as such a switch means that the underlying GARCH process is not stationary.

Another, perhaps more puzzling, observation is related to exponential GARCH models to be considered in Section 4. Malmsten (2004) noticed that if a GARCH(1,1) model is fitted to a time series generated by a stationary first-order exponential GARCH model (see Section 4), the probability of the estimated sum  $\alpha_1 + \beta_1$  exceeding unity can sometimes be rather large. In short, if the estimated sum of these two parameters in a standard GARCH(1,1) model is close to unity, imposing the restriction  $\alpha_1 + \beta_1 = 1$  without further investigation may not necessarily be the most reasonable action to take.

Assuming  $p = q = 1$ , the GARCH( $p, q$ ) equation (4) can also be written in the "ARMA(1,1) form" by adding  $\varepsilon_t^2$  to both sides and moving  $h_t$  to the right-hand side:

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + \nu_t - \beta_1\nu_{t-1} \quad (27)$$

where  $\{\nu_t\} = \{\varepsilon_t^2 - h_t\}$  is a martingale difference sequence with respect to  $h_t$ . For the IGARCH process, (27) has the "ARIMA(0,1,1) form"

$$(1 - L)\varepsilon_t^2 = \alpha_0 + \nu_t - \beta_1\nu_{t-1}. \quad (28)$$

Equation (28) has served as a starting-point for the fractionally integrated GARCH (FIGARCH) model. The FIGARCH(1, $d$ ,0) model is obtained from (28) by replacing the difference operator by a fractional difference operator:

$$(1 - L)^d \varepsilon_t^2 = \alpha_0 + \nu_t - \beta_1\nu_{t-1}. \quad (29)$$

The FIGARCH equation (29) can be written as an infinite-order ARCH model by applying the definition  $\nu_t = \varepsilon_t^2 - h_t$  to it. This yields

$$h_t = \alpha_0(1 - \beta_1)^{-1} + \lambda(L)\varepsilon_t^2$$

where  $\lambda(L) = \{1 - (1 - L)^d(1 - \beta_1 L)^{-1}\}\varepsilon_t^2 = \sum_{j=1}^{\infty} \lambda_j L^j \varepsilon_t^2$ , and  $\lambda_j \geq 0$  for all  $j$ . Expanding the fractional difference operator into an infinite sum yields the result that for long lags  $j$ ,

$$\lambda_j = \{(1 - \beta_1)\Gamma(d)^{-1}\}j^{-(1-d)} = cj^{-(1-d)}, c > 0 \quad (30)$$

where  $d \in (0, 1)$  and  $\Gamma(d)$  is the gamma function. From (30) it is seen that the effect of the lagged  $\varepsilon_t^2$  on the conditional variance decays hyperbolically as a function of the lag length. This is the reason why Baillie, Bollerslev and Mikkelsen (1996) introduced the FIGARCH model, as it would conveniently explain the apparent slow decay in autocorrelation functions of squared observations of many daily return series. The FIGARCH model thus offers a competing view to the one according to which changes in parameters in a GARCH model are the main cause of the slow decay in the autocorrelations. The first-order FIGARCH model (29) can of course be generalized into a FIGARCH( $p, d, q$ ) model.

The probabilistic properties of FIGARCH processes such as stationarity, still an open question, are quite complex, see, for example, Davidson (2004) and 2007GIRAITIS for discussion. The hyperbolic GARCH model introduced in the first-mentioned paper contains the standard GARCH and the FIGARCH models as two extreme special cases; for details see Davidson (2004).

### 3.7 Semi- and nonparametric ARCH models

The ARCH decomposition of returns (2) has also been used in a semi- or nonparametric approach. The semiparametric approach is typically employed in situations where the distribution of  $z_t$  is left unspecified and is estimated nonparametrically. In nonparametric models, the issue is the estimation of the functional form of the relationship between  $\varepsilon_t^2$  and  $\varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2$ . Semi- and nonparametric ARCH models are considered in detail in Linton (2007).

### 3.8 GARCH-in-mean model

GARCH models are often used for predicting the risk of a portfolio at a given point of time. From this it follows that the GARCH type conditional variance could be useful as a representation of the time-varying risk premium

in explaining excess returns, that is, returns compared to the return of a riskless asset. An excess return would be a combination of the unforecastable difference  $\varepsilon_t$  between the ex ante and ex post rates of return and a function of the conditional variance of the portfolio. Thus, if  $y_t$  is the excess return at time  $t$ ,

$$y_t = \varepsilon_t + \beta + g(h_t) - \mathbb{E}g(h_t) \quad (31)$$

where  $h_t$  is defined as a GARCH process (4) and  $g(h_t)$  is a positive-valued function. Engle, Lilien and Robins (1987) originally defined  $g(h_t) = \delta h_t^{1/2}$ ,  $\delta > 0$ , which corresponds to the assumption that changes in the conditional standard deviation appear less than proportionally in the mean. The alternative  $g(h_t) = \delta h_t$  has also appeared in the literature. Equations (31) and (4) form the GARCH-in-mean or GARCH-M model. It has been quite frequently applied in the applied econometrics and finance literature. Glosten et al. (1993) developed their asymmetric GARCH model as a generalization of the GARCH-M model.

