PROPERTIES AND EVALUATION OF VOLATILITY MODELS

Hans Malmsten

AKADEMISK AVHANDLING

som för avläggande av ekonomie doktorsexamen
vid Handelshögskolan i Stockholm
framläggs för offentlig granskning
Torsdagen den 3 Juni 2004
kl. 10.15 i sal KAW
Handelshögskolan, Sveavägen 65

Properties and evaluation of volatility models

STOCKHOLM SCHOOL
OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

EFI Mission
EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.
EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization
The research activities are organized in twenty Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

ORGANIZATION AND MANAGEMENT	
Management and Organisation; (A)	Prof Sven-Erik Sjöstrand
Center for Ethics and Economics; (CEE)	Adj Prof Hans de Geer
Center for Entrepreneurship and Business Creation; (E)	Prof Carin Holmquist
Public Management; (F)	Prof Nils Brunsson
Information Management; (I)	Prof Mats Lundeberg
Center for People and Organization; (PMO)	Prof Jan Löwstedt
Center for Innovation and Operations Management; (T)	Prof Christer Karlsson
ECONOMIC PSYCHOLOGY	
Center for Risk Research; (CFR)	Prof Lennart Sjöberg
Economic Psychology; (P)	Prof Guje Sevón
MARKETING	
Center for Consumer Marketing; (CCM)	Acting Prof Magnus Söderlund
Center for Information and Communication Research; (CIC)	Adj Prof Bertil Thorngren
Marketing, Distribution and Industrial Dynamics; (D)	Prof Björn Axelsson
ACCOUNTING, CONTROL AND CORPORATE FINANCE	
Accounting and Managerial Finance; (B)	Prof Lars Östman
Managerial Economics; (C)	Prof Peter Jennergren
FINANCE	
Finance; (FI)	Prof Clas Bergström
ECONOMICS	
Center for Health Economics; (CHE)	Prof Bengt Jönsson
International Economics and Geography; (IEG)	Prof Mats Lundahl
Economics; (S)	Prof Lars Bergman
ECONOMIC STATISTICS	
Economic Statistics; (ES)	Prof Anders Westlund
LAW	
Law; (RV)	Prof Erik Nerep

Chairman of the Board: Prof Carin Holmquist. *Director:* Associate Prof Bo Sellstedt.

Adress
EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

PROPERTIES AND EVALUATION OF VOLATILITY MODELS

Hans Malmsten

STOCKHOLM SCHOOL
OF ECONOMICS
HANDELSHÖGSKOLAN I STOCKHOLM

EFI, The Economic Research Institute

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics 2004

ISBN 91-7258-641-9

Keywords: Volatility models, GARCH, EGARCH

Printed by:
Elanders Gotab, Stockholm 2004

Distributed by:
EFI, The Economic Research Institute
Stockholm School of Economics
P O Box 6501, SE-113 83 Stockholm, Sweden
www.hhs.se/efi

To Lena and Saga

Contents

Acknowledgments

Under the process of writing this thesis a number of people have contributed in various ways.

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Timo Teräsvirta. Throughout these years he has always found the time to listen and discuss ideas, read the drafts of the papers, and giving insightful and detailed comments, meanwhile showing great patience with my progress.

I am also grateful to my co-advisor Changli He.

I would like to thank Professor David F Hendry and Professor Tony Hall for inviting me to Nuffield College, Oxford University and University of Technology Sydney, respectively.

Finally, I would like to thank my wife Lena, for her love and support.

Part I

Summary of Thesis

Introduction

The volatility of financial time series varies over time. Volatility of returns is a key issue for analysts in financial markets. The prices of stocks and other assets depend on the expected volatility of returns. For capital asset pricing model, the option-pricing formulas and the value-at-risk model, modelling volatility of asset returns is essential.

In time series econometrics, model builders generally parameterize the conditional mean of a variable where the error term is a sequence of uncorrelated random variables with mean zero. When estimating the parameters it is typically assumed that the unconditional variance of the error term is constant. Engle (1982) considered that, while the unconditional error variance, if it exists, is constant, the conditional error variance is time-varying. This revolutionary notion made it possible to explain systematic features in the movements of variance over time. In Engle's autoregressive conditional heteroskedasticity (ARCH) model the conditional variance is a function of past values of squared errors. Among other things, Engle presented a test for the hypothesis of no ARCH (equal conditional and unconditional variance) in the errors and derived the properties of the maximum likelihood estimator of the parameters of the ARCH model. He also applied the model to financial time series such as return series. Engle's work on time-varying volatility earned him the Bank of Sweden Prize of Economic Sciences in Memory of Alfred Nobel 2003.

In practice, the autocorrelation function of squared returns tends to decay slowly. An adequate characterization of this stylized fact requires an ARCH model with many lags. By adding lags of the the conditional variance (one is often enough), the resulting model can be formulated with only a small number of parameters and still display a slowly decaying autocorrelation function for the squared errors. Bollerslev (1986) introduced this generalized ARCH (GARCH) model. The GARCH model has attracted plenty of attention. As the GARCH model is suitable for modelling financial time series, it can be used to forecast volatility. For a overview of GARCH models see, for example, Bollerslev, Engle and Nelson (1994), Palm (1996), Shephard (1996), Gouriéroux (1997), and Poon and Granger (2003).

What is parameterized in the GARCH model is the conditional variance. An alternative is to parametrize the conditional standard deviation. A more general model would contain both alternatives, the GARCH and the absolute value GARCH (AVGARCH), as special case. Such a model, called the asymmetric power ARCH (A-PARCH) model, was introduced by Ding, Granger and Engle (1993). Engle's original idea inspired for different parametrization. The most commonly applied is the exponential GARCH (EGARCH) model of Nelson (1991), where the logarithm of the conditional variance has a parametric form. This was the first asymmetric GARCH model. Another innovation was that while ordinary GARCH models require parameter restrictions for conditional variance to be positive, such restrictions are not needed in the EGARCH model.

Another model that deserves mention in this context is the autoregressive stochastic volatility (ARSV) model. It differs from GARCH models in that the logarithm of the conditional variance is itself a stochastic process. In fact, the ARSV model is strictly speaking not a model of conditional heteroskedasticity because its conditioning set differs from the one in ARCH and GARCH models.

The general theme of this thesis is theoretical properties and evaluation of volatility models. In the first chapter the moment structure of the EGARCH model is derived. The second chapter contains new results on the A-PARCH model. The third chapter is about certain stylized facts of financial time series and the idea is to investigate how well the first-order GARCH, EGARCH and ARSV models are able to reproduce these characteristics. The fourth chapter is about evaluating the EGARCH model and fills a gap in the literature following Nelson (1991). A more detailed overview of the chapters follows next.

Main results of the Chapters

Chaper 1. Moment structure of a family of first-order exponential GARCH models[1]

The statistical properties of GARCH models have attracted plenty of attention. Among other things, their moment structure has been investigated in a number of papers; see, for example, He and Teräsvirta (1999), He and Teräsvirta (1999*a*) and Karanasos (1999). On the other hand, the moment structure of the EGARCH model has not been fully worked out. In Chapter 1 we derive the condition for the existence of moments, the expression for the kurtosis and the one for the autocorrelation function of positive powers of the absolute-valued observations for the first-order EGARCH model.

The results of the paper are useful, for example, if we want to compare the EGARCH model with the GARCH model. They reveal certain differences in the moment structure between these models. While the autocorrelations of the squared observations decay exponentially in the first-order GARCH model, the decay rate is not exponential in the first-order EGARCH model. While for the GARCH model the conditions for parameters allowing the existence of higher-order moments become more and more stringent for each even moment this is not the case for the EGARCH model. The explicit expressions of the autocorrelation structure of the positive powers of the absolute-valued observations of the model are particularly important in the considerations of Chapter 3 of the thesis.

The results of the paper only cover first-order EGARCH models. Working out the moment structure of higher-order EGARCH models is done in He (2000).

[1]This chapter has appeared in Econometric Theory, 18, 2002, 868-885, and is joint work with Changli He and Timo Teräsvirta. Reprinted with permission from Cambridge University Press.

Chapter 2. Higher-order dependence in the general power ARCH process and the role of the power parameter

The A-PARCH model contains as special cases a large number of well-known models, among other the ARCH, the GARCH and the AVGARCH models. The A-PARCH model contains a particular positive power parameter. By letting the power parameter approach zero, the A-PARCH family of models also includes a family of EGARCH models as a special case. In this chapter we derive the autocorrelation function of squared and logarithmed observations for this family and show that it may be obtained as a limiting case of a general power ARCH (GPARCH) model. A notable feature of the A-PARCH model is that, due to its parameterization, it is only possible to find analytically certain fractional moments of the absolute values of the original process related to the power parameter. An interesting thing to notice is that the autocorrelation structure of this GPARCH process, if it exists, is exponential, and that this property is retained at the limit as the power parameter approaches zero, which means that the autocorrelation function of the process of logarithms of squared observations also decay exponentially. While this is true for the logarithmed squared observations of an EGARCH(1,1) process it cannot simultaneously be true for the untransformed observations defined by these processes as we in Chapter 1 have demonstrated.

In order to explain the role of the power parameter we present a detailed analysis of how the autocorrelation functions of the squared observations differ across members of the GPARCH models. We demonstrate the fact that the power parameter lowers the autocorrelations of squared observations compared to the corresponding autocorrelations implied, other things equal, by the GARCH model. This fact may explain the regularities in estimation results in papers in which GPARCH models have been fitted to stock return series. In an empirical example we consider return series of 30 most actively traded stocks of the Stockholm Stock Exchange. We report the maximum likelihood estimates of the power parameter. Most of the estimates lie close to the mean value of 1.40. We find that the estimates are remarkably similar to the ones obtained by othor authors. We also show that the estimated power parameter considerably improves the correspondence between the estimated autocorrelations on the one hand and the autocorrelation estimates from the model on the other.

Chapter 3. Stylized facts of financial time series and three popular models of volatility

Financial time series seem to share a number of characteristic features, sometimes called stylized facts. One is the stylized fact of high kurtosis and relative low autocorrelations of squared observations. Another is the fact that the autocorrelation function of absolute-valued returns raised to a positive power tends to peak when the power is unity, that is, for the absolute-valued observations. He and Teräsvirta (1999) defined the corresponding theoretical property and called it the Taylor property. Given a set of stylized facts, one may ask the following question: "Have popular volatility models been parameterized in such a way that they can accommodate and explain the most common stylized facts visible in the data?" Models for which the answer is positive may be viewed as suitable for practical use. In Chapter 3, possible answers to this question for the three popular models of volatility, GARCH, EGARCH and ARSV models are investigated.

In this chapter we show that there exist possibilities of parameterizing all three models in such a way that they can accommodate and explain many of the stylized facts visible in the data. Some stylized facts may in certain cases remain unexplained, however. For example, it appears that the standard GARCH(1,1) model may not particularly often generate series that display the Taylor effect. This is due to the fact that this model does not appear to satisfy the corresponding theoretical property, the Taylor property. On the contrary, this property is approximately satisfied for a relevant subset of EGARCH(1,1) and ARSV(1) models and, albeit very narrowly, for a subset of AVGARCH models.

Many researchers observed quite early on that for GARCH models, assuming normal errors is too strong a restriction, and they have suggested leptokurtic error distributions in their stead. The results in this paper show how these distributions add to the flexibility of the GARCH model and help the model to reproduce the stylized fact of high kurtosis and relative low autocorrelations of squared observations. As a drawback it may be noted that the parameterization of the first-order stochastic volatility model becomes very restrictive when the amount of the leptokurtosis in the error distribution increases, and the model therefore cannot accommodate easy situations with relatively low kurtosis and high autocorrelations of squared observations.

The present investigation is only concerned with first-order models, and it may be asked if adding more lags would enhance the flexibility of the models. Such additions would certainly help to generate and reproduce more elaborate autocorrelation patterns for the squared observations than is the case with first-order models. It is far from certain, however, that they would also improve reproduction of the stylized facts considered in this study.

Chapter 4. Evaluating exponential GARCH models

Model evaluation is an important part of modelling not only for the conditional mean models but for the conditional variance specifications as well. Recently, Lundbergh and Teräsvirta (2002) presented a unified framework for testing the adequacy of an estimated GARCH model. It appears that less work has been done for the evaluation of EGARCH model. In Chapter 4 we consider misspecification tests for an EGARCH model. We derive two new misspecification tests for an EGARCH model. Because the tests of an EGARCH model against a higher-order EGARCH model and testing parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This is useful for a model builder who wants to find out about possible weaknesses in the estimated specification. It may also give him/her useful ideas of how the model could be further improved. They can be recommended as standard tools when it comes to testing the adequacy of an estimated EGARCH model.

Furthermore, we investigate various ways of testing the EGARCH model against GARCH ones as another check of model adequacy. The literature on testing non-nested hypotheses for volatility models includes Chen and Kuan (2002), Lee and Brorsen (1997); see also Engle and Ng (1993). Their tests are considered in the present framework, and the small-sample properties of the tests are investigated by simulation. Our simulations show that the size of the test may be a problem in applying the test suggested by Chen and Kuan (2002).

An empirical example shows based on investigating daily return series of 30 most actively traded stocks of the Stockholm Stock Exchange that there is substantial evidence for parameter nonconstancy in these return series. Turning to choosing between EGARCH and GARCH, the tests indicate that both models fit the data more or less equally well.

Bibliography

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, Amsterdam: Elsevier, pp. 2961–3039.

Chen, Y.-T. and Kuan, C.-M. (2002), 'The pseudo-true score encompassing test for non-nested hypotheses', *Journal of Econometrics* **106**, 271–295.

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.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of the united kingdom inflation', *Econometrica* **50**, 987–1007.

Engle, R. F. and Ng, V. K. (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance* **48**, 1749–1778.

Gouriéroux, C. (1997), *ARCH Models and Financial Applications*, New York: Springer.

He, C. (2000), Moments and the autocorrelation structure of the exponential GARCH(p,q) process. Working Paper Series in Economics and Finance, Stockholm School of Economics.

He, C. and Teräsvirta, T. (1999*a*), 'Fourth moment structure of the GARCH(p,q) process', *Econometric Theory* **15**, 824–846.

He, C. and Teräsvirta, T. (1999*b*), 'Properties of moments of a family of GARCH processes', *Journal of Econometrics* **92**, 173–192.

Karanasos, M. (1999), 'The second moment and the autocovariance function of the squared errors of the GARCH model', *Journal of Econometrics* **90**, 63–76.

Lee, J.-H. and Brorsen, B. W. (1997), 'A non-nested test of GARCH vs. EGARCH models', *Applied Economics Letters* 4, 765–768.

Lundbergh, S. and Teräsvirta, T. (2002), 'Evaluating GARCH models', *Journal of Econometrics* **110**, 417–435.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**, 347–370.

Palm, F. C. (1996), GARCH models of volatility, *in* G. S. Maddal and C. R. Rao, eds, 'Handbook of Statistics', Vol. 14, *Statistical Methods in Finance*, Amsterdam: Elsevier, pp. 209–240.

Poon, S. H. and Granger, C. W. J. (2003), 'Forecasting volatility in financial markets', *Journal of Economic Literature* **XLI**, 478–539.

Shephard, N. (1996), Statistical aspects of ARCH and stochastic volatility, *in* D. R. Cox and O. E. Barndorff-Nielsen and D. V. Hinkley, eds, *Time series models: in econometrics, finance and other fields*, London: Chapman and Hall, chapter 1, pp. 1–67.

Part II

The Chapters

Chapter 1

Moment structure of a family of first-order exponential GARCH models

1.1 Introduction

The moment structure of generalized autoregressive conditional heteroskedasticity (GARCH) models is a topic that has recently attracted plenty of attention. Bollerslev (1986, 1988) already derived conditions for weak stationarity and the existence of higher order moments in the standard GARCH(1,1) model under the assumption of normal errors. Several authors have studied conditions for the existence of higher order moments in GARCH models; see, for example, Ling and Li (1997), An and Chen (1998), Carrasco and Chen (2000), and Giraitis, Kokoszka and Leipus (2000). He and Teräsvirta (1999a) concentrate on the fourth moment structure of a family of first-order GARCH models and derived analytic expressions of the fourth moment, the kurtosis and the autocorrelation function of squared observations for this family. He and Teräsvirta (1999b) do all this for the standard GARCH(p,q) model; see also Karanasos (1999).

Some results also exist for the exponential GARCH (EGARCH) model which has become a frequently applied GARCH model. Nelson (1991) derives the autocorrelation function of the logarithm of the conditional variance of the EGARCH process. Breidt, Crato and de Lima (1998) obtain the autocorrelation function of the squared and logarithmed observations obeying the EGARCH model. Both Nelson and Breidt et al. base their considerations on the infinite moving average representation of the logarithm of the conditional variance. For the EGARCH(1,1) process, Chapter 2 shows how the autocorrelation function of the squared and logarithmed observations may be obtained as a limiting case from the asymmetric power ARCH model of Ding, Granger and Engle (1993). This autocorrelation function was shown to decay exponentially from the first lag, and He and Teräsvirta derive analytic expressions for these autocorrelations.

Despite this progress, the moment structure of the EGARCH model has not been fully worked out yet. It would be interesting to know the properties of the autocorrelation function of, say, the squared observations themselves, not just their logarithms. This would be useful when comparing the EGARCH model with the standard GARCH model family. In particular, possible differences in the moment structure of these models might help explain the success of the EGARCH model in applications.[1]

In this paper we derive the moment structure of a class of EGARCH(1,1) models without assuming normality of the errors. Choosing the first-order

[1] Recently, Demos (2000) has worked out the autocorrelation function of squared observations for the EGARCH(1,1) process assuming normality and using a technique different from ours.

model here is motivated by the fact that it is the most widely applied exponential GARCH model. Besides, as in the GARCH case, considering higher-order models is more tedious and is not done here. The paper is organized as follows. In Section 2 we introduce notation and define our class of exponential GARCH models. In Section 3 we derive the condition for the existence of moments of members of this class. Section 4 contains the expression for the kurtosis and Section 5 the one for the autocorrelation function of positive powers of the absolute-valued observations. The interesting properties of the autocorrelation function are highlighted and discussed. Section 6 contains conclusions.

1.2 Preliminaries

Let $\{\varepsilon_t\}$ be a real-valued discrete time stochastic process generated by

$$\varepsilon_t = z_t h_t \tag{1.1}$$

where $\{z_t\}$ is a sequence of independent and identically distributed random variables with mean zero and h_t is a positive with probability one, $\mathcal{F}_{t-1}$-measurable function, where $\mathcal{F}_{t-1}$ is the sigma-algebra generated by $\{z_{t-1}, z_{t-2}, z_{t-3}, ...\}$. Let

$$\ln h_t^2 = \alpha_0 + g(z_{t-1}) + \beta \ln h_{t-1}^2 \tag{1.2}$$

where $g(z_t)$ is a well-defined function of z_t. Equations (1.1) and (1.2) define a family of first-order exponential GARCH (EGARCH(1,1)) models. For example, setting $g(z_t) = \phi z_t + \psi(|z_t| - \mathsf{E}\,|z_t|)$ in (1.2) yields

$$\ln h_t^2 = \alpha_0 + \phi z_{t-1} + \psi(|z_{t-1}| - \mathsf{E}\,|z_{t-1}|) + \beta \ln h_{t-1}^2, \tag{1.3}$$

which, jointly with (1.1), defines the EGARCH(1,1) model of Nelson (1991). On the other hand, setting

$$g(z_{t-1}) = \alpha \ln z_{t-1}^2 = \alpha(\ln \varepsilon_{t-1}^2 - \ln h_{t-1}^2)$$

one obtains the logarithmic GARCH (LGARCH) model of Geweke (1986) and Pantula (1986).

First we establish a stationarity condition for the EGARCH model class (1.1) and (1.2). We have the following result.

Theorem 1 *Assume that $\nu_2 = \mathsf{E} z_t^2 < \infty$ and $var(g(z_t)) < \infty$. Then the EGARCH(1,1) process (1.1) and (1.2) is strictly stationary if and only if*

$$|\beta| < 1. \tag{1.4}$$

Proof. See the Appendix.

Remark 1. Nelson (1991) derives a necessary and sufficient condition for the strict stationarity of the EGARCH process (1.1) and (1.3) when $\ln h_t^2$ has an infinite moving average representation. Hentschel (1995) also obtains the same result. Theorem 1 gives this result for the general EGARCH class (1.1) and (1.2).

1.3 Existence of unconditional moments of the EGARCH process

In this section we consider the existence of unconditional moments of ε_t defined by (1.1) and (1.2). We concentrate ourselves to the case where $\{\varepsilon_t\}$ is a strictly stationary sequence. From definition (1.2) it follows that

$$h_t^2 = \exp\{\alpha_0 + g(z_{t-1}) + \beta \ln h_{t-1}^2\} = [\exp\{\alpha_0\} \exp\{g(z_{t-1})\}]\, h_{t-1}^{2\beta}. \qquad (1.5)$$

Recursions on equation (1.5) ultimately give the $2m$-th unconditional moment of ε_t whose expression is stated in the following theorem.

Theorem 2 *Consider the EGARCH(1,1) process (1.1) and (1.2) with* $|\beta| < 1$. *Assume that* $\mathsf{E}\exp\{mg(z_t)\} < \infty$ *for an arbitrary* $m > 0$ *and* $\nu_{2m} = \mathsf{E}\,|z_t|^{2m} < \infty$. *Then for this process, the* $2m$*-th unconditional moment of* ε_t *exists and has the form*

$$\mu_{2m} = \mathsf{E}\varepsilon_t^{2m} = \nu_{2m} \exp\{m\alpha_0(1-\beta)^{-1}\} \prod_{i=1}^{\infty} \mathsf{E}\exp\{m\beta^{i-1} g(z_t)\}. \qquad (1.6)$$

Proof. See the Appendix.

Remark 2. Note that (1.4) does not depend on m. Thus, if $|\beta| < 1$ and all moments of z_t and $\exp\{g(z_t)\}$ exist then all moments of ε_t in (1.1) exist as well. The situation is quite different from that in standard GARCH processes, where the condition $\mathsf{E}z_t^{2m} < \infty$ is necessary but far from sufficient for $\mathsf{E}\varepsilon_t^{2m} < \infty$, $m \geq 1$. In that case, the conditions for parameters allowing the existence of higher-order moments become more and more stringent for each even moment. It should be noted, however, that this property of the EGARCH process is not exclusively a consequence of its exponential structure. Processes with a moving average conditional heteroskedasticity (MACH) structure that the EGARCH process shares also have this property. For example, the simple MACH(1,1) process (1.1) with

$$h_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \beta h_{t-1}^2 \qquad (1.7)$$

has the same existence condition (1.4) for its unconditional $2m$-th moment as the EGARCH (1,1) process. Engle and Ng (1993) called (1.1) and (1.7) a VGARCH(1,1) model.

Remark 3. The moments μ_{2m} in (1.6) may be difficult to compute in the general case. We have to approximate the infinite product in (1.6) by a product with a finite number of terms. In some special cases, however, an analytical expression for the product is available. Assume, for example, that $z_t \sim$nid(0,1), $g(z_t) = \phi z_t$ and $|\beta| < 1$. Then

$$\begin{aligned}\prod_{i=1}^{\infty} \mathsf{E}\exp\{m\beta^{i-1}\phi z_t\} &= \prod_{i=1}^{\infty} \exp\{0.5(m\phi)^2\beta^{2(i-1)}\} \\ &= \exp\{0.5(m\phi)^2(1-\beta^2)^{-1}\}. \end{aligned} \tag{1.8}$$

1.4 The kurtosis of the family of EGARCH models

As discussed in the Introduction, we want to obtain a closed-form expression for the kurtosis of ε_t. Such an expression follows directly from Theorem 1.

Corollary 4.1 *Consider the EGARCH(1,1) model (1.1) and (1.2) and assume that* $\nu_4 < \infty$ *and, furthermore, that* $\mathsf{E}\exp\{g(z_t)\}$ *and* $\mathsf{E}\exp\{2g(z_t)\}$ *exist and the moment condition (1.4) holds. Then the kurtosis of* ε_t *exists and is given by*

$$\kappa_4 = \mathsf{E}\varepsilon_t^4/(\mathsf{E}\varepsilon_t^2)^2 = \kappa_4(z_t)\prod_{i=1}^{\infty}\frac{\mathsf{E}\exp\{2\beta^{i-1}g(z_t)\}}{[\mathsf{E}\exp\{\beta\{\beta^{i-1}g(z_t)\}]^2}, \tag{1.9}$$

where $\kappa_4(z_t) = \nu_4/\nu_2^2$ *is the kurtosis of* z_t.

Proof. Set $m = 1, 2$ in equation (1.6). ■

In particular, the kurtosis for the EGARCH(1,1) process (1.1) and (1.2), with $z_t \sim$nid(0,1) and $g(z_t) = \phi z_t + \psi(|z_t| - \sqrt{2/\pi})$ as in (1.3) is the special case of (1.9). It has the form

$$\begin{aligned}\kappa_4 = {} & 3\exp\left\{\frac{(\psi+\phi)^2}{1-\beta^2}\right\} \\ & \times \prod_{i=1}^{\infty}\frac{\Phi(2\beta^{i-1}(\psi+\phi)) + \exp\{-8\beta^{2(i-1)}\psi\phi\}\Phi(2\beta^{i-1}(\psi-\phi))}{[\Phi(\beta^{i-1}(\psi+\phi)) + \exp\{-2\beta^{2(i-1)}\psi\phi\}\Phi(\beta^{i-1}(\psi-\phi))]^2},\end{aligned} \tag{1.10}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. This result follows from equations (A.11) and (A.12) in the Appendix. Setting $\psi = 0$ in (1.10) yields a simple formula:

$$\kappa_4 = 3\exp\{\phi^2(1-\beta^2)^{-1}\}. \tag{1.11}$$

This is because the product

$$\frac{\prod_{i=1}^{M}[\Phi(2\beta^{i-1}\phi)+\Phi(-2\beta^{i-1}\phi)]}{[\Phi(\beta^{i-1}\phi)+\Phi(-\beta^{i-1}\phi)]^2}=1$$

for any $M \geq 1$ as $\Phi(-x)+\Phi(x)=1$. On the other hand, for $\psi \neq 0$ and $\phi \neq 0$ this product does not converge to 1 as $M \to \infty$.

It follows from Corollary 4.1 that the kurtosis of ε_t for our family of EGARCH(1,1) models has the following property.

Corollary 4.2 *Consider the EGARCH(1,1) model (1.1) and (1.2) and assume that (1.4) holds. Then the kurtosis of ε_t satisfies the inequality*

$$\kappa_4 > \kappa_4(z_t), \tag{1.12}$$

where $\kappa_4(z_t)=\nu_4/\nu_2^2$ is the kurtosis of z_t.

Proof. Apply Jensen's inequality to (1.9). ∎

Thus the unconditional distribution of ε_t in the EGARCH(1,1) is always leptokurtic if the error distribution is not platykurtic.

1.5 The autocorrelation function of positive powers of absolute observations

As in stochastic volatility models (for a survey, see Ghysels, Harvey, and Renault, 1996), it is possible to derive the autocorrelation function for any $|\varepsilon_t|^{2m}$, $m > 0$, when $\{\varepsilon_t\}$ obeys an EGARCH(1,1) model (1.1) and (1.2). This autocorrelation function is obtained by applying Theorem 2 and a recursion that is somewhat different from the one required in the proof of Theorem 2. For notational simplicity, let $g = g(z_t)$. We have the following theorem.

Theorem 3 *Consider the family of EGARCH(1,1) models (1.1) and (1.2). Assume that $\nu_{4m} < \infty$, $\mathsf{E}|z_t|^{2m}\exp\{mg\} < \infty$ and $\mathsf{E}\exp\{2mg\} < \infty$ for $0 < m < \infty$ and that condition (1.4) holds. Then the autocorrelation function of $\{|\varepsilon_t|^{2m}\}$ when $\{\varepsilon_t\}$ is generated by any member of the EGARCH family* *defined by (1.1) and (1.2) equals*

$$\rho_n(m) = \Bigg(\nu_{2m}\mathsf{E}\,|z_t|^{2m}\exp\{m\beta^{n-1}g\}\prod_{i=1}^{n-1}\mathsf{E}\exp\{m\beta^{i-1}g\}$$
$$\times\prod_{i=1}^{\infty}\mathsf{E}\exp\{m(1+\beta^n)\beta^{i-1}g\}-\nu_{2m}^2\prod_{i=1}^{\infty}(\mathsf{E}\exp\{m\beta^{i-1}g\})^2\Bigg)\Bigg/$$
$$\left(\nu_{4m}\prod_{i=1}^{\infty}\mathsf{E}\exp\{2m\beta^{i-1}g\}-\nu_{2m}^2\prod_{i=1}^{\infty}(\mathsf{E}\exp\{m\beta^{i-1}g\})^2\right) \quad (1.13)$$

for $n \geq 1$.

Proof. See the Appendix.

The following special case is of interest.

Corollary 5.1 *Consider the asymmetric EGARCH (1,1) model (1.1) and (1.3) and assume* $z_t \sim nid(0,1)$. *Then the autocorrelation function of* $\{|\varepsilon_t|^{2m}\}$ *when* $\{\varepsilon_t\}$ *is generated by (1.1) and (1.3) has the form*

$$\rho_n(m) = \frac{A\exp\left\{\frac{m^2(\psi+\phi)^2(\beta^{2(n-1)}(\beta^2-1)/4+\beta^n)}{1-\beta^2}\right\}D(\cdot)\prod_{i=1}^{n-1}\Phi_{1i}\prod_{i=1}^{\infty}\Phi_{2i}-\prod_{i=1}^{\infty}\Phi_{1i}^2}{\frac{\pi^{1/2}\Gamma(2m+1/2)}{(\Gamma(m+1/2))^2}\exp\left\{\frac{m^2(\psi+\phi)^2}{1-\beta^2}\right\}\prod_{i=1}^{\infty}\Phi_{3i}-\prod_{i=1}^{\infty}\Phi_{1i}^2} \quad (1.14)$$

where

$$\begin{aligned}
A &= \frac{\Gamma(2m+1)}{2^{m+1/2}\Gamma(m+1/2)},\\
D(\cdot) &= D_{-(2m+1)}[-m\beta^{n-1}(\psi+\phi)]\\
&\quad+\exp\{-m^2\beta^{2(n-1)}\psi\phi\}D_{-(2m+1)}[-m\beta^{n-1}(\psi-\phi)],\\
\Phi_{1i} &= \Phi(m\beta^{i-1}(\psi+\phi))+\exp\{-2m^2\beta^{2(i-1)}\psi\phi\}\Phi(m\beta^{i-1}(\psi-\phi)),\\
\Phi_{2i} &= \Phi(m\beta^{i-1}(1+\beta^n)(\psi+\phi))\\
&\quad+\exp\{-2m^2\beta^{2(i-1)}(1+\beta^n)^2\psi\phi\}\Phi(m\beta^{i-1}(1+\beta^n)(\psi-\phi)),
\end{aligned}$$

and

$$\Phi_{3i} = \Phi(2m\beta^{i-1}(\psi+\phi))+\exp\{-8m^2\beta^{2(i-1)}\psi\phi\}\Phi(2m\beta^{i-1}(\psi-\phi)).$$

Furthermore, $\Phi(\cdot)$ *is the cumulative distribution function of the standard normal distribution, and*

$$D_{(-p)}[q] = \frac{\exp\{-q^2/4\}}{\Gamma(p)}\int_0^{\infty}x^{p-1}\exp\{-qx-x^2/2\}dx, \quad p>0,$$

is the parabolic cylinder function (Gradshteyn and Ryzhik, 1980), where $\Gamma(\cdot)$ *is the gamma function.*

Proof. See the Appendix.

If $\phi = 0$ or $\psi = 0$ in (1.3) and, in addition, the errors are standard normal, the resulting autocorrelation function becomes quite simple. These results are given in the following corollaries.

Corollary 5.2 *Consider the asymmetric EGARCH (1,1) model (1.1) and (1.3) and set* $\phi = 0$*. The n-th order autocorrelation of* $\{|\varepsilon_t|^{2m}\}$ *has the form*

$$\rho_n(m) = \frac{A\exp\left\{\dfrac{m^2\psi^2(\beta^{2(n-1)}(\beta^2-1)/4+\beta^n)}{1-\beta^2}\right\} D(\cdot)\prod_{i=1}^{n-1}\Phi_{1i}\prod_{i=1}^{\infty}\Phi_{2i} - \prod_{i=1}^{\infty}\Phi_{1i}^2}{\dfrac{\pi^{1/2}\Gamma(2m+1/2)}{2(\Gamma(m+1/2))^2}\exp\left\{\dfrac{m^2\psi^2}{1-\beta^2}\right\}\prod_{i=1}^{\infty}\Phi_{3i} - \prod_{i=1}^{\infty}\Phi_{1i}^2}, \tag{1.15}$$

where

$$\begin{aligned} A &= \frac{\Gamma(2m+1)}{2^{m-1/2}\Gamma(m+1/2)}, \\ D(\cdot) &= D_{-(2m+1)}[-m\beta^{n-1}\psi], \\ \Phi_{1i} &= \Phi(m\beta^{i-1}\psi), \\ \Phi_{2i} &= \Phi(m\beta^{i-1}(1+\beta^n)\psi), \end{aligned}$$

and

$$\Phi_{3i} = \Phi(2m\beta^{i-1}(\psi+\phi)).$$

Proof. Setting $\phi = 0$ in equation (1.14) and simplifying the expression yields equation (1.15). ■

Corollary 5.3 *Consider the asymmetric EGARCH (1,1) model (1.1) with (1.3) and set* $\psi = 0$*. The n-th order autocorrelation of* $\{|\varepsilon_t|^{2m}\}$ *is given by*

$$\rho_n(m) = \frac{A\exp\left\{\dfrac{m^2\phi^2(\beta^{2(n-1)}(\beta^2-1)/4+\beta^n)}{1-\beta^2}\right\} D(\cdot) - 1}{\dfrac{\pi^{1/2}\Gamma(2m+1/2)}{(\Gamma(m+1/2))^2}\exp\left\{\dfrac{m^2\phi^2}{1-\beta^2}\right\} - 1}, \tag{1.16}$$

where

$$\begin{aligned} A &= \frac{\Gamma(2m+1)}{2^{m+1/2}\Gamma(m+1/2)}, \\ D(\cdot) &= D_{-(2m+1)}[-m\beta^{n-1}\phi] + D_{-(2m+1)}[m\beta^{n-1}\phi]. \end{aligned}$$

The n-th order autocorrelation of the squared observations ($m = 1$) *has the simple form*

$$\rho_n(1) = \frac{(1+\phi^2\beta^{2(n-1)})\exp\{\phi^2\beta^n(1-\beta^2)^{-1}\}-1}{3\exp\{\phi^2(1-\beta^2)^{-1}\}-1}. \tag{1.17}$$

The corresponding autocorrelation for the n-th order of the absolute-valued observations ($m = 1/2$) *equals*

$$\begin{aligned}\rho_n(1/2) = \Big(\Big\{\left(\sqrt{2\pi}/4\right)\phi\beta^{n-1}\left[2\Phi\left(\phi\beta^{n-1}/2\right)-1\right] \\ +\exp\left\{-\phi^2\beta^{2(n-1)}/8\right\}\Big\}\exp\left\{\left(\phi^2\beta^n/4\right)\left(1-\beta^2\right)^{-1}\right\}-1\Big)\Big/ \\ \left(\Gamma(3/2)\sqrt{\pi}\exp\left\{(\phi^2/4)(1-\beta^2)^{-1}\right\}-1\right).\end{aligned} \tag{1.18}$$

Proof. See the Appendix.

