STATISTICAL PROPERTIES OF GARCH PROCESSES

STOCKHOLM SCHOOL OF ECONOMICS
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.

EFI conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. EFI also encourages and assists in the publication and distribution of its research findings, and is 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 Centres		Directors
A	Management and Organisation Theory	Prof. Sven-Erik Sjöstrand
B	Accounting and Managerial Finance	Prof. Lars Östman
C	Mangerial Economics	Prof. Peter Jennergren
CEE	Centre for Ethics and Economics	Adjunct Prof. Hans de Geer
CFR	Centre for Risk Research	Prof. Lennart Sjöberg
CHE	Centre for Health Economics	Prof. Bengt Jönsson
D	Marketing, Distribution and Industry Dynamics	Prof. Lars-Gunnar Mattsson
ES	Economic Statistics	Prof. Anders Westlund
F	Public Management	Prof. Nils Brunsson
FDR	The Foundation for Distribution Research	Acting Prof. Richard Wahlund
FI	Finance	Prof. Clas Bergström
I	Information Management	Prof.. Mats Lundeberg
IEG	International Economics and Geography	Prof.. Mats Lundahl
P	Economic Psychology	Prof. Lennart Sjöberg
PMO	Man and Organisation	Prof. Bengt Stymne
PSC	Policy Science	Adjunct Prof. Brita Schwarz
RV	Law	Prof. Bertil Wiman
S	Economics	Prof. Lars Bergman
T	Industrial Production	Prof. Christer Kärlsson

Chairperson: Prof. Sven-Erik Sjöstrand
Director: Ekon Lic. Rune Castenäs

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

STATISTICAL PROPERTIES OF GARCH PROCESSES

Changli He

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

A Dissertation for the
Doctor's Degree in Philosophy
Stockholm School of Economics

ISBN 91-7258-460-2

Distributed by
EFI, Stockholm School of Economics
Box 6501, S-113 83 Stockholm
Sweden

Contents

Preface

This dissertation contains five chapters. An introduction and a summary of my research are given in Chapter 1. The other chapters present theoretical results on the moment structure of GARCH processes. Some chapters also contain empirical examples in order to illustrate applications of the theory. The focus, however, is on statistical theory. Responsibility for any shortcomings or errors in this dissertation remains solely mine.

While working with GARCH I have greatly benefited from my supervisor, Professor Timo Teräsvirta, his initiation, support and supervision. I wish to express my deep gratitude to Professor Teräsvirta for his consistent support and encouragement as well as invaluable guidance for my research. All of what I have received from him has been and will continue to be of great importance to my work.

Also, I wish to thank all the members of the Department of Economic Statistics for creating a pleasant working environment. In particular, I would like to thank Professor Anders Westlund for his help and encouragement. I am indebted to Joakin Skalin for his help with computer problems, and to Monica Peijne and Carina Morton-Fincham for their help with all administrative matters.

Finally, I want to thank the Bank of Sweden Tercentenary Foundation and the Swedish Council for Research in the Humanities and Social Sciences for financial support.

Changli He
Stockholm, October 1997

Abstract

Properties of Moments of a Family of GARCH Processes

Chapter 2 considers the moments of a family of first-order GARCH processes. First, a general condition of the existence of any integer moment of the absolute values of the observations is given. Second, a general expression for this moment as a function of lower-order moments is derived. Third, the kurtosis and the autocorrelation function of the squared and absolute-valued observations are derived. The results apply to a host of different GARCH parameterizations. Finally, the existence, or the lack of it, of a theoretical counterpart to the so-called Taylor effect for some members of this GARCH family is discussed. Possibilities of extending some of the results to higher-order GARCH processed are indicated and potential applications of the statistical theory proposed.

Statistical Properties of the Asymmetric Power ARCH Process

The asymmetric power ARCH model is a recent addition to time series models that may be used for predicting volatility. Its performance is compared with that of standard models of conditional heteroskedasticity such as GARCH. This has previously been done empirically. In Chapter 3 the same issue is studied theoretically using unconditional fractional moments for the A-PARCH model that are derived for the purpose. The role of the

heteroskedasticity parameter of the A-PARCH process is highlighted and compared with corresponding empirical results involving autocorrelation functions of power-transformed absolute-valued return series.

Fourth Moment Structure of the GARCH(p, q) Process

In Chapter 4, a necessary and sufficient condition for the existence of the unconditional fourth moment of the GARCH(p, q) process is given as well as an expression for the moment itself. Furthermore, the autocorrelation function of the centred and squared observations of this process is derived. The statistical theory is further illustrated by a few special cases such as the GARCH(2,2) process and the ARCH(q) process.