The GARCH-M process has an interesting moment structure. Assume that  $\mathbb{E}z_t^3 = 0$  and  $\mathbb{E}\varepsilon_t^4 < \infty$ . From (31) it follows that the  $k$ th order autocovariance

$$\mathbb{E}(y_t - \mathbb{E}y_t)(y_{t-k} - \mathbb{E}y_t) = \mathbb{E}\varepsilon_{t-k}g(h_t) + \text{cov}(g(h_t), g(h_{t-k})) \neq 0.$$

This means that there is forecastable structure in  $y_t$ , which may contradict some economic theory if  $y_t$  is a return series. Hong (1991) showed this in a special case where  $g(h_t) = \delta h_t$ ,  $\mathbb{E}\varepsilon_t^4 < \infty$ , and  $h_t$  follows a GARCH( $p, q$ ) model. In that situation, all autocorrelations of  $y_t$  are nonzero. Furthermore,

$$\mathbb{E}(y_t - \mathbb{E}y_t)^3 = 3\mathbb{E}h_t\{g(h_t) - \mathbb{E}g(h_t)\} + \mathbb{E}\{g(h_t) - \mathbb{E}g(h_t)\}^3 \neq 0. \quad (32)$$

It follows from (32) that a GARCH-M model implies postulating a skewed marginal distribution for  $y_t$  unless  $g(h_t) \equiv \text{constant}$ . For example, if  $g(h_t) = \delta h_t^{1/2}$ ,  $\delta < 0$ , this marginal distribution is negatively skewed. If the model builder is not prepared to make this assumption or the one of forecastable structure in  $y_t$ , the GARCH-M model, despite its theoretical motivation, does not seem an appropriate alternative to use. For more discussion of this situation, see He, Silvennoinen and Teräsvirta (2006).

### 3.9 Stylized facts and the first-order GARCH model

As already mentioned, financial time series such as high-frequency return series constitute the most common field of applications for GARCH models. These series typically display rather high kurtosis. At the same time, the



autocorrelations of the absolute values or squares of the observations are low and decay slowly. These two features are sometimes called stylized facts of financial time series. Granger and Ding (1995) listed a few more such features. Among them is the empirical observation that in a remarkable number of financial series, the autocorrelations of the powers of observations,  $|\varepsilon_t|^k$ , peak around  $k = 1$ . Granger and Ding called this stylized fact the Taylor effect as Taylor (1986) was the first to draw attention to it (by comparing the autocorrelations of  $\varepsilon_t^2$  and  $|\varepsilon_t|$ ).

One way of evaluating the adequacy of GARCH models is to ask how well they can be expected to capture the features or stylized facts present in the series to be modelled. The expressions for kurtosis and the autocorrelation function of absolute values and squared observations are available for the purpose. They allow one to find out, for example, whether or not a GARCH(1,1) model is capable of producing realizations with high kurtosis and low, slowly decaying autocorrelations. The results of Malmsten and Teräsvirta (2004) who have used these expressions, illustrate the well known fact, see, for example, Bollerslev, Engle and Nelson (1994), that a GARCH model with normally distributed errors does not seem to be a sufficiently flexible model for explaining these two features in financial return series. This is shown in Figure 1. The panels contain a number of isoquants for which the sum  $\alpha_1 + \beta_1$  remains constant as a function of the kurtosis and the first-order autocorrelation of squared observations. Note that  $\alpha_1 + \beta_1$  is the decay rate of the autocorrelation function, that is, the  $j$ th autocorrelation  $\rho_j = (\alpha_1 + \beta_1)^{j-1} \rho_1$  for  $j \geq 1$ . They also contain combinations of the kurtosis and the first-order autocorrelation estimated directly from time series. It is seen that very often the kurtosis/autocorrelation combinations do not tend to lie in the vicinity of these isoquants even when  $\alpha_1 + \beta_1$  is very close to one. The isoquants are of course only defined for combinations of  $\alpha_1$  and  $\beta_1$  for which  $E\varepsilon_t^4 < \infty$ .

Malmsten and Teräsvirta (2004) also demonstrated how the situation can be improved, as is customary in practice, by replacing the normal error distribution by a more fat-tailed one. In Figure 2 it is seen how increasing the "baseline kurtosis", that is, the kurtosis of the distribution of  $z_t$ , the error, helps the GARCH(1,1) model to capture the stylized fact of high kurtosis/low autocorrelation. The isoquants are moved to the right because the baseline kurtosis increases. At the same time it seems that the level of the autocorrelations decreases. But then, this does not simultaneously affect the decay rate  $\alpha_1 + \beta_1$  of the autocorrelations.