To illustrate the preceding theory, consider the case $0 < \beta < 1$. The autocorrelation function (16) then appears to have the property that the decay rate $0 < \rho_{n+1}(m)/\rho_n(m) < \beta$ for any finite n and $\lim_{n\to\infty}(\rho_{n+1}(m)/\rho_n(m)) = \beta$. For the special case (1.17) this can be shown analytically, but in the general case it is just a conjecture based on numerical calculations. Results of such calculations are shown in Table 1.1, in which we report the starting decay rate $\rho_2(m)/\rho_1(m)$ and the ratio $\rho_{51}(m)/\rho_{50}(m)$ for a number of EGARCH models and $m = 1$, $1/2$, $1/4$, assuming normally distributed errors. Table 1.1 also contains the kurtosis and the first-order autocorrelations $\rho_1(m)$, $m = 1$, $1/2$, $1/4$, for the parameter combinations considered. When ϕ decreases (becomes more negative) and/or ψ increases the kurtosis of the EGARCH(1,1) process increases. High values of the kurtosis seem to be combined with relatively rapidly decaying autocorrelations of the squares. As for the absolute values or their square roots, their first-order autocorrelations are higher than those of the squared observations. When the kurtosis of the EGARCH model is high, $\rho_1(1/4)$ is high and clearly higher than $\rho_1(1/2)$. For the lowest values of kurtosis in the table, $\rho_1(1/2) > \rho_1(1/4)$.

The situation is further illustrated in Figure 1.1 which also contains a comparison with a standard GARCH(1,1) model. The GARCH(1,1) model is defined by (1.1) and

$$h_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2, \tag{1.19}$$

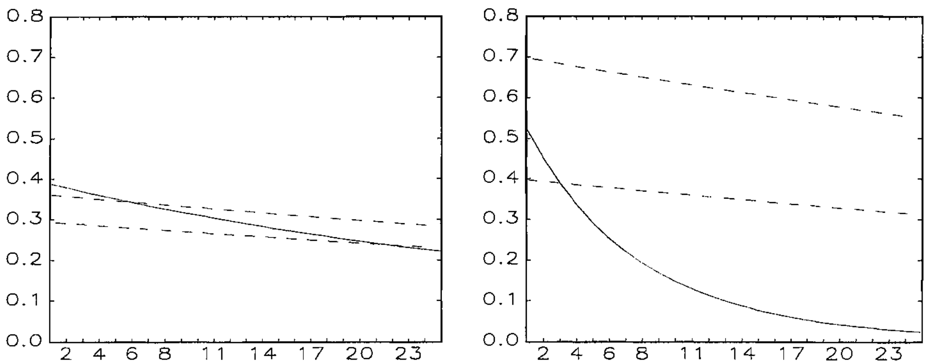
where $\alpha_0 > 0$, $\alpha_1, \beta_1 \geq 0$. In the figure, $\alpha_1 + \beta_1 = 0.99$, which is the exponential decay rate of the autocorrelations of $\{\varepsilon_t^2\}$ when $\mathsf{E}z_t^2 = 1$. The figure depicts these autocorrelations for both the GARCH(1,1) and the symmetric

Table 1.1 Kurtosis κ_4, first-order autocorrelation of $\left\{|\varepsilon_t|^{2m}\right\}$, when $m = 1, 1/2, 1/4$ and two local decay rates ρ_2/ρ_1 and ρ_{51}/ρ_{50} for a number of EGARCH(1,1) models with standard normal errors and $\beta = 0.99$.

$\psi = \mathbf{0}$		$\phi = 0$			-0.2			-0.4	
	$m = 1$	0.5	0.25	$m = 1$	0.5	0.25	$m = 1$	0.5	0.25
κ_4					22.4			9310	
ρ_1				0.309	0.409	0.393	0.357	0.602	0.682
$\frac{\rho_2}{\rho_1}$				0.976	0.987	0.990	0.921	0.977	0.987
$\frac{\rho_{51}}{\rho_{50}}$				0.983	0.988	0.990	0.951	0.983	0.988
$\psi = \mathbf{0.3}$									
κ_4		19.2			483			6×10^6	
ρ_1	0.385	0.455	0.421	0.417	0.587	0.591	0.422	0.658	0.751
$\frac{\rho_2}{\rho_1}$	0.975	0.987	0.989	0.940	0.982	0.988	0.840	0.966	0.985
$\frac{\rho_{51}}{\rho_{50}}$	0.982	0.988	0.990	0.965	0.986	0.989	0.911	0.977	0.987
$\psi = \mathbf{0.6}$									
κ_4		1×10^4			2×10^6			4×10^{12}	
ρ_1	0.497	0.686	0.726	0.485	0.699	0.765	0.457	0.694	0.818
$\frac{\rho_2}{\rho_1}$	0.907	0.976	0.987	0.844	0.968	0.986	0.699	0.942	0.982
$\frac{\rho_{51}}{\rho_{50}}$	0.946	0.983	0.988	0.916	0.979	0.988	0.834	0.966	0.986

($\phi = 0$) EGARCH(1,1) model when the error distribution is standard normal. Furthermore, the autocorrelation function of $\{\ln \varepsilon_t^2\}$ for the EGARCH(1,1) process, given in Lemma 2 of Chapter 2, is included in the figure also. This function also decays exponentially from the first lag, the decay rate $\beta = 0.99$ being equal to that in the GARCH(1,1) process. The autocorrelations are shown for two kurtosis values, $\kappa_4 = 20, 10^6$. It is seen that for $\kappa_4 = 20$, the autocorrelations of the squared observations for the EGARCH process decay somewhat faster in the beginning than those for the standard GARCH model but are of the same order of magnitude. The first autocorrelations of $\{\ln \varepsilon_t^2\}$ are somewhat lower than those of $\{\varepsilon_t^2\}$. For $\kappa_4 = 10^6$, the autocorrelation function of the EGARCH process has a radically different shape: the decay rate is now much faster in the beginning. On the other hand, the corresponding autocorrelation function of the GARCH model is little affected by this remarkable change in the kurtosis of the process. Finally, it is seen that the level of the autocorrelations of $\{\ln \varepsilon_t^2\}$ in the symmetric EGARCH model can be very sensitive to changes in the kurtosis or, in other words, in ψ, when β remains unchanged in (1.3).

Figure 1.1 The first 25 autocorrelations of squared observations and the logarithms of squared observations for an EGARCH(1,1) model (1.1) and (1.3) with $\beta = 0.99$, $\phi = 0$, and (left panel, $\kappa_4 = 20$) $\psi = 0.3032$, (right panel, $\kappa_4 = 10^6$) $\psi = 0.7225$, and of squared observations for a standard GARCH(1,1) model (1.1) and (1.19) with $\alpha_1 + \beta_1 = 0.99$, and (left panel, $\kappa_4 = 20$) $\alpha_1 = 0.09196$, (right panel, $\kappa_4 = 10^6$) $\alpha_1 = 0.09975$. Both models have standard normal independent errors. EGARCH(1,1), squared observations: solid line; logarithms of squared observations: dashed-dotted line; and GARCH(1,1), squared observations: dashed line.



For some distributions of $\{z_t\}$ such as the Student t with a finite number of degrees of freedom, $\mathsf{E}\exp\{g\}$ typically has no finite unconditional moments. To improve the fit in applications when the normality assumption is unsatisfactory, Nelson (1991) proposed using more flexible parametric families of distributions such as the generalized error distribution (GED) (Harvey, 1981, Box and Tiao, 1973). The density of a GED random variable is given by

$$f(z) = \frac{v\exp\{-0.5\,|z/\lambda|^v\}}{\lambda 2^{(1+1/v)}\Gamma(1/v)}, \quad -\infty < z < \infty < z < \infty, \quad 0 < v \leq \infty, \quad (1.20)$$

where $\lambda = [2^{-2/v}\Gamma(1/v)/\Gamma(3/v)]^{1/2}$ and v is a tail-thickness parameter (for discussion, see Nelson, 1991, and Bolleslev, Engle, and Nelson, 1994). The following corollary gives the autocorrelation function of $\{|\varepsilon_t|^{2m}\}$ when $z_t \sim$GED(v).

Corollary 5.4 *Consider the asymmetric EGARCH (1,1) model (1.1) with (1.3) and assume $z_t \sim$iid GED(v) with* $\mathsf{E}z_t = 0$, $\mathsf{E}z_t^2 = 1$ and $v > 1$. *Then* *the autocorrelation function of* $\{|\varepsilon_t|^{2m}\}$, *when* $\{\varepsilon_t\}$ *is generated by (1.1) and (1.3), has the form*

$$\rho_n(m) = \frac{\dfrac{\lambda^{2m-1}\Gamma(1/v)}{\Gamma((2m+1)/v)} S_n \prod_{i=1}^{n-1} S_{1i} \prod_{i=1}^{\infty} S_{2i} - \prod_{i=1}^{\infty} S_{1i}^2}{\dfrac{\Gamma(1/v)\Gamma((4m+1)/v))}{\lambda\left(\Gamma((2m+1)/v)\right)^2} \prod_{i=1}^{\infty} S_{3i} - \prod_{i=1}^{\infty} S_{1i}^2}, \tag{1.21}$$

where

$$\begin{aligned}
S_n &= \sum_{k=0}^{\infty} (m\beta^{n-1})^k \omega_k(v,m),\\
S_{1i} &= \sum_{k=0}^{\infty} (m\beta^{i-1})^k \omega_k(v,0),\\
S_{2i} &= \sum_{k=0}^{\infty} [m(1+\beta^n)\beta^{i-1}]^k \omega_k(v,0),\\
S_{3i} &= \sum_{k=0}^{\infty} (2m\beta^{i-1})^k \omega_k(v,0),
\end{aligned}$$

and for $v > 1$ *and* $m > 0$,

$$\omega_k(v,m) = (2^{1/v}\lambda)^k[(\psi-\phi)^k + (\psi+\phi)^k]\frac{\Gamma((k+2m+1)/v)}{2\Gamma(1/v)\Gamma(k+1)}, \quad k = 0,1,2,\cdots. \tag{1.22}$$

Proof. See the Appendix.

Remark 4. When $v = 2$ in (1.20), z_t has a standard normal distribution. By Formula 3.381 #4 of Gradshteyn and Ryzhik (1980) it can be seen that

$$\lim_{K\to\infty} \sum_{k=0}^{K} b^k \omega_k(2,0) = \exp\{b^2(\psi-\phi)^2/2\}\Phi(b(\psi-\phi)) + \exp\{b^2(\psi+\phi)^2/2\}\Phi(b(\psi+\phi)). \tag{1.23}$$

When $v > 1$ but $v \neq 2$, the limit (1.23) does not have an analytic expression. On the other hand, if $v \leq 1$, $\mathsf{E}\exp\{bg\}$ in expression (1.13) does not exist unless restrictions on ψ and ϕ are proposed. For a detailed discussion, see Nelson (1991).

Remark 5. Autocorrelation functions for logarithmic GARCH models of types (1.1) and (1.4) can also be obtained from formula (1.13), but we omit them from this presentation.

1.6 Conclusions

In this paper we have presented a characterization of the moment structure of the first-order exponential GARCH model, which is the most widely applied EGARCH model. The results of the paper are useful, for example, if we want to compare the EGARCH model with the standard GARCH model. They reveal certain differences in the moment structure between these models. Although the autocorrelations of the squared observations decay exponentially in the first-order GARCH model, the decay rate is not exponential in the first-order EGARCH model. Furthermore, in the case of the standard GARCH model, a high value of the kurtosis implies a slowly decaying autocorrelation function. This relationship does not hold for the EGARCH model class as a whole, although it may be valid for many individual members of the class.

The results of the paper only cover a family of EGARCH(1,1) models. Working out the moment structure of higher order EGARCH models is a more complicated task and is left for further work.

Bibliography

[1] An, H. and M. Chen (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica* **8**, 505-510.

[2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

[3] Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of Time Series Analysis* **9**, 121-131.

[4] Bollerslev, T., R.F. Engle and D.B. Nelson (1994). ARCH models. In: R.F. Engle and D.L. McFadden, eds. *Handbook of Econometrics*, Vol. **4**, 2959-3038. Amsterdam: Elsevier.

[5] Box, G. E. P. and G. C. Tiao (1973). *Bayesian Inference in Statistical Analysis.* Reading, Massachusetts: Addison-Wesley.

[6] Breidt, F.J., N. Crato and P. de Lima (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* **83**, 325-348.

[7] Carrasco, M. and X. Chen (2002). Mixing and moment properties of various GARCH and Stochastic Volatility models. *Econometric Theory* **18**, 17-39.

[8] Demos, A. (2000). The autocorrelation function of conditionally heteroskedastic in mean models. Working Paper, Athens University of Economics and Business.

[9] Ding, Z., C.W.J. Granger and R.F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1**, 83-106.

[10] Engle, R.F. and V.K. Ng (1993). Measuring and testing the impact of news on volatility. *Journal of Finance* **48**, 1749-1778.

[11] Geweke, J. (1986). Modelling persistence of conditional variances: Comment. *Econometric Reviews* **5**, 57-61.

[12] Ghysels, E., A.C. Harvey and E. Renault (1996). Stochastic Volatility. In: G. S. Maddala and C R. Rao, eds. *Handbook of Statistics*, Vol. **14**, 119-189. Amsterdam: Elsevier.

[13] Giraitis, L., P. Kokoszka and R. Leipus (2000). Stationary ARCH models: Dependence structure and central limit theorem. *Econometric Theory* **16**, 3-22.

[14] Gradshteyn, I. S. and I. M. Ryzhik (1980). *Tables of Integrals, Series and Products*. New York: Academic Press.

[15] Harvey, A. C. (1981). *The Econometric Analysis of Time Series*. Oxford: Philip Allan.

[16] He, C. and T. Teräsvirta (1999a). Properties of moments of a family of GARCH processes. *Journal of Econometrics* **92**, 173-192.

[17] He, C. and T. Teräsvirta (1999b). Fourth moment structure of the GARCH(p,q) process. *Econometric Theory* **15**, 824-846.

[18] Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* **39**, 71-104.

[19] Karanasos, M. (1999). The second moment and the autocovariance function of the squared errors of the GARCH model. *Journal of Econometrics* **90**, 63-76.

[20] Ling, S. and W.K. Li (1997). On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *Journal of the American Statistical Association* **92**, 1184-1194.

[21] Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**, 347-370.

[22] Pantula, S.G. (1986). Modelling persistence of conditional variances: Comment. *Econometric Reviews* **5**, 71-74.

Appendix A

Proofs

A.1 Theorem 1

Note that $|\beta| < 1$ implies that the infinite sum of the squared sequence $\{\beta_k\}$ defined in (2.1) of Nelson (1991) is finite. It follows directly from Theorem 2.1 of Nelson (1991) that when $\mathrm{var}(g(z))$ is finite, $\{\varepsilon_t\}$ defined in (1.1) and (1.2) is strictly stationary if and only if $|\beta| < 1$. ■

A.2 Theorem 2

When $m = 2$, $\mathsf{E}\exp\{2g(z_t)\} < \infty$ implies $\mathsf{E}(g(z_t))^2 < \infty$. It follows from Theorem 1 that under (1.4), $\{\varepsilon_t\}$ defined in (1.1) and (1.2) is strictly stationary. Then raising both sides of equation (1.5) to the power $m > 0$ yields

$$h_t^{2m} = [\exp\{m\alpha_0\}\exp\{mg(z_{t-1})\}]\, h_{t-1}^{2m\beta}. \tag{A.1}$$

Applying (1.5) to the right-hand side of (A.1) one obtains

$$h_t^{2m} = (\exp\{m\alpha_0(1+\beta)\}\exp\{m\,[g(z_{t-1}) + \beta g(z_{t-2})]\})h_{t-2}^{2m\beta^2}.$$

Further recursion gives

$$h_t^{2m} = \left[\exp\left\{m\alpha_0\sum_{i=0}^{n-1}\beta^{i-1}\right\}\exp\left\{m\sum_{i=1}^{n}\beta^{i-1}g(z_{t-i})\right\}\right] h_{t-n}^{2m\beta^n}. \tag{A.2}$$

The strict stationarity condition $|\beta| < 1$ and letting $n \to \infty$ in (A.2) lead to

$$h_t^{2m} = \exp\{m\alpha_0(1-\beta)^{-1}\}\exp\left\{m\sum_{i=1}^{\infty}\beta^{i-1}g(z_{t-i})\right\}. \tag{A.3}$$

Next define

$$a_i = \exp\{m\beta^{i-1}g(z_{t-i})\}.$$

Assume furthermore that $\mathsf{E}\exp\{mg(z_t)\} < \infty$ for $m > 0$. Then it follows that $\mathsf{E}a_i < \infty$ for $i = 0, 1, 2, \cdots$. A sufficient condition for $\prod_{i=1}^{\infty} \mathsf{E}a_i < \infty$ is that $\sum_{i=1}^{\infty} |\mathsf{E}a_i - 1| < \infty$. To show the validity of this condition, expand a_i into a Taylor series around $\beta^{i-1} = 0$. This yields

$$a_i = 1 + mg(z_t)\beta^{i-1} + O(|\beta^{i-1}|^2).$$

Then

$$\begin{aligned}
\sum_{i=1}^{\infty} |\mathsf{E}a_i - 1| &= \sum_{i=1}^{\infty} \left| m\beta^{i-1}\mathsf{E}g(z_t) + O(|\beta^{i-1}|^2) \right| \\
&\leq m\mathsf{E}g(z_t)\sum_{i=1}^{\infty} |\beta^{i-1}| + \sum_{i=1}^{\infty} \left| O(|\beta^{i-1}|^2) \right| < \infty
\end{aligned}$$

because $|\beta| < 1$. Thus, from (A.3)

$$\begin{aligned}
\mu_{2m} = \mathsf{E}\,|\varepsilon_t|^{2m} &= \nu_{2m}\exp\{m\alpha_0(1-\beta)^{-1}\} \\
&\times \prod_{i=1}^{\infty} \mathsf{E}\exp\{m\beta^{i-1}g(z_t)\}. \qquad \text{(A.4)}
\end{aligned}$$

■

A.3 Theorem 3

Applying (A.2) gives

$$\begin{aligned}
h_t^{2m} h_{t-n}^{2m} &= \exp\{m\alpha_0(1-\beta^n)(1-\beta)^{-1}\} \\
&\times \left[\prod_{i=1}^{n} \exp\{m\beta^{i-1}g(z_{t-i})\} \right] h_{t-n}^{2m(\beta^n+1)}. \qquad \text{(A.5)}
\end{aligned}$$

Multiplying both sides of (A.5) by $z_t^{2m} z_{t-n}^{2m}$, applying condition (1.4) and taking expectations yields

$$\begin{aligned}
\mathsf{E}(\varepsilon_t^{2m}\varepsilon_{t-n}^{2m}) &= \exp\{m\alpha_0(1-\beta^n)(1-\beta)^{-1}\}\mathsf{E}z_t^{2m}\mathsf{E}z_{t-n}^{2m}\exp\{m\beta^{n-1}g(z_{t-n})\} \\
&\times \left[\prod_{i=1}^{n-1} \mathsf{E}\exp\{m\beta^{i-1}g(z_{t-i})\} \right] \mathsf{E}h_{t-n}^{2m(\beta^n+1)} \\
&= \nu_{2m}\exp\{m\alpha_0(1-\beta^n)(1-\beta)^{-1}\}\mathsf{E}(z_t^{2m}\exp\{m\beta^{n-1}g(z_t)\}) \\
&\times \left[\prod_{i=1}^{n-1} \mathsf{E}\exp\{m\beta^{i-1}g(z_t)\} \right] \mathsf{E}h_t^{2m(\beta^n+1)}. \qquad \text{(A.6)}
\end{aligned}$$

Furthermore, from Theorem 2 it follows that

$$\mathsf{E}h_t^{2m(\beta^n+1)} = \exp\{m\alpha_0(1+\beta^n)(1-\beta)^{-1}\}\prod_{i=1}^{\infty}\mathsf{E}\exp\{m(\beta^n+1)\beta^{i-1}g(z_t)\}. \quad \text{(A.7)}$$

Expression (13) is obtained by inserting (A.7) in (A.6), defining $\rho_n(m) = (\mathsf{E}(\varepsilon_t^{2m}\varepsilon_{t-n}^{2m}) - (\mathsf{E}\varepsilon_t^{2m})^2)/(\mathsf{E}\varepsilon_t^{4m} - (\mathsf{E}\varepsilon_t^{2m})^2)$, and applying Theorem 2. ■

A.4 Lemma 1

Let $z_t \sim nid(0,1)$ and set $g = g(z_t)$. For any finite, real number b and $a > 0$,

$$\begin{aligned}\mathsf{E}\,|z_t|^a \exp\{bg\} =& (2\pi)^{-1/2}\Gamma(a+1)\exp\{-b\psi(2/\pi)^{1/2}\} \\ & \times (D_{-(a+1)}[-b(\psi+\phi)] + \exp\{-b^2\psi\phi\}D_{-(a+1)}[-b(\psi-\phi)]). \quad \text{(A.8)}\end{aligned}$$

Proof. Nelson (1991) gives the expression for $\mathsf{E}z_t^a\exp\{bg\}$ when a is a positive integer. It follows from Formula 3.462 #1 in Gradshteyn and Ryzhik (1980) that (A.8) holds when a is a positive number. The finiteness of (A.8) follows as a special case from the proof of Theorem A1.2 of Nelson (1991). ■

A.5 Corollary 5.1

Note that under condition (1.4),

$$\begin{aligned}\prod_{i=1}^{\infty}\exp\{m^2\beta^{2(i-1)}(\phi+\psi)^2\} &= \exp\{m^2(\phi+\psi)^2(1-\beta^2)^{-1}\}, \\ \prod_{i=1}^{\infty}\exp\{-2m\beta^{i-1}\psi\sqrt{2/\pi}\} &= \exp\{-2m\psi\sqrt{2/\pi}(1-\beta)^{-1}\}.\end{aligned}$$

It follows from Theorem A1.1 of Nelson (1991) and Lemma 1 that

$$\begin{aligned}\mathsf{E}\,|z_t|^{2m}\exp\{m\beta^{n-1}g\} = & \sqrt{2/\pi}\Gamma(2m+1)\exp\left\{-\sqrt{2/\pi}m\psi\beta^{n-1}\right\} \\ & \times \exp\{(1/4)m^2(\phi+\psi)^2\beta^{2(n-1)}\} \\ & \times \{D_{-(2m+1)}[-m\beta^{n-1}(\phi+\psi)] + \exp\{-m^2\beta^{2(n-1)}\phi\psi\} \\ & \times D_{-(2m+1)}[-m\beta^{n-1}(\psi-\phi)]\}, \quad \text{(A.9)}\end{aligned}$$

$$\begin{aligned}\prod_{i=1}^{n-1} \mathsf{E}\exp\{m\beta^{i-1}g\} &= \exp\left\{-\sqrt{2/\pi}m\psi(1-\beta^{n-1})(1-\beta)^{-1}\right\} \\ &\times \exp\{(1/2)m^2(\psi+\phi)^2(1-\beta^{2(n-1)})(1-\beta^2)^{-1}\} \\ &\times \prod_{i=1}^{n-1}[\Phi(m\beta^{i-1}(\psi+\phi)) \\ &+ \exp\{-2m^2\beta^{2(i-1)}\psi\phi\}\Phi(m\beta^{i-1}(\psi-\phi))], \quad \text{(A.10)}\end{aligned}$$

$$\begin{aligned}\prod_{i=1}^{\infty} \mathsf{E}\exp\{m(\beta^n+1)\beta^{i-1}g\} &= \exp\left\{-\sqrt{2/\pi}m\psi(1+\beta^{n})(1-\beta)^{-1}\right\} \\ &\times \exp\{(1/2)m^2(\psi+\phi)^2(1+\beta^{n})^2(1-\beta^2)^{-1}\} \\ &\times \prod_{i=1}^{\infty}[\Phi(m\beta^{i-1}(1+\beta^n)(\psi+\phi)) \\ &+ \exp\{-2m^2\beta^{2(i-1)}(1+\beta^n)^2\psi\phi\} \\ &\times\Phi(m\beta^{i-1}(1+\beta^n)(\psi-\phi))], \quad \text{(A.11)}\end{aligned}$$

$$\begin{aligned}\prod_{i=1}^{\infty} (\mathsf{E}\exp\{m\beta^{i-1}g\})^2 &= \exp\left\{-2\sqrt{2/\pi}m\psi(1-\beta)^{-1}\right\} \\ &\times \exp\{m^2(\psi+\phi)^2(1-\beta^2)^{-1}\} \\ &\times \prod_{i=1}^{\infty}[\Phi(m\beta^{i-1}(\psi+\phi)) \\ &+ \exp\{-2m^2\beta^{2(i-1)}\psi\phi\}\Phi(m\beta^{i-1}(\psi-\phi))]^2, \quad \text{(A.12)}\end{aligned}$$

$$\begin{aligned}\prod_{i=1}^{\infty} \mathsf{E}\exp\{2m\beta^{i-1}g\} &= \exp\left\{-2\sqrt{2/\pi}m\psi(1-\beta)^{-1}\right\} \\ &\times \exp\{2m^2(\psi+\phi)^2(1-\beta^2)^{-1}\} \\ &\times \prod_{i=1}^{\infty}[\Phi(2m\beta^{i-1}(\psi+\phi)) \\ &+ \exp\{-8m^2\beta^{2(i-1)}\psi\phi\}\Phi(2m\beta^{i-1}(\psi-\phi))], \quad \text{(A.13)}\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and $D_q[\cdot]$ is the parabolic cylinder function. On the other hand, for any $m > 0$,

$$\mathsf{E}\,|z_t|^{2m} = \pi^{-1/2}2^m\Gamma(m+1/2). \quad \text{(A.14)}$$

Inserting (A.9)-(A.14) into (1.13), manipulating the equation further, and defining $D(\cdot)$ and Φ_{ji}, $j = 1, 2, 3$, through (A.9)-(A.13), respectively, gives (1.14). ■

A.6 Corollary 5.3

It follows from (1.14) and the fact $\Phi(x) + \Phi(-x) = 1$ that (1.16) holds. Applying

$$\int_0^\infty x^2 e^{-qx-x^2/2} dx = -q + \sqrt{2\pi}(1+q^2)\Phi(-q)e^{q^2/2}$$

and

$$\int_0^\infty x e^{-qx-x^2/2} dx = 1 - \sqrt{2\pi} q \Phi(-q) e^{q^2/2},$$

where q is a finite real number, in turn to equation (1.16) yields expressions (1.17) and (1.18). ■

A.7 Lemma 2

Let z_t ~iid $GED(v)$ such that $\mathsf{E}z_t = 0$, $\mathsf{E}z_t^2 = 1$, and $v > 1$. For any $|b| < \infty$ and $a \geq 0$,

$$\begin{aligned}\mathsf{E}\,|z_t|^a \exp\{bg\} &= 2^{a/v}\lambda^a \exp\{-b\psi\nu_1\} \\ &\times \sum_{k=0}^{\infty} (2^{1/v}\lambda b)^k [(\psi-\phi)^k + (\psi+\phi)^k] \frac{\Gamma((k+a+1)/v)}{2\Gamma(1/v)\Gamma(k+1)}, \quad \text{(A.15)}\end{aligned}$$

where $\nu_1 = \mathsf{E}\,|z_t|$.

Proof. Nelson (1991) considers the expectation $\mathsf{E}z_t^a \exp\{bg\}$ when a is a nonnegative integer. We extend his result to the case of real $a \geq 0$. It follows from (1.20) that

$$\begin{aligned}\mathsf{E}\,|z_t|^a \exp\{bg\} &= 2^{(a/v)-1}\lambda^a \Gamma(1/v) \exp\{-b\psi\nu_1\} \\ &\times \int_0^\infty y^{[(a+1)/v]-1} \exp\{-y\} [\exp\{b\lambda(\psi-\phi)(2y)^{1/v}\} \\ &+ \exp\{b\lambda(\psi+\phi)(2y)^{1/v}\}] dy. \quad \text{(A.16)}\end{aligned}$$

Applying $\exp\{x\} = \sum_{k=0}^{\infty} (x^k/\Gamma(k+1))$, $|x| < \infty$, to $\exp\{b\lambda(\psi \pm \phi)(2y)^{1/v}\}$ in (A.16) gives

$$\begin{aligned}\mathsf{E}\,|z_t|^a \exp\{bg\} &= 2^{(a/v)-1}\lambda^a \Gamma(1/v) \exp\{-b\psi\nu_1\} \\ \times \int_0^\infty \sum_{k=0}^{\infty} &(2^{1/v}\lambda b)^k [(\psi-\phi)^k + (\psi+\phi)^k] \frac{y^{[(a+1+k)/v]-1}\exp\{-y\}}{\Gamma(k+1)} dy. \quad \text{(A.17)}\end{aligned}$$

It follows from Theorem A1.2 of Nelson (1991) that the order of summation and integration in (A.17) can be interchanged. Applying Formula 3.381 #4 in Gradshteyn and Ryzhik (1980) to (A.17) then gives (A.15). ■

A.8 Corollary 5.4

Note that if $z_t \sim \mathrm{GED}(\upsilon)$, then, for any $a > 0$,

$$\mathsf{E}\,|z_t|^a = 2^{a/\upsilon}\lambda\frac{\Gamma((a+1)/\upsilon)}{\Gamma(1/\upsilon)}. \tag{A.18}$$

It follows from Lemma 2 that

$$\begin{aligned}\mathsf{E}\,|z_t|^{2m}\exp\{m\beta^{n-1}g\} &= 2^{2m/\upsilon}\lambda^{2m}\exp\{-m\beta^{n-1}\psi\nu_1\} \\ &\quad\times\sum_{k=0}^{\infty}(m\beta^{n-1})^k\omega_k(\upsilon,m),\end{aligned} \tag{A.19}$$

$$\begin{aligned}\prod_{i=1}^{n-1}\mathsf{E}\exp\{m\beta^{i-1}g\} &= \exp\{-m\psi\nu_1(1-\beta^{n-1})(1-\beta)^{-1}\} \\ &\quad\times\prod_{i=1}^{n-1}\left[\sum_{k=0}^{\infty}(m\beta^{i-1})^k\omega_k(\upsilon,0)\right],\end{aligned} \tag{A.20}$$

$$\begin{aligned}\prod_{i=1}^{\infty}\mathsf{E}\exp\{m(1+\beta^n)\beta^{i-1}g\} &= \exp\{-m\psi\nu_1(1+\beta^n)(1-\beta)^{-1}\}\times \\ &\quad\prod_{i=1}^{\infty}\left[\sum_{k=0}^{\infty}(m(1+\beta^n)\beta^{i-1})^k\omega_k(\upsilon,0)\right]\end{aligned} \tag{A.21}$$

$$\begin{aligned}\prod_{i=1}^{\infty}(\mathsf{E}\exp\{m\beta^{i-1}g\})^2 &= \exp\{-2m\psi\nu_1(1-\beta)^{-1}\} \\ &\quad\times\prod_{i=1}^{\infty}\left[\sum_{k=0}^{\infty}(m\beta^{i-1})^k\omega_k(\upsilon,0)\right]^2,\end{aligned} \tag{A.22}$$

$$\begin{aligned}\prod_{i=1}^{\infty}\mathsf{E}\exp\{2m\beta^{i-1}g\} &= \exp\{-2m\psi\nu_1(1-\beta)^{-1}\} \\ &\quad\times\prod_{i=1}^{\infty}\left[\sum_{k=0}^{\infty}(2m\beta^{i-1})^k\omega_k(\upsilon,0)\right],\end{aligned} \tag{A.23}$$

where $\omega_k(\upsilon, m)$ is defined in (1.22). Applying (A.18)-(A.23) to equation (1.13), manipulating the equation further, and defining S_n and S_{ji}, $j = 1, 2, 3$, through (A.19)-(A.23), respectively, gives (1.21). ■

Chapter 2

Higher-order dependence in the general power ARCH process and the role of the power parameter

2.1 Introduction

In a recent paper, Ding et al. (1993) introduced a class of autoregressive conditional heteroskedastic models called Asymmetric Power Autoregressive Conditional Heteroskedastic (A-PARCH) models. The authors showed that this class contains as special cases a large number of well-known ARCH and GARCH models. The A-PARCH model contains a particular power parameter that makes the conditional variance equation nonlinear in parameters. Among other things, Ding, Granger and Engle showed that by letting the power parameter approach zero, the A-PARCH family of models also includes the logarithmic GARCH model as a special case. Hentschel (1995) defined a slightly extended A-PARCH model and showed that after this extension, the A-PARCH model also contains the exponential GARCH (EGARCH) model of Nelson (1991) as a special case as the power parameter approaches zero. Allowing this to happen in a general A-PARCH model forms a starting-point for our investigation.

A notable feature of the A-PARCH model is that, due to its parameterization, it is only possible to find analytically certain fractional moments of the absolute values of the original process related to the power parameter. Expressions for such moments were derived in He and Teräsvirta (1999*c*). In this paper we first define a slight generalization of the class of EGARCH models. Then we derive the autocorrelation function of squared and logarithmed observations for this class of models. For Nelson's EGARCH model it is possible to reconcile our results with those in Breidt, Crato and de Lima (1998). Furthermore, we show that this autocorrelation function follows as a limiting case from the autocorrelation function of some fractional powers of the absolute values of the original observations.

On the other hand, if we want to derive the autocorrelation function of squares of the original observations and not their logarithms for the EGARCH model then the techniques applied in this paper do not apply. The solution to that problem can be found in Chapter 1. The autocorrelation functions can be used for evaluating an estimated model by checking how well the model is able to reproduce stylized facts; see Chapter 3 for an example. This means estimating the autocorrelation function from the data and comparing it with the corresponding autocorrelations obtained by plugging in the parameter estimates from the PARCH model into the theoretical expressions of the autocorrelations.

This approach cannot be applied if the autocorrelations compared are, say, autocorrelations of squared observations instead of autocorrelations of suitable fractional moments of their absolute values. This is the case for example when one wants to compare autocorrelations of squares implied by two different models, for instance a standard GARCH and a symmetric PARCH model, with each other. The only possibility is to estimate the autocorrelations of squared observations for the PARCH model by simulation. This becomes an issue in this paper, for the role of the power parameter in the PARCH model will be an object of our investigation.

Applications of the A-PARCH model to return series of stocks and exchange rates have revealed some regularities in the estimated values of the power parameter; see Ding, Granger and Engle (1993), Brooks, Faff, McKenzie and Mitchell (2000) and McKenzie and Mitchell (2002). We add to these results by fitting symmetric first-order PARCH models to return series of 30 most actively traded stocks of the Stockholm Stock Index. Our results agree with the previous ones and suggest that the power parameter lowers the autocorrelations of squared observations compared to the corresponding autocorrelations implied, other things equal, by the standard first-order GARCH model. In the present situation this means estimating the autocorrelation function of the squared observations from the data and comparing that with the corresponding values obtained by plugging the parameter estimates into the theoretical expressions of the autocorrelations. Another example can be found in He and Teräsvirta (1999*c*).

The plan of the paper is as follows. Section 2 defines the class of models of interest and introduces notation. The main theoretical results appear in Section 3. Section 4 contains a comparison of autocorrelation functions of squared observations for different models and Section 5 a discussion of empirical examples. Finally, conclusion appear in Section 6. All proofs can be found in Appendix.