Recently, Kim and White (2004) suggested that the standard estimator of kurtosis exaggerates the true kurtosis and that robust measures yield more reliable results. It follows that high kurtosis values estimated from return

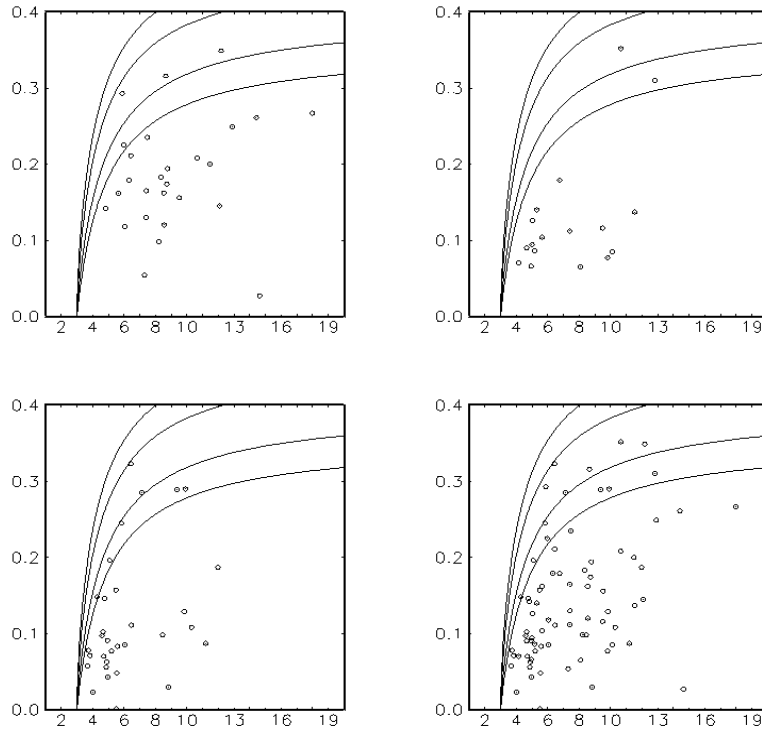


Figure 1: Kurtosis/first-order autocorrelation isoquants for the GARCH(1,1) model, from highest to lowest:  $\alpha + \beta = 0.9, 0.95, 0.99, 0.999$ , and corresponding combinations estimated from data: Upper left panel: Daily returns of the 27 most actively traded stocks at the Stockholm Stock Exchange; Upper right panel: Returns of five major daily exchange rates, divided into 34 subseries; Lower left panel: Daily returns of the S&P 500 index from 3 January 1928 to 19 September 2001, divided into 20 equally long subseries; Lower right panel: All observations

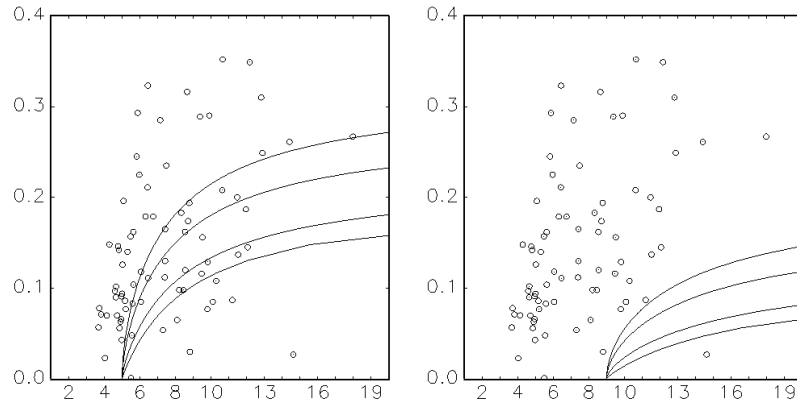


Figure 2: Isoquants of pairs of kurtosis and first-order autocorrelation of squared observations in the GARCH(1,1) model with  $t(7)$ -distributed (left-hand panel) and  $t(5)$ -distributed errors (right-hand panel), for (from above)  $\alpha + \beta = 0.90, 0.95, 0.99$  and  $0.999$ , and corresponding observations (the same ones as in the lower right panel of Figure 1)

series are a result of a limited number of outliers. If this is the case, then the use of a non-normal (heavy-tailed) error distribution may not necessarily be an optimal extension to the standard normal-error GARCH model. However, Teräsvirta and Zhao (2006) recently studied 160 daily return series and, following Kim and White (2004), used robust kurtosis and autocorrelation estimates instead of standard ones. Their results indicate that leptokurtic distributions for  $z_t$  are needed in capturing the kurtosis-autocorrelation stylized fact even when the influence of extreme observations is dampened by the use of robust estimates.

As to the Taylor effect, He and Teräsvirta (1999a) defined a corresponding theoretical property, the Taylor property, as follows. Let  $\rho(|\varepsilon_t|^k, |\varepsilon_{t-j}|^k)$  be the  $j$ th order autocorrelation of  $\{|\varepsilon_t|^k\}$ . The stochastic process has the Taylor property when  $\rho(|\varepsilon_t|^k, |\varepsilon_{t-j}|^k)$  is maximized for  $k = 1$  for  $j = 1, 2, \dots$ . In practice, He and Teräsvirta (1999a) were able to find analytical results for the AVGARCH(1,1) model, but they were restricted to comparing the first-order autocorrelations for  $k = 1$  and  $k = 2$ . For this model,  $\rho(|\varepsilon_t|, |\varepsilon_{t-1}|) > \rho(\varepsilon_t^2, \varepsilon_{t-1}^2)$  when the kurtosis of the process is sufficiently high. The corresponding results for the standard GARCH(1,1) model (4) with  $p = q = 1$  and normal errors are not available as the autocorrelation function of  $\{|\varepsilon_t|\}$  cannot be derived analytically. Simulations conducted by He and Teräsvirta (1999a) showed that the GARCH(1,1) model probably

does not possess the Taylor property, which may seem disappointing. But then, the results of Teräsvirta and Zhao (2006) show that if the standard kurtosis and autocorrelation estimates are replaced by robust ones, the evidence of the Taylor effect completely disappears. This stylized fact may thus be a consequence of just a small number of extreme observations in the series.

## 4 Family of Exponential GARCH models

### 4.1 Definition and properties

The Exponential GARCH (EGARCH) model is another popular GARCH model. Nelson (1991) who introduced it had three criticisms of the standard GARCH model in mind. First, parameter restrictions are required to ensure positivity of the conditional variance at every point of time. Second, the standard GARCH model does not allow an asymmetric response to shocks. Third, if the model is an IGARCH one, measuring the persistence is difficult since this model is strongly but not weakly stationary. Shocks may be viewed persistent as the IGARCH process looks like a random walk. However, the IGARCH model with  $\alpha_0 > 0$  is strictly stationary and ergodic, and when  $\alpha_0 = 0$ , the realizations collapse into zero almost surely, as already indicated in Section 3.6. The second drawback has since been removed as asymmetric GARCH models such as GJR-GARCH (Glosten et al. (1993)) or smooth transition GARCH have become available. A family of EGARCH( $p, q$ ) models may be defined as in (2) with

$$\ln h_t = \alpha_0 + \sum_{j=1}^q g_j(z_{t-j}) + \sum_{j=1}^p \beta_j \ln h_{t-j}. \quad (33)$$

When  $g_j(z_{t-j}) = \alpha_j z_{t-j} + \psi_j(|z_{t-j}| - \mathbb{E}|z_{t-j}|)$ ,  $j = 1, \dots, q$ , (33) becomes the EGARCH model of Nelson (1991). It is seen from (33) that no parameter restrictions are necessary to ensure positivity of  $h_t$ . Parameters  $\alpha_j, j = 1, \dots, q$ , make an asymmetric response to shocks possible.

When  $g_j(z_{t-j}) = \alpha_j \ln z_{t-j}^2$ ,  $j = 1, \dots, q$ , (2) and (33) form the logarithmic GARCH (LGARCH) model that Geweke (1986) and Pantula (1986) proposed. The LGARCH model has not become popular among practitioners. A principal reason for this may be that for parameter values encountered in practice, the theoretical values of the first few autocorrelations of  $\{\varepsilon_t^2\}$  at short lags tend to be so high that such autocorrelations can hardly be found in financial series such as return series. This being the case, the LGARCH model cannot be expected to provide an acceptable fit when applied to fi-

nancial series. Another reason are the occasional small values of  $\ln \varepsilon_t^2$  that complicate the estimation of parameters.

As in the standard GARCH case, the first-order model is the most popular EGARCH model in practice. Nelson (1991) derived existence conditions for moments of the general infinite-order Exponential ARCH model. Translated to the case of the EGARCH model (2) and (33) such that  $g_j(z_{t-j}) = \alpha_j z_{t-j} + \psi_j(|z_{t-j}| - \mathbb{E}|z_{t-j}|)$ ,  $j = 1, \dots, q$ , where not all  $\alpha_j$  and  $\psi_j$  equal zero, the conditions imply that if the error process  $\{z_t\}$  has all moments and  $\sum_{j=1}^p \beta_j^2 < 1$  in (33), then all moments for the EGARCH process  $\{\varepsilon_t\}$  exist. For example, if  $\{z_t\}$  is a sequence of independent standard normal variables then the restrictions on  $\beta_j$ ,  $j = 1, \dots, p$ , are necessary and sufficient for the existence of all moments simultaneously. This is different from the family (12) of GARCH models considered in Section 3.2. For those models, the moment conditions become more and more stringent for higher and higher even moments. The expressions for moments of the first-order EGARCH process can be found in He, Teräsvirta and Malmsten (2002); for the more general case, see He (2000).