2.2 The model

Let $\{\varepsilon_t\}$ be a real-valued discrete time stochastic process generated by

$$\varepsilon_t = z_t h_t \tag{2.1}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with mean zero and unit variance, and h_t is a $\mathcal{F}_{t-1}$-measurable function, where $\mathcal{F}_{t-1}$ is the sigma-algebra generated by $\{z_{t-1}, z_{t-2}, z_{t-3}, \ldots\}$, and positive with probability one. Let

$$h_t^{2\delta} = \alpha_0 + c_\delta(z_{t-1})h_{t-1}^{2\delta}, \ \delta > 0 \tag{2.2}$$

where α_0 is a positive scalar and $c_{\delta t} = c_\delta(z_t)$ is a well-defined function of z_t. The sequence $\{c_{\delta t}\}$ is a sequence of independent identically distributed random variables such that each $c_{\delta t}$ is stochastically independent of $h_t^{2\delta}$. Function $c_{\delta t}$ contains parameters that determine the moment structure of $\{\varepsilon_t\}$. Constrains on these parameters are necessary to guarantee that $h_t^{2\delta}$ remains positive with probability 1. We call (2.1) and (2.2) a general power ARCH (GPARCH(δ, 1, 1)) model. This model appeared in He and Teräsvirta (1999*c*) in a slightly more general form with $\alpha_0 = g(z_t)$ being a stochastic variable.

Setting $c_\delta(z_{t-1}) = \alpha(|z_t| - \phi z_t)^{2\delta} + \beta$ in equation (2.2), defines, together with equation (2.1), the Asymmetric Power ARCH (A-PARCH) (1,1) model of Ding et al. (1993). Note that these authors use δ in place of 2δ in equation (2.2) but that does not affect the results. Hentschel (1995) also defined a parametric family of GARCH models similar to (2.1) and (2.2) for highlighting relations between different GARCH models and their treatment of asymmetry.

In this paper we are interested in the limiting case $\delta \to 0$. Taking logarithms of (2.1) yields

$$\ln \varepsilon_t^2 = \ln z_t^2 + \ln h_t^2. \tag{2.3}$$

On the other hand, equation (2.1) can be modified such that it relates the Box-Cox transformed ε_t^2, that is, $\varphi_\delta(\varepsilon_t^2) = (\varepsilon_t^{2\delta} - 1)/\delta$, to $(h_t^{2\delta} - 1)/\delta$. Then by applying l'Hôpital's rule it can be shown that letting $\delta \to 0$ in the modified equation also leads to (2.3). This entitles us to consider certain exponential GARCH models as limiting cases of the power ARCH model (2.1) and (2.2). In order to see that, rewrite (2.2) in terms of $(h_t^{2\delta} - 1)/\delta$ and define $c_\delta(z_t) = \delta g(z_t) + \beta$. It can be shown that under certain conditions, as $\delta \to 0$, equation (2.2) becomes

$$\ln h_t^2 = \alpha_0 + g(z_{t-1}) + \beta \ln h_{t-1}^2 \tag{2.4}$$

where $g(z_t)$ is a well-defined function of z_t. Equation (2.4) is thus nested in (2.2). Equations (2.1) and (2.4) or (2.3) and (2.4) define a class of GPARCH(0,1,1) models that contains certain well-known models as special cases. For example, setting $g(z_t) = \phi z_t + \psi(|z_t| - \mathsf{E}\,|z_t|)$ in (2.4) yields

$$\ln h_t^2 = \alpha_0 + \phi z_{t-1} + \psi(|z_{t-1}| - \mathsf{E}\,|z_{t-1}|) + \beta \ln h_{t-1}^2 \tag{2.5}$$

which, jointly with (2.1), defines the EGARCH(1,1) model of Nelson (1991). Similarly, we may set $c_\delta(z_t) = \alpha g_1^\delta(z_t) + \beta$ where $g_1(z_t) > 0$ for all t with probability one. Then, by l'Hôpital's rule, (2.2) converges to

$$\ln h_t^2 = \alpha_0 + \alpha \ln g_1(z_{t-1}) + (\alpha + \beta) \ln h_{t-1}^2 \tag{2.6}$$

as $\delta \to 0$. Equations (2.1) and (2.6) define a class of logarithmic GARCH (LGARCH(1,1)) models. Setting $g_1(z_t) = z_t^2$ in (2.6) yields

$$\ln h_t^2 = \alpha_0 + \alpha \ln \varepsilon_{t-1}^2 + \beta \ln h_{t-1}^2 \tag{2.7}$$

which is the LGARCH(1,1) model of Geweke (1986) and Pantula (1986). Since (2.4) and (2.6) have a similar structure we mainly consider results for the GPARCH(0,1,1) model (2.1) and (2.4). They can be easily modified to apply to the class of LGARCH(1,1) models.

2.3 The limiting results

In this section we derive the asymptotic moment structure of the GPARCH(1,1) model (2.1) and (2.2) as $\delta \to 0$ under the Box-Cox transformation. We first give the moment structure of (2.1) and (2.2) for $\delta > 0$. Having done that we derive the moment structure of model (2.3) with (2.4). Finally, we show that this result may be also obtained as a limiting case of model (2.1) with (2.2) as $\delta \to 0$.

To formulate our first result let $\gamma_\delta = \mathsf{E} c_{\delta t}$ and $\gamma_{2\delta} = \mathsf{E} c_{\delta t}^2$. We have

Lemma 1 *For the GPARCH(δ,1,1) model (2.1) with (2.2), a necessary and sufficient condition for the existence of the 4δ-th unconditional moment $\mu_{4\delta} = \mathsf{E}\left|\varepsilon_t\right|^{4\delta}$ of $\{\varepsilon_t\}$ is*

$$\gamma_{2\delta} < 1. \tag{2.8}$$

If (2.8) holds, then

$$\mu_{4\delta} = \alpha_0^2 \nu_{4\delta}(1+\gamma_\delta)/\{(1-\gamma_\delta)(1-\gamma_{2\delta})\} \tag{2.9}$$

where $\nu_{2\psi} = \mathsf{E}\left|z_t\right|^{2\psi}$, $\psi>0$. The autocorrelation function $\rho_n(\delta) = \rho(\left|\varepsilon_t\right|^{2\delta}, \left|\varepsilon_{t-n}\right|^{2\delta})$, $n \geq 1$, of $\{\left|\varepsilon_t\right|^{2\delta}\}$ has the form

$$\rho_1(\delta) = \frac{\nu_{2\delta}[\overline{\gamma}_\delta(1-\gamma_\delta^2) - \nu_{2\delta}\gamma_\delta(1-\gamma_{2\delta})]}{\nu_{4\delta}(1-\gamma_\delta^2) - \nu_{2\delta}^2(1-\gamma_{2\delta})} \tag{2.10}$$

where $\overline{\gamma}_\delta = \mathsf{E}(\left|z_t\right|^{2\delta} c_{\delta t})$, and $\rho_n(\delta) = \gamma_\delta \rho_{n-1}(\delta)$, $n \geq 2$.

Let $\mathcal{M}_\delta(\mu_{4\delta}, \rho_n(\delta))$ be the analytic moment structure defined by Lemma 1 for the GPARCH(δ,1,1) model (2.1) and (2.2). It is seen that $\mathcal{M}_\delta(\cdot)$ is a function of power parameter δ. Note that the autocorrelation function of $\{\left|\varepsilon_t\right|^{2\delta}\}$ is decaying exponentially with the discount factor γ_δ. In particular, setting $\delta = 1$ in equations (2.8) and (2.10) yields the existence condition of the fourth moment and the autocorrelation function of the squared observations of the standard GARCH(1,1) model (Bollerslev (1986)) with non-normal errors.

It is customary to also consider the kurtosis of any given GARCH process, see, for example, Bollerslev (1986) or He and Teräsvirta (1999). In this case, the kurtosis of $|\varepsilon_t|^\delta$ or $\varphi_\delta(|\varepsilon_t|) = (|\varepsilon_t|^\delta - 1)/\delta$ may be defined as

$$\begin{aligned}\kappa_4(\delta) &= \frac{\mathsf{E}(|\varepsilon_t|^\delta - \mathsf{E}\,|\varepsilon_t|^\delta)^4}{\{\mathsf{E}(|\varepsilon_t|^\delta - \mathsf{E}\,|\varepsilon_t|^\delta)^2\}^2} \\ &= \frac{\mathsf{E}(\varphi_\delta(|\varepsilon_t|) - \mathsf{E}\varphi_\delta(|\varepsilon_t|))^4}{\{\mathsf{E}(\varphi_\delta(|\varepsilon_t|) - \mathsf{E}\varphi_\delta(|\varepsilon_t|))^2\}^2}\end{aligned}$$

so that the limiting case

$$\lim_{\delta\to 0}\kappa_4(\delta) = \frac{\mathsf{E}(\ln|\varepsilon_t| - \mathsf{E}\ln|\varepsilon_t|)^4}{\{\mathsf{E}(\ln|\varepsilon_t| - \mathsf{E}\ln|\varepsilon_t|)^2\}^2}. \tag{2.11}$$

The kurtosis (2.11) is thus the limiting case of the kurtosis of the absolute-valued process $\{|\varepsilon_t|^\delta\}$. Computing it would require the expectations $\mathsf{E}(\ln|\varepsilon_t|)^4$ and $\mathsf{E}(\ln|\varepsilon_t|)^3$ or, alternatively, $\mathsf{E}(\ln\varepsilon_t^2)^4$ and $\mathsf{E}(\ln\varepsilon_t^2)^3$ for which no analytical expressions have been derived above. The kurtosis of $\ln|\varepsilon_t|$ is a concept quite different from that of ε_t, and for this reason it is not considered any further here.

For the GPARCH(0,1,1) process we obtain the following result:

Lemma 2 *For the GPARCH(0,1,1) process (2.3) and (2.4), assume that variances of* $(\ln z_t^2)^2$ *and* $(g(z_t))^2$ *are finite for any t. Then the second unconditional moment of* $\ln\varepsilon_t^2$ *exists if and only if*

$$|\beta| < 1. \tag{2.12}$$

When (2.12) holds, this second moment can be expressed as

$$\mu_0 = \mathsf{E}(\ln\varepsilon_t^2)^2 = \frac{\Delta}{(1-\beta)(1-\beta^2)} \tag{2.13}$$

where $\Delta = \gamma_{(\ln z^2)^2}(1-\beta)(1-\beta^2) + 2\gamma_{\ln z^2}(\alpha_0+\gamma_g)(1-\beta^2) + [\alpha_0^2(1+\beta) + 2\alpha_0(1+\beta)\gamma_g + 2\beta\gamma_g^2 + (1-\beta)\gamma_{g^2}]$ *and* $\gamma_{(\ln z^2)^2} = \mathsf{E}(\ln z_t^2)^2$, $\gamma_{\ln z^2} = \mathsf{E}\ln z_t^2$, $\gamma_g = \mathsf{E}g(z_t)$ *and* $\gamma_{g^2} = \mathsf{E}(g(z_t))^2$. *Furthermore, the autocorrelation function* $\rho_n^0 = \rho(\ln\varepsilon_t^2, \ln\varepsilon_{t-n}^2)$, $n \geq 1$, *of* $\{\ln\varepsilon_t^2\}$ *has the form*

$$\begin{aligned}\rho_1^0 &= \frac{(1-\beta^2)(\gamma_{g\ln z^2} - \gamma_g\gamma_{\ln z^2}) + \beta(\gamma_{g^2} - \gamma_g^2)}{(1-\beta^2)(\gamma_{(\ln z^2)^2} - \gamma_{\ln z^2}^2) + (\gamma_{g^2} - \gamma_g^2)}, \\ \rho_n^0 &= \rho_1^0\beta^{n-1}, n \geq 2,\end{aligned} \tag{2.14}$$

where $\gamma_{g\ln z^2} = \mathsf{E}(g(z_t)\ln z_t^2)$.

Nelson (1991) derived the autocovariance function of the logarithm of the conditional variance of the EGARCH process. Breidt et al. (1998) obtained the autocorrelation function of $\{\ln \varepsilon_t^2\}$ for the EGARCH model. In both articles the authors made use of the infinite moving average representation of the logarithm of the conditional variance. Lemma 2 gives the corresponding result for the first-order process directly in terms of the parameters of the original model, which is practical for model evaluation purposes.

Let $\mathcal{M}_0(\mu_0, \rho_n^0)$ be the moment structure defined by Lemma 2 for the GPARCH(0,1,1) process (2.3) and (2.4). Next, let $I_\delta = (0, l)$, $l > 0$, be an open interval such that $\delta_0 \in I_\delta$ if and only if $\gamma_{2\delta_0} < 1$. We have

Theorem *Assume that $\mathcal{M}_\delta(\cdot)$ is defined on I_δ and the corresponding functions in $\mathcal{M}_\delta(\cdot)$ are continuous on I_δ and are twice differentiable with respect to δ. Then under the transformation $\varphi_\delta(\varepsilon_t^2) = (\varepsilon_t^{2\delta} - 1)/\delta$,*

$$\mathcal{M}_\delta(\mu_{4\delta}, \rho_n(\delta)) \to \mathcal{M}_0(\mu_0, \rho_n^0) \tag{2.15}$$

as $\delta \to 0$.

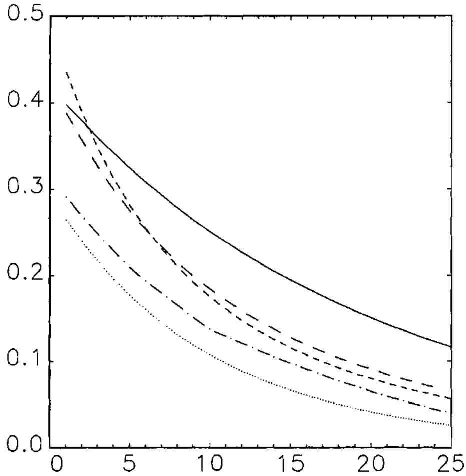
Remark. It has been pointed out above that, under the Box-Cox transformation $\varphi_\delta(\varepsilon_t^2) = (\varepsilon_t^{2\delta} - 1)/\delta$, equation (2.1), when appropriately modified, converges to equation (2.3) as $\delta \to 0$. The theorem then says that under this transformation the moment structure of the GPARCH(δ,1,1) model (2.1) and (2.2) approaches the moment structure of the GPARCH(0,1,1) model as $\delta \to 0$: $\mu_{4\delta} \to \mu_0$ and $\rho_n(\delta) \to \rho_n^0$. This convergence shows that the moment structure $\mathcal{M}_0(\cdot)$ belongs to the class of structures $\mathcal{M}_\delta(\cdot)$ as a boundary case. These moment structures are thus isomorphic. Besides, the parameter δ in the GPARCH(δ,1,1) process defines a value for which the autocorrelation function $\rho(|\varepsilon_t|^\delta, |\varepsilon_{t-k}|^\delta)$, $k \geq 1$, decays exponentially with k. To consider the practical value of these results suppose, for example, that $\gamma_4 < 1$. Then we have a class of GPARCH($\delta, 1, 1$) models with the same parameter values such that the available $\mathcal{M}_\delta(\cdot)$ is defined on $[0, 1]$, that is, $\gamma_{2\delta} < 1$, $\delta \leq 1$. Practitioners may want to use these results to see what kind of moment implications GPARCH models they estimate may have. Results in $\mathcal{M}_\delta(\cdot)$ defined on $[0, 1]$ may also be useful in checking how well different GPARCH models represent the reality, which is done by comparing parametric moment estimates from a GPARCH($\delta, 1, 1$) model with corresponding nonparametric ones obtained directly from the data. First-order LGARCH and EGARCH models may thus be compared with, say, a standard GARCH(1,1) model in this respect if both are estimated using the same data.

2.4 Autocorrelation functions of squared observations

In this section we show how the autocorrelation function of ε_t^2 varies with δ across GPARCH(δ,1,1) models with normal errors. We demonstrate how the power parameter increases the flexibility of the specification compared to the GARCH model. We also include the symmetric first-order EGARCH model and LGARCH models in our comparison. The three parameters in the GARCH model are selected such that the unconditional variance equals unity, the kurtosis equals 12, and the decay rate of the autocorrelations of ε_t^2 equals 0.95. For the GARCH(1,1) model, this decay rate is obtained by setting $\alpha+\beta = 0.95$. The parameter values for the EGARCH model and the LGARCH are chosen such as to make the models as comparable with the GARCH model as possible. Thus, $\beta = 0.95$ in the EGARCH model and $\alpha + \beta = 0.95$ in the LGARCH model correspond to $\alpha + \beta = 0.95$ in the GARCH model, because β and $\alpha + \beta$, respectively, control the decay of the autocorrelation function of the squared observations in these two models. Note, however, that while the decay rate of the autocorrelation of ε_t^2 in the GARCH(1,1) model equals $\alpha+\beta$, it only approaches β with increasing lag length in the EGARCH(1,1) model and $\alpha+\beta$ from below in the LGARCH(1,1) model. The individual parameters are chosen such that the unconditional variance and the kurtosis are the same in all three models as well. This can be done using the analytic expressions for the relevant moments of the EGARCH(1,1) model in Chapter 1 and the LGARCH(1,1) model in Appendix (Lemma 3). For the EGARCH model, the moment restrictions are the same as in Chapter 4.

In order to illustrate the role of δ, we consider the GPARCH(1,1) model with $\delta = 1.5$ and $\delta = 1$ under the assumption that the other parameters are the same as in the GARCH(1,1) model. For $\delta = 1.5$ the autocorrelations of ε_t^2 cannot be obtained analytically, and we have computed them by simulation from 1,000 series of 100,000 observations each. For $\delta = 1$, they are available from He and Teräsvirta (1999) where this special case is considered under the name absolute-valued GARCH model. It can be seen from Figure 2.1 that $\delta < 2$ reduces the autocorrelations of ε_t^2 (other things equal) compared to $\delta = 2$ (the GARCH model). The difference in autocorrelations between $\delta = 1.5$ and $\delta = 1$ is smaller than the corresponding difference between $\delta = 2$ and $\delta = 1.5$ which is quite large. The autocorrelations of the EGARCH model and the LGARCH are different from the ones of the GARCH model. As already mentioned, the decay is exponential for the autocorrelations of the GARCH model but faster than exponential and exponential only asymptotically (as a

Figure 2.1 Autocorrelation functions of squares for five first-order GARCH models, EGARCH (dashed line, short dashes), GARCH (solid line), LGARCH (dashed line, long dashes), PGARCH $\delta = 1.5$ (dashed-dotted line) and PGARCH $\delta = 1$ (dotted line).



function of the lag length), both for the EGARCH model and the LGARCH model. This is also in fact the case for the GPARCH models for which $\delta < 2$.

2.5 Empirical examples

Ding et al. (1993) demonstrated the potential of the GPARCH model by fitting the model with normal errors to the long S&P 500 daily stock return series from January 3, 1928 to August 30, 1991, 17,055 observations in all. The estimate of the power parameter δ was equal to 1.43 and significantly different from two (the GARCH model). Brooks, Faff, McKenzie and Mitchell (2000) applied the GPARCH(1,1) model with t-distributed errors to national stock market returns for 10 countries plus a world index for the period February 1989 to December 1996, which amounted to a total of 2062 daily observations. Except for three extremes cases, the power parameter estimates were between 1.17 and 1.45, with most values close to the mean value 1.36. The authors concluded that in the absence of leverage effects there is moderate evidence supporting the need for the power parameter. In the case of six countries plus the world index, the standard GARCH model could not be rejected in favour of the symmetric GPARCH model at the 5% significance level. The evidence against the standard GARCH model was, however, much stronger in case of a combination of leverage and power effects. More specifically, with the exception of one national return series, the GARCH model was strongly rejected in favour of the asymmetric GPARCH model.

McKenzie and Mitchell (2002) applied the GPARCH model to daily return series of 17 heavily traded bilateral exchange rates and found the estimated power parameter equal to 1.37 on average. All power parameter estimates were between one and two. For seven of the estimated models the power parameter was significantly different from two. The results were thus quite similar to the ones Brooks et al. (2000) reported.

Tse and Tsui (1997) fitted the A-PARCH model to two exchange rate return series, the Malaysian ringgit and the Singaporean dollar. Their results do not fit the pattern just described. The most notable fact is the large change in the estimated value of δ when the t-distributed errors are substituted for the normal ones. The estimated degrees of freedom of the t-distribution are in both cases so low that they alone exclude the existence of the finite fourth moment for the underlying GARCH process.

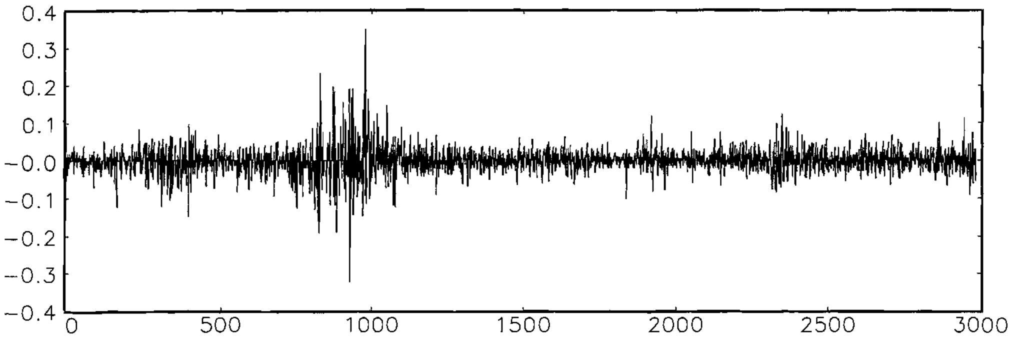
In order to further explore the role of the power parameter in practice we consider daily return series of the 30 most actively traded stocks in Stockholm Stock Exchange and estimate a symmetric GPARCH(δ,1,1) model (with normal errors), a standard GARCH model and an EGARCH model for these series. The names of the stocks can be found in Table 2.1 together with information about the length of the series. The period investigated ends April 24, 2001. The return series have been obtained from Datastream and are also used in Chapter 4.

In Table 2.1 we report the maximum likelihood estimates of the power parameter δ. The estimates of δ lie between 1.21 and 1.49, most of them close to the mean value of 1.40. We find that the estimates are remarkably similar (around a mean value of 1.40) to the ones Brooks et al. (2000) obtained for their return series. The estimates of δ are significantly different from two in a majority of cases, see Table 2.1, where the p-value of the test is less than 0.01 in 15 cases out of 29. It should be noted, however, that some of the estimated autocorrelations may not actually have a theoretic counterpart because the moment condition $\gamma_{2\delta} < 1$ appearing in Lemma 1 is not satisfied. This does not mean that the 2δth moment of ε_t cannot exist, because $\widehat{\gamma}_{2\delta}$ is an estimate, but empirical support for the existence of this moment cannot be argued to be strong. If we merely compare standard GARCH and EGARCH models using tests of non-nested hypotheses, the results of Chapter 4 indicate that both models fit the 30 series more or less equally well.

Table 2.1 The stocks, the estimates of the power parameter, length of the series, p-values of the likelihood ratio test of GARCH against GPARCH, and the estimated left-hand side of the moment condition of Lemma 1.

y	δ	T	p	$\widehat{\gamma}_{2\delta}$
ABB	1.41	3717	0.008	1.013
Assa A.	1.37	1617	0.382	0.733
Assi D.	1.40	1769	0.003	1.015
Astra	1.37	3591	3×10^{-7}	1.029
Atlas C.	1.40	2915	0.001	0.976
Autoliv	1.30	1690	9×10^{-5}	1.055
Electrolux	1.42	4577	0.1648	0.959
Ericsson	1.42	4576	1×10^{-6}	0.962
FSB	1.49	1470	0.1915	0.622
Gambro	1.41	2454	0.037	1.004
Holmen	1.42	4568	0.022	0.959
Industriv.	1.43	2061	0.200	0.905
Investor	1.42	4146	0.009	0.944
Nokia	1.41	2907	5×10^{-5}	0.993
OMG	1.43	2084	9×10^{-6}	0.962
Pharmacia	1.44	1370	0.339	0.826
Sandvik	1.38	4576	9×10^{-6}	1.037
Scania	1.38	1268	2×10^{-14}	1.173
Securitas	1.33	2461	9×10^{-5}	1.025
Skandia	1.42	4566	0.314	0.959
SEB	1.43	2984	0.003	0.957
Skanska	1.41	4337	0.173	0.984
SKF	1.43	4578	0.012	0.939
SSAB	1.45	2963	0.719	0.784
Stora	1.42	3263	0.197	0.964
SCA	1.39	4576	0.290	1.019
SHB	1.43	2612	5×10^{-5}	0.920
Sw. Match	1.21	1239	0.045	1.039
VOLVO	1.37	5324	2×10^{-8}	1.033

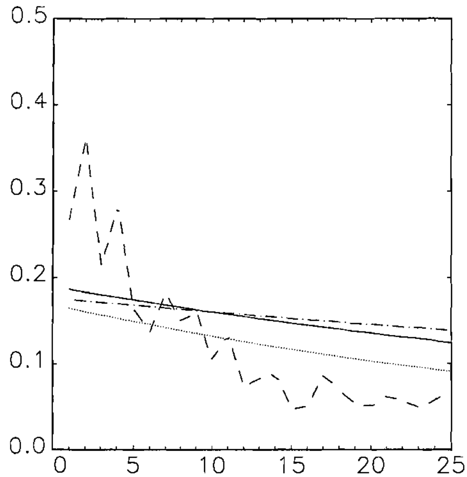
Figure 2.2 Daily returns of the stock SEB, from 1989 to April 2001.



As a detailed example we consider the return series of SEB which is plotted in Figure 2.2. For this series we estimate the autocorrelation function of the squared observations from the data and compare them with the autocorrelations obtained by plugging the parameter estimates for the three estimated models into the theoretical expressions of the autocorrelations. Note that from the GARCH(1,1) model estimated for this series one obtains $\widehat{\gamma}_4 < 1$, so the fourth-moment condition is satisfied and we can discuss the autocorrelation function of squares of the GPARCH(1,1) model with some confidence. It is seen from Figure 2.3 that for all models the discrepancy between the autocorrelation functions and the autocorrelations estimated directly from data is large at small lags. For long lags, the gap between the two is much smaller for the GPARCH(δ,1,1) model than for the two other models. As already noted, augmenting the GARCH model by the power parameter δ, other things equal, reduces autocorrelations of squared observations compared to the two other models. This probably explains the results obtained by Brooks et al. (2000) and McKenzie and Mitchell (2002).

The present example shows that the estimated power parameter considerably improves the correspondence between the estimated autocorrelations on the one hand and the autocorrelation estimates from the model on the other. But then, the rapid decrease of the autocorrelations at first lags is not accounted for by any of the models; a higher-order model is required for the purpose. He and Teräsvirta (1999*b*) showed how a second-order GARCH model already can have an autocorrelation function of squared observations that is much more flexible than the corresponding autocorrelation function for the GARCH(1,1) model.

Figure 2.3 Autocorrelation functions of squared observations, estimated for the SEB (dashed line), and computed from three estimated first-order GARCH models; GARCH (solid line), EGARCH (dashed-dotted line), and GPARCH (dotted line).



2.6 Final remarks

In this chapter we derive the autocorrelation structure of the logarithms of squared observations of a class of power ARCH processes and show that this structure may be obtained as a limiting case of a general power ARCH model. An interesting thing to notice is that the autocorrelation structure of the δth power of absolute-valued observations of this first-order GPARCH process is exponential for all GPARCH$(\delta, 1, 1)$ processes such that the 4δth fractional moment exists. This property is retained at the limit as the power parameter approaches zero, which means that the autocorrelation function of the process of logarithms of squared observations also decay exponentially. While this is true for the logarithmed squared observations of an LGARCH(1,1) or EGARCH(1,1) process it cannot simultaneously be true for the untransformed observations defined by these processes as we have demonstrated in Chapter 1 for the EGARCH(1,1) case.

Conversely, if we have the original GARCH(1,1) [GPARCH(1,1,1)] process of Bollerslev (1986) with the autocorrelations of $\{\varepsilon_t^2\}$ decaying exponentially, the autocorrelation function of $\{\ln \varepsilon_t^2\}$ does not have this property. The practical value of these facts when discriminating between GARCH(1,1) and EGARCH(1,1) models is not clear, but they illustrate the theoretical differences in the higher-order dynamics between these two classes of models. Note that possible asymmetry is not an issue here. Nelson's EGARCH(1,1) model is a member of the GPARCH(0, 1, 1) family independent of the value of the asymmetry parameter. Likewise, if the standard GARCH(1,1) process is generalized to an asymmetric GJR-GARCH(1,1) (Glosten, Jagannathan and Runkle (1993)) process the argument remains the same. This is because the GJR-GARCH model is still a member of the GPARCH(1,1,1) class; see Ding et al. (1993) and He and Teräsvirta (1999) for more discussion.

In order to explain the role of the power parameter we present a detailed analysis of how the autocorrelation function of ε_t^2 differ across members of the GPARCH(δ,1,1) models. We demonstrate that $\delta < 2$ reduces the autocorrelations of ε_t^2 (other things equal) compared to $\delta = 2$ (the GARCH model). This fact may explain the regularities in estimation results in papers in which GPARCH models have been fitted to stock return series. In an empirical example we show that the estimated power parameter considerably improves the correspondence between the estimated autocorrelations on the one hand and the autocorrelation estimates from the model on the other.

Bibliography

Billingsley, P. (1986), *Probability and Measure*, Wiley, New York.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**, 307–327.

Breidt, F. J., Crato, N. and de Lima, P. (1998), 'The detection and estimation of long memory in stochastic volatility', *Journal of Econometrics* **83**, 325–348.

Brooks, R., Faff, R. W., McKenzie, M. D. and Mitchell, H. (2000), 'A multi-country study of power ARCH models and national stock market returns', *Journal of International Money and Finance* **19**, 377–397.

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.

Geweke, J. (1986), 'Modelling persistence of conditional variances: Comment', *Econometric Reviews* **5**, 57–61.

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.

He, C. and Teräsvirta, T. (1999*a*), 'Properties of moments of a family of GARCH processes', *Journal of Econometrics* **92**, 173–192.

He, C. and Teräsvirta, T. (1999*b*), 'Properties of the autocorrelation function of squared obeservations for second-order GARCH processes under two sets of parameter constraints', *Journal of Time Series Analysis* **20**, 23–30.

He, C. and Teräsvirta, T. (1999*c*), Statistical properties of the asymmetric power ARCH process, *in* R. F. Engle and H. White, eds, 'Cointegration, Causality, and Forecasting. Festschrift in Honour of Clive W J Granger', Oxford University Press, Oxford, pp. 462–474.

He, C., Teräsvirta, T. and Malmsten, H. (1999), Fourth moment structure of a family of first-order exponential GARCH models. Working Paper Series in Economics and Finance, Stockholm School of Economics, No.345.

Hentschel, L. (1995), 'All in the family. Nesting symmetric and asymmetric GARCH models', *Journal of Financial Economics* **39**, 71–104.

McKenzie, M. D. and Mitchell, H. (2002), 'Generalized asymmetric power ARCH modelling of exchange rate volatility', *Applied Financial Economics* **12**, 555–564.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**, 347–370.

Pantula, S. G. (1986), 'Modelling persistence of conditional variances: Comment', *Econometric Reviews* **5**, 71–74.

Tse, Y. K. and Tsui, A. K. C. (1997), 'Conditional volatility in foreign exchange rates: Evidence from the Malaysian ringgit and Singapore dollar', *Pacific-Basin Finance Journal* **5**, 345–356.

Appendix A

Proofs

A.1 Lemma 1

(i) We shall show that $\{\varepsilon_t\}$ defined in (2.1) and (2.2) is strictly stationary if $\gamma_\delta < 1$. Note that under $\gamma_\delta < 1$ (2.2) has a representation

$$h_t^{2\delta} = \alpha_0 + \sum_{i=1}^{\infty} c_\delta(z_{t-i}).$$

Since $\{c_\delta(z_t)\}$ is a sequence of iid and $var(c_\delta(z_t))$ is finite, $\sum_{i=1}^{\infty} Var(c_\delta(z_{t-i})) < \infty$. It follows from Billingsley (1986, Theorem 22.6)) that $\{h_t^{2\delta}\}$ is finite almost surely. This, combined with Theorem 2.1 of Nelson (1991) and $\gamma_{2\delta} < 1$ implies that $\{\varepsilon_t\}$ in (2.1) and (2.2) is strictly stationary.

(ii) That $\{h_t^{4\delta}\}$ is finite almost surely follows by the fact that $\gamma_{2\delta} < 1$ and (2.2) is strictly stationary. Thus, the results in He and Teräsvirta (1999a,b) apply and thus (2.9) and (2.10) hold.

(iii) It follows from Theorems 22.3 and 22.8 in Billingsley (1986) that if $\sum_{i=1}^{\infty} Var(c_\delta(z_{t-i})) = \infty$, then $\{h_t^{4\delta}\} = \infty$ almost surely. Thus the necessary condition (2.8) holds. ■

A.2 Lemma 2

(i) Similarly to (i) of Lemma 1, strict stationarity of $\{\ln \varepsilon_t^2\}$ in (2.3) and (2.4) follows from the fact that $|\beta| < 1$, and $Var(\ln z_t^2)^2$ and $Var(g(z_t))^2$ are finite.

(ii) As (ii) and (iii) in Lemma 1, under the assumptions of Lemma 2, $\{(\ln \varepsilon_t^2)^2\}$ is finite almost surely if and only $|\beta| < 1$.

(iii) To compute μ_0 under (2.12), we repeatedly apply (2.4) to $\ln h_t^2$, which yields

$$\ln h_t^2 = \alpha_0 \sum_{i=1}^{k+1} \beta^{i-1} + \sum_{i=1}^{k+1} \beta^{i-1} g(z_{t-1}) + \beta^{k+1} \ln h_{t-(k+1)}^2. \tag{A.1}$$

Taking expectations of both sides of (A.1) and letting $k \to \infty$ yield

$$\mathsf{E}(\ln h_t^2) = (\alpha_0 + \gamma_g)/(1-\beta). \tag{A.2}$$

Similarly, repeated application of

$$(\ln h_t^2)^2 = (\alpha_0 + g(z_{t-1}))^2 + 2\beta(\alpha_0 + g(z_{t-1})) \ln h_{t-1}^2 + \beta^2 (\ln h_{t-1}^2)^2 \quad \text{(A.3)}$$

to $(\ln h_t^2)^2$ in (A.3) yields

$$(\ln h_t^2)^2 = \sum_{i=1}^{k} [\beta^{2(i-1)} (\alpha_0 + g(z_{t-i}))^2] + 2\beta \sum_{i=1}^{k} \beta^{2(i-1)} (\alpha_0 + g(z_{t-i}))(\ln h_{t-i}^2) + \beta^{2k} (\ln h_{t-k}^2)^2.$$

Thus, under (2.12) by letting $k \to \infty$ and taking expectations

$$\mathsf{E}(\ln h_t^2)^2 = [\alpha_0^2(1+\beta) + 2\alpha_0(1+\beta)\gamma_g + 2\beta\gamma_g^2 + (1-\beta)\gamma_{g^2}]/[(1-\beta)(1-\beta^2)]. \quad \text{(A.3)}$$

It follows from formulas (2.3), (A.2) and (A.3) that expression (2.13) is valid.