## 4.2 Stylized facts and the first-order EGARCH model

In Section 3.9 we considered the capability of first-order GARCH models to characterize certain stylized facts in financial time series. It is instructive to do the same for EGARCH models. For the first-order EGARCH model, the decay of autocorrelations of squared observations is faster than exponential in the beginning before it slows down towards an exponential rate; see He et al. (2002). Thus it does not appear possible to use a standard EGARCH(1,1) model to characterize processes with very slowly decaying autocorrelations. Malmsten and Teräsvirta (2004) showed that the symmetric EGARCH(1,1) model with normal errors is not sufficiently flexible either for characterizing series with high kurtosis and slowly decaying autocorrelations. As in the standard GARCH case, assuming normal errors means that the first-order autocorrelation of squared observations increases quite rapidly as a function of kurtosis for any fixed  $\beta_1$  before the increase slows down. Analogously to GARCH, the observed kurtosis/autocorrelation combinations cannot be reached by the EGARCH(1,1) model with standard normal errors. The asymmetry parameter is unlikely to change things much.

Nelson (1991) recommended the use of the so-called Generalized Error Distribution (GED) for the errors. It contains the normal distribution as a special case but also allows heavier tails than the ones in the normal distribution. Nelson (1991) also pointed out that a  $t$ -distribution for the errors may imply infinite unconditional variance for  $\{\varepsilon_t\}$ . As in the case of the

GARCH(1,1) model, an error distribution with fatter tails than the normal one helps to increase the kurtosis and, at the same time, lower the autocorrelations of squared observations or absolute values.

Because of analytical expressions of the autocorrelations for  $k > 0$  given in He et al. (2002) it is possible to study the existence of the Taylor property in EGARCH models. Using the formulas for the autocorrelations of  $\{|\varepsilon_t|^k\}$ ,  $k > 0$ , it is possible to find parameter combinations for which these autocorrelations peak in a neighbourhood of  $k = 1$ . A subset of first-order EGARCH models thus has the Taylor property. This subset is also a relevant one in practice in the sense that it contains EGARCH processes with the kurtosis of the magnitude frequently found in financial time series. For more discussion on stylized facts and the EGARCH(1,1) model, see Malmsten and Teräsvirta (2004).

### 4.3 Stochastic volatility

The EGARCH equation may be modified by replacing  $g_j(z_{t-j})$  by  $g_j(s_{t-j})$  where  $\{s_t\}$  is a sequence of continuous unobservable independent random variables that are often assumed independent of  $z_t$  at all lags. Typically in applications,  $p = q = 1$  and  $g_1(s_{t-1}) = \delta s_{t-1}$  where  $\delta$  is a parameter to be estimated. This generalization is called the autoregressive stochastic volatility (SV) model, and it substantially increases the flexibility of the EGARCH parameterization. For evidence of this, see Malmsten and Teräsvirta (2004) and Carnero, Peña and Ruiz (2004). A disadvantage is that model evaluation becomes more complicated than that of EGARCH models because the estimation does not yield residuals. Several articles in this *Handbook* are devoted to SV models.

## 5 Comparing EGARCH with GARCH

The standard GARCH model is probably the most frequently applied parameterization of conditional heteroskedasticity. This being the case, it is natural to evaluate an estimated EGARCH model by testing it against the corresponding GARCH model. Since the EGARCH model can characterize asymmetric responses to shocks, a GARCH model with the same property, such as the GJR-GARCH or the smooth transition GARCH model, would be a natural counterpart in such a comparison. If the aim of the comparison is to choose between these models, they may be compared by an appropriate model selection criterion as in Shephard (1996). Since the GJR-GARCH and the EGARCH model of the same order have equally many parameters, this

amounts to comparing their maximized likelihoods.

If the investigator has a preferred model or is just interested in knowing if there are significant differences in the fit between the two, the models may be tested against each other. The testing problem is a non-standard one because the two models do not nest each other. Several approaches have been suggested for this situation. Engle and Ng (1993) proposed combining the two models into an encompassing model. If the GARCH model is an GJR-GARCH( $p, q$ ) one (both models can account for asymmetries), this leads to the following specification of the conditional variance:

$$\begin{aligned} \ln h_t = & \sum_{j=1}^q \{ \alpha_j^* z_{t-j} + \psi_j^* (|z_{t-j}| - \mathbb{E}|z_{t-j}|) \} + \sum_{j=1}^p \beta_j^* \ln h_{t-j} \\ & + \ln(\alpha_0 + \sum_{j=1}^q \{ \alpha_j + \omega_j I(\varepsilon_{t-j}) \} \varepsilon_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j}). \end{aligned} \quad (34)$$

Setting  $(\alpha_j, \omega_j) = (0, 0)$ ,  $j = 1, \dots, q$ ; and  $\beta_j = 0$ ,  $j = 1, \dots, p$ , in (34) yields an EGARCH( $p, q$ ) model. Correspondingly, the restrictions  $(\alpha_j^*, \psi_j^*) = (0, 0)$ ,  $j = 1, \dots, q$ , and  $\beta_j^* = 0$ ,  $j = 1, \dots, p$ , define the GJR-GARCH( $p, q$ ) model. Testing the models against each other amounts to testing the appropriate restrictions in (34). A Lagrange Multiplier test may be constructed for the purpose. The test may also be viewed as another misspecification test and not only as a test against the alternative model.