Next, consider the n-th order autocorrelation of $\{\ln \varepsilon_t^2\}$

$$\rho_n^0 = \frac{\mathsf{E}(\ln \varepsilon_t^2 \ln \varepsilon_{t-n}^2) - (\mathsf{E}(\ln \varepsilon_t^2))^2}{\mathsf{E}(\ln \varepsilon_t^2)^2 - (\mathsf{E}(\ln \varepsilon_t^2))^2}. \tag{A.4}$$

We have

$$\ln \varepsilon_t^2 \ln \varepsilon_{t-n}^2 = \ln z_t^2 \ln z_{t-n}^2 + \ln z_t^2 \ln h_{t-n}^2 + \ln h_t^2 \ln z_{t-n}^2 + \ln h_t^2 \ln h_{t-n}^2. \tag{A.5}$$

It follows from (A.1) that

$$(\ln h_t^2)(\ln h_{t-n}^2) = \alpha_0 \sum_{i=1}^{n} \beta^{i-1} \ln h_{t-n}^2 + (\sum_{i=1}^{n} \beta^{i-1} g(z_{t-1})) \ln h_{t-n}^2 + \beta^n (\ln h_{t-n}^2)^2 \tag{A.6}$$

and

$$\begin{aligned}\ln h_t^2 \ln z_{t-n}^2 &= \alpha_0 \sum_{i=1}^{n} \beta^{i-1} \ln z_{t-n}^2 + (\sum_{i=1}^{n-1} \beta^{i-1} g(z_{t-1})) \ln z_{t-n}^2 \\ &\quad + \beta^{n-1} g(z_{t-n}) \ln z_{t-n}^2 + \beta^n \ln z_{t-n}^2 \ln h_{t-n}^2. \qquad \text{(A.7)}\end{aligned}$$

The expectation of (A.5) is obtained by taking expectations of both sides of (A.6) and (A.7) and inserting them to (A.5). Applying this expectation to (A.4) yields (2.14). ■

A.3 Theorem

For the ease of exposition, write (2.2) as

$$h_t^{2\delta} = \alpha_0^* + c_\delta(z_{t-1}) h_{t-1}^{2\delta} \qquad \text{(A.8)}$$

where $c_\delta(z_{t-1}) = \delta g(z_{t-1}) + \beta$. Following Ding, Granger and Engle (1993), decompose α_0^* as

$$\alpha_0^* = (1 - \gamma_\delta)\omega^\delta \qquad \text{(A.9)}$$

where $\gamma_\delta = \delta\gamma_g + \beta$ and $\omega^\delta = \mathsf{E} h_t^{2\delta}$, $\omega > 0$. Rewrite (A.8) as

$$\begin{aligned}(h_t^{2\delta} - 1)/\delta &= (\alpha_0^* + \beta - 1)/\delta + g(z_{t-1}) h_{t-1}^{2\delta} \\ &\quad + \beta(h_{t-1}^{2\delta} - 1)/\delta. \qquad \text{(A.10)}\end{aligned}$$

Insert (A.9) into (A.10) and let $\delta \to 0$ on both sides of (A.10). Then, by l'Hôpital's rule (A.10) converges to (2.4). In particular,

$$(\alpha_0^* + \beta - 1)/\delta \to \alpha_0, \qquad \text{(A.11)}$$

where $\alpha_0 = (1 - \beta)(\mathsf{E} \ln h_t^2) - \gamma_g$ is the constant term in (2.4). Besides, from (A.9) we have, as $\delta \to 0$,

$$\alpha_0^* \to 1 - \beta. \qquad \text{(A.12)}$$

The convergence results (A.11) and (A.12) are used to prove the following results.

(i) We shall show that $\mu_{4\delta} \to \mu_0$ as $\delta \to 0$ under the Box-Cox transformation. From Lemma 1 we obtain

$$\mu_{2\delta} = \mathsf{E}\varepsilon_t^{2\delta} = \frac{\alpha_0^* \nu_{2\delta}}{1 - \gamma_\delta}. \qquad \text{(A.13)}$$

From (A.13) it follows for the Box-Cox transformed $\varepsilon_t^{2\delta}$ that

$$\mathsf{E}\varphi_\delta(\varepsilon_t^2) = \frac{[\alpha_0^*(\nu_{2\delta}-1)+(\alpha_0^*+\beta-1)]/\delta+\gamma_g}{1-\gamma_\delta}. \tag{A.14}$$

Letting $\delta \to 0$ on both sides of (A.14) and applying (A.11) and (A.12) to the right-hand side of (A.14) gives

$$\mu_2^0 = \mathsf{E}\ln\varepsilon_t^2 = \frac{\gamma_{\ln z^2}(1-\beta)+(\alpha_0+\gamma_g)}{1-\beta}. \tag{A.15}$$

From (2.9) it follows that

$$\mathsf{E}(\varphi_\delta(\varepsilon_t^2))^2 = [\frac{\alpha_0^{*2}\nu_{4\delta}(1+\gamma_\delta)}{(1-\gamma_\delta)(1-\gamma_{2\delta})} - 2\mathsf{E}\varepsilon_t^{2\delta}+1]/\delta^2. \tag{A.16}$$

Applying (A.13) to the right-hand side of expression (A.16) it is seen that (A.16) is equivalent to

$$\begin{aligned}\mathsf{E}(\varphi_\delta(\varepsilon_t^2))^2 \quad = \quad & \frac{1}{(1-\gamma_\delta)(1-\gamma_{2\delta})}\{\frac{1}{\delta^2}[\alpha_0^{*2}\nu_{4\delta}(1+\beta)-2\alpha_0^*\nu_{4\delta}(1-\beta^2)\\ & +1-\beta-\beta^2+\beta^3]+\frac{1}{\delta}[\alpha_0^{*2}\nu_{4\delta}\gamma_g+4\alpha_0^*\beta\nu_{2\delta}\gamma_g\\ & -\gamma_g(1+2\beta-3\beta^2)]+[2\alpha_0^*\beta\nu_{2\delta}\gamma_{g^2}-\gamma_{g^2}(1-\beta)\\ & +2\beta\gamma_g^2]+\delta\gamma_g\gamma_{g^2}\}. \end{aligned} \tag{A.17}$$

Note that, as $\delta \to 0$, $\mathsf{E}(\varphi_\delta(\varepsilon_t^2))^2 \to \mathsf{E}(\ln\varepsilon_t^2)^2$, $(\nu_{4\delta}-2\nu_{2\delta}+1)/\delta^2 \to \mathsf{E}(\ln z_t^2)^2$ and $(\nu_{2\delta}-1)/\delta \to \mathsf{E}(\ln z_t^2)$. Apply those facts and (A.11) and (A.12) to the right-hand side of (A.17) while letting $\delta \to 0$ on both sides of (A.17). It follows from l'Hôpital's rule that (A.17) converges to

$$\mu_0 = \mathsf{E}(\ln\varepsilon_t^2)^2 = \frac{\Delta}{(1-\beta)(1-\beta^2)}. \tag{A.18}$$

Then $\mu_{4\delta} \to \mu_0$ holds in (2.15).

(ii) We shall now prove that $\lim_{\delta\to 0}\rho_n(\delta) = \rho_n^0$. Since $\lim_{\delta\to 0}\rho_n(\delta) = \lim_{\delta\to 0}\rho_1(\delta)\gamma_\delta^{n-1}$ $= \beta^{n-1}\lim_{\delta\to 0}\rho_1(\delta)$, we have to prove that $\lim_{\delta\to 0}\rho_1(\delta) \to \rho_1^0$.

Let $\rho_1(\delta) = u/v$ in (2.10) where $u = \nu_{2\delta}[\overline{\gamma}_\delta(1-\gamma_\delta^2)-\nu_{2\delta}\gamma_\delta(1-\gamma_{2\delta})]$ and $v = \nu_{4\delta}(1-\gamma_\delta^2)-\nu_{2\delta}^2(1-\gamma_{2\delta})$. Since $\lim_{\delta\to 0} u = 0$ and $\lim_{\delta\to 0} v = 0$ we need to apply l'Hôpital's rule in order to obtain $\lim_{\delta\to 0}\rho_1(\delta)$. Note that

$$\begin{aligned}
\frac{\partial}{\partial\delta}u &= (\nu_{2\delta} - \beta^2\nu_{2\delta} - \nu_{2\delta}^2 + \beta^2\nu_{2\delta}^2)\overline{\gamma}_g \\
&+\delta\frac{\partial}{\partial\delta}(\nu_{2\delta} - \beta^2\nu_{2\delta} - \nu_{2\delta}^2 + \beta^2\nu_{2\delta}^2)\overline{\gamma}_g \\
&+\frac{\partial}{\partial\delta}(-\delta^3\nu_{2\delta}\overline{\gamma}_g\gamma_g^2 - \delta^2\beta\nu_{2\delta}^2\gamma_g^2 - 2\delta^2\beta\nu_{2\delta}\overline{\gamma}_g\gamma_g \\
&+\delta^3\nu_{2\delta}^2\gamma_g^2\gamma_{g^2} + \delta^2\beta\nu_{2\delta}^2\gamma_{g^2} + 2\delta^2\beta\gamma_g^2)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial\delta}v &= \frac{\partial}{\partial\delta}(\nu_{4\delta}(1 - \delta^2\gamma_g^2 - 2\delta\beta\gamma_g - \beta^2) \\
&-\nu_{2\delta}^2(1 - \delta^2\gamma_{g^2} - 2\delta\beta\gamma_g - \beta^2))
\end{aligned}$$

imply that $\lim_{\delta\to 0}\frac{\partial}{\partial\delta}u = 0$ and $\lim_{\delta\to 0}\frac{\partial}{\partial\delta}v = 0$. Thus we have to calculate $\frac{\partial^2}{\partial\delta^2}u$ and $\frac{\partial^2}{\partial\delta^2}v$. We obtain

$$\begin{aligned}
\frac{\partial^2}{\partial\delta^2}u &= \frac{\partial}{\partial\delta}[(\nu_{2\delta} - \beta^2\nu_{2\delta} - \nu_{2\delta}^2 + \beta^2\nu_{2\delta}^2)\overline{\gamma}_g] \\
&+\frac{\partial}{\partial\delta}(\nu_{2\delta}\overline{\gamma}_g - \beta^2\nu_{2\delta}\overline{\gamma}_g - \nu_{2\delta}^2\overline{\gamma}_g + \beta^2\nu_{2\delta}^2\overline{\gamma}_g) \\
&+\delta\frac{\partial^2}{\partial\delta^2}(\nu_{2\delta}\overline{\gamma}_g - \beta^2\nu_{2\delta}\overline{\gamma}_g - \nu_{2\delta}^2\overline{\gamma}_g + \beta^2\nu_{2\delta}^2\overline{\gamma}_g) \\
&+\frac{\partial^2}{\partial\delta^2}(-\delta^3\nu_{2\delta}\overline{\gamma}_g\gamma_g^2 - \delta^2\beta\nu_{2\delta}^2\gamma_g^2 - 2\delta^2\beta\nu_{2\delta}\overline{\gamma}_g\gamma_g \\
&+\delta^3\nu_{2\delta}^2\gamma_g^2\gamma_{g^2} + \delta^2\beta\nu_{2\delta}^2\gamma_{g^2} + 2\delta^2\beta\gamma_g^2)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2}{\partial\delta^2}v &= (1 - \delta^2\gamma_g^2 - 2\delta\beta\gamma_g - \beta^2)(\frac{\partial^2}{\partial\delta^2}\nu_{4\delta}) \\
&-4(\delta\gamma_g^2 - \beta\gamma_g)(\frac{\partial}{\partial\delta}\nu_{4\delta}) - 2\gamma_g^2\nu_{4\delta} \\
&-2\nu_\delta(1 - \delta^2\gamma_{g^2} - 2\delta\beta\gamma_g - \beta^2)(\frac{\partial^2}{\partial\delta^2}\nu_{2\delta}) \\
&-2(1 - \delta^2\gamma_{g^2} - 2\delta\beta\gamma_g - \beta^2)(\frac{\partial}{\partial\delta}\nu_{2\delta})^2 \\
&+4\nu_\delta(\delta\gamma_{g^2} - \beta\gamma_g)(\frac{\partial}{\partial\delta}\nu_{2\delta}) + 2\nu_{2\delta}^2\gamma_{g^2}.
\end{aligned}$$

Note that

$$\begin{aligned}
\lim_{\delta\to 0}\frac{\partial}{\partial\delta}\nu_{2\delta} &= \lim_{\delta\to 0}(\frac{\partial}{\partial\delta}\mathsf{E}z_t^{2\delta}) = \lim_{\delta\to 0}\int\frac{\partial}{\partial\delta}x^{2\delta}f(x)dx \\
&= \int\lim_{\delta\to 0}x^{2\delta}(\ln x^2)f(x)dx = \mathsf{E}(\ln z_t^2). \qquad (A.19) \\
\lim_{\delta\to 0}\frac{\partial^2}{\partial\delta^2}\nu_{2\delta} &= \lim_{\delta\to 0}(\frac{\partial^2}{\partial\delta^2}\mathsf{E}z_t^{2\delta}) = \lim_{\delta\to 0}\int\frac{\partial^2}{\partial\delta^2}x^{2\delta}f(x)dx \\
&= \int\lim_{\delta\to 0}x^{2\delta}(\ln x^2)^2f(x)dx = \mathsf{E}(\ln z_t^2)^2. \qquad (A.20) \\
\lim_{\delta\to 0}\frac{\partial}{\partial\delta}\overline{\gamma}_g &= \lim_{\delta\to 0}\frac{\partial}{\partial\delta}\mathsf{E}(z_t^{2\delta}g(z_t)) = \lim_{\delta\to 0}\int\frac{\partial}{\partial\delta}(x^{2\delta}g(x))f(x)dx \\
&= \int\lim_{\delta\to 0}(x^{2\delta}g(x)\ln x^2)f(x)dx = \mathsf{E}(g(z_t)\ln z_t^2). \qquad (A.21)
\end{aligned}$$

Applying (A.19)- (A.21) gives $\lim_{\delta\to 0}\frac{\partial^2}{\partial\delta^2}u$ and $\lim_{\delta\to 0}\frac{\partial^2}{\partial\delta^2}v$, respectively. We see that $\rho_1(\delta)\to\rho_1^0$ as $\delta\to 0$. ∎

A.4 Lemma 3

Consider the LGARCH (1,1) model (2.1) and (2.7) and assume $z_t \sim$nid$(0,1)$. Then the autocorrelation function of squared observations has the form

$$\rho_n = \frac{\dfrac{2\Gamma(1+0.5\alpha(\alpha+\beta)^{n-1})}{B(1+0.5\alpha(\alpha+\beta)^{n-1},0.5)}\prod_{i=1}^{n-1}\dfrac{\Gamma_{1i}(.)}{B_{1i}(.)}\prod_{i=1}^{\infty}\dfrac{\Gamma_{2i}(.)}{B_{2i}(.)} - \prod_{i=1}^{\infty}(\dfrac{\Gamma_{1i}(.)}{B_{1i}(.)})^2}{3\prod_{i=1}^{\infty}\dfrac{\Gamma_{3i}(.)}{B_{3i}(.)} - \prod_{i=1}^{\infty}(\dfrac{\Gamma_{1i}(.)}{B_{1i}(.)})^2} \qquad (A.22)$$

where

$$\begin{aligned}
\Gamma_{1i}(.)/B_{1i}(.) &= \Gamma(0.5\alpha(\alpha+\beta)^{i-1})/B(0.5\alpha(\alpha+\beta)^{i-1},0.5) \\
\Gamma_{2i}(.)/B_{2i}(.) &= \Gamma(0.5\alpha(1+(\alpha+\beta)^n)(\alpha+\beta)^{i-1})/ \\
&\quad B(0.5\alpha(1+(\alpha+\beta)^n(\alpha+\beta)^{i-1},0.5) \\
\Gamma_{3i}(.)/B_{3i}(.) &= \Gamma(\alpha(\alpha+\beta)^{i-1})/B(\alpha(\alpha+\beta)^{i-1},0.5)
\end{aligned}$$

and $\Gamma(\cdot)$ and $B(\cdot)$ are the Gamma function and the Beta function, respectively.

Proof. See He, Teräsvirta and Malmsten (1999). ∎

Chapter 3

Stylized facts of financial time series and three popular models of volatility

3.1 Introduction

Modelling volatility of financial series such as stock returns has become common practice, as the demand for volatility forecasts has increased. Several types of models such as models of autoregressive conditional heteroskedasticity and stochastic volatility models have been applied for the purpose. A practitioner can thus choose between a variety of models. A popular way of comparing volatility models has been to estimate a number of models by maximum likelihood and observe which one has the highest log-likelihood value; see Shephard (1996) for an example. If the models under comparison do not have the same number of parameters, one may want to favour parsimony and apply a suitable model selection criterion, such as AIC or BIC, for the purpose. It is also possible to choose a model after actually applying it to forecasting. Poon and Granger (2003) provide a survey of papers that contain results of such comparisons.

Another way of comparing models is to submit estimated models to misspecification tests and see how well they pass the tests. This also paves the way for building models within the same family of models. One can extend a failed model by estimating the alternative it has been tested against and subject that model to new misspecification tests. Such tests have been derived for generalized autoregressive conditional heteroskedasticity (GARCH) models; see, for example, Engle and Ng (1993), Chu (1995), Lin and Yang (1999), and Lundbergh and Teräsvirta (2002). Similar devices for the exponential GARCH model (EGARCH) of Nelson (1991) who already suggested such tests, are presented in Chapter 4. In addition, nonnested models can be tested against each other. Kim, Shephard and Chib (1998) considered testing GARCH against the autoregressive stochastic volatility (ARSV) model and Lee and Brorsen (1997) suggested the simulated likelihood ratio test for choosing between GARCH and EGARCH. The pseudo-score test of Chen and Kuan (2002) can be applied to this problem as well. Small sample properties of some of the available tests for that testing problem are considered in Chapter 4. It should be noted, however, that testing two models against each other does not necessary lead to a unique choice of a model. Neither model may be rejected against the other or both may be rejected against each other. For a discussion of conceptual differences between the model selection and testing approaches, see Granger, King and White (1995).

The purpose of this paper is to compare volatility models from another angle. Financial time series of sufficiently high frequency such as daily or weekly or even intradaily stock or exchange rate return series seem to share a number of characteristic features, sometimes called stylized facts. Granger and

Ding (1995a) and Granger, Ding and Spear (2000), among others, pointed out such features and investigated their presence in financial time series. Given a set of characteristic features or stylized facts, one may ask the following question: "Have popular volatility models been parameterized in such a way that they can accommodate and explain the most common stylized facts visible in the data?" Models for which the answer is positive may be viewed as suitable for practical use. The other parameterizations may be regarded as less useful in practice.

There exists some work towards answering this question. Teräsvirta (1996) considered the ability of the GARCH model to reproduce series with high kurtosis and, at the same time, positive but low and slowly decreasing autocorrelations of squared observations. Liesenfeld and Jung (2000) discussed this stylized fact in connection with the ARSV model whereas Andersson (2001) focussed on the ARSV model based on the normal inverse Gaussian distribution. Carnero, Peña and Ruiz (2001) compared the ARSV model and the GARCH model using the kurtosis-autocorrelation relationship as their benchmark. The work of Rydén, Teräsvirta and Åsbrink (1998) on the hidden Markov model for the variance may also be mentioned in this context.

Answering the question by using the approach of this paper is only possible in the case of rather simple models. On the other hand, a vast majority of popular models such as GARCH, EGARCH and ARSV models used in applications are first-order models. Higher-order models, although theoretically well-defined, are rather seldom used in practice. This suggests that restricting the considerations to simple parameterizations does not render the results useless.

The plan of the paper is as follows. The stylized facts are defined in Section 2 and the models are discussed in Section 3. Section 4 considers the kurtosis-autocorrelation relationship. In Section 5 the Taylor effect are discussed. In Section 6 the kurtosis-autocorrelation relationship is reconsidered using confidence regions. Finally, conclusions appear in Section 7.

3.2 Stylized facts

The stylized facts to be discussed in this paper are illuminated by Figure 3.1. The first panel depicts the return series of the S&P 500 stock index (daily first differences r_t of logarithms of the index; 19261 observations) from 3 January 1928 to 24 April 2001. The marginal distribution of r_t appears leptokurtic and a number of volatility clusters are clearly visible. The volatility models considered in this study are designed for parameterizing this type of variation. The second panel shows the autocorrelation function of $|r_t|^m$, $m = 0.25, 0.5, 0.75, 1$

and the third one the corresponding function for $m = 1, 1.25, 1.5, 1.75, 2$ for the first 500 lags. It is seen that the first autocorrelations have positive but relatively small values and that the autocorrelations decay slowly. A similar figure can be found in Ding et al. (1993), but here the time series has been extended to cover ten more years from 1992 to 2001.

The first stylized fact illustrated by Figure 3.1 and typical of a large amount of return series is the combination of relatively high kurtosis and rather low autocorrelations of $|r_t|^m$. In the case of the standard GARCH model, we restrict ourselves to inspect the combination of kurtosis and the autocorrelations of r_t^2 because in that case, an analytic expression for the autocorrelation function is available. The second stylized fact to be considered is the fact that the autocorrelations as a function of m tend to peak for $m = 1$. This is the so-called Taylor effect that has been found in a large number of financial time series; see Granger and Ding (1995a) and Granger et al. (2000). In the GARCH framework, this stylized fact can only be investigated using analytic expressions when the GARCH model is the so-called absolute-value GARCH (AVGARCH) model and $m = 1$ or $m = 2$. This is because no analytical expressions for $\rho(|r_t|^m, |r_{t-j}|^m)$ exist when $m < 2$ and the model is the standard GARCH model. For the AVGARCH model, they are available for both $m = 1$ and $m = 2$ but not for non-integer values of m.

Yet another fact discernible in Figure 3.1 is that the decay rate of the autocorrelations is very low, apparently lower than the exponential rate. This has prompted some investigators to introduce the fractionally integrated GARCH (FIGARCH) model; see Baillie, Bollerslev and Mikkelsen (1996). In this paper, this slow decay is not included among the stylized facts under consideration. To illustrate the reason, we split the S&P 500 return series into 20 subseries of 980 observations each, and estimate the autocorrelations $\rho(|r_t|, |r_{t-j}|)$, $j = 1, ..., 500$, for these subseries. The fourth panel of Figure 3.1 contains these autocorrelations for the whole series and the mean of the corresponding autocorrelations of the 20 subseries together with the plus/minus one standard deviation band. It is seen that the decay of autocorrelations in the subseries on the average is substantially faster than in the original series and roughly exponential. This lack of self-similarity in autocorrelations can be taken as evidence against the FIGARCH model in this particular case, but that is beside the point. We merely want to argue that the very slow decay rate of the autocorrelations of $|r_t|$ or r_t^2 may not necessarily be a feature typical of series with a couple of thousand observations. As such series are most often modelled by one of the standard models of interest in this study, we do not consider very slow decay of autocorrelations a stylized fact in our discussion.

3.3 The models and their fourth-moment structure

3.3.1 GARCH model

Suppose an error term or an observable variable can be decomposed as follows:

$$\varepsilon_t = z_t h_t^{1/2}, \tag{3.1}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with zero mean. Furthermore, assume that

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{3.2}$$

Equations (3.1) and (3.2) define the standard GARCH(p,q) model of Bollerslev (1986). Parameter restrictions are required to ensure positiveness of the conditional variance h_t in (3.2). Assuming $\alpha_j \geqslant 0$, $j = 1, ..., q$, and $\beta_j \geqslant 0$, $j = 1, ..., p$, is sufficient for this. Both necessary and sufficient conditions were derived by Nelson and Cao (1992). In this paper we shall concentrate on (3.1) with (3.2) assuming $p = q = 1$. This is done for two reasons. First, the GARCH(1,1) model is by far the most frequently applied GARCH specification. Second, we want to keep our considerations simple.

The GARCH(1,1) model is covariance stationary if

$$\alpha_1 \nu_2 + \beta_1 < 1, \tag{3.3}$$

where $\nu_2 = \mathsf{E} z_t^2 < \infty$. For the discussion of stylized facts we need moment condition and fourth moments of $\{\varepsilon_t\}$. Assuming $\nu_4 = \mathsf{E} z_t^4 < \infty$, the unconditional fourth moment for the GARCH(1,1) model exists if and only if

$$\alpha_1^2 \nu_4 + 2\alpha_1 \beta_1 \nu_2 + \beta_1^2 < 1. \tag{3.4}$$

Under (3.4) the kurtosis of ε_t equals

$$\kappa_4 = \frac{\kappa_4(z_t)\{1 - (\alpha_1 \nu_2 + \beta_1)^2\}}{1 - (\alpha_1^2 \nu_4 + 2\alpha_1 \beta_1 \nu_2 + \beta_1^2)}, \tag{3.5}$$

where $\kappa_4(z_t) = \nu_4/\nu_2^2$ is the kurtosis of z_t. Assuming normality, one obtains the following well-known result:

$$\kappa_4 = 3\frac{1 - (\alpha_1 + \beta_1)^2}{1 - (3\alpha_1^2 + 2\alpha_1 \beta_1 + \beta_1^2)} > 3. \tag{3.6}$$

Furthermore, when (3.4) holds, the autocorrelation function of $\{\varepsilon_t^2\}$ is defined as follows:

$$\rho_n = (\alpha_1\nu_2 + \beta_1)^{n-1}\frac{\alpha_1\nu_2(1-\beta_1^2-\beta_1\alpha_1\nu_2)}{1-\beta_1^2-2\beta_1\alpha_1\nu_2}\quad n \geq 1. \tag{3.7}$$

The autocorrelation function of $\{\varepsilon_t^2\}$ is dominated by an exponential decay from the first lag with decay rate $\alpha_1\nu_2 + \beta_1$. Setting $\nu_2 = 1$ and $\nu_4 = 3$ (normality) in (3.7) gives the result in Bollerslev (1988). Note that the existence of the autocorrelation function does depend on the existence of ν_4 although (3.7) is not a function of ν_4. The necessary and sufficient conditions for the existence of the unconditional fourth moments of the GARCH(p,q) process and the expressions (3.5) and (3.7) are special cases of results in He and Teräsvirta (1999*a*).

3.3.2 EGARCH model

Nelson (1991) who introduced the EGARCH model listed three drawbacks with the GARCH models. First, the lack of asymmetry in the response of shocks. Secondly, the GARCH models impose parameter restrictions to ensure positivity of the conditional variance. Finally, measuring the persistence is difficult. Consider (3.1) with

$$\ln h_t = \alpha_0 + \sum_{j=1}^{q}(\phi_j z_{t-j} + \psi_j(|z_{t-j}| - \mathsf{E}\,|z_{t-j}|)) + \sum_{j=1}^{p}\beta_j \ln h_{t-j} \tag{3.8}$$

which defines the EGARCH(p,q) model of Nelson (1991). It is seen from (3.8) that no parameter restrictions are necessary to ensure positivity of h_t. The moment structure of the EGARCH(p,q) model has been worked out in He (2000). As in the GARCH case, the first-order model is the most popular EGARCH model. The term $\psi(|z_{t-1}| - \mathsf{E}\,|z_{t-1}|)$ represents a magnitude effect in the spirit of the GARCH(1,1) model. The term ϕz_t represents the asymmetry effect. Nelson (1991) derived existence conditions for moments of the EGARCH(1,1) model. They can be summarized by saying that if the error process $\{z_t\}$ has all moments then all moment for the EGARCH(1,1) process exists if and only if

$$|\beta| < 1. \tag{3.9}$$

For example, if $\{z_t\}$ is standard normal then the restriction (3.9) guarantee the existence of all moments simultaneously. This is different from the

GARCH model. For that model, the moment conditions become more and more restrictive for higher moments.

Another difference between the GARCH models and the EGARCH model is that for the latter analytical expressions exists for all moments of $|\varepsilon_t|^{2m}$, $m > 0$. The moments can be found in Chapter 1; see also Nelson (1991). If (3.9) holds, then the kurtosis of ε_t, with $z_t \sim$nid(0,1), is given by

$$\kappa_4 = 3\exp\left\{\frac{(\psi+\phi)^2}{1-\beta^2}\right\}\times$$
$$\prod_{i=1}^{\infty}\frac{\Phi(2\beta^{i-1}(\psi+\phi))+\exp\{-8\beta^{2(i-1)}\psi\phi\}\Phi(2\beta^{i-1}(\psi-\phi))}{[\Phi(\beta^{i-1}(\psi+\phi))+\exp\{-2\beta^{2(i-1)}\psi\phi\}\Phi(\beta^{i-1}(\psi-\phi))]^2} > 3, \quad (3.10)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The expression contain infinite products, and care is therefore required in computing them (selecting the number of terms in the product). Setting $\psi = 0$ in (3.10) yields a simple formula

$$\kappa_4 = 3\exp\{\phi^2(1-\beta^2)^{-1}\} > 3. \quad (3.11)$$

If (3.9) holds, the autocorrelation function for $|\varepsilon_t|^{2m}$, with $z_t \sim$nid(0,1), has the form

$$\rho_n(m) = \frac{A\exp\left\{\frac{m^2(\psi+\phi)^2(\beta^{2(n-1)}(\beta^2-1)/4+\beta^n)}{1-\beta^2}\right\}D(\cdot)\prod_{i=1}^{n-1}\Phi_{1i}\prod_{i=1}^{\infty}\Phi_{2i}-\prod_{i=1}^{\infty}\Phi_{1i}^2}{\frac{\pi^{1/2}\Gamma(2m+1/2)}{(\Gamma(m+1/2))^2}\exp\left\{\frac{m^2(\psi+\phi)^2}{1-\beta^2}\right\}\prod_{i=1}^{\infty}\Phi_{3i}-\prod_{i=1}^{\infty}\Phi_{1i}^2}, n \geq 1 \quad (3.12)$$

where

$$\begin{aligned}
A &= \frac{\Gamma(2m+1)}{2^{m+1/2}\Gamma(m+1/2)},\\
D(\cdot) &= D_{-(2m+1)}[-m\beta^{n-1}(\psi+\phi)] +\\
&\quad \exp\{-m^2\beta^{2(n-1)}\psi\phi\}D_{-(2m+1)}[-m\beta^{n-1}(\psi-\phi)],\\
\Phi_{1i} &= \Phi(m\beta^{i-1}(\psi+\phi)) + \exp\{-2m^2\beta^{2(i-1)}\psi\phi\}\Phi(m\beta^{i-1}(\psi-\phi)),\\
\Phi_{2i} &= \Phi(m\beta^{i-1}(1+\beta^n)(\psi+\phi)) +\\
&\quad \exp\{-2m^2\beta^{2(i-1)}(1+\beta^n)^2\psi\phi\}\Phi(m\beta^{i-1}(1+\beta^n)(\psi-\phi)),
\end{aligned}$$

and

$$\Phi_{3i} = \Phi(2m\beta^{i-1}(\psi+\phi)) + \exp\{-8m^2\beta^{2(i-1)}\psi\phi\}\Phi(2m\beta^{i-1}(\psi-\phi)).$$

Furthermore, $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution and

$$D_{(-p)}[q] = \frac{\exp\{-q^2/4\}}{\Gamma(p)} \int_0^\infty x^{p-1} \exp\{-qx - x^2/2\} dx, \; p > 0,$$

is the parabolic cylinder function where $\Gamma(\cdot)$ is the gamma function. If $\phi = 0$ or $\psi = 0$ in the EGARCH(1,1) model the resulting autocorrelation function becomes quite simple; see Chapter 1. The autocorrelation function of the squared observations ($m = 1$), when $\psi = 0$, has the simple form

$$\rho_n(1) = \frac{(1 + \phi^2\beta^{2(n-1)}) \exp\{\phi^2\beta^n(1-\beta^2)^{-1}\} - 1}{3\exp\{\phi^2(1-\beta^2)^{-1}\} - 1}, n \geq 1. \tag{3.13}$$

To illustrate the above theory, consider the case $0 < \beta < 1$. The decay of the autocorrelations is controlled by the parameter β. The autocorrelation function of $\{|\varepsilon_t|^{2m}\}$ then appear to have the property that the decay rate is faster than exponential at short lags and stabilizes to β as the lag length increases. For the special case (3.13) this can be shown analytically, but in the general case it is just a conjecture based on numerical calculations; see the table in Chapter 1.

3.3.3 ARSV model

The ARSV model offers yet another way of characterizing conditional heteroskedasticity. See Ghysels, Harvey and Renault (1996) for a survey on the properties of the ARSV model. It bears certain resemblance to the EGARCH model. As with the EGARCH model, defining the dynamic structure using $\ln h_t$ and its lags ensures that h_t is always positive, but the difference to the GARCH model and the EGARCH model is that it does not depend on past observations but on some unobserved latent variable instead. The simplest and most popular ARSV(1) model, Taylor (1986), is given by

$$\varepsilon_t = \sigma z_t h_t^{1/2}, \tag{3.14}$$

where σ is a scale parameter, which removes the need for a constant term in the first-order autoregressive process

$$\ln h_{t+1} = \beta \ln h_t + \eta_t. \tag{3.15}$$

In (3.15), $\{\eta_t\}$ is a sequence of independent normal distributed random variables with mean zero and a known variance σ_η^2. The error processes $\{z_t\}$

and $\{\eta_t\}$ are assumed to be mutually independent. One motivation for the EGARCH model was the need to capture the non-symmetric response to the sign of the shock. If z_t and η_t are assumed to be correlated with each other, the ARSV(1) model also allows for asymmetry. The model can be generalized so that $\ln h_t$ follows an ARMA(p, q) process, but in this work we only consider the ARSV(1) model.

As η_t is normally distributed, $\ln h_t$ is also normally distributed. From standard theory we know that all moments of $\ln h_t$ exist if and only if

$$|\beta| < 1 \tag{3.16}$$

in (3.15). Thus, if $|\beta| < 1$ and all moments of z_t exist then all moment of ε_t in (3.14) exist as well, as they do in the EGARCH(1,1) model. If condition (3.16) is satisfied, the kurtosis of ε_t is given by

$$\kappa_4 = \kappa_4(z_t) \exp\{\sigma_h^2\}, \tag{3.17}$$

where $\sigma_h^2 = \sigma_\eta^2/(1-\beta^2)$ is the variance of $\ln h_t$. Thus $\kappa_4 > \kappa_4(z_t)$, so that if $z_t \sim$nid(0,1), ε_t is leptokurtic. Formula (3.17) bears considerable resemblance to (3.11). In the ARSV(1) model (3.14) and (3.15), z_t and η_t are independent. The same is true for z_t and z_{t-1} in the EGARCH(1,1) model. When $\psi_1 = 0$ in the latter model, the moment expressions for the two models therefore look alike.

As in EGARCH models it is possible to derive the autocorrelation function for any $|\varepsilon_t|^{2m}$, $m > 0$, when $\{\varepsilon_t\}$ obeys an ARSV(1) model (3.14) and (3.15). When (3.16) holds, then the autocorrelation function of $\{|\varepsilon_t|^{2m}\}$ is defined as follows, see Ghysels et al. (1996):

$$\rho_n(m) = \frac{\exp(m^2\sigma_h^2\beta^n) - 1}{\kappa_m \exp(m^2\sigma_h^2) - 1}, n \geqslant 1, \tag{3.18}$$

where κ_m is

$$\kappa_m = E\,|z_t|^{4m} / (E\,|z_t|^{2m})^2. \tag{3.19}$$

The autocorrelation function of $\{|\varepsilon_t|^{2m}\}$ has the property that the decay rate is faster than exponential at short lags and stabilizes to β as the lag length increases, analogously to the EGARCH model. Thus, the decay of the autocorrelations is controlled by β only.