Another way of testing the EGARCH model against GARCH consists of forming the likelihood ratio statistic despite the fact that the null model is not nested in the alternative. This is discussed in Lee and Brorsen (1997) and Kim, Shephard and Chib (1998). Let  $\mathcal{M}_0$  be the EGARCH model and  $\mathcal{M}_1$  the GARCH one, and let the corresponding log-likelihoods be  $L_T(\varepsilon; \mathcal{M}_0, \theta_0)$  and  $L_T(\varepsilon; \mathcal{M}_1, \theta_1)$ , respectively. The test statistic is

$$LR = 2\{L_T(\varepsilon; \mathcal{M}_1, \hat{\theta}_1) - L_T(\varepsilon; \mathcal{M}_0, \tilde{\theta}_0)\}. \quad (35)$$

The asymptotic null distribution of (35) is unknown but can be approximated by simulation. Assuming that the EGARCH model is the null model and that  $\tilde{\theta}_0$  is the true parameter, one generates  $N$  realizations of  $T$  observations from this model and estimates both models and calculates the value of (35) using each realization. Ranking the  $N$  values gives an empirical distribution with which one compares the original value of (35). The true value of  $\theta_0$  is unknown, but the approximation error due to the use of  $\tilde{\theta}_0$  as a replacement vanishes asymptotically as  $T \rightarrow \infty$ . If the value of (35) exceeds the  $100(1 - \alpha)\%$  quantile of the empirical distribution, the null model is rejected at significance level  $\alpha$ . Note that the models under comparison need

not have the same number of parameters, and the value of the statistic can also be negative. Reversing the roles of the models, one can test GARCH models against EGARCH ones.

Chen and Kuan (2002) proposed yet another method based on the pseudo-score, whose estimator under the null hypothesis and assuming the customary regularity conditions is asymptotically normally distributed. This result forms the basis for a  $\chi^2$ -distributed test statistic; see Chen and Kuan (2002) for details.

Results of small-sample simulations in Malmsten (2004) indicate that the pseudo-score test tends to be oversized. Furthermore, the Monte Carlo likelihood ratio statistic seems to have consistently higher power than the encompassing test, which suggests that the former rather than the latter should be applied in practice.

## 6 Final remarks and further reading

The literature on univariate GARCH models is quite voluminous, and it is not possible to incorporate all developments and extensions of the original model in the present text. Several articles of this *Handbook* provide detailed analyses of various aspects of GARCH models. Modern econometrics texts contain accounts of conditional heteroskedasticity. A number of surveys of GARCH models exist as well. Bollerslev et al. (1994), Diebold and Lopez (1995), Palm (1996), and Guégan (1994, Ch. 5), survey developments till the early 1990s; see Giraitis, Leipus and Surgailis (2006) for a very recent survey. Shephard (1996) considers both univariate GARCH and stochastic volatility models. The focus in Gouriéroux (1996) lies on both univariate and multivariate ARCH models. The survey by Bollerslev et al. (1992) also reviews applications to financial series. The focus in Straumann (2004) is on estimation in models of conditional heteroskedasticity. Theoretical results on time series models with conditional heteroskedasticity are also reviewed in Li, Ling and McAleer (2002). Engle (1995) contains a selection of the most important articles on ARCH and GARCH models up until 1993.

Multivariate GARCH models are not included in this article. There exists a recent survey by Bauwens, Laurent and Rombouts (2006), and these models are also considered in Silvennoinen and Teräsvirta (2007).



## References

- Andersen, T. G. and Bollerslev, T.: 1998, Answering the skeptics: Yes, standard volatility models provide accurate forecasts, *International Economic Review* **39**, 885–905.
- Anderson, H. M., Nam, K. and Vahid, F.: 1999, Asymmetric nonlinear smooth transition GARCH models, in P. Rothman (ed.), *Nonlinear time series analysis of economic and financial data*, Kluwer, Boston, pp. 191–207.
- Audrino, F. and Bühlmann, P.: 2001, Tree-structured generalized autoregressive conditional heteroscedastic models, *Journal of the Royal Statistical Society B* **63**, 727–744.
- Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O.: 1996, Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **74**, 3–30.
- Bauwens, L., Laurent, S. and Rombouts, J. V. K.: 2006, Multivariate GARCH models: A survey, *Journal of Applied Econometrics* **21**, 79–109.
- Bollerslev, T.: 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**, 307–327.
- Bollerslev, T., Chou, R. Y. and Kroner, K. F.: 1992, ARCH modeling in finance. A review of the theory and empirical evidence, *Journal of Econometrics* **52**, 5–59.
- Bollerslev, T., Engle, R. F. and Nelson, D. B.: 1994, ARCH models, in R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, North-Holland, Amsterdam, pp. 2959–3038.
- Cai, J.: 1994, A Markov model of switching-regime ARCH, *Journal of Business and Economic Statistics* **12**, 309–316.
- Carnero, M. A., Peña, D. and Ruiz, E.: 2004, Persistence and kurtosis in GARCH and stochastic volatility models, *Journal of Financial Econometrics* **2**, 319–342.
- Chen, Y.-T. and Kuan, C.-M.: 2002, The pseudo-true score encompassing test for non-nested hypotheses, *Journal of Econometrics* **106**, 271–295.