3.4 Kurtosis-autocorrelation relationship

3.4.1 GARCH(1,1) model

The above results make it possible to consider how well the models fits the first stylized fact of financial time series mentioned in Section 2: leptokurtosis and low but rather persistent autocorrelation of the squared observations or errors. Consider GARCH(1,1) model with normal errors and express the autocorrelation function (3.7) as a function of the kurtosis (3.5). This yields

$$\rho_n = (\alpha_1 + \beta_1)^{n-1}\left(\frac{\beta_1(1 - 3\kappa_4^{-1})}{3(1 - \kappa_4^{-1})} + \alpha_1\right), n \geq 1. \tag{3.20}$$

Figure 3.2 illuminates the relationship between the kurtosis κ_4 and the autocorrelation ρ_1. It contains isoquants, curves defined by sets of points for which the sum $\alpha_1 + \beta_1$ has the same value. The kurtosis and the first-order autocorrelation of squared observations are both increasing functions of α_1 when $\alpha_1 + \beta_1$ equals a constant. They all start at $\kappa_4 = 3$ and $\rho_1 = 0$ where $\alpha_1 = 0$ and the GARCH(1,1) model is unidentified (the conditional variance equals unity). For previous examples of similar figures, see Teräsvirta (1996), Liesenfeld and Jung (2000) and Andersson (2001). It is seen from the present figure that the first-order autocorrelation first increases rapidly as a function of the kurtosis (and α_1) and that the increase gradually slows down. It is also clear that the autocorrelation decreases as a function of $\alpha_1 + \beta_1$ when the kurtosis is held constant. Nevertheless, low autocorrelations cannot exist with high kurtosis.

This figure offers a useful background for studying the observed kurtosis-autocorrelation combinations. Figure 3.3 contains the same isoquants as the Figure 3.2, together with kurtosis-autocorrelation combinations estimated from observed time series. The upper-left panel contains them for 27 return series of the most frequently traded stocks in the Stockholm stock exchange, see also Chapter 4. There seems to be large variation among the series. A large majority have an unreachable combination of κ_4 and ρ_1 in the sense that the combinations do not correspond to a GARCH(1,1) with a finite variance ($\alpha_1 + \beta_1 < 1$). Only four observations appear in the area defined by $\alpha_1 + \beta_1 < 0.999$. The upper-right panel gives a less variable picture. The rates of return are the 20 subperiods of the return series of the S&P 500 discussed in Section 2. Three of them do not appear in the panel. The subseries containing the October 1987 crash do not appear in the panel because the kurtosis is 87.3. 15 of them lie out of reach for the GARCH(1,1) model with normal errors. The lower-left panel tells a similar story. The rates of return are 34 subseries of five major exchange rates, the Japanese yen, the German mark, the English pound, the Canadian dollar, and the Australian dollar, all against the U.S. dollar, from 2 April 1973 to 10 September 2001. One of them, the first subseries of the Canadian dollar, does not appear in the panel because the autocorrelation is 0.456. The lower-right panel contains all data-points in the three other panels. It is seen from the figure that a majority of the points lie even below the lowest isoquant $\alpha_1 + \beta_1 = 0.999$. An obvious conclusion is that the GARCH(1,1) model with normal errors cannot in a satisfactory fashion reproduce the stylized fact of high kurtosis and low-starting autocorrelation of squares observed in a large number of financial series. This is true at least if we require the existence of the unconditional fourth moment of ε_t. We shall return to this point in Section 6.

The first-order autocorrelation of ε_t^2 does decrease with $\alpha_1 + \beta_1$ when the kurtosis is kept constant. This may suggest that an integrated GARCH model of Engle and Bollerslev (1986) could offer an adequate description of the stylized fact. The first-order IGARCH model is obtained by setting $\alpha_1 + \beta_1 = 1$ in (3.2), which implies that the GARCH process does not have a finite variance. As there are no moment results to rely on, this idea has been investigated by simulation. Figure 3.4 contains the same isoquants as before, completed with 100 kurtosis-autocorrelation combinations obtained by simulating the first-order IGARCH with $\beta_1 = 0.9$. The number of observations increases from $T = 100$ in the upper-left panel to 10000 in the lower-right one. It is quite clear that for $T = 100$, it is difficult to even argue that the observations come from a GARCH model. For about the half of the observations, the process appears platykurtic and for a third, the first-order autocorrelation of squared observations is negative. One conclusion is that when the null of no conditional heteroskedasticity is rejected for the errors of a macroeconomic equation, estimated using a small number of quarterly observations, fitting an ARCH or a GARCH model to the errors is hardly a sensible thing to do.

Another conclusion, relevant for our stylized fact considerations, is that when the number of observations increases, the point cloud in the figure moves to the right. This is what it should do as the fourth moment of ε_t does not exist. However, the points follow the isoquants on their way out of the frame, and they do not cross the area where most of the observations were found in Figure 3.3. The conclusion from this small simulation experiment therefore is that the IGARCH model is not the solution to the problem that the GARCH(1,1) model with normal errors does not accord with this particular stylized fact.

Most researchers nowadays do not assume normal errors for z_t in (3.1) but rather make use of a leptokurtic error distribution such as the t-distribution.

Why this is the case can be seen from Figure 3.5. It contains the same isoquants as before, measured by $\alpha_1\nu_2+\beta_1$. This is the condition for covariance stationarity just as $\alpha_1+\beta_1<1$ in Figure 3.1 is in the case of normal errors. It depends on the degrees of freedom of the t-distribution through ν_2. In the left-hand panel the t-distribution has seven degrees of freedom so that $\kappa_4=5$ and in the right-hand panel five, in which case $\kappa_4=9$. Figure 3.5 also contains the kurtosis/autocorrelation combinations for the series shown in the fourth panel of Figure 3.3 but now under the assumption that the errors have a t-distribution with seven (left panel) and five degrees of freedom (right panel). It is seen how the baseline kurtosis now increases from three to five (left panel) and nine (right panel). The observations now fall inside the fan of isoquants, and the corresponding GARCH(1,1) model with the finite fourth moment is sufficiently flexible to characterize the stylized fact of high kurtosis and low autocorrelation of squared observations.

3.4.2 EGARCH(1,1) model

The GARCH(1,1) model with normal errors does not adequately describe the stylized fact of high kurtosis/low autocorrelation of squares combinations. In this section we consider the situation in the symmetric EGARCH(1,1) model. The relationship between κ_4 and ρ_1 for three symmetric EGARCH(1,1) models, $\phi=0$, with normal errors with different persistence measured by β is depicted in Figure 3.7[1]. The isoquants now contain the points with β being a constant, while ψ is changing. The kurtosis is a monotonically increasing function of ψ. This figure shows that large values of κ_4 and low values of ρ_1 cannot exist simultaneously for the symmetric EGARCH(1,1) model either. The lowest values for ρ_1 are obtained when β is close to one but these values are not sufficiently low to reach down where the data-points are.

Nelson (1991) recommends the use of generalized error distribution (GED(υ)) for the errors. Granger et al. (2000) used the double exponential (Laplace) distribution. The GED(υ) includes both the normal distribution, $\upsilon=2$, and the Laplace distribution, $\upsilon=1$, as special cases. If $\upsilon\leq 1$, restrictions on ψ (and ϕ) are needed to guarantee finite moments. Note that the t-distribution for the errors may imply an infinite unconditional variance for $\{\varepsilon_t\}$. For a detailed discussion, see Nelson (1991). The autocorrelations of $\{|\varepsilon_t|^{2m}\}$ with $z_t\sim$GED(υ) can be found in Chapter 1.

[1]For the EGARCH(1,1) model with $\psi=0$ we can express the first-order autocorrelation of squares as function of kurtosis: $\rho_n(1)=\frac{(1+\phi^2)(\kappa_4/3)^{\beta^n}-1}{\kappa_4-1}$.

3.4.3 ARSV(1) model

In order to complete our scrutiny of the kurtosis/auocorrelation relationship we consider the first-order ARSV model. Carnero et al. (2001) have also done similar work. The autocorrelation function of $\{\varepsilon_t^2\}$ of the ARSV(1) model can be expressed as a function of the kurtosis as follows:

$$\rho_n(1) = \frac{(\kappa_4/\kappa_4(z_t))^{\beta^n}}{\kappa_4 - 1}, n \geq 1. \tag{3.21}$$

Figure 3.8 contains a plot of the relationship between κ_4 and $\rho_1(1)$ for three ARSV(1) models with normal errors ($\kappa_4(z_t) = 3$) with different persistence measures β. The isoquants now consist of the points with β being 0.95, 0.99, 0.999, respectively, while σ_η^2 is changing. The kurtosis is a monotonically increasing function of σ_η^2. An important difference between the EGARCH(1,1) model and the ARSV(1) model lies in the behaviour of the first-order autocorrelation when the kurtosis is held constant. In the EGARCH(1,1) model, the value of the autocorrelation decreases as a function of β_1, the parameter that controls the decay rate of the autocorrelations. In the ARSV(1) model this value increases as a function of the corresponding parameter β. Thus, contrary to the EGARCH model, a low first-order autocorrelation and high persistence can coexist in the ARSV model. In general, the first-order autocorrelations, given the kurtosis, are lower in the ARSV than the EGARCH model with normal errors. This may at least partly explain the fact that in some applications the ARSV(1) model seems to fit the data better than its EGARCH or GARCH counterpart. It may also explain the stylized fact mentioned in Shephard (1996) that β estimated from an ARSV(1) model tends to be lower than the sum $\alpha_1 + \beta_1$ estimated from a GARCH(1,1) model.

In Figure 3.6 the errors have a t-distribution with seven (left panel) and five degrees of freedom (right panel). It is seen that when the number of degrees of freedom in the t-distribution decreases, the first-order autocorrelation becomes practically independent of the persistence parameter β. At the same time, the value of the autocorrelation rapidly decreases with the number of degrees of freedom for any given σ_η^2. Compared to the GARCH(1,1) model, the difference is quite striking.

3.5 Taylor effect

3.5.1 GARCH(1,1) model

As discussed in Section 2, a large number of financial series display an autocorrelation structure such that the autocorrelation of $|\varepsilon_t|^{2m}$ decay slowly and the autocorrelations as a function of $m > 0$ peak around $m = 0.5$. He and Teräsvirta (1999) defined the corresponding theoretical property and called it the Taylor property. From the results in Section 3 it follows that the existence of the Taylor property in the EGARCH(1,1) and ARSV(1) models can be considered analytically because the analytic expressions for $\mathsf{E}\,|\varepsilon_t|^{2m}$ exist for any $m > 0$. This is not true for the GARCH models, however, because analytic expressions are available only for integer moments. An exception is the power-GARCH model of Ding et al. (1993). For this model, certain non-integer moments have an analytic definition, but then, the integer moments generally do not; see Chapter 2.

One can think of considering a more restricted form of definition that only concerns the first and second moment. The model is then said to have the Taylor property if

$$\rho(|\varepsilon_t|, |\varepsilon_{t-n}|) > \rho(|\varepsilon_t|^2, |\varepsilon_{t-n}|^2), n \geqslant 1. \tag{3.22}$$

This can be defended by referring to the original discussion in Taylor (1986). The problem is that for the standard GARCH model, an analytic definition of $\mathsf{E}\,|\varepsilon_t|$ as a function of the parameters is not available. On the other hand, it exists for the AVGARCH(1,1) model defined by Taylor (1986) and Schwert (1989). This prompted He and Teräsvirta (1999) to discuss the existence of the Taylor property in the AVGARCH(1,1) model. Their conclusion, based on considerations with $n = 1$ in (3.22), was that the AVGARCH model possesses the Taylor property if the kurtosis of the model is sufficiently large. However, the difference between the autocorrelations of $|\varepsilon_t|$ and ε_t^2 remains very small even when the kurtosis is very large. These authors also investigated the existence of the Taylor property in the standard GARCH(1,1) model by simulation, and their results suggested that this model does not have the Taylor property. Of course, due to sample uncertainty, the GARCH model can still generate realizations displaying the Taylor effect, at least when the number of observations is relatively small. This would not, however, happen at the frequency with which the Taylor effect is found in financial series; see Granger and Ding (1995a).

3.5.2 EGARCH(1,1) model

We extend the considerations in He and Teräsvirta (1999) to the EGARCH(1,1) and ARSV(1) model. For the other two model, the situation is different. The results of Section 3 allow us to say something about the capability of the EGARCH(1,1) model to generate series with the Taylor property. Figure 3.9 contains a description of the relationship between κ_4 and the two first-order autocorrelations $\rho_1(m)$, $m = 1, 0.5$, for $\beta = 0.95$ and $\beta = 0.99$. It is seen that the Taylor property is present at high values of the kurtosis. The value of the kurtosis where the Taylor property is present decreases as a function of β. The difference between the two first-order autocorrelations is substantially greater than in the AVGARCH(1,1) model.

As analytical expressions for non-integer moments of $\mathsf{E}\left|\varepsilon_t\right|^{2m}, m > 0$, exist for the EGARCH model, we can extend our considerations by use of them. Figure 3.10 contains graphs showing the first-order autocorrelation as a function of the exponent m for $\beta = 0.95$ and $\beta = 0.99$ at three different kurtosis values. It turns out that for the symmetric EGARCH process, with kurtosis of the magnitude found in financial time series, the maximum appears to be attained for m around 0.5. The conclusion is that the Taylor property is satisfied for an empirically relevant subset of EGARCH(1,1) models.

3.5.3 ARSV(1) model

In order to complete our discussion about Taylor effect we consider the ARSV model. Figure 3.12 illustrates the relationship between κ_4 and the two first-order autocorrelations $\rho_1(m)$, $m = 1, 0.5$, for $\beta = 0.95$ and $\beta = 0.99$. It is seen that the Taylor property is present already at low values of the kurtosis.

Analogously to the preceding subsection, Figure 3.11 contains a graph showing the first-order autocorrelation as a function of m for $\beta = 0.95$ and $\beta = 0.99$ and the three different kurtosis values. There is a difference between the EGARCH(1,1) model and the ARSV(1) model regarding the peak value of $\rho_1(m)$ when the persistence parameter changes. In the EGARCH(1,1) model, the peak of the autocorrelation moves to left with higher β_1. In the ARSV(1) model, increasing β shifts the peak of the autocorrelation to the right. This feature demonstrates the difference in the relationship between the persistence and the first-order autocorrelation in these two models. Nevertheless, the general conclusion even here is that for the ARSV(1) model, there exists an empirically relevant subset of these models such that the definition of the Taylor property is approximately satisfied. Thus both the ARSV(1) and the EGARCH(1,1) model appear to reproduce this stylized fact considerably better than the first-order GARCH model.

3.6 Confidence regions for the kurtosis/autocorrelation combination

When the kurtosis-autocorrelation combination and volatility models were discussed in Section 4, the observations were treated as fixed for simplicity. In reality, they are estimates based on time series. This being the case, it would be useful to account for the uncertainty of these estimates and see whether that would change the conclusions offered in Section 4. For this purpose it becomes necessary to estimate confidence regions for kurtosis-autocorrelation combinations.

First it may be noted that it is not possible to obtain these confidence regions analytically. The kurtosis and first-order autocorrelation of squared observations are nonlinear functions of the parameters of the model, be that a GARCH, an EGARCH or an ARSV model. Furthermore, there is no one-to-one mapping between the two parameters of interest and the parameters in the three models. This implies that the confidence regions have to be obtained by simulation. As an example, suppose that the true model generating the time series is a GARCH(1,1) one with a finite fourth moment and fit this model to the series. Use the formulas (3.5) and (3.7) to obtain the plug-in estimate of the kurtosis-autocorrelation combinations. Next, use the asymptotic distribution of the maximum likelihood estimator of the parameters and the same formulas to obtain a random sample of kurtosis-autocorrelation combinations from this distributions. The element that fail the fourth-order moment conditions are discarded and the remaining ones used for construting confidence intervals.

In order to illustrate the situation, consider Figure 3.13 that contains 200 kurtosis-autocorrelation combinations generated from an estimated GARCH(1,1) model. The original time series has been generated from a GARCH(1,1) model with parameters $\alpha_0 = 0.05, \alpha_1 = 0.19121, \beta_1 = 0.75879$. A striking feature is that the point cloud has a form of a sausage that appears to be shaped by the isoquants also included in the figure. This feature has an important consequence: estimating the joint density function of the two variables, kurtosis and autocorrelation estimators, is hardly possible by applying a bivariate kernel estimator based on a linear grid. The problem is that the linear grid would cover areas where no observations are located. Kernel estimation can instead be carried out by replacing the linear grid by a particular nonlinear one that makes use of the isoquants; see Eklund (2004) for details. Desired confidence intervals are then obtained as highest density regions; for computational details, see Hyndman (1996).

As an application we consider two daily return series of stocks traded in the Stockholm stock exchange. For the stock Assi D the kurtosis equals 5.8, and the first-order autocorrelation of squared returns equals 0.305. The solid square in Figures 14, 15 and 16 represents this kurtosis/autocorrelation pair. After estimating the three models, the plug-in estimate of the kurtosis/autocorrelation pair can be obtained for each model, and the solid circle represents the estimated pair in the three figures. To estimate the ARSV model we use the quasi-maximum likelihood estimator suggested in Ghysels et al. (1996). Finally, the solid lines define the 90% confidence regions of the true kurtosis/autocorrelation pair.

For the GARCH model in Figure 3.14 the deviation of the plug-in estimated kurtosis/autocorrelation point from the directly estimated pair is small, and the directly estimated combination remains inside the 90% confidence region. For the EGARCH model in Figure 3.15 and for the ARSV model in Figure 16 the autocorrelation is underestimated. However, for the EGARCH(1,1) model in Figure 3.15 the directly estimated combination remains inside the 90% confidence region, whereas this is not the case for the ARSV(1) model, see Figure 3.16. This is probably due to the fact, discussed in Section 4.3, that the persistence parameter β does not play a large role in the determination of the autocorrelation of squared observations.

Next we consider a more typical return series that has a combination of kurtosis and first-order autocorrelation of squares that lies below even the lowest isoquants for the GARCH model in Figure 3.3 and the EGARCH model in Figure 3.7. This is the return series for the stock SEB that has kurtosis 18.0 and the autocorrelation of squares 0.267. For the GARCH model in Figure 3.17 the kurtosis/autocorrelation combination is heavily underestimated. The 90% confidence region does not cover the directly estimated kurtosis/autocorrelation combination. The same is true for the EGARCH model and the ARSV model, see Figure 3.18 (the latter not reported here). A conclusion of this small application, under the assumption that the observations have been generated from a member of the family of models in question, is that the GARCH(1,1) model and the EGARCH(1,1) model cannot reproduce the stylized fact of high kurtosis and low-starting autocorrelation of squares even if we account for the uncertainty. For low kurtosis the GARCH(1,1) and the EGARCH(1,1) model appear to work better than the first-order ARSV model.

3.7 Conclusions

In this paper we have shown that there exist possibilities of parameterizing all three models in such a way that they can accommodate and explain many of the stylized facts visible in the data. Some stylized facts may in certain cases remain unexplained, however. For example, it appears that the standard GARCH(1,1) model may not particularly often generate series that display the Taylor effect. This is due to the fact that this model does not appear to satisfy the corresponding theoretical property, the Taylor property. On the contrary, this property is approximately satisfied for a relevant subset of EGARCH(1,1) and ARSV(1) models and, albeit very narrowly, for a subset of absolute-valued GARCH models.

Many researchers observed quite early on that for GARCH models, assuming normal errors is too strong a restriction, and they have suggested leptokurtic error distributions in their stead. The results in this paper show how these distributions add to the flexibility of the GARCH model and help the model to reproduce the stylized fact of high kurtosis and relative low autocorrelations of squared observations. It is also shown that the IGARCH model with normal errors does not rescue the normality assumption. As a drawback it may be noted that the parameterization of the first-order stochastic volatility model becomes very restrictive when the amount of the leptokurtosis in the error distribution increases, and the model therefore cannot accommodate 'easy' situations with relatively low kurtosis and high autocorrelations of squared observations.

The paper contains an application of a novel method of obtaining confidence regions for the kurtosis/autocorrelation combinations. The brief application of this method to stock returns indicates, not surprisingly, that when normality of errors is assumed, the GARCH model as well as the EGARCH model are at their best when it comes to characterizing models based on time series with relatively low kurtosis and high first-order autocorrelation of squares. Time series displaying a combination of high kurtosis and high autocorrelation are better modelled using an ARSV(1) model. While this observation may serve as a rough guide when one wants to select one of these models, nonnested tests are also available for comparing them. Examples of such tests have already been mentioned in the Introduction.

Another observation that emerges from the empirical example is that the estimated kurtosis/autocorrelation combination is often an underestimate compared to the one estimated directly from the data without a model. This is the case when the kurtosis is high. This fact may be interpreted as support to the notion that a leptokurtic error distribution is a necessity when using GARCH models. But then, it may also indicate that daily return series contain truly exceptional observations in the sense that they cannot be satisfactorily explained by the members of the standard GARCH or EGARCH family of models.

The present investigation is only concerned with first-order models, and it may be asked if adding more lags would enhance the flexibility of the models. Such additions would certainly help to generate and reproduce more elaborate autocorrelation patterns for the squared observations than is the case with first-order models. It is far from certain, however, that they would also improve reproduction of the stylized facts considered in this study.

Bibliography

Andersson, J. (2001), 'On the normal inverse Gaussian stochastic volatility model', *Journal of Business Economic Statistics* **19**, 44–54.

Baillie, R. T., Bollerslev, T. and Mikkelsen, H.-O. (1996), 'Fractionally integrated generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **74**, 3–30.

Bollerslev, T. (1988), 'On the correlation structure for the generalized autoregressive conditional heteroskedastic process', *Journal of Time Series Analysis* **9**, 121–131.

Carnero, M. A., Peña, D. and Ruiz, E. (2001), 'Is stochastic volatility more flexible than GARCH?'. Working Paper 01-08, Universidad Carlos III de Madrid.

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.

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.

Eklund, B. (2004), 'Estimating confidence regions over bounded domains', *Computational Statistics and Data Analysis* (forthcoming).

Engle, R. F. and Bollerslev, T. (1986), 'Modelling the persistence of conditional variances', *Econometric Reviews* **5**, 1–50.

Engle, R. F. and Ng, V. K. (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance* **48**, 1749–1778.

Ghysels, E., Harvey, A. C. and Renault, E. (1996), 'Stochastic volatility', *Handbook of Statistics* **14**, 119–191.

Granger, C. W. J. and Ding, Z. (1995a), 'Some properties of absolute returns: An alternative measure of risk', *Annales d'économie et de statistique* **40**, 67–95.

Granger, C. W. J., Ding, Z. and Spear, S. (2000), *Stylized Facts on the Temporal and Distributional Properties of Absolute Returns: An Update*, Statistics and Finance: An Interface. Proceedings of the Hong Kong International Workshop on Statistics in Finance, Chan W-S, W K Li and H Tong (eds). Imperial College Press: London, pp. 97–120.

Granger, C. W. J., King, M. L. and White, H. (1995), 'Comments on testing economic theories and the use of model selection criteria', *Journal of Econometrics* **67:1**, 173–187.

He, C. (2000), Moments and the autocorrelation structure of the exponential GARCH(p,q) process. Working Paper Series in Economics and Finance, Stockholm School of Economics.

He, C. and Teräsvirta, T. (1999*a*), 'Fourth moment structure of the GARCH(p,q) process', *Econometric Theory* **15**, 824–846.

He, C. and Teräsvirta, T. (1999*b*), 'Properties of moments of a family of GARCH processes', *Journal of Econometrics* **92**, 173–192.

Hyndman (1996), 'Computing and graphing highest density regions', *the American Statistician* **50**, 120–126.

Kim, S., Shephard, N. and Chib, S. (1998), 'Stochastic volatility: Likelihood inference and comparison with ARCH models', *Review of Economic Studies* **65**, 361–93.

Lee, J.-H. and Brorsen, B. W. (1997), 'A non-nested test of GARCH vs. EGARCH models', *Applied Economics Letters* **4**, 765–768.

Liesenfeld, R. and Jung, R. C. (2000), 'Stochastic volatility models: Conditional normality versus heavy-tailed distributions', *Journal of Applied Econometrics* **15**, 137–160.

Lin, S.-J. and Yang, J. (1999), 'Testing shift in financial models with conditional heteroskedasticity: An empirical distribution function approach'.

Research Paper 30, University of Technology Sydney, Quantitative Finance Research Group.

Lundbergh, S. and Teräsvirta, T. (2002), 'Evaluating GARCH models', *Journal of Econometrics* **110**, 417–435.

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.

Poon, S. H. and Granger, C. W. J. (2003), 'Forecasting volatility in financial markets', *Journal of Economic Literature* **XLI**, 478–539.

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. (1989), 'Why does stock market volatility change over time?', *Journal of Finance* **45**, 1129–1155.

Shephard, N. (1996), Statistical aspects of ARCH and stochastic volatility, *in* D. R. Cox and O. E. Barndorff-Nielsen and D. V. Hinkley, eds, *Time series models: in econometrics, finance and other fields*, London: Chapman and Hall, chapter 1, pp. 1–67.

Taylor, S. (1986), *Modelling Financial Time Series*, Wiley, New York.

Teräsvirta, T. (1996), Two stylized facts and the GARCH(1,1) model. Working Paper Series in Economics and Finance, Stockholm School of Economics.

Figures

Figure 3.1 First Panel, the S&P 500 index 3 January 1928 to 19 September 2001. Second panel, shows the autocorrelation function of $|r_t|^m$, $m = 0.25, 0.5, 0.75, 1$, from low to high, for the S&P 500 index. Third panel, shows the autocorrelation function of $|r_t|^m$, $m = 1, 1.25, 1.5, 1.75, 2$, from high to low, for the S&P 500 index. Fourth panel shows the autocorrelation function of $|r_t|$ for the whole series and the mean of the corresponding autocorrelations of the 20 equally long subseries of the S&P 500 index together with the plus/minus one standard deviation band.

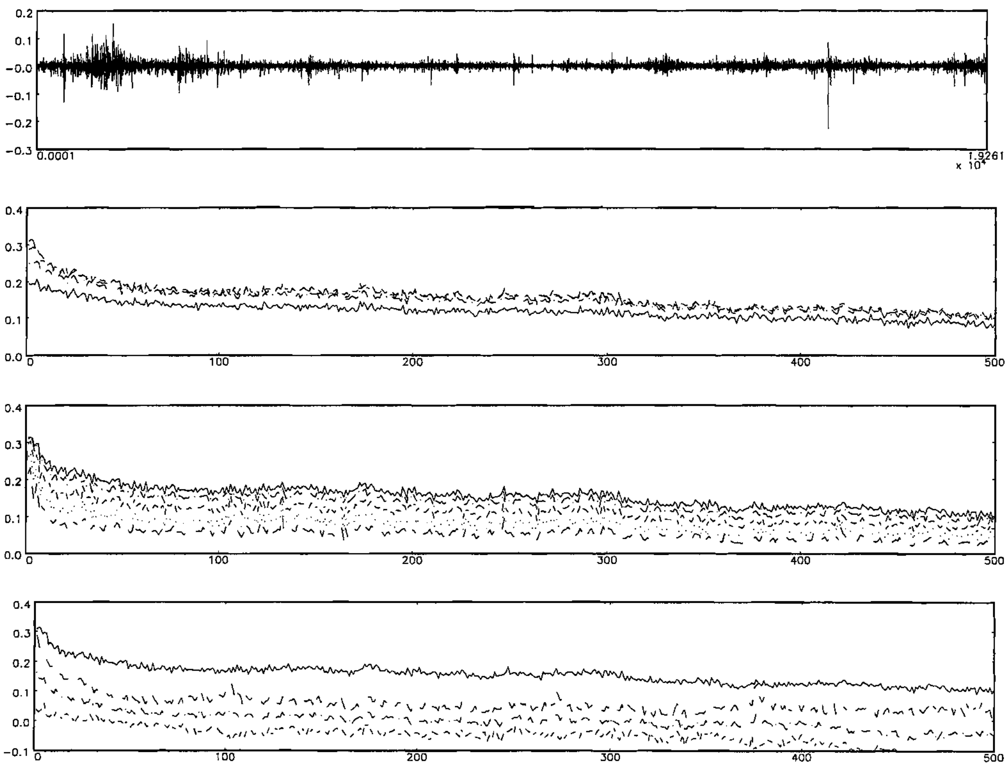


Figure 3.2 Combinations of the first-order autocorrelation of squared observations and kurtosis for the GARCH(1,1) model with normal errors for various values of $\alpha + \beta$:$\alpha + \beta = 0.999$ (solid line), $\alpha + \beta = 0.99$ (dashed-dotted line) and $\alpha + \beta = 0.95$ (dashed line).

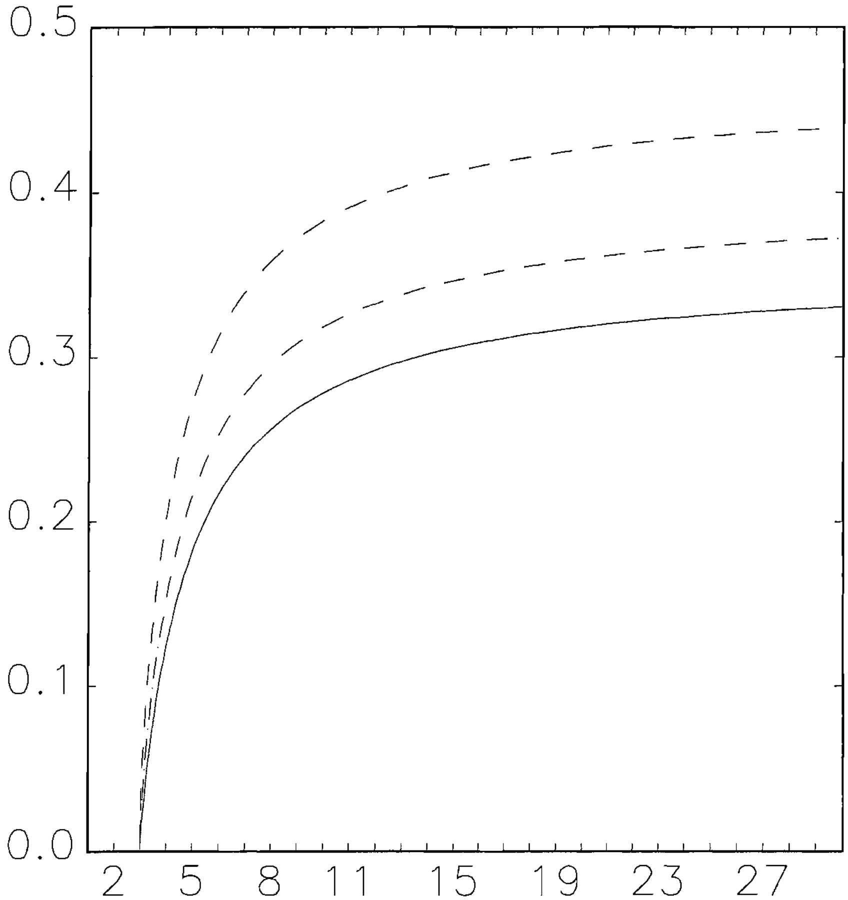


Figure 3.3 Combinations of the first-order autocorrelation of squared observations and kurtosis for the GARCH(1,1) model with normal errors for various values of $\alpha + \beta$ together with observed combiantions of daily rates of return: First panel, the 27 most traded stocks at the Stockholm Stock Exchange. Second panel, the S&P 500 index 3 January 1928 to 19 September 2001, divided to 20 equally long subsereies. Third panel, five major exchange rates series divided to 34 subseries. Fourth panel, all data points. The isoquants: $\beta = 0.999$ (solid line),$\beta = 0.99$ (dashed-dotted line), $\beta = 0.95$ (dotted line) and $\beta = 0.9$ (dashed line).

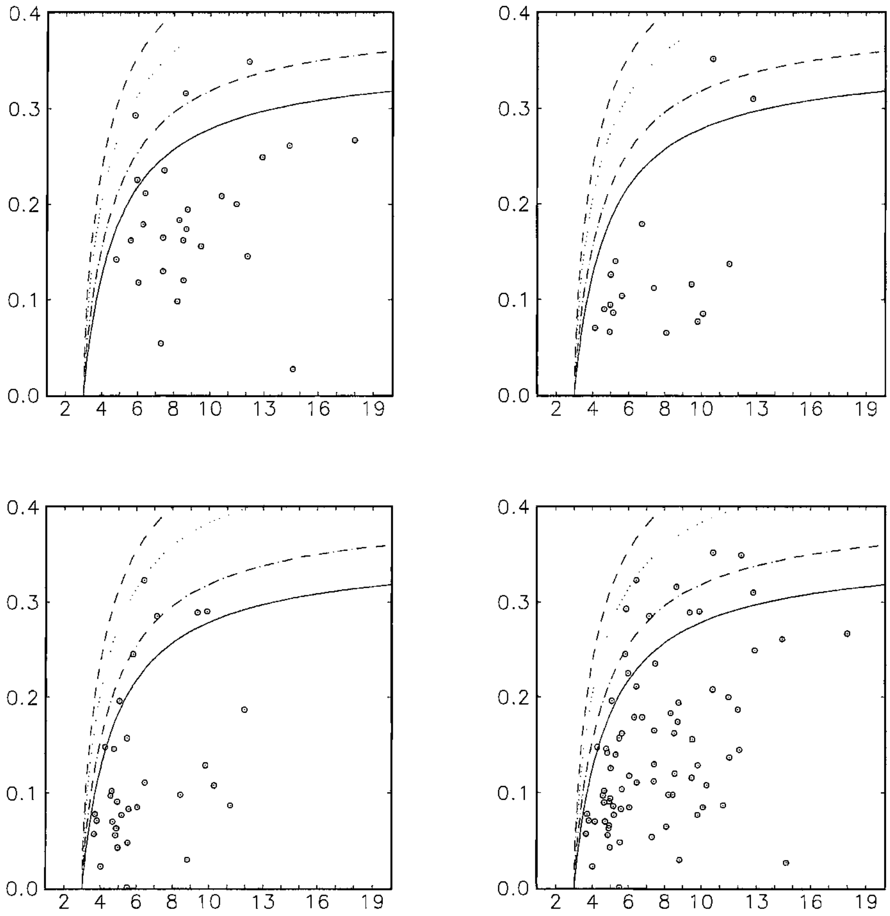


Figure 3.4 Combinations of the first-order autocorrelation of squared observations and kurtosis for the GARCH(1,1) model with normal errors for various values of $\alpha + \beta$ together with 100 realizations based on T simulated observations from an IGARCH(1,1) model with $\alpha_0 = \alpha = 0.1$: $T = 100$ (first panel), 500 (second panel), 1000 (third panel) and 2000 (fourth panel).