- Chu, C.-S. J.: 1995, Detecting parameter shift in GARCH models, *Econometric Reviews* **14**, 241–266.
- Davidson, J.: 2004, Moment and memory properties of linear conditional heteroscedasticity models, and a new model, *Journal of Business and Economic Statistics* **22**, 16–29.
- Diebold, F. X.: 1986, Modeling the persistence of conditional variances: A comment, *Econometric Reviews* **5**, 51–56.
- Diebold, F. X. and Lopez, J. A.: 1995, Modeling volatility dynamics, in K. D. Hoover (ed.), *Macroeconometrics: Developments, Tensions, and Prospects*, Kluwer, Boston, pp. 427–472.
- 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.
- Duan, J.-C.: 1997, Augmented GARCH(p,q) process and its diffusion limit, *Journal of Econometrics* **79**, 97–127.
- Engle, R. F.: 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**, 987–1007.
- Engle, R. F.: 2004, Risk and volatility: Econometric models and financial practice, *American Economic Review* **94**, 405–420.
- Engle, R. F. and Bollerslev, T.: 1986, Modeling the persistence of conditional variances, *Econometric Reviews* **5**, 1–50.
- Engle, R. F. (ed.): 1995, *ARCH. Selected Readings*, Oxford University Press, Oxford.
- Engle, R. F., Lilien, D. M. and Robins, R. P.: 1987, Estimating time-varying risk premia in the term structure: The ARCH-M model, *Econometrica* **55**, 391–407.
- Engle, R. F. and Mezrich, J.: 1996, GARCH for groups, *Risk* **9**(8), 36–40.
- Engle, R. F. and Ng, V. K.: 1993, Measuring and testing the impact of news on volatility, *Journal of Finance* **48**, 1749–1777.
- Friedman, M.: 1977, Nobel lecture: Inflation and unemployment, *Journal of Political Economy* **85**, 451–472.

- Geweke, J.: 1986, Modeling the persistence of conditional variances: A comment, *Econometric Reviews* **5**, 57–61.
- Giraitis, L., Kokoszka, P. and Leipus, R.: 2000, Stationary ARCH models: Dependence structure and central limit theorem, *Econometric Theory* **16**, 3–22.
- Giraitis, L., Leipus, R. and Surgailis, D.: 2006, Recent advances in ARCH modelling, in A. Kirman and G. Teyssiere (eds), *Long Memory in Economics*, Springer, Berlin.
- 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.
- Gonzalez-Rivera, G.: 1998, Smooth transition GARCH models, *Studies in Nonlinear Dynamics and Econometrics* **3**, 161–178.
- Gouriéroux, C.: 1996, *ARCH Models and Financial Applications*, Springer, Berlin.
- Granger, C. W. J. and Ding, Z.: 1995, Some properties of absolute returns. An alternative measure of risk, *Annales d'économie et de statistique* **40**, 67–92.
- Gray, S. F.: 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* **42**, 27–62.
- Guégan, D.: 1994, *Séries chronologiques non linéaires à temps discret*, Economica, Paris.
- Haas, M., Mittnik, S. and Paolella, M. S.: 2004, A new approach to Markov-switching GARCH models, *Journal of Financial Econometrics* **4**, 493–530.
- Hagerud, G.: 1997, *A New Non-Linear GARCH Model*, EFI Economic Research Institute, Stockholm.
- Hamilton, J. D. and Susmel, R.: 1994, Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics* **64**, 307–333.
- He, C.: 2000, Moments and the autocorrelation structure of the exponential GARCH( $p, q$ ) process., *SSE/EFI Working Paper Series in Economics and Finance* 359, Stockholm School of Economics.