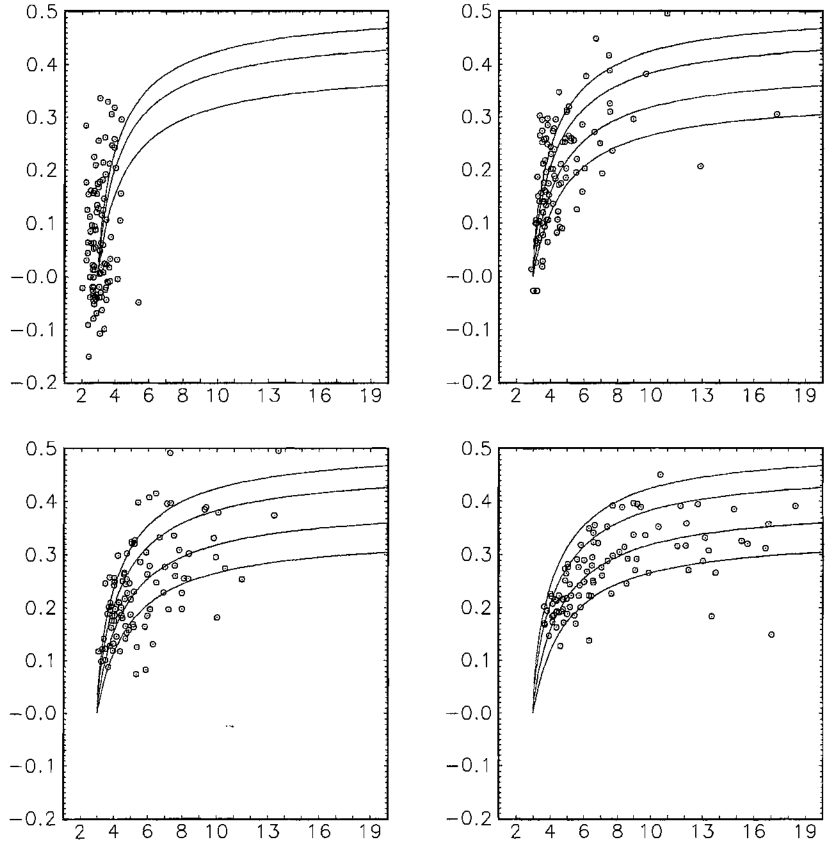


Figure 3.5 Combinations of the first-order autocorrelation of squared observations and kurtosis for the GARCH(1,1) model with t-distributed errors for various values of $\alpha\nu_2 + \beta$: t(7) left panel and t(5) right panel, the isoquants: $\beta = 0.999$ (solid line),$\beta = 0.99$ (dashed-dotted line), $\beta = 0.95$ (dotted line) and $\beta = 0.9$ (dashed line). The observed combinations are all the data from the fourth panel in Figure 3.1.

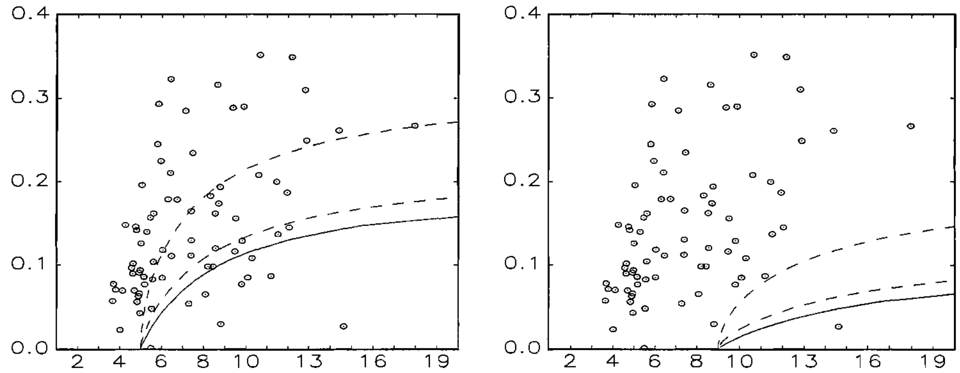


Figure 3.6 Combinations of the first-order autocorrelation of squared observations and kurtosis for the ARSV(1) model with t-distributed errors for various values of β: t(7) left panel and t(5) right panel, the isoquants: $\beta = 0.999$ (solid line),$\beta = 0.99$ (dashed-dotted line), $\beta = 0.95$ (dotted line) and $\beta = 0.9$ (dashed line). The observed combinations are all the data from the fourth panel in Figure 3.1.

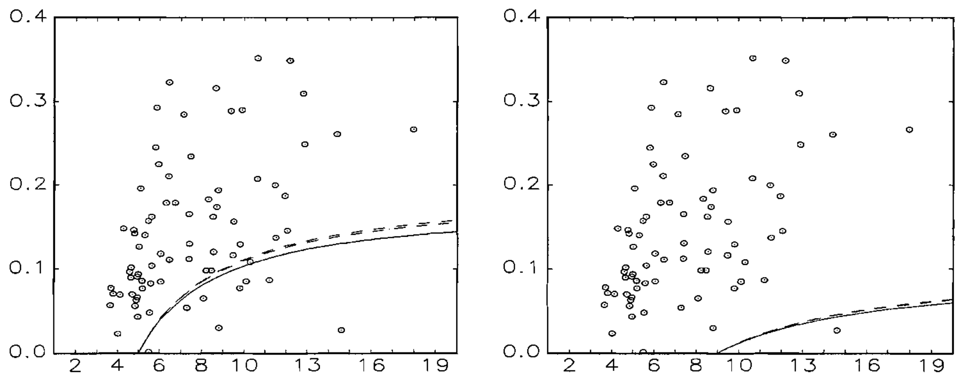


Figure 3.7 Combinations of the first-order autocorrelation of squared observations and kurtosis for the EGARCH(1,1) model with normal errors for various values of β:β = 0.95 (solid line),β = 0.99 (dashed-dotted line) and $\beta = 0.999$ (dashed line). The observed combinations are all the data from the fourth panel in Figure 3.1.

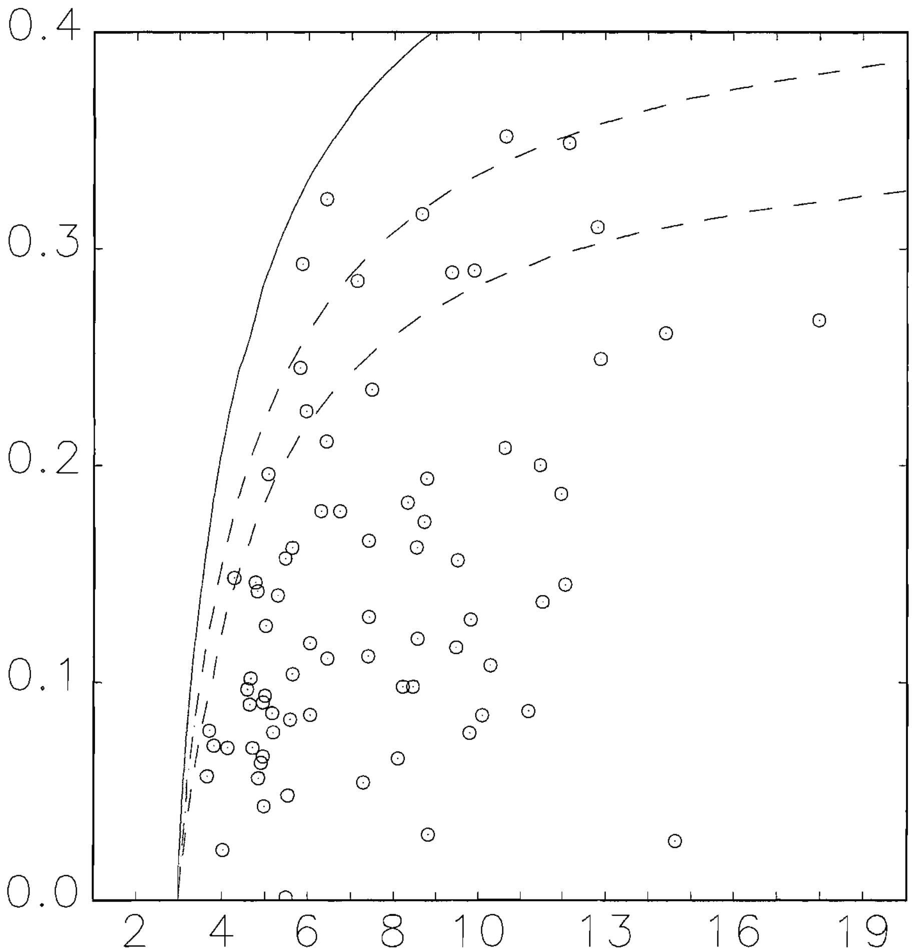


Figure 3.8 Combinations of the first-order autocorrelation of squared observations and kurtosis for the ARSV(1) model with normal errors for various values of β:$\beta = 0.999$ (dashed line),$\beta = 0.99$ (dashed-dotted line) and $\beta = 0.95$ (solid line). The observed combinations are all the data from the fourth panel in Figure 3.1.

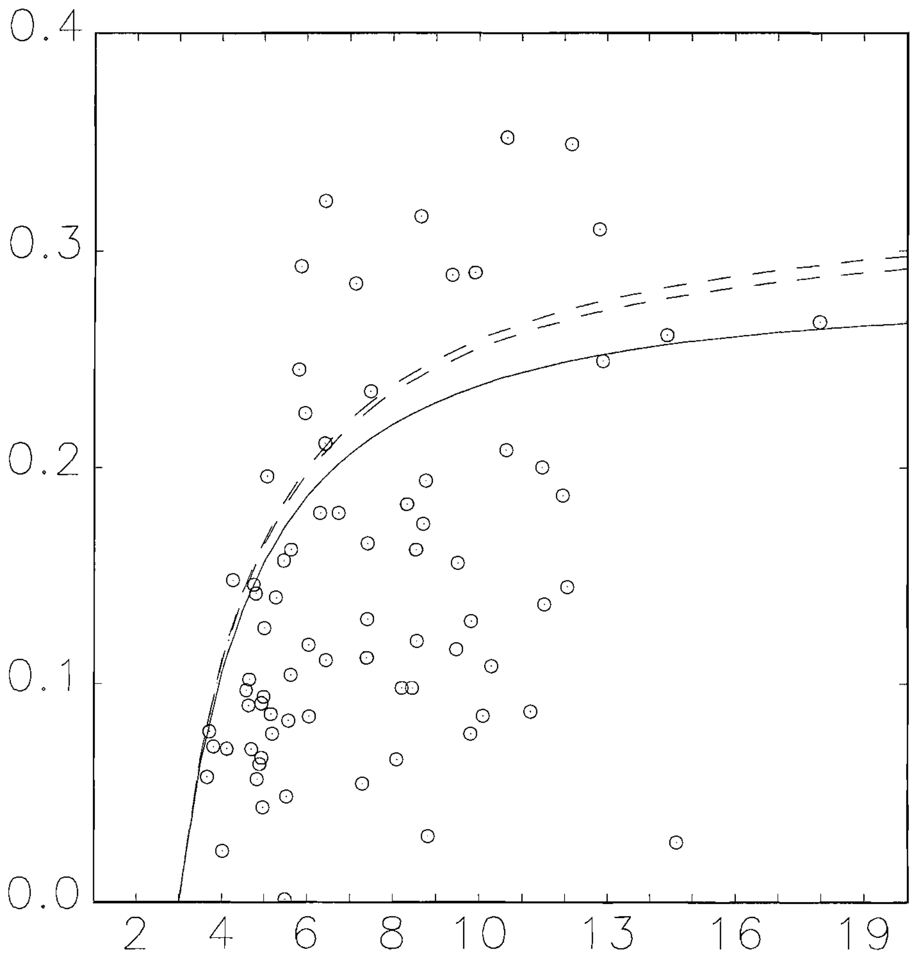


Figure 3.9 Combinations of two first-order autocorrelations, the squared observations (dashed line) and the absolute observations (solid line), and corresponding kurtosis for the EGARCH(1,1) model with normal errors for $\beta = 0.95$ (high) and $\beta = 0.99$ (low).

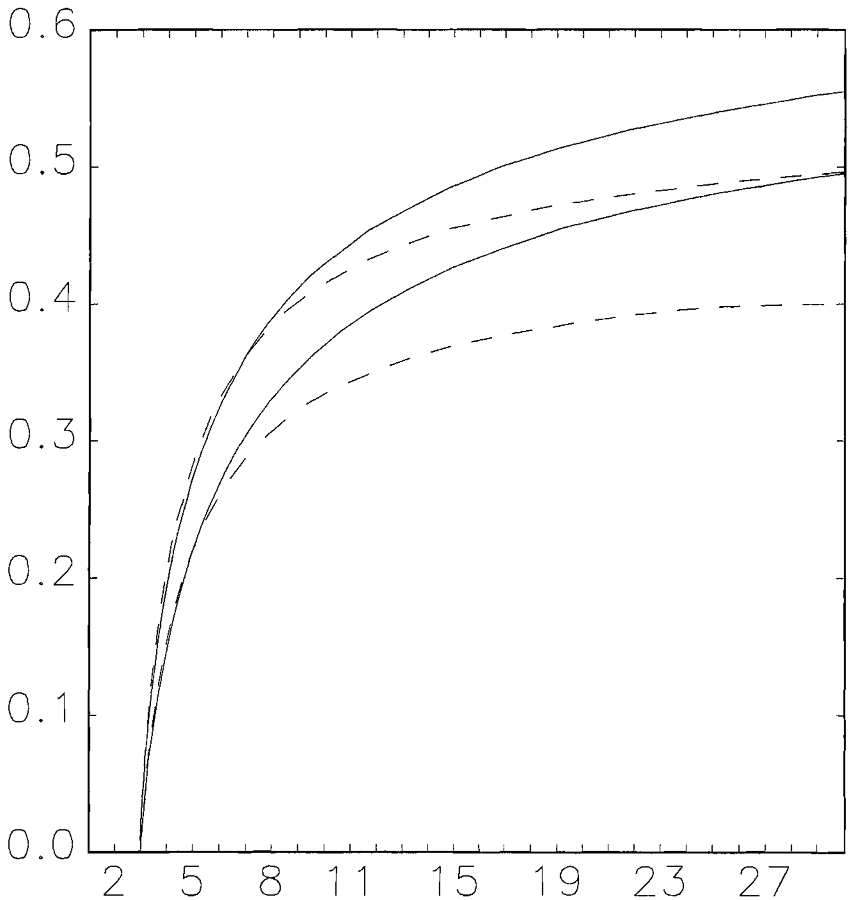


Figure 3.10 Combinations of first-order autocorrelation as a function of the exponent m for the EGARCH(1,1) model with normal errors for $\beta = 0.95$ (left panel) and $\beta = 0.99$ (right panel) at three different kurtosis values. $\kappa_4 = 6$ (solid line),$\kappa_4 = 12$ (dashed-dotted line) and $\kappa_4 = 24$ (dotted line).

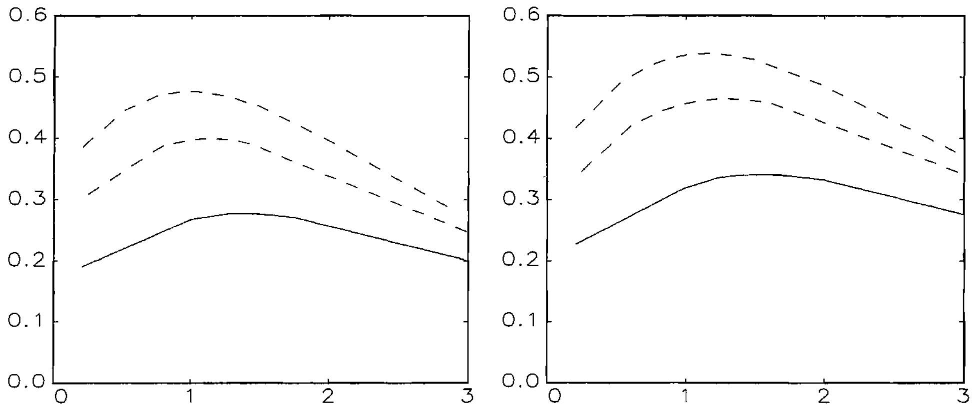


Figure 3.11 Combinations of first-order autocorrelation as a function of the exponent m for the ARSV(1) model with normal errors for $\beta = 0.95$ (left panel) and $\beta = 0.99$ (right panel) at three different kurtosis values. $\kappa_4 = 6$ (solid line),$\kappa_4 = 12$ (dashed-dotted line) and $\kappa_4 = 24$ (dotted line).

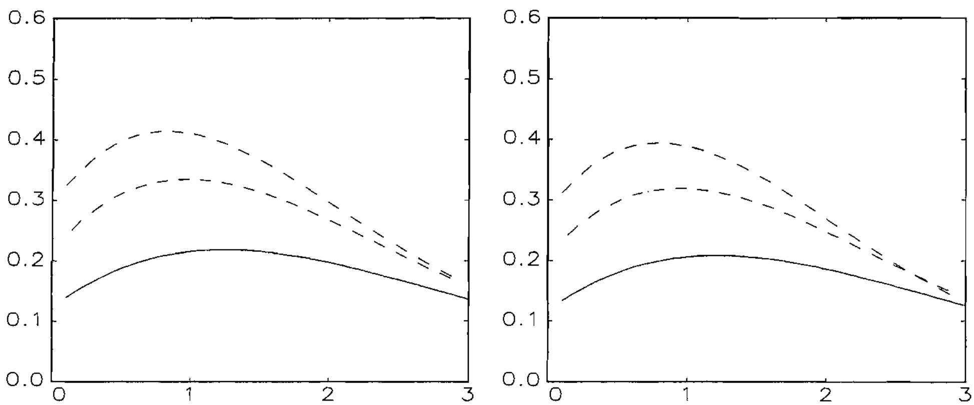


Figure 3.12 Combinations of two first-order autocorrelations, the squared observations (dashed line) and the absolute observations (solid line), and corresponding kurtosis for the ARSV(1) model with normal errors for $\beta = 0.95$ (low) and $\beta = 0.99$ (high).

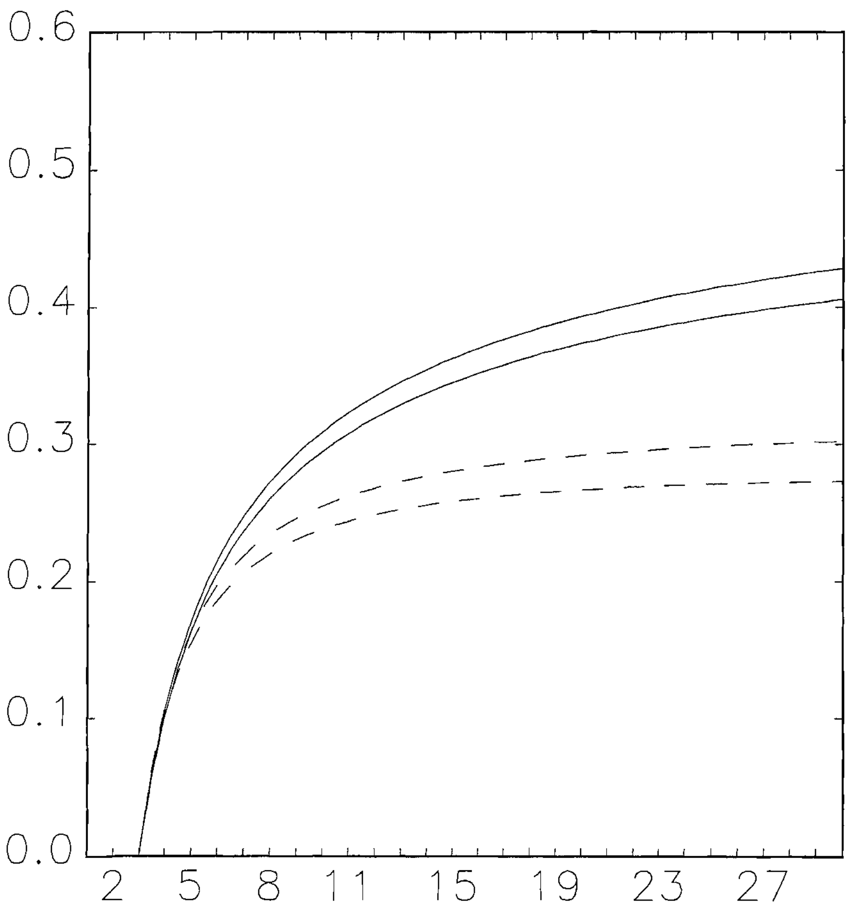


Figure 3.13 Simulated kurtosis/autocorrelation combinations for the GARCH(1,1) with $(\alpha_0, \alpha_1, \beta) = (0.05, 0.19121, 0.75879)$, and approximative 50%, 60%, 70%, 80%, and 90% confidence intervals of the true value, 1000 observations and 200 realizations, solid square shows the true value, solid circle represents the estimated pair, the empty circles represents the generated pairs.

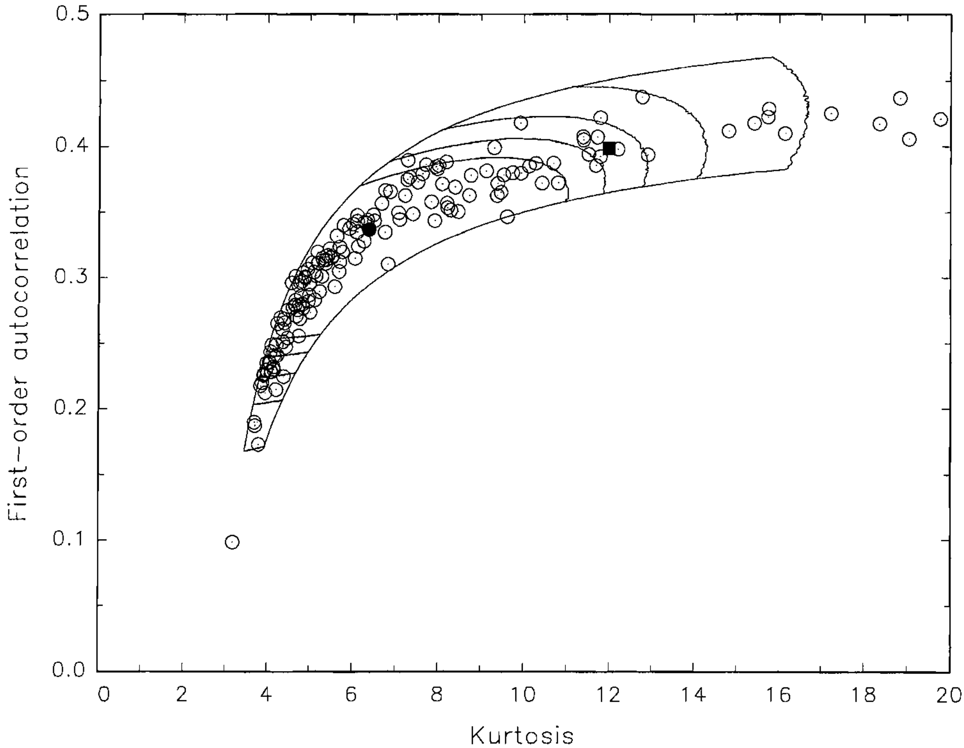


Figure 3.14 Approximative 90% confidence intervals of the true kurtosis/autocorrelation combination for the stock Assi D assuming that the observations have been generated by a GARCH(1,1) model. Figure based on 200 realizations, solid square shows the directly estimated value, solid circle represents the estimated pair.

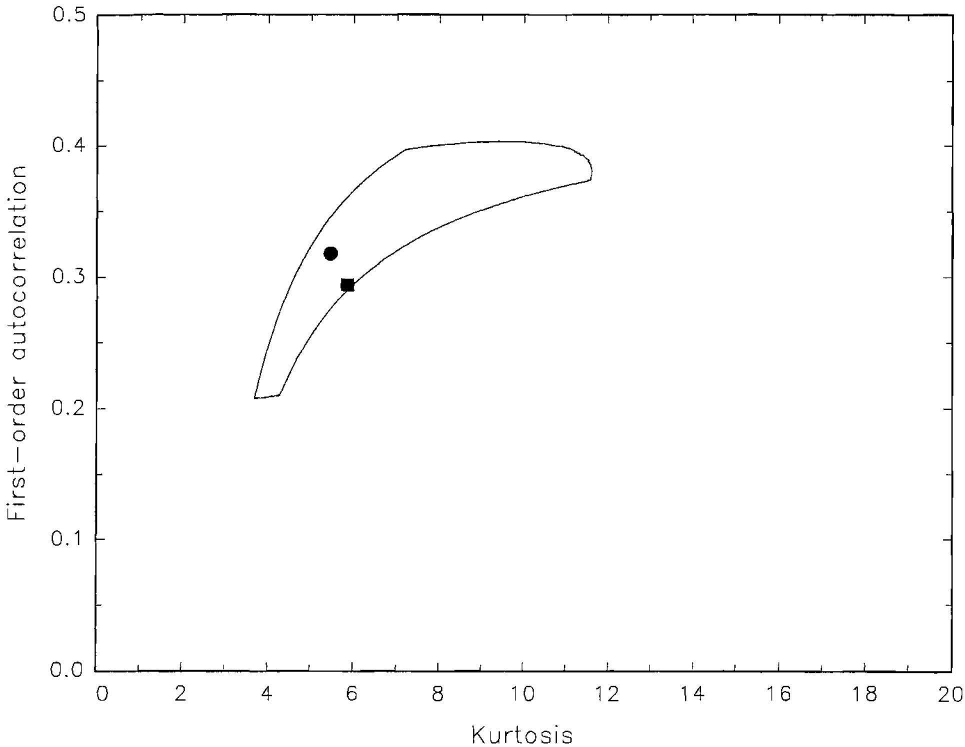


Figure 3.15 Approximative 90% confidence intervals of the true kurtosis/autocorrelation combination for the stock Assi D assuming that the observations have been generated by an EGARCH(1,1) model. Figure based on 200 realizations, solid square shows the directly estimated value, solid circle represents the estimated pair.

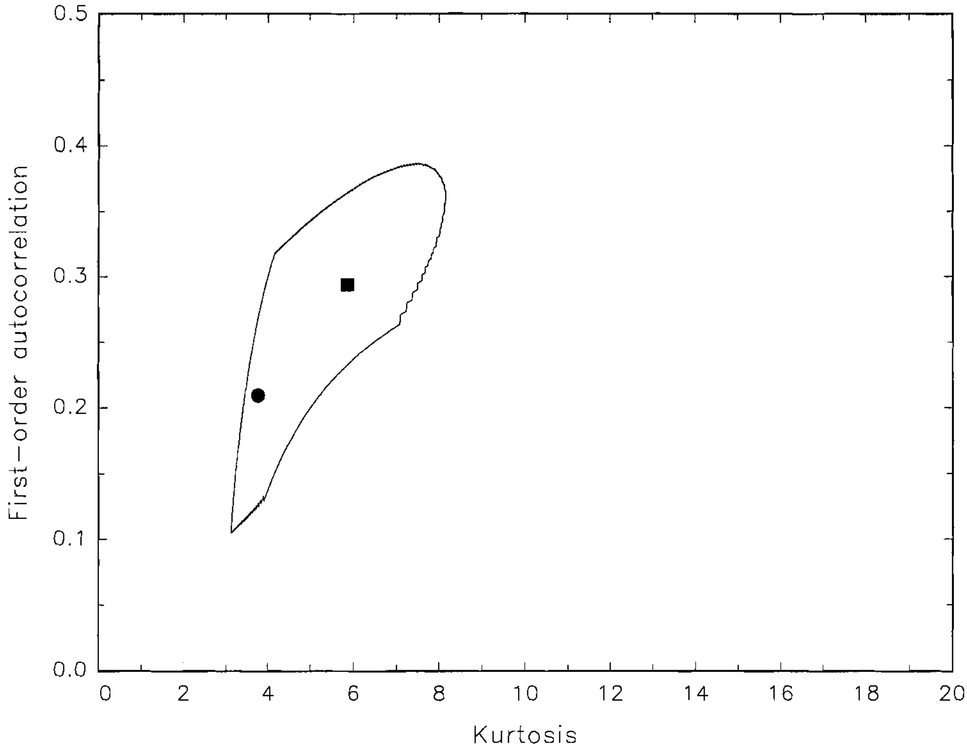


Figure 3.16 Approximative 90% confidence intervals of the true kurtosis/autocorrelation combination for the stock Assi D assuming that the observations have been generated by an ARSV(1) model. Figure based on 200 realizations, solid square shows the directly estimated value, solid circle represents the estimated pair.

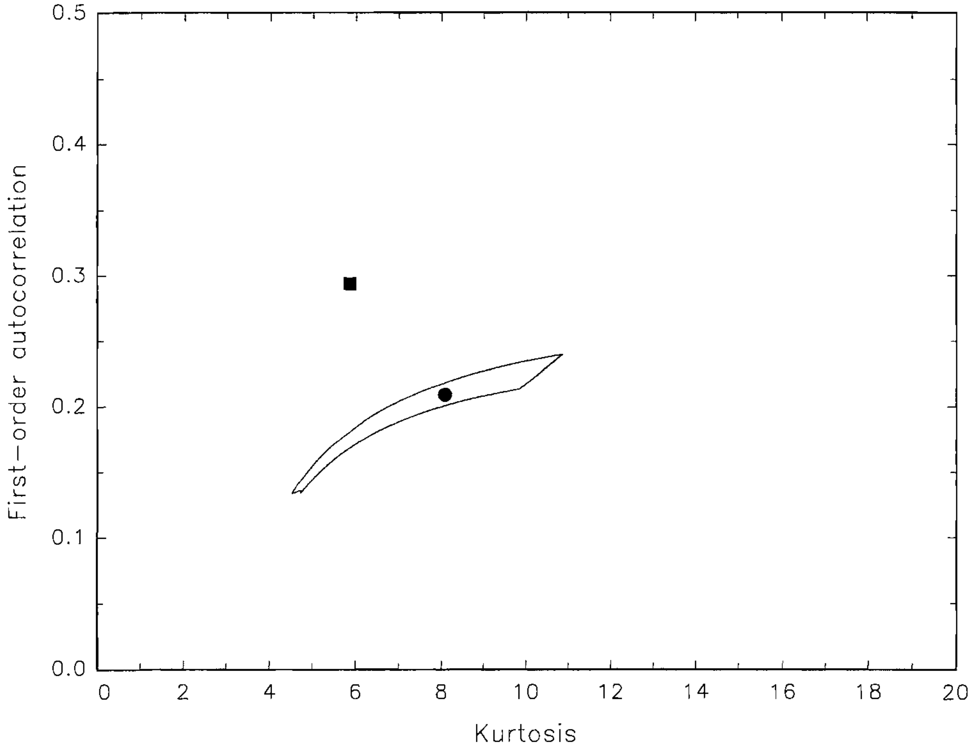


Figure 3.17 Approximative 90% confidence intervals of the true kurtosis/autocorrelation combination for the stock SEB assuming that the observations have been generated by a GARCH(1,1) model. Figure based on 200 realizations, solid square shows the directly estimated value, solid circle represents the estimated pair.

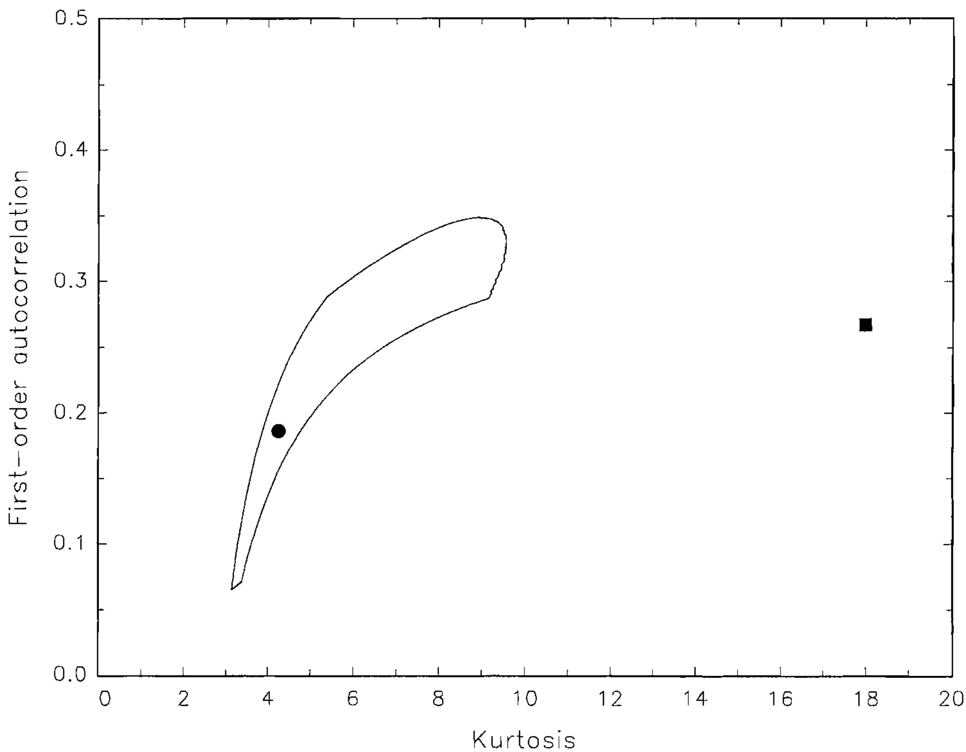
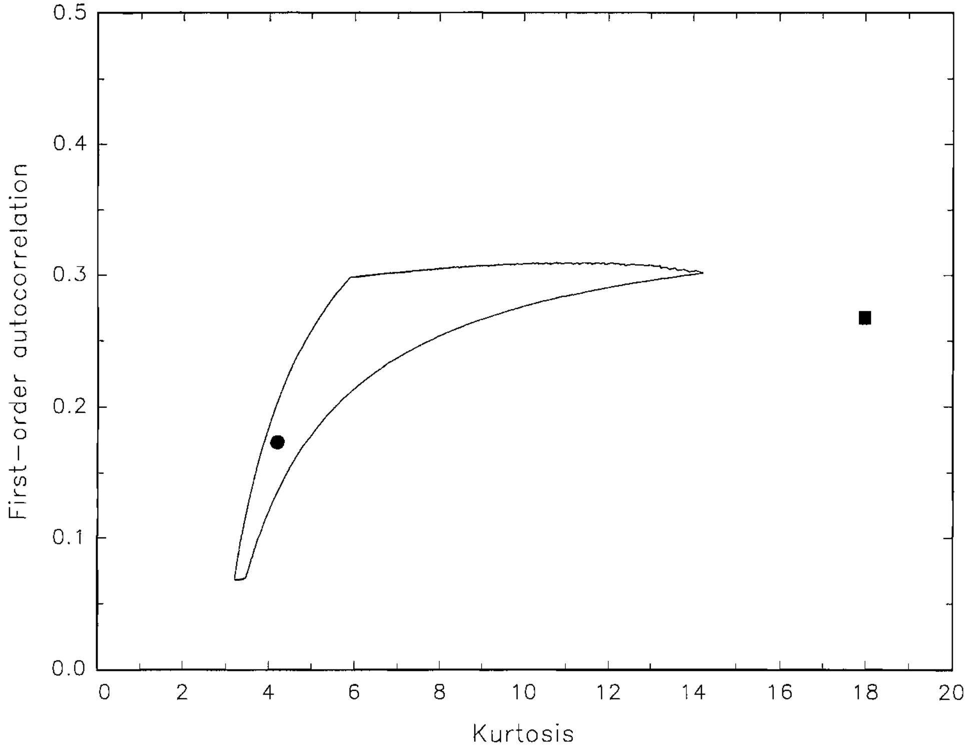


Figure 3.18 Approximative 90% confidence intervals of the true kurtosis/autocorrelation combination for the stock SEB assuming that the observations have been generated by an EGARCH(1,1) model. Figure based on 200 realizations, solid square shows the directly estimated value, solid circle represents the estimated pair.



Chapter 4

Evaluating exponential GARCH models

4.1 Introduction

Model evaluation is an important part of modelling not only for the conditional mean models but for the conditional variance specifications as well. It is useful, one could argue even necessary, to carry out an in-sample evaluation of a volatility model before it is used for forecasting. Engle and Ng (1993), Li and Mak (1994) and Chu (1995), to name a few examples, derived misspecification tests for generalized autoregressive heteroskedasticity (GARCH) models. Recently, Lundbergh and Teräsvirta (2002) presented a unified framework for testing the adequacy of an estimated GARCH model. Their framework covers, among other things, testing the null of no ARCH in the standardized errors, testing symmetry against a smooth transition GARCH (STGARCH) model and a test of parameter constancy against smoothly changing parameters.

It appears that less work has been done for the evaluation of exponential GARCH (EGARCH) model by Nelson (1991). In fact, Nelson already suggested several tests based on orthogonality conditions that the errors of the model satisfy under the null hypothesis, but not much has happened since, nor have Nelson's tests been regularly applied in empirical work. In this paper we continue Nelson's work and consider a number of misspecification tests for the EGARCH model. They are Lagrange multiplier or Lagrange multiplier type tests and include testing an EGARCH model against a higher-order one and testing parameter constancy. Furthermore, we investigate various ways of testing the EGARCH model against GARCH ones as another check of model adequacy. The literature on testing non-nested hypotheses for volatility models includes Chen and Kuan (2002), Kim et al. (1998) and Lee and Brorsen (1997); see also Engle and Ng (1993). Their tests are considered in the present framework, and the small-sample properties of the tests are investigated by simulation.

The plan of the paper is as follows. The model is defined in Section 2 and the estimation of parameters is discussed briefly in Section 3. Section 4 considers testing an EGARCH model against a higher order one and testing parameter constancy. In Section 5 non-nested tests for testing EGARCH and GARCH models against each other are discussed. Section 6 contains results of a simulation experiment and Section 7 an empirical example. Finally, conclusion appear in Section 8.

4.2 The model

Let

$$y_t = f(\mathbf{w}_t; \boldsymbol{\varphi}) + \varepsilon_t, t = 1, ..., T \tag{4.1}$$

where f is at least twice continuously differentiable function of $\boldsymbol{\varphi}$, with $\mathbf{w}_t = (1, y_{t-1}, ..., y_{t-n}, x_{1t}, ..., x_{kt})'$. The error process is parameterized as

$$\varepsilon_t = z_t h_t^{1/2}, t = 1, ..., T \tag{4.2}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with zero mean and unit variance. A family of EGARCH(p,q) models may be defined as a combination of (4.2) and

$$\ln h_t = \alpha_0 + \sum_{j=1}^{q} g_j(z_{t-j}) + \sum_{j=1}^{p} \beta_j \ln h_{t-j}. \tag{4.3}$$

The conditional variance is constrained to be non-negative by the assumption that the logarithm of h_t is a function of past z_t's. Equations (4.1) and (4.3) define a class of EGARCH(p,q) models. Setting

$$g_j(z_{t-j}) = \alpha_j z_{t-j} + \psi_j(|z_{t-j}| - E\,|z_t|), j = 1, ..., q, \tag{4.4}$$

in (4.3) yields the EGARCH(p,q) model proposed by Nelson (1991). The overwhelmingly most popular EGARCH model in application has been (4.4) with $p = q = 1$. When $g_j(z_{t-j}) = \alpha_j \ln z_{t-j}^2, j = 1, ..., q$, (4.2) and (4.3) define the logarithmic GARCH (LGARCH) model that Geweke (1986) and Pantula (1986) proposed. The specification in (4.4) amounts to $g(z_t)$ being a function of both the magnitude and sign of z_t. This enables h_t to respond asymmetrically to positive and negative values of ε_t, which is believed to be important for example in modelling the behaviour of stock returns. As to the distribution of z_t, we assume it to be symmetric around zero, which implies $\mathsf{E}z_t^3 = 0$. This assumption together with $\mathsf{E}z_t^3 = 0$ guarantees block diagonality of the information matrix of the log-likelihood function. Block diagonality turn allows us to concentrate of the conditional variance function (4.3) without simultaneously considering (4.1).

4.3 Estimation of parameters

Before considering misspecification tests we briefly discuss parameter estimation. If we complete the previous assumptions about z_t by assuming normality, the log-likelihood function of the EGARCH(p, q) model is

$$L_t = c - (1/2)\sum_{t=1}^{T} \ln h_t - (1/2)\sum_{t=1}^{T}(\varepsilon_t^2/h_t) \tag{4.5}$$

with

$$\ln h_t = \alpha_0 + \sum_{j=1}^{q}\{\alpha_j z_{t-j} + \psi_j(|z_{t-j}| - E\,|z_t|)\} + \sum_{j=1}^{p}\beta_j \ln h_{t-j}. \tag{4.6}$$

Let $\boldsymbol{\beta} = (\alpha_0, \alpha_1, ..., \alpha_q, \psi_1, ..., \psi_q, \beta_1, ..., \beta_p)'$. Nelson (1991) discussed maximum likelihood estimation under the assumption that the errors have a generalized error distribution, but we do not follow his path here. The first partial derivatives with respect to the EGARCH parameters are

$$\sum_{t=1}^{T}\frac{\partial l_t}{\partial \boldsymbol{\beta}} = (1/2)\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{h_t} - 1)\frac{\partial \ln h_t}{\partial \boldsymbol{\beta}} \tag{4.7}$$

where

$$\frac{\partial \ln h_t}{\partial \boldsymbol{\beta}} = \mathbf{x}_{\beta t} - (1/2)\sum_{j=1}^{q}\{\alpha_j z_{t-j} + \psi_j\,|z_{t-j}|\}\frac{\partial \ln h_{t-j}}{\partial \boldsymbol{\beta}} + \sum_{j=1}^{p}\beta_j\frac{\partial \ln h_{t-j}}{\partial \boldsymbol{\beta}} \tag{4.8}$$

where $\mathbf{x}_{\beta t} = (1, z_{t-1}, ..., z_{t-q}, |z_{t-1}| - E\,|z_t|\,, ..., |z_{t-q}| - E\,|z_t|\,, \ln h_{t-1}, ..., \ln h_{t-p})'$.

The parameters of (4.1) with EGARCH errors (4.6) may be estimated jointly by maximum likelihood. The normality assumption guarantees block diagonality of the information matrix such that the off-diagonal blocks involving partial derivatives with respect to both mean and variance parameters are null matrices. Thus the parameters of the conditional mean defined by (4.1) can be estimated separately without asymptotic loss of efficiency. This implies that maximum likelihood estimates for the parameters in (4.6) can be obtained numerically from the first-order conditions defined by setting (4.7) equal to zero.

Under sufficient regularity conditions, the maximum likelihood estimators can be expected to be consistent and asymptotically normal. It appears, however, that these conditions have not yet been verified in the present situation. Verifying them in the GARCH case has been a demanding task, and things do not appear to be any easier in the case of EGARCH models. In what follows, it is assumed that the maximum likelihood estimators are consistent and asymptotically normal.

It is seen from (4.8) that parameter estimation implies a number of recursions, and starting-values for parameters are therefore necessary. Nelson

(1991) discussed the role of starting-values and concluded that in his simulations the use of other starting-values than the unconditional mean of $\ln h_t$ very rapidly led to values of h_t obtained by starting from the estimate of $\mathsf{E}\ln h_t$.

4.4 Evaluation of EGARCH models

4.4.1 Testing against a higher-order EGARCH model

In this section our starting-point is that the parameters of the EGARCH(p,q) model have been estimated by maximum likelihood, assuming that the errors are standard normal and independent. If, in addition to independence, it is only assumed that $\mathsf{E}z_t = 0$, $\mathsf{E}z_t^2 = 1$ and $\mathsf{E}z_t^3 = 0$, the estimators are quasi maximum likelihood estimators. First we consider testing an EGARCH(p,q) model against a higher-order model, either an EGARCH($p+r$,q) or EGARCH(p,$q+r$), $r > 0$. This is analogous to Bollerslev's test of GARCH(p,q) against GARCH($p+r$,q) or GARCH(p,$q+r$), $r > 0$. Consider now an augmented version of model (4.2),

$$\varepsilon_t = z_t h_t^{1/2} g_t^{1/2} \tag{4.9}$$

where

$$\ln g_t = \sum_{j=1}^{r} \{\alpha_{q+j} z_{t-q-j} + \psi_{q+j}(|z_{t-q-j}| - E\,|z_t|)\} \tag{4.10}$$

The null hypothesis $\mathrm{H}_0 : (\alpha_j, \psi_j) = (0,0), j = q+1, ..., q+r$. Under this hypothesis, $g_t \equiv 1$, and the model collapses into a EGARCH(p,q). Assume now that the alternative H_1 : at least one $\psi_j \neq 0, j = q+1, ..., q+r$. The log-likelihood function of the model is

$$L_T = c - (1/2)\sum_{t=1}^{T}(\ln h_t + \ln g_t) - (1/2)\sum_{t=1}^{T}[\varepsilon_t^2/(h_t g_t)]. \tag{4.11}$$

Let $\boldsymbol{\beta}_r = (\alpha_{q+1}, ..., \alpha_{q+r}, \psi_{q+1}, ..., \psi_{q+r})'$. The block of the score vector containing the partial derivatives with respect to $\boldsymbol{\beta}_r$ has the form

$$\frac{\partial L_T}{\partial \boldsymbol{\beta}_r} = (1/2)\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{h_t g_t} - 1)\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r} \tag{4.12}$$

where

$$\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r} = \mathbf{x}_{\beta_r t} - (1/2)\sum_{j=1}^{r}\{\alpha_j z_{t-q-j} + \psi_j(|z_{t-q-j}| - E\,|z_t|)\}\frac{\partial \ln g_{t-q-j}}{\partial \boldsymbol{\beta}_r} \tag{4.13}$$

with $\mathbf{x}_{\beta_r t} = (z_{t-q-1}, |z_{t-q-1}| - E|z_t|, ..., z_{t-q-r}, |z_{t-q-r}| - E|z_t|)'$. Let $\widehat{h}_t$ and $\partial \ln \widehat{h_t}/\partial \boldsymbol{\beta}$ be the conditional variance h_t and $\partial \ln h_t/\partial \boldsymbol{\beta}$, respectively, estimated under H_0, and let

$$\widehat{\mathbf{v}}_t = (|\widehat{\varepsilon}_{t-q-1}| / \widehat{h}_{t-q-1}, ..., |\widehat{\varepsilon}_{t-q-r}| / \widehat{h}_{t-q-r}, \widehat{\varepsilon}_{t-q-1}/\widehat{h}_{t-q-1}, ..., \widehat{\varepsilon}_{t-q-r}/\widehat{h}_{t-q-r})'.$$

Assume, furthermore, that $1\text{-}\sum_{j=1}^{p} \beta_j L^j$ has its roots outside the unit circle. This, together with the assumption of normality for z_t, guarantees that ε_t has all moments see Nelson (1991) and He, Teräsvirta and Malmsten (2002). Thus the moment conditions required for the asymptotic distribution theory of the LM test statistic are satisfied. The LM test can be carried out in the TR^2 form as follows:

1. Estimate the parameters of the EGARCH(p,q) model and compute the squared standardized residuals $\widehat{\varepsilon}_t^2/\widehat{h}_t - 1$, $t = 1, ..., T$, and the "residual sum of squares" $SSR_0^* = \sum_{t=1}^{T} (\widehat{\varepsilon}_t^2/\widehat{h}_t - 1)^2$.

2. Regress $\widehat{\varepsilon}_t^2/\widehat{h}_t - 1$ on $\partial \ln \widehat{h_t}/\partial \boldsymbol{\beta}$ and $\widehat{\mathbf{v}}_t$ and compute the sum of squared residuals, SSR_1^*.

3. Compute the value of the test statistic

$$LM_{addEGARCH} = T\frac{SSR_0^* - SSR_1^*}{SSR_0^*} \tag{4.14}$$

that has an asymptotic χ^2 distribution with $2r$ degrees of freedom under the null hypothesis.

If the normality assumption does not hold, this distribution theory is not valid. Nevertheless, it is possible to robustify the test against non-normal errors following Wooldridge (1991). Assuming that $\mathsf{E}|\varepsilon_t|^3 < \infty$, the robust version of the test is carried out as follows:

1. Regress $\widehat{\mathbf{v}}_t$ on $\partial \ln \widehat{h_t}/\partial \boldsymbol{\beta}$, and compute the $(2r{\times}1)$ residual vectors $\mathbf{r}_t$, $t = 1, ..., T$.

2. Regress 1 on $(\widehat{\varepsilon}_t^2/\widehat{h}_t - 1)\mathbf{r}_t$ and compute the residual sum of squares SSR^* from that regression. The test statistic is

$$LM_{addEGARCH-R} = T - SSR^* \tag{4.15}$$

and has an asymptotic χ^2 distribution with $2r$ degrees of freedom under the null hypothesis.

When $p = q = 0$, the test collapses into a test of no EARCH against EARCH(r). The test against H_1 : "at least one $\boldsymbol{\beta}_j \neq 0, j = q+1, ..., q+r$", is constructed analogously by redefining vector $\widehat{\mathbf{v}}_t$. Note that when $p = q = 0$, the corresponding test would be meaningless.

4.4.2 Testing parameter constancy

Testing parameter constancy is important in its own right but also because nonconstancy signals an apparent lack of covariance stationarity. Here we assume that the alternative to constant parameters in the conditional variance is that the parameters, or a subset of them, change smoothly over time. This test may be viewed as the EGARCH counterpart of the test for parameter constancy against smooth continuous change in parameters for the GARCH model in Lundbergh and Teräsvirta (2002). Lin and Teräsvirta (1994) applied the same idea to testing parameter constancy in the conditional mean. Consider now the augmented model (4.9) where

$$\ln g_t = (\pi_0 + \sum_{j=1}^{q} \{\pi_{1j} z_{t-j} + \pi_{2j} |z_{t-j}|\} + \sum_{j=1}^{p} \pi_{3j} \ln h_{t-j}) G_n(t; \gamma, \mathbf{c}) \quad (4.16)$$

with the transition function

$$G_n(t; \gamma, \mathbf{c}) = \left[1 + \exp(-\gamma \prod_{i=1}^{n} (t - c_i)) \right]^{-1}, \gamma > 0, c_1 \leq \; ... \; \leq c_n. \quad (4.17)$$

In (4.17) γ is a slope parameter, and $\mathbf{c} = (c_1, ..., c_n)$ a location vector. Conditions $\gamma > 0$ and $c_1 \leq ... \leq c_n$ are identifying conditions. When $\gamma = 0$, $G_n(t; \gamma, \mathbf{c}) \equiv 1/2$. Typically in practice, $n = 1$ or $n = 2$. The former choice yields a standard logistic function. When the slope parameter $\gamma \to \infty$, (4.17) with $n = 1$ becomes a step function whose value equals one for $t > c_1$ and zero otherwise. This special case represents a single structural break in the model at $t = c_1$. When $n = 2$, (4.17) is symmetric about $(c_1 + c_2)/2$, and its minimum value, achieved at this point, lies between zero and $1/2$. The value of the function approaches unity as $t \to \pm\infty$. When $\gamma \to \infty$, function (4.17) becomes a "double step" function that obtains value zero for $c_1 \leq t \leq c_2$ and unity otherwise.

In order to consider the testing problem, let $\overline{G}_n = G_n - 1/2$. This transformation simplifies notation in deriving the test but does not effect the generality of the arguments. The smooth transition alternative poses an identification problem. The null hypothesis can be expressed as $\mathrm{H}_0 : \gamma = 0$ in

$\overline{G}_n$. It can be seen from (4.16) and (4.17) that when the null hypothesis holds $\pi_0, \pi_{1j}, \pi_{2j}, j = 1, ..., q$, and $\pi_{3j}, j = 1, ..., p$, in (4.16) and $c_1, ..., c_n$ in (4.17) are unidetified nuisance parameters. The standard asymptotic distribution theory is thus not available in this situation, for a general discussion see Hansen (1996).

We circumvent the identification problem by following Luukkonen, Saikkonen and Teräsvirta (1988), see also Lundbergh and Teräsvirta (2002). This is done by expanding the transition function $\overline{G}_n$ into a first-order Taylor series around $\gamma = 0$, replacing the transition function (4.17) with this Taylor approximation in (4.16) and rearranging terms. This results in

$$\ln g_t = \sum_{i=1}^{n} \boldsymbol{\delta}_i' \mathbf{v}_{it} + R \tag{4.18}$$

where $\boldsymbol{\delta}_i = \gamma \widetilde{\boldsymbol{\delta}}_i$, $\widetilde{\boldsymbol{\delta}}_i \neq \mathbf{0}$, $\mathbf{v}_{it} = t^i \mathbf{x}_{\beta t}$, $i = 1, ..., n$, and R is the remainder. The new null hypothesis based on (4.18) is equals $\boldsymbol{\delta}_1 = ... = \boldsymbol{\delta}_n = \mathbf{0}$. Note that under $\mathrm{H}_0 : R = 0$ so that the remainder does not affect the asymptotic distribution theory. The test can be carried out in TR^2 form via an auxiliary regression exactly as in the previous section. Vectors $\mathbf{v}_{it}$ now contains the additional variables that appear in the auxiliary regression such that $\widehat{\mathbf{v}}_t = (\mathbf{v}_{1t}', ..., \mathbf{v}_{nt}')'$ with $\widehat{\mathbf{v}}_{it} = t^i \widehat{\mathbf{x}}_{\beta t}$, $i = 1, ..., n$. The test can easily be modified to concern only a subset of parameters. A number of terms in the auxiliary equation now contains trending variables. Nevertheless, applying the results of Lin and Teräsvirta (1994), it can be shown that the asymptotic null distribution even in this case is a chi-squared one. The number of degrees of freedom in the test statistic equals $n(p + q + 1)$. The test can be robustified against non-normality in the same way as the previous one.

It is also possible to construct a test against a single structural break by adapting the test of Chu (1995) to the EGARCH case (see Hansen (1996) for obtaining critical values), but that has not been done here.

4.5 Testing EGARCH against GARCH

As the GARCH model of Bollerslev (1986) and Taylor (1986) is a very popular alternative to the EGARCH model, it would be useful in practice to also compare the estimated EGARCH model with its GARCH counterpart in order to see if one is to be preferred to the other. In this section we discuss three non-nested tests for testing EGARCH and GARCH models against each other. The question we pose is whether or not the GARCH model characterizes some features in the data that the EGARCH model is unable to capture. The tests can thus been seen as misspecification tests of the EGARCH model against the GARCH model or vice versa, depending on which one of the models is the null model. In the GARCH(p,q) model, the conditional variance is

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{4.19}$$

A sufficient condition for the conditional variance to be positive is $\alpha_0 > 0$, $\alpha_j \geq 0,\ j = 1, ..., q, \beta_j \geq 0, j = 1, ..., p$. The necessary and sufficient conditions for positivity of the conditional variance in higher-order GARCH models are complicated; see Nelson and Cao (1992).

The standard GARCH model has been extended to characterize asymmetric responses to shocks. The GJR-GARCH model (Glosten et al. (1993)) is obtained by adding $\sum_{j=1}^{q} \omega_j I(\varepsilon_{t-j})\varepsilon_{t-j}^2$ to the GARCH specification (4.19) where $I(\varepsilon_{t-1}) = 1$ if $\varepsilon_{t-1} < 0$, and $I(\varepsilon_{t-1}) = 0$ otherwise. A useful nonlinear version of the GJR-GARCH model is obtained by making the transition between regimes smooth. A smooth transition GARCH (STGARCH) model may be defined as (4.2) with

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_{1j} \varepsilon_{t-j}^2 + \sum_{j=1}^{q} \alpha_{2j} \varepsilon_{t-j}^2 G_n(\varepsilon_{t-j}; \gamma, \mathbf{c}) + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{4.20}$$

where ε_{t-j} is the transition variable. When $n = 1$, G_1 is the logistic function that controls the change of the coefficient of ε_{t-j}^2 from α_j to $\alpha_j + \omega_j$ as a function of ε_{t-j}. In that case, letting $\gamma \to \infty$ yields the GJR-GARCH model. For discussions of the STGARCH model, see Hagerud (1997), González-Rivera (1998), Anderson, Nam and Vahid (1999), Lanne and Saikkonen (2002) and Lundbergh and Teräsvirta (2002). The EGARCH model does not nest these models, and next we shall present three nonnested tests for testing EGARCH against GARCH. In particular, we are intrested in the case where the alternative is a GJR-GARCH model.

4.5.1 The encompassing test

In this subsection we consider an LM test suggested in Engle and Ng (1993). It is based on a minimal nesting model; see Mizon and Richard (1986). The idea is to construct model that encompasses both alternatives. Thus, decomposing $\ln h_t$ into two components

$$\ln h_t = \ln k_t + \ln g_t \tag{4.21}$$

where $k_t = \boldsymbol{\theta}'\mathbf{z}_t$ and $g_t = \exp(\boldsymbol{\phi}'\mathbf{x}_t)$ both k_t and g_t are functions of lags of h_t, $\mathbf{z}_t$ is a $k \times 1$ and $\mathbf{x}_t$ is a $m \times 1$ vector of explanatory variables, and $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are parameter vectors, yields such a model. Equation (4.21) shows that the model is another special case of the augmented EGARCH model (4.9). This model is the smallest model which encompasses both the EGARCH and GARCH models and can be used for testing EGARCH and GARCH models against each other. For example, if $\ln k_t = \alpha^* z_{t-1} + \psi^* |z_{t-1}| + \beta^* \ln h_{t-1}$ and $\ln g_t = \ln(\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \omega_1 I(\varepsilon_{t-1})\varepsilon_{t-1}^2 + \beta_1 h_{t-1})$, then the model encompassing the EGARCH(1,1) and GJR-GARCH(1,1) ones is

$$\begin{aligned} \ln h_t &= \alpha^* z_{t-1} + \psi^* |z_{t-1}| + \beta^* \ln h_{t-1} \\ &\quad + \ln(\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \omega_1 I(\varepsilon_{t-1})\varepsilon_{t-1}^2 + \beta_1 h_{t-1}) \end{aligned} \tag{4.22}$$

Setting $\alpha_1 = \omega_1 = \beta_1 = 0$ yields the EGARCH(1,1) model. On the other hand, $\alpha^* = \psi^* = \beta^* = 0$ corresponds to the GJR-GARCH(1,1) model. A test of the latter restrictions can also been seen as a misspecification test of the GJR-GARCH model against the EGARCH model. Considering the former case implies the null hypothesis $\mathrm{H}_0 : \alpha_1 = \omega_1 = \beta_1 = 0$. The relevant block of the score vector evaluated under H_0 has the form

$$\frac{\partial L_T}{\partial \boldsymbol{\beta}_r}|_{\mathrm{H}_0} = (1/2)\sum_{t=1}^{T}(\frac{\widehat{\varepsilon}_t^2}{\widehat{h}_t} - 1)\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r}|_{\mathrm{H}_0} \tag{4.23}$$

where

$$\frac{\partial \ln g_t}{\partial \boldsymbol{\beta}_r}|_{\mathrm{H}_0} = \frac{1}{\widehat{h}_t}(\widehat{\varepsilon}_{t-1}^2, I(\widehat{\varepsilon}_{t-1})\widehat{\varepsilon}_{t-1}^2, \widehat{h}_{t-1})' \tag{4.24}$$

and $\boldsymbol{\beta}_r = (\alpha_1, \omega, \beta_1)'$. Using previous notation, $\widehat{\mathbf{v}}_t = (\widehat{\varepsilon}_{t-1}^2/\widehat{h}_t, I(\widehat{\varepsilon}_{t-1})\widehat{\varepsilon}_{t-1}^2/\widehat{h}_t, \widehat{h}_{t-1}/\widehat{h}_t)'$ in the auxiliary TR^2 regression. Test statistic (4.14) has an asymptotic χ^2 distribution with three degrees of freedom when the null hypothesis is valid. The test can be robustified against non-normal errors in the same way as the tests in Section 4. It should be noted that when the null hypothesis is rejected, the rejection is against the encompassing model and not the GJR-GARCH one.

4.5.2 The Pseudo-Score Test

In this subsection we briefly describe a competing test suggested by Chen and Kuan (2002). It is based on the finite sample counterpart of the pseudo-true score, the limit of the expected value of the score function from the alternative model, where the expectation is taken with respect to the null model. Suppose that the null model is a EGARCH model, whereas the alternative model is a GARCH model. Then the pseudo-true score function is

$$PS_{GARCH} = \lim_{T\to\infty} E_{EGARCH(\boldsymbol{\theta})}\frac{1}{2T}\sum_{t=1}^{T}(\frac{\varepsilon_t^2}{h_t}-1)\frac{\partial \ln h_t}{\partial \boldsymbol{\beta}} \tag{4.25}$$

where $E_{EGARCH(\boldsymbol{\theta})}$ denotes the expectation taken with respect to the EGARCH model. When the GARCH model is the correct one, expectation (4.25) equals zero. The test can be constructed by checking if the estimate of (4.25) is sufficiently close to zero. As shown in Chen & Kuan (2002) the finite sample counterpart of the pseudo-true score equals

$$\widehat{PS}_{GARCH} = \frac{1}{2}\sum_{t=1}^{T}\frac{(\widehat{h}_{EGARCH,t}-\widehat{h}_t)}{\widehat{h}_t}\partial \ln \widehat{h}_t/\partial \boldsymbol{\beta} \tag{4.26}$$

where $\widehat{h}_{EGARCH,t}$ is the estimate of the conditional variance under the EGARCH model. The test statistic is

$$CK = T\widehat{PS}'_{GARCH}\widehat{\boldsymbol{\Omega}}^{-}\widehat{PS}_{GARCH} \tag{4.27}$$

where $\widehat{\boldsymbol{\Omega}}$ is a consistent estimator of the information matrix $\boldsymbol{\Omega}$, and $\widehat{\boldsymbol{\Omega}}^{-}$ is its generalized inverse. The test statistic is asymptotically distributed as chi-square with r degrees of freedom when the null hypothesis is true, where r is the rank of $\widehat{\boldsymbol{\Omega}}$. Note that $\boldsymbol{\Omega}$ is not of full rank. In our simulations we estimate $\boldsymbol{\Omega}$ using the estimate of $cov(\widehat{PS}_{GARCH})$ given in Chen and Kuan (2000).

4.5.3 Simulated likelihood ratio statistic

Our remaining test is the one proposed by Lee and Brorsen (1997) and Kim et al. (1998). The latter authors suggested it for testing the GARCH model against the autoregressive stochastic volatility model or vice versa. The test is based on the log likelihood ratio

$$LR = L_T(\widehat{\boldsymbol{\theta}}_{GARCH}) - L_T(\widehat{\boldsymbol{\theta}}_{EGARCH}) \tag{4.28}$$

where $L_T(\widehat{\boldsymbol{\theta}}_{GARCH})$ and $L_T(\widehat{\boldsymbol{\theta}}_{EGARCH})$ are the maximized log-likelihood function under the GARCH model and under the EGARCH model, respectively. The asymptotic distribution of LR under the hypothesis that the EGARCH model is the true model or under the hypothesis that the GARCH model is the true one is unknown and an empirical distribution is constructed by simulation. Under the assumption that the EGARCH model is true and that its parameter vector is $\widehat{\boldsymbol{\theta}}_{EGARCH}$, we generate N time series from the "true" model. For each simulated series we estimate the parameters of the GARCH and EGARCH models and record the value of LR^i, $i = 1, ..., N$. The resulting values of LR^i are a sample from the exact distribution of LR under the EGARCH model. This gives us the critical value LR_α of the test to which LR is compared. If $LR > LR_\alpha$ the null hypothesis is rejected. For a general discussion of Monte Carlo tests of this type; see Ripley (1987).

4.6 Simulation experiment

The above distribution theory is asymptotic, and we have to find out how our tests behave in finite samples. This is done by simulation. For all simulations we used the following data generating process (DGP)

$$\begin{aligned} y_t &= \varepsilon_t \\ \varepsilon_t &= z_t h_t^{1/2} \end{aligned} \tag{4.29}$$

where the definition of the conditional variance h_t depends on the test statistic to be simulated. Under the null hypothesis h_t is the conditional variance of the standard EGARCH(1,1) model (4.4) with $p = q = 1$. The random numbers, z_t, have been generated by the random number generator in GAUSS 3.2. The distribution for the random numbers sampled is either standard normal or a standardized (unit variance) generalized error $GED(\nu)$ distribution, see Nelson (1991). In the latter case, parameter ν is chosen such that the kurtosis $\mathsf{E}z_t^4 = 5$. The first 1000 observations of each generated series have been discarded to avoid initialization effects. Size experiments are performed with series of 1000 and 3000 observations. The empirical power of the tests is investigated using series of 1000 observations. We use 1000 replications in each experiment. Both the normality-based and the robust version of each test are considered.

4.6.1 Evaluation of EGARCH models

Testing against a higher-order EGARCH model

First, we consider the test against a higher-order EGARCH model. We define a DGP such that the conditional variance follows a symmetric EGARCH(2,2) process. Thus,

$$\begin{aligned} \ln h_t &= -0.00127 + 0.11605(|z_{t-1}| - E\,|z_{t-1}|)\} + 0.95 \ln h_{t-1} + \\ &\quad + \psi_2(|z_{t-2}| - E\,|z_{t-2}|) + \beta_2 \ln h_{t-2}. \end{aligned} \tag{4.30}$$

The values of ψ_2 and β_2 are chosen being varied in simulations. The moment structure of the EGARCH(p,q) model has been worked out in He (2000). For $\psi_2 = \beta_2 = 0$ the DGP reduces to a symmetric EGARCH(1,1) model. In the simulations these tests were all computed with a singel parameter in the alternative. That is, for different values of ψ_2 we choose $\widehat{\mathbf{v}}_t = (|\widehat{\varepsilon}_{t-2}| / \widehat{h}_{t-2})$. For different values of β_2 we choose $\widehat{\mathbf{v}}_t = \ln \widehat{h}_{t-2}$. The asymptotic null distribution is thus $\chi^2(1)$. The actual rejection frequencies based on the significant level 0.05 under the asymptotic distribution are reported.

The results of both size and power simulations can be found in Table 4.1. They indicate that both test is well sized for $T = 1000$. When the errors are normal, the nonrobust test is somewhat more powerful than the robust one. When the error distribution is a $GED(5)$ one, the robust test is more powerful than the nonrobust one. A tentative recommendation would be to always use the robustified test unless there is strong evidence of the errors being normally distributed.

Testing parameter constancy

We consider two cases of parameter nonconstancy for the symmetric EGARCH(1,1) model: the DGP is a EGARCH with either (a) a single or (b) double structural break in the intercept. Our test is computed using $n = 1$ in case (a), and $n = 2$ in case (b).

We consider the following symmetric models with a break in the constant term

$$\begin{aligned} \ln h_t &= -0.00127 + 0.11605(|z_{t-1}| - E\,|z_{t-1}|) + 0.95 \ln h_{t-1}, \\ \text{(a) } t &< \eta T,\ \text{(b) } t < \eta_1 T,\ t > \eta_2 T, \\ \ln h_t &= -0.03593 + 0.11605(|z_{t-1}| - E\,|z_{t-1}|) + 0.95 \ln h_{t-1}, \\ \text{(a) } t &\geqslant \eta T,\ \text{(b) } \eta_1 T \le t \le \eta_2 T \end{aligned} \tag{4.31}$$

where T is the sample size and $0 \le \eta, \eta_1, \eta_2 \le 1$. The parameters under the null hypothesis are chosen to mimic one of the sets of parameter values considered in Engle and Ng (1993), see below. In the simulation experiment, the unconditional variance is halved at ηT and $\eta_1 T$. Even if only the intercept changes in (4.31), we assume that under the alternative, the break affects all three parameters. The asymptotic null distribution is thus $\chi^2(3)$.

The results of both size and power simulations can be found in Table 4.2. They indicate that both tests are somewhat oversized for $T = 1000$ but well-sized for $T = 3000$. When the errors are normal, the nonrobust test is more powerful than the robust one for both a single and a double structural break in the intercept. When the error distribution is a $GED(5)$ one, the robust test is more powerful than the nonrobust one. A tentative recommendation would be similar to the previous one: use the robustified test unless there is strong evidence favouring normal errors.

4.6.2 Testing EGARCH against GARCH

In this section we consider the small-sample performance of the nonnested tests of testing EGARCH against GARCH. First we consider symmetric, then asymmetric models.

Symmetric models

We consider six pairs of parameter vectors for the GARCH(1,1) and the symmetric EGARCH(1,1) model. They can be found in Table 4.3. For GARCH, $\alpha_1 + \beta_1$ is the exponential decay rate of the autocorrelations of squared observations, which has been used as a measure for persistence in volatility. We choose three different values of the persistence. Engle and Ng (1993) used the same values in their simulation experiments. The three parameters, α_0,α_1 and β_1, are selected such that the unconditional variance $\mathsf{E}\varepsilon_t^2$ equals unity but the kurtosis equals either 6 or 12. These parameter values are obtained from the analytic expressions of the second moment, the kurtosis and the autocorrelation function of squared observations of a family of GARCH models with normal or t distributed errors that are available in He and Teräsvirta (1999). This family includes the GJR-GARCH model.

The parameter values for the EGARCH model are chosen to be as comparable with the ones for the GARCH models as possible. Thus, β is set equal to $\alpha_1 + \beta_1$ in the corresponding GARCH model, as β controls the decay of the autocorrelation function of the squared observations in the EGARCH model. Note, however, that while the decay rate of the autocorrelation of ε_t^2 in the GARCH(1,1) model equals $\alpha_1 + \beta_1$, it only approaches β from below with increasing lag length in the EGARCH(1,1) model. Parameters α_0 and ψ are chosen such that the unconditional variance and the kurtosis are the same in both models as well. This can be done using the analytic expressions for the relevant moments of the EGARCH(1,1) model in Chapter 1. The parameters of the EGARCH model (4.31) under the null hypothesis are chosen to mimic one of the sets of GARCH parameter values considered in Engle and Ng (1993). They can be found in Table 4.3. In simulating the LR statistic we use 99 replications to construct the empirical null distribution.

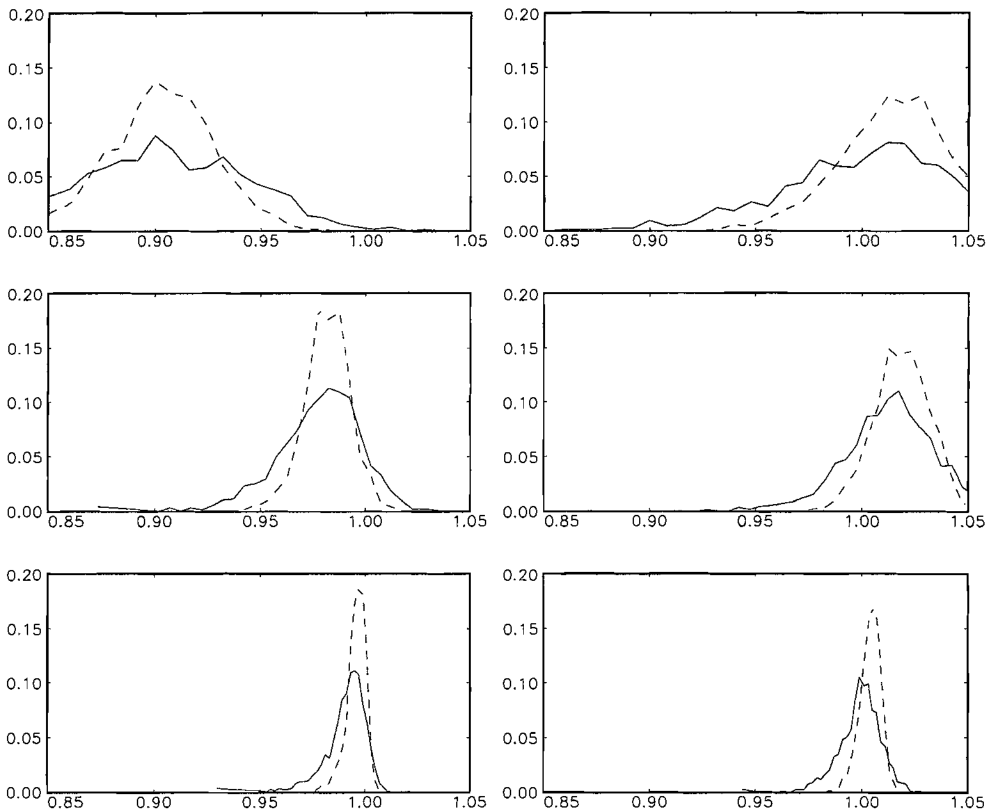
A general result valid for all our simulation experiments is that the size of both the encompassing test and the simulated LR test is close to the nominal size already at $T = 1000$, see Tables 4.4 and 4.5. As to the pseudo-score statistic, it is oversized even for $T = 3000$. This is due to the estimated matrix $\widehat{\mathbf{\Omega}}$. In our experiment the rank of the (3×3) matrix $\mathbf{\Omega}$ equals two, but the estimated matrix $\widehat{\mathbf{\Omega}}$ is not seriously ill-conditioned. This causes the test statistic to be oversized. In fact, assuming rank equal to three, that is, using the $\chi^2(3)$ distribution instead of $\chi^2(2)$ as the null distribution, would not be such a bad idea for $T = 1000$. The test would be conservative (undersized), but not overly so.