- He, C., Silvennoinen, A. and Teräsvirta, T.: 2006, Parameterizing unconditional skewness in models for financial time series, *Unpublished paper*, Stockholm School of Economics.
- He, C. and Teräsvirta, T.: 1999a, Properties of moments of a family of GARCH processes, *Journal of Econometrics* **92**, 173–192.
- He, C. and Teräsvirta, T.: 1999b, Properties of the autocorrelation function of squared observations for second order GARCH processes under two sets of parameter constraints, *Journal of Time Series Analysis* **20**, 23–30.
- He, C., Teräsvirta, T. and Malmsten, H.: 2002, Moment structure of a family of first-order exponential GARCH models, *Econometric Theory* **18**, 868–885.
- Hentschel, L.: 1995, All in the family. Nesting symmetric and asymmetric GARCH models, *Journal of Financial Economics* **39**, 71–104.
- Hong, E. P.: 1991, The autocorrelation structure for the GARCH-M process, *Economics Letters* **37**, 129–132.
- Kim, S., Shephard, N. and Chib, S.: 1998, Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies* **65**, 361–393.
- Kim, T.-H. and White, H.: 2004, On more robust estimation of skewness and kurtosis, *Finance Research Letters* **1**, 56–73.
- Klaassen, F.: 2002, Improving GARCH volatility forecasts with regime-switching GARCH, *Empirical Economics* **27**, 363–394.
- Lamoureux, C. G. and Lastrapes, W. G.: 1990, Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**, 225–234.
- Lange, T. and Rahbek, A.: 2007, An introduction to regime switching time series, in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch (eds), *Handbook of Financial Time Series*, Springer, New York.
- Lanne, M. and Saikkonen, P.: 2005, Nonlinear GARCH models for highly persistent volatility, *Econometrics Journal* **8**, 251–276.
- Lee, J.-H. and Brorsen, B. W.: 1997, A non-nested test of GARCH vs. EGARCH models, *Applied Economics Letters* **4**, 765–768.

- Li, C. W. and Li, W. K.: 1996, On a double threshold autoregressive heteroskedasticity time series model, *Journal of Applied Econometrics* **11**, 253–274.
- Li, W. K., Ling, S. and McAleer, M.: 2002, Recent theoretical results for time series models with GARCH errors, *Journal of Economic Surveys* **16**, 245–269.
- Linton, O.: 2007, Semi- and nonparametric ARCH/GARCH-Modelling, in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch (eds), *Handbook of Financial Time Series*, Springer, New York.
- Liu, J.-C.: 2006, Stationarity of a Markov-switching GARCH model, *Journal of Financial Econometrics* **4**, 573–593.
- Malmsten, H.: 2004, Evaluating Exponential GARCH models, *SSE/EFI Working Paper Series in Economics and Finance* 564, Stockholm School of Economics.
- Malmsten, H. and Teräsvirta, T.: 2004, Stylized facts of financial time series and three popular models of volatility, *SSE/EFI Working Paper Series in Economics and Finance* 563, Stockholm School of Economics.
- Mandelbrot, B.: 1963, The variation of certain speculative prices, *Journal of Business* **36**, 394–419.
- Mandelbrot, B. and Taylor, H.: 1967, On the distribution of stock price differences, *Operations Research* **15**, 1057–1062.
- Mikosch, T. and Starica, C.: 2004, Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects, *Review of Economics and Statistics* **86**, 378–390.
- Nelson, D. B.: 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* **59**, 347–370.
- Nelson, D. B. and Cao, C. Q.: 1992, Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics* **10**, 229–235.
- Palm, F. C.: 1996, GARCH models of volatility, in G. Maddala and C. Rao (eds), *Handbook of Statistics: Statistical Methods in Finance*, Vol. 14, Elsevier, Amsterdam, pp. 209–240.

- Pantula, S. G.: 1986, Modeling the persistence of conditional variances: A comment, *Econometric Reviews* **5**, 71–74.
- Rabemananjara, R. and Zakoïan, J. M.: 1993, Threshold ARCH models and asymmetries in volatility, *Journal of Applied Econometrics* **8**, 31–49.
- Rachev, S. and Mitnik, S.: 2000, *Stable Paretian Models in Finance*, Wiley, Chichester.
- Rydén, T., Teräsvirta, T. and Åsbrink, S.: 1998, Stylized facts of daily return series and the hidden Markov model, *Journal of Applied Econometrics* **13**, 217–244.
- Schwert, G. W.: 1990, Stock volatility and the crash of '87, *Review of Financial Studies* **3**, 77–102.
- Sentana, E.: 1995, Quadratic ARCH models, *Review of Economic Studies* **62**, 639–661.
- Shephard, N. G.: 1996, Statistical aspects of ARCH and stochastic volatility, in D. R. Cox, D. V. Hinkley and O. E. Barndorff-Nielsen (eds), *Time Series Models. In Econometrics, Finance and Other Fields*, Chapman and Hall, London, pp. 1–67.
- Silvennoinen, A. and Teräsvirta, T.: 2007, Multivariate GARCH models, in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch (eds), *Handbook of Financial Time Series*, Springer, New York.
- Straumann, D.: 2004, *Estimation in Conditionally Heteroscedastic Time Series Models*, Springer, New York.
- Taylor, S.: 1986, *Modelling Financial Time Series*, Wiley, Chichester.
- Teräsvirta, T. and Zhao, Z.: 2006, Stylized facts of return series, robust estimates, and three popular models of volatility, *Unpublished paper*, Stockholm School of Economics.
- Tong, H.: 1990, *Non-Linear Time Series. A Dynamical System Approach*, Oxford University Press, Oxford.
- Zakoïan, J.-M.: 1994, Threshold heteroskedastic models, *Journal of Economic Dynamics and Control* **18**, 931–955.