The case of $\alpha_1 + \beta_1 = 1$ in the GARCH model has received attention in the literature. Engle and Bollerslev (1986) called the model with this restriction the integrated GARCH (IGARCH) model. The behaviour of $\widehat{\alpha}_1 + \widehat{\beta}_1$ when the true model is GARCH(1,1) with $\alpha_1 + \beta_1 < 1$ has also received attention. In that case there is a substantial probability of estimating this persistence parameter to be greater than one when T is small; see Shephard (1996). Figure 4.1 contains the estimated density of $\widehat{\alpha}_1 + \widehat{\beta}_1$ when the true model is a symmetric EGARCH(1,1). If we generate data from a stationary EGARCH(1,1) model with normal errors and fit a GARCH(1,1) model with normal errors to the observations, there is a large probability of 'finding IGARCH', that is, ending up with $\widehat{\alpha}_1 + \widehat{\beta}_1 \geq 1$. Furthermore, this probability increases with the sample size. Lamoureux and Lastrapes (1990) obtained a similar result when they generated data with a GARCH(1,1) model with a structural break, but here the DGP is a constant-parameter stationary EGARCH model. In simulating the LR statistic we only use the replications with $\widehat{\alpha}_1 + \widehat{\beta}_1 < 1$. We discard the rest and add new ones until there are 1000 replications in each experiment.

Another result valid for all our simulations is that the simulated LR test is more powerful than the encompassing test, see Table 4.6. Because of the size problems, the power of the pseudo-score test is not comparable. For all tests, the power is higher for models with low than with high persistence.

We use the EGARCH model (4.31) under the null hypothesis and the corresponding GARCH model in the simulation experiments of the robust version of the tests. They can be found in Table 4.3. Because of size problems, the behaviour of the pseudo-score test is not investigated. Our results indicate that the nonrobust version of the simulated LR test is undersized when the error distribution is a *GED*, see Table 4.7. The robust version of the encompassing test is undersized for $T = 1000$ but well-sized for $T = 3000$, see Table 4.8.

Figure 4.1 Estimated density of $\widehat{\alpha}_1 + \widehat{\beta}_1$ for the GARCH model when DGP is the symmetric EGARCH model. Upper panel, left: E5, upper panel, right: E6, middle panel, left: E3, middle panel, right: E4, lower panel, left: E1, lower panel, right: E2. Number of observations 1000 (solid) and 3000 (dashed).

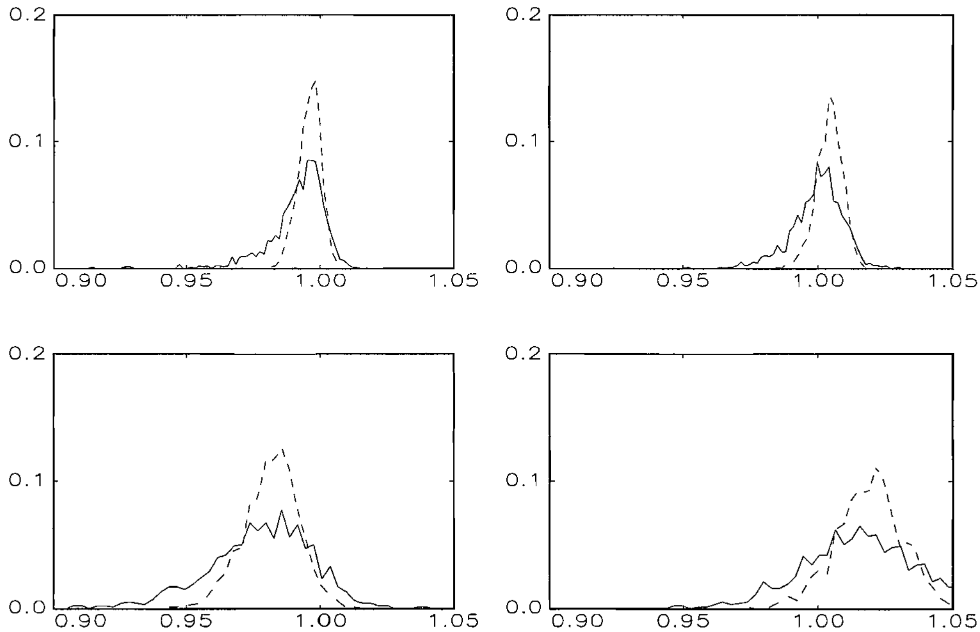


In Table 4.9 we report the result of the power simulations. When the errors are normal, the robust tests perform as well as the nonrobust ones. When the error distribution is a *GED*, the robust tests are always more powerful than the nonrobust ones. The simulated LR test is more powerful than the encompassing test.

Asymmetric models

We now turn to asymmetric models. We add the asymmetric component ϕz_{t-1} to four of the symmetric EGARCH models assuming $\phi = -0.04$. The asymmetry introduced through ϕ tends to increase kurtosis and, at the same time, reduce the the first-order autocorrelation of squared observations. We choose α_0 and ψ such that the unconditional variance equals unity and the kurtosis is 6 or 12 even here. The parameter values of the asymmetric EGARCH models are reported in Table 4.3. We use the reduction in the first-order autocorrelation of squared observations, in percentage, due to the asymmetry as an additional condition to identify the sets of parameters in the GJR-GARCH model. The corresponding GJR-GARCH models are also found in Table 4.3.

Figure 4.2 Estimated density of $\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\omega}/2$ for the GJR-GARCH model when DGP is the EGARCH model. Upper panel, left: asE3, upper panel, right: asE4, lower panel, left: asE1, lower panel, right: asE2. Number of observations 1000 (solid) and 3000 (dashed).



The results in the asymmetric case are similar to the symmetric case. The probability of finding GJR-IGARCH model in which $\alpha_1 + \beta_1 + \omega_1/2 = 1$ when the observations have been generated by an EGARCH(1,1) model is quite large, see Figure 4.2. The size of the encompassing test is close to the nominal size already at $T = 1000$, see Table 4.10. As to the pseudo-score statistic, it is oversized in the case of asymmetric models for $T = 3000$; see Table 4.11.

In Table 4.12 we report result of the power simulations. Note that the results appearing in the tables are not based on size-adjusted tests. For both tests, the power is higher for models with low than the ones with high persistence.

A conclusion of our simulations is that the simulated LR is more powerful than the encompassing test. The pseudo-score test in the form applied in this paper cannot be recommended because of the size problems pointed out.

4.7 Empirical example

In this section we apply our tests to daily return series of the 29 most actively traded stocks in Stockholm stock exchange. The list of stocks appears in Table 4.16 together with information about the length of the series. The period investigated ends April 24, 2001. The return series are continuously compounded returns calculated from the closing prices obtained from Datastream.

In Table 4.13 we report results of the test against a higher-order model. There is some evidence of a need for an EGARCH(1,2) model. The p-value of the test lies between 0.01 and 0.05 in 13 cases out of 29, but there is only one occasion in which it does not exceed 0.01. An EGARCH(2,1) model is not a likely alternative. As a whole it seems that the need for higher-order EGARCH models is not very strong. Table 4.14 containes results of the parameter constancy test. In almost about half of the cases, there is strong evidence of time-varying parameters. It seems that nonconstancy of the intercept is often a strong reason for rejection. This suggests that the unconditional variance of the series changes over time. The dynamic behaviour of the conditional variance may be less prone to change in time.

Turning to choosing between EGARCH and GARCH, Table 4.15 contains results based on the robust version of the encompassing test and the simulated LR test for testing GJR-GARCH(1,1) and EGARCH(1,1) models against each other. They indicate that both models fit the data more or less equally well. In most cases there is no clear difference between the models. The encompassing test does not reject either model in 16 cases. The simulated LR test does not reject either model in 9 cases. It is rare that both models are rejected simultaneously. For the encompassing test this happens only once. The EGARCH model is rejected more often than the GJR-GARCH model. Because of the size problems, the pseudo-score test is not applied to these series.

The main conclusion of the empirical example is that there is substantial evidence for parameter nonconstancy. Rejections, measured in p-values, are generally weaker for the other tests applied to the estimated models.

4.8 Conclusions

In this paper we consider misspecification tests for an EGARCH(p,q) model. We derive two new misspecification tests for an EGARCH model. Since both test statistics are asymptotically χ^2-distributed under the null hypothesis, possible misspecification of an EGARCH model can be detected at low computational cost. Because the tests of an EGARCH model against a higher-order EGARCH model and testing parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This is useful for a model builder who wants to find out possible weakness of estimated specification. It may also give him/her useful ideas of how the model could be further improved. These tests may be viewed as the EGARCH counterpart of the tests for the GARCH model in Lundbergh and Teräsvirta (2002).

Furthermore, we investigate various way of testing the EGARCH model against GARCH ones as another check of model adequacy. Our simulations show that the simulated LR test is more powerful than the encompassing test and that the size of the test may be a problem in applying the pseudo-score test.

Finally, the simulation results indicate that in practice, the robust versions of our tests should be preferred to nonrobust ones. They can be recommended as standard tools when it comes to testing the adequacy of an estimated EGARCH (p,q) model.

Bibliography

Anderson, H. M., Nam, K. and Vahid, F. (1999), *Asymmetric Nonlinear Smooth Transition GARCH Models*, In Rothman, P (Ed.), Nonlinear time series analysis of economic and financial data. Kluwer, Boston.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**, 307–327.

Chen, Y.-T. and Kuan, C.-M. (2000), The pseudo-true score encompassing test for non-nested hypotheses. Working Paper Academia Sinica, Taipei, Taiwan.

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.

Engle, R. F. and Bollerslev, T. (1986), 'Modelling the persistence of conditional variances', *Econometric Reviews* **5**, 1–50.

Engle, R. F. and Ng, V. K. (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance* **48**, 1749–1778.

Geweke, J. (1986), 'Modelling persistence of conditional variances: Comment', *Econometric Reviews* **5**, 57–61.

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.

González-Rivera, G. (1998), 'Smooth transition GARCH models', *Studies in Non-linear Dynamics and Econometrics* **3**, 61–78.

Hagerud, G. E. (1997), *A New Non-Linear GARCH Model*, Stockholm: EFI Economic Research Institute.

Hansen, B. (1996), 'Inference when a nuisance parameter is not identified under the null hypothesis', *Econometrica* **64**, 413–430.

He, C. (2000), Moments and the autocorrelation structure of the exponential GARCH(p,q) process. Working Paper Series in Economics and Finance, Stockholm School of Economics.

He, C. and Teräsvirta, T. (1999), 'Properties of moments of a family of GARCH processes', *Journal of Econometrics* **92**, 173–192.

He, C., Teräsvirta, T. and Malmsten, H. (2002), 'Moment structure of a family of first-order exponential GARCH models', *Econometric Theory* **18**, 868–885.

Kim, S., Shephard, N. and Chib, S. (1998), 'Stochastic volatility: Likelihood inference and comparison with ARCH models', *Review of Economic Studies* **65**, 361–93.

Lamoureux, C. G. and Lastrapes, W. D. (1990), 'Persistence in variance, structural change and the GARCH model', *Journal of Business and Economic Statistics* **8:2**, 225–234.

Lanne, M. and Saikkonen, P. (2002), Nonlinear GARCH models for highly persistent volatility. Discussion Paper 20/2002 Humboldt-Universität zu Berlin.

Lee, J.-H. and Brorsen, B. W. (1997), 'A non-nested test of GARCH vs. EGARCH models', *Applied Economics Letters* **4**, 765–768.

Li, W. K. and Mak, T. K. (1994), 'On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity', *Journal of Time Series Analysis* **15**, 627–636.

Lin, C.-F. and Teräsvirta, T. (1994), 'Testing the constancy of regression parameters against continous structural change', *Journal of Econometrics* **62**, 211–228.

Lundbergh, S. and Teräsvirta, T. (2002), 'Evaluating GARCH models', *Journal of Econometrics* **110**, 417–435.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988), 'Testing linearity against smooth transition autoregressive models', *Biometrika* **75**, 491–499.

Mizon, G. E. and Richard, J. F. (1986), 'The encompassing principle and its application to testing non-nested hypotheses', *Econometrica* **54**, 657–678.

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.

Pantula, S. G. (1986), 'Modelling persistence of conditional variances: Comment', *Econometric Reviews* **5**, 71–74.

Ripley, B. D. (1987), *Stochastic Simulation*, Wiley, New York.

Taylor, S. (1986), *Modelling Financial Time Series*, Wiley, New York.

Wooldridge, J. (1991), 'On the application of robust-based diagnostics to model of conditional means and conditional variances', *Journal of Econometrics* **47**, 5–46.

Tables

Table 4.1 Empirical p-values of the nonrobust and robustified test against higher-order model when the observations have been generated from model (4.30) using normal and GED(5) errors, for 1000 and 3000 observations, based on 1000 replications. The nominal significance level equals 0.05.

Error distribution			Normal(0,1)		GED(5)	
ψ_2	$-\beta_2$	T	Nonrobust	Robustified	Nonrobust	Robustified
0	0	1000	4.3	4.2	4.6	4.9
0	0	3000	5.2	4.7	5.0	4.7
0.1	0	1000	5.5	5.4	6.9	18.8
0.2	0	1000	13.5	10.6	11.9	25.8
0	0.1	1000	6.2	6.1	5.3	18.3

Table 4.2 Empirical p-values of the nonrobust and robustified parameter constancy test when the observations have been generated from model (4.31) using normal and GED(5) errors, for 1000 and 3000 observations, based on 1000 replications. The nominal significance level equals 0.05.

Error distribution				Normal(0,1)		GED(5)	
η	η_1	η_2	T	Nonrobust	Robustified	Nonrobust	Robustified
1	–	–	1000	8.2	7.4	6.8	6.2
1	–	–	3000	6.6	5.3	6.6	5.5
0.25	–	–	1000	55.8	50.8	32.7	36.0
0.5	–	–	1000	80.8	75.9	54.3	59.2
–	0.333	0.667	1000	31.2	23.5	16.2	25.7
–	0.25	0.75	1000	63.4	57.4	32.1	45.2

Table 4.3 This table reports the DGP's in the simulations. The first six pairs of parameters are the symmetric models in the simulation experiment. The next four are the asymmetric models. The last pair is for the simulation experiments of the robust version of the tests.

	GARCH					EGARCH			
Model	α_0	α_1	β	ω	Model	$-\alpha_0^*$	ψ	β^*	$-\phi$
G1	0.01	0.0705	0.9195	-	E1	0.0033	0.1876	0.99	-
G2	0.01	0.0864	0.9036	-	E2	0.0065	0.2614	0.99	-
G3	0.05	0.1561	0.7939	-	E3	0.0155	0.3974	0.95	-
G4	0.05	0.1912	0.7588	-	E4	0.0299	0.5453	0.95	-
G5	0.2	0.3	0.5	-	E5	0.0565	0.7095	0.8	-
G6	0.2	0.3674	0.4326	-	E6	0.1064	0.9555	0.8	-
GJR1	0.01	0.0422	0.9251	0.0453	asE1	0.0032	0.1731	0.99	0.04
GJR2	0.01	0.0562	0.9112	0.0495	asE2	0.0064	0.2504	0.99	0.04
GJR3	0.05	0.1202	0.7986	0.0624	asE3	0.0154	0.3891	0.95	0.04
GJR4	0.05	0.1515	0.8036	0.0698	asE4	0.0297	0.5383	0.95	0.04
RG	0.05	0.05	0.9	-	RE	0.0013	0.1161	0.95	-

Table 4.4 Empirical p-values from size simulations of three tests of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Simulated LR		Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%	5%	10%
G1	5.7	11.2	5.0	9.6	6.1	13.0
G2	5.6	10.6	4.8	9.7	6.4	13.0
G3	5.7	11.4	4.4	9.6	8.4	15.1
G4	5.9	11.7	4.9	9.5	8.1	16.1
G5	5.6	11.1	4.8	9.4	7.4	17.1
G6	5.1	10.7	4.5	8.8	8.4	16.4
E1	5.5	11.1	5.4	10.5	6.5	13.1
E2	4.8	9.7	5.7	11.0	7.0	14.2
E3	5.9	12.0	4.8	10.0	9.4	15.4
E4	5.4	10.5	5.0	10.1	9.8	16.0
E5	5.1	9.9	5.3	10.1	11.0	18.3
E6	5.0	9.7	4.9	9.6	11.0	17.7

Table 4.5 Empirical p-values from size simulations of two tests of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 3000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
G1	5.2	9.8	5.7	11.4
G2	4.9	10.0	5.2	11.7
G3	5.1	9.7	5.5	11.6
G4	5.3	10.1	6.2	11.7
G5	5.1	10.2	7.4	15.5
G6	4.1	10.0	7.8	15.6
E1	5.2	10.4	6.0	13.0
E2	5.1	10.3	7.7	15.5
E3	4.5	9.6	6.6	13.8
E4	5.8	10.3	7.0	13.9
E5	4.4	9.3	5.9	12.6
E6	4.7	10.0	5.3	13.2

Table 4.6 Empirical p-values from power simulations of three tests of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Simulated LR		Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%	5%	10%
G1	42.3	53.8	13.2	22.2	21.9	33.2
G2	45.8	57.2	12.4	20.7	26.0	39.3
G3	56.7	68.4	22.2	34.0	64.5	76.2
G4	65.8	74.2	25.0	37.9	89.3	93.8
G5	87.2	97.3	57.3	71.4	21.9	33.2
G6	94.1	99.1	74.7	84.9	26.0	39.3
E1	48.0	61.4	17.4	26.0	20.0	29.8
E2	39.0	51.8	20.7	32.8	24.9	38.6
E3	69.9	80.7	26.1	35.0	44.4	57.8
E4	71.7	79.1	30.3	40.7	61.1	70.8
E5	89.4	99.1	43.2	53.4	64.5	76.2
E6	97.8	99.4	51.0	61.1	89.3	93.8

Table 4.7 Empirical p-values from size simulations of two tests, nonrobust and robust version, of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05.

	Normal(0,1)				t(7) or GED(5)			
	Engle-Ng		Simulated LR		Engle-Ng		Simulated LR	
True model	Nonr.	Robust	Nonr.	Robust	Nonr.	Robust	Nonr.	Robust
RG	4.8	4.6	5.5	5.9	3.3	3.1	2.3	5.4
RE	4.4	4.5	5.8	6.0	3.6	3.3	2.1	5.9

Table 4.8 Empirical p-values from size simulations of two tests, nonrobust and robust version, of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 3000 observations, based on 1000 replications. The nominal significance level equals 0.05.

	Normal(0,1)		t(7) or GED(5)	
	Engle-Ng		Engle-Ng	
True model	Nonr.	Robust	Nonr.	Robust
RG	4.9	4.6	4.1	5.0
RE	4.2	4.2	3.2	4.3

Table 4.9 Empirical p-values from power simulations of two tests, nonrobust and robust version, of testing symmetric EGARCH(1,1) against GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05.

	Normal(0,1)				t(7) or GED(5)			
	Engle-Ng		Simulated LR		Engle-Ng		Simulated LR	
True model	Nonr.	Robust.	Nonr.	Robust.	Nonr.	Robust.	Nonr.	Robust
RG	43.9	37.9	50.1	53.1	19.6	14.8	13.6	40.2
RE	20.2	25.1	45.3	43.2	17.0	23.4	23.8	48.8

Table 4.10 Empirical p-values from size simulations of two tests of testing EGARCH(1,1) against GJR-GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
GJR1	5.8	11.4	11.3	19.0
GJR2	5.7	10.8	11.9	19.0
GJR3	5.8	11.3	10.8	19.3
GJR4	6.4	11.7	12.9	21.8
asE1	5.8	11.8	7.8	16.6
asE2	5.7	11.0	7.6	14.9
asE3	5.6	11.2	10.4	18.3
asE4	5.2	10.7	10.5	18.4

Table 4.11 Empirical p-values from size simulations of two tests of testing EGARCH(1,1) against GJR-GARCH(1,1) and vice versa, 3000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
GJR1	5.7	10.7	5.8	12.4
GJR2	5.5	10.4	7.7	14.0
GJR3	5.6	10.4	7.4	13.6
GJR4	5.7	10.6	8.4	16.3
asE1	5.6	11.2	4.3	10.0
asE2	5.4	10.6	5.7	12.2
asE3	5.4	10.7	5.7	12.1
asE4	5.0	10.2	5.8	12.4

Table 4.12 Empirical p-values from power simulations of two tests of testing EGARCH(1,1) against GJR-GARCH(1,1) and vice versa, 1000 observations, based on 1000 replications. The nominal significance level equals 0.05 and 0.1.

	Engle-Ng		Pseudo-Score	
True model	5%	10%	5%	10%
GJR1	23.2	34.7	19.8	31.1
GJR2	20.5	32.0	26.0	37.2
GJR3	30.2	45.0	43.1	55.2
GJR4	34.9	50.4	62.5	73.0
asE1	20.1	36.1	11.8	21.1
asE2	24.1	39.4	14.1	22.4
asE3	32.7	44.4	19.1	29.6
asE4	41.3	53.1	22.1	34.2

Table 4.13 Rejection frequencies of the robust version of the test against a higher-order model.

	$p \leq 0.01$	$0.01 < p \leq 0.05$	$0.05 < p \leq 0.1$	$p > 0.1$
EGARCH(2,1)	0	0	2	27
EGARCH(1,2)	1	12	10	6

Table 4.14 Rejection frequencies of the robust version of the parameter constancy test with n=2.

	$p \leq 0.001$	$0.001 < p \leq 0.01$	$0.01 < p \leq 0.1$	$p > 0.1$
Intercept parameter	9	3	7	10
All four parameters	12	2	7	8

Table 4.15 Rejection frequencies of the EGARCH(1,1) and GJR-GARCH(1,1) models from the encompassing test and the simulated LR (between brackets) test for the 29 series. 1 simulation of the LR test do not converge.

		H_0:EGARCH			
		$p \leq 0.01$	$0.01<p\leq 0.05$	$p > 0.05$	$\sum$
H_0:GARCH	$p \leq 0.01$	1(0)	1(0)	2(2)	4(2)
	$0.01<p\leq 0.05$	0(0)	0(1)	0(4)	0(5)
	$p > 0.05$	6(6)	3(6)	16(9)	25(21)
	$\sum$	7(6)	4(7)	18(15)	29(28)

Table 4.16 This table lists the 29 stocks investigated. The column labeled "T", reports the number of observations.

y	T
ABB	3717
Assa A.	1617
Assi D.	1769
Astra	3591
Atlas C.	2915
Autoliv	1690
Electrolux	4577
Ericsson	4576
FSB	1470
Gambro	2454
Holmen	4568
Industriv.	2061
Investor	4146
Nokia	2907
OMG	2084
Pharmacia	1370
Sandvik	4576
Scania	1268
Securitas	2461
Skandia	4566
SEB	2984
Skanska	4337
SKF	4578
SSAB	2963
Stora	3263
SCA	4576
SHB	2612
Sw. Match	1239
VOLVO	5324

EFI
The Economic Research Institute

Reports since 1999
A complete publication list can be found at www.hhs.se/efi

Published in the language indicated by the title

2004
Jutterström, M., Att påverka beslut – företag i EUs regelsättande.
Larsson, P., Förändringens villkor. En studie av organisatoriskt lärande och förändring inom skolan.
Salabasis, M., Bayesian Time Series and Panel Models – Unit Roots, Dynamics and Random Effects.
Skallsjö, S., Essays on Term Structure and Monetary Policy.
Söderström, J., Från Produkt till Tjänst. Utveckling av affärs- och miljöstrategier i produktorienterade företag.

2003
Andersson, H., Valuation and Hedging of Long-Term Asset-Linked Contracts.
Bergman, M., Essays on Human Capital and Wage Formation.
Damsgaard, N., Deregulation and Regulation of Electricity Markets.
Eklund, B., Four Contributions to Statistical Inference in Econometrics.
Hakkala, K., Essays on Restructuring and Production Decisions in Multi-Plant Firms.
Holgersson, C., Rekrytering av företagsledare. En studie i hoimosocialitet.
Ivaschenko, I, Essays on Corporate Risk, U.S. Business Cycles, International Spillovers of Stock Returns, and Dual Listing.
Lange, F., Brand Choice in Goal-derived Categories - What are the Determinants?
Le Coq, C., Quantity Choices and Market Power in Electricity Market.
Magnusson, P.R., Customer-Oriented Product Development - Experiments Involving Users in Service Innovation.
Meisiek, S., Beyond the Emotional Work Event Social Sharing of Emotion in Organizations.
Mårtensson, A.. Managing Mission-Critical IT in the Financial Industry.
Nilsson, G., Processorientering och styrning – Regler, mål eller värderingar?
Sandberg, R., Corporate Consulting for Customer Solutions Bridging Diverging Business Logics.
Sturluson, J.T., Topics in the Industrial Organization of Electricity Markets.
Tillberg, U., Ledarskap och samarbete – En jämförande fallstudie i tre skolor.
Waldenström, D., Essays in Historical Finance.
Wallén, U., Effektivitet i grundskolan i anslutning till en stadsdelsnämndsreform.
Ögren, A., Empirical Studies in Money, Credit and Banking - The Swedish Credit Market in Transition under the Silver and the Gold Standards, 1834 – 1913.

2002
Barinaga, E., Levelling Vagueness – A study of cultural diversity in an international project group.
Berglund, J., De otillräckliga - En studie av personalspecialisternas kamp för erkännande och status.
Bolander, P., Anställningsbilder och rekryteringsbeslut.

Damjanovic, T., Essays in Public Finance.
Ekman, M., Studies in Health Economics – Modelling and Data Analysis of Costs and Survival.
Företagerskan – Om kvinnor och entreprenörskap. Holmquist, C. och Sundin, E (red)
Heyman, F., Empirical Studies on Wages, Firm Performance and Job Turnover.
Kallifatides, M., Modern företagsledning och omoderna företagsledare.
Kaplan, M., Acquisition of Electronic Commerce Capability - The Cases of Compaq and Dell in Sweden.
Mähring, M., IT Project Governance.
Nilsson, M., Essays in Empirical Corporate Finance and Governance.
Rekrytering av koncernstyrelsen – Nomineringsförfaranden och styrelsesammansättning med focus på kvinnors ställning och möjligheter. Sjöstrand, S-E. och Petrelius, P.,(red)
Scener ur ett företag – Organiseringsteori för kunskapssamhället. Löwstedt, J. Stymne, B.,(red).
Schenkel, A., Communities of Practice or Communities of Discipline - Managing Deviations at the Øresund Bridge.
Schuster, W., Företagets Valutarisk – En studie av horisontella och vertikala styrprocesser.
Skogsvik, S., Redovisningsmått, värderelevans och informationseffektivitet.
Sundén, D., The Dynamics of Pension Reform.
Ternström, I., The Management of Common-Pool Resources - Theoretical Essays and Empirical Evidence.
Tullberg, J., Reciprocitet – Etiska normer och praktiskt samarbete.
Westling, G., Balancing Innovation and Control – The Role of Face-to-face Meetings in Complex Product Development Projects.
Viklund, M., Risk Policy – Trust, Risk Perception, and Attitudes.
Vlachos, J., Risk Matters - Studies in Finance, Trade and Politics.

2001

Adolfson, M., Monetary Policy and Exchange Rates – Breakthrough of Pass-Through.
Andersson, P., Expertise in Credit Granting: Studies on Judgment and Decision-Making behavior.
Björklund, C., Work Motivation - Studies of its Determinants and Outcomes.
Center for Management and Organization 50 (1951-2001).
Charpentier, C., Uppföljning av kultur- och fritidsförvaltningen efter stadsdelsnämnds-reformen.
Dahlén, M., Marketing on the Web - Empirical Studies of Advertising and Promotion Effectiveness.
Eckerlund, I., Essays on the Economics of Medical Practice Variations.
Ekelund, M., Competition and Innovation in the Swedish Pharmaceutical Market.
Engström, S., Success Factors in Asset Management.
Ericsson, D., Kreativitetsmysteriet – Ledtrådar till arbetslivets kreativisering och skrivandets metafysik.
Eriksson, R., Price Responses to Changes in Costs and Demand.
Frisell, L., Information and Politics.
Giordani, P., Essays in Monetary Economics and Applied Econometrics.
Gustavsson, P., Essays on Trade, Growth and Applied Econometrics.

Hedlund, A., Konsumentens erfarenhet – och dess inverkan på livsmedelsinköp på Internet.
Hill, M., Essays on Environmental Policy Analysis: Computable General Equilibrium Approaches Applied to Sweden.
Hvenmark, J., Varför slocknar elden? Om utbrändhet bland chefer i ideella organisationer.
Hägglund, P.B., Företaget som investeringsobjekt – Hur placerare och analytiker arbetar med att ta fram ett investeringsobjekt.
Höök, P., Stridspiloter i vida kjolar, om ledarutveckling och jämställdhet.
Johansson, C., Styrning för samordning.
Josephson, J., Evolution and Learning in Games.
Kjellberg, H., Organising Distribution - Hakonbolaget and the efforts to rationalise food distribution, 1940-1960.
Lange, F. och Wahlund, R., Category Management – När konsumenten är manager.
Liljenberg, A., Customer-geared competition – A socio-Austrian explanation of Tertius Gaudens.
Lindkvist, B., Kunskapsöverföring mellan produktutvecklingsprojekt.
Ljunggren, U., Nyckeltal i grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen.
Läkemedel – Kostnad eller resurs för sjukvården? Jönsson, B.,(red).
Löf, M., On Seasonality and Cointegration.
Martensen, K., Essays on Entry Externalities and Market Segmentation.
Matros, A., Stochastic Stability and Equilibrium Selection in Games.
Mårtensson, P., Management Processes – An Information Perspective on Managerial Work.
Nilsson, A., Market Transparency.
Norberg, P., Finansmarknadens amoralitet och det kalvinska kyrkorummet – En studie i ekonomisk mentalitet och etik.
Persson, B., Essays on Altruism and Health Care Markets.
Rech, G., Modelling and Forecasting Economic Time Series with Single Hidden-layer Feedforward Autoregressive Artificial Neural Networks.
Skoglund, J., Essays on Random Effects Models and GARCH.
Strand, N., Empirical Studies of Pricing.
Thorén, B., Stadsdelsnämndsreformen och det ekonomiska styrsystemet - Om budgetavvikelser.

2000

Berg-Suurwee, U., Styrning före och efter stadsdelsnämndsreform inom kultur och fritid – Resultat från intervjuer och enkät.
Bergkvist, L., Advertising Effectiveness Measurement: Intermediate Constructs and Measures.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Koncernchefen och ägarna.
Bornefalk, A., Essays on Social Conflict and Reform.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform.
Edman, J., Information Use and Decision Making in Groups.
Emling, E., Svenskt familjeföretagande.
Ericson, M., Strategi, kalkyl, känsla.
Gunnarsson, J., Wahlund, R., Flink, H., Finansiella strategier i förändring: segment och beteenden bland svenska hushåll.

Hellman, N., Investor Behaviour – An Empirical Study of How Large Swedish Institutional Investors Make Equity Investment Decisions.
Hyll, M., Essays on the Term Structure of Interest Rates.
Håkansson, P., Beyond Private Label – The Strategic View on Distributor Own Brands.
I huvudet på kunden. Söderlund, M (red). *EFI och Liber Förlag.*
Karlsson Stider, A., Familjen och firman.
Ljunggren, U., Styrning av grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen – Resultat från intervjuer och enkät.
Ludvigsen, J., The International Networking between European Logistical Operators.
Nittmar, H., Produktutveckling i samarbete – Strukturförändring vid införande av nya Informationssystem.
Robertsson, G., International Portfolio Choice and Trading Behavior.
Schwarz, B., Weinberg, S., Serviceproduktion och kostnader – att söka orsaker till kommunala skillnader.
Stenström, E., Konstiga företag.
Styrning av team och processer – Teoretiska perspektiv och fallstudier. Bengtsson, L., Lind, J., Samuelson, L.A., (red).
Sweet, S., Industrial Change Towards Environmental Sustainability – The Case of Replacing Chloroflouorocarbons.
Tamm Hallström, K., Kampen för auktoritet – standardiseringsorganisationer i arbete.

1999

Adler, N., Managing Complex Product Development.
Allgulin, M., Supervision and Monetary Incentives.
Andersson, P., Experto Credite: Three Papers on Experienced Decision Makers.
Ekman, G., Från text till batong – Om poliser, busar och svennar.
Eliasson, A-C., Smooth Transitions in Macroeconomic Relationships.
Flink, H., Gunnarsson, J., Wahlund, R., Svenska hushållens sparande och skuldsättning– ett konsumentbeteende-perspektiv.
Gunnarsson, J., Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
Hamrefors, S., Spontaneous Environmental Scanning.
Helgesson, C-F., Making a Natural Monopoly: The Configuration of a Techno-Economic Order in Swedish Telecommunications.
Japanese Production Management in Sunrise or Sunset. Karlsson, C., (red).
Jönsson, B., Jönsson, L., Kobelt, G., Modelling Disease Progression and the Effect of Treatment in Secondary Progressive MS. Research Report.
Lindé, J., Essays on the Effects of Fiscal and Monetary Policy.
Ljunggren, U., Indikatorer i grundskolan i Stockholms stad före stadsdelsnämndsreformen – en kartläggning.
Ljunggren, U., En utvärdering av metoder för att mäta produktivitet och effektivitet i skolan – Med tillämpning i Stockholms stads grundskolor.
Lundbergh, S., Modelling Economic High-Frequency Time Series.
Mägi, A., Store Loyalty? An Empirical Study of Grocery Shopping.
Mölleryd, B.G., Entrepreneurship in Technological Systems – the Development of Mobile Telephony in Sweden.
Nilsson, K., Ledtider för ledningsinformation.
Osynlig Företagsledning. Sjöstrand, S-E., Sandberg, J., Tyrstrup, M., (red).
Rognes, J., Telecommuting – Organisational Impact of Home Based – Telecommuting.
Sandström, M., Evaluating the Benefits and Effectiveness of Public Policy.

Skalin, J., Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.

Spagnolo, G., Essays on Managerial Incentives and Product-Market Competition.

Strauss, T., Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.

Svedberg Nilsson, K., Effektiva företag? En studie av hur privatiserade organisationer konstrueras.

Söderström, U., Monetary Policy under Uncertainty.

Werr, A., The Language of Change The Roles of Methods in the Work of Management Consultants.

Wijkström, F., Svenskt organisationsliv – Framväxten av en ideell sektor.