Properties of the Autocorrelation Function of Squared Observations for Second Order GARCH Processes under Two Sets of Parameter Constraints

Nonnegativity constraints on the parameters of the GARCH(p, q) model may be relaxed without giving up the requirement of the conditional variance remaining nonnegative with probability one. Chapter 5 looks into the consequences of adopting these less severe constraints in the GARCH(2,2) case and its two second-order special cases, GARCH(2,1) and GARCH(1,2). This is done by comparing the autocorrelation function of squared observations under these two sets of constraints. The less severe constraints allow more flexibility in the shape of the autocorrelation function than the constraints restricting the parameters to be nonnegative. The theory is illustrated by an empirical example.

1
Introduction and Summary

In addition to this introduction, the dissertation contains four chapters. All of them present theoretical results on the moment structure of GARCH processes. Some chapters also contain empirical examples in order to illustrate applications of the theory. The focus, however, is mainly on statistical theory. An introduction and a summary of this research follows.

1.1 Introduction

A popular way of modelling volatility is to allow the conditional variance to change over time as a function of past errors and to parameterize this process. This line of research started with Engle's (1982) seminal article on the Autoregressive Conditional Heteroskedasticity (ARCH). Bollerslev (1986) and Taylor (1986) expanded Engle's model into the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. Later developments include a host of modifications and extensions to the original ARCH and GARCH model. The classical symmetric GARCH model has found wide use. It has been successfully applied to modelling time-varying volatility in financial time series, when the time series display high kurtosis and slowly decaying autocorrelations of squared observations but no or little linear autocorrelation. Furthermore, a number of asymmetric GARCH models has been introduced in order to capture the leverage effect (Black, 1976), when the conditional variance responds asymmetrically to positive and negative residuals. For recent surveys on ARCH/GARCH

models see, for example, Bera and Higgins (1993), Bollerslev, Engle and Nelson (1994), Diebold and Lopez (1995), Engle (1995), Gouriéroux (1997), Guégan (1994), Palm (1996) and Shephard (1996).

Several papers such as Bollerslev (1986, 1988), Engle (1982), Fornari and Mele (1996, 1997), Granger and Ding (1996), Kunst (1997), Milhϕj (1985), Nelson (1990, 1991) and Sentana (1995) have paid attention to statistical properties of some of those ARCH/GARCH models. Among them, importantly, Milhϕj (1985) found the moment structure of the ARCH(q) process. Bollerslev (1986, 1988) derived a necessary and sufficient condition for the existence of the fourth unconditional moment of the standard GARCH(1,1) model as well as the expressions of the moment and the autocorrelation function of the squared process under the normality assumption. In this dissertation these results are generalized to symmetric non-normal error distributions and various GARCH(1,1) parameterizations. Furthermore, they are generalized to the GARCH(p, q) case. This is done without the normality assumption.

Results on the existence of unconditional moments for GARCH models are not only of statistical interest. Practitioners may want to use them to see what kind of moment implications GARCH models they estimate may have. In particular, the existence of the unconditional fourth moment of stochastic processes generating, say, financial return data has interested researchers. Given both the existence conditions and suitable expression for the fourth moment, the investigator would be able to check what his estimated model implies about this moment. Obtaining estimates of the kurtosis and the autocorrelations of centred and squared observations from the model would also be useful. That would enable one to see how well the kurtosis and autocorrelation implied by the estimated model match the estimates obtained directly from the data. Such a comparison would give an indication of how well the model fits the data, because high kurtosis and slowly decaying autocorrelations of the centred and squared observations (or errors) are characteristic features of many financial high-frequency series. The results in this work provide the necessary tools for carrying out such investigations.

On the other hand, a unified framework for considering statistical properties of various GARCH models has been lacking in the literature. Ding, Granger and Engle (1993) designed one which they called the Asymmetric Power ARCH model. The motivation for this model lies in observed autocorrelation functions of the power of absolute returns. This model contains as special cases a number of previous GARCH models: GARCH, GJR-GARCH (Glosten, Jagannathan and Runkle, 1993), Non-linear ARCH (Higgins and Bera, 1992) and log-ARCH (Geweke, 1986, and Pantula, 1986); but not the EGARCH (Nelson, 1991) one, for example. Hentschel (1995) defined another model class, but the main purpose of that paper was not to consider statistical properties of GARCH models. Besides, models such as VS-GARCH (Fornari and Mele, 1996, 1997) and GQARCH (Sentana, 1995)

are not included in Hentschel's family of GARCH models. This dissertation provides a unifying framework to considering statistical properties of many, if not all, GARCH models. In this framework, existence conditions for moments and analytic expressions for existing moments of various GARCH processes as special cases of a single unifying model are obtained.

1.2 Summary of findings

In this section I sum up the main findings contained in this dissertation.

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 identically distributed random variables with a symmetric density about origin, and h_t is positive with probability one, $\mathcal{F}_{t-1}$-measurable function, where $\mathcal{F}_{t-1}$ is the sigma-algebra generated by $\{z_{t-1}, z_{t-2}, z_{t-3}, ...\}$. Assume that h_t^δ, $\delta > 0$, has the following dynamic definition:

$$h_t^\delta = g_\lambda(z_{t-1}) + c_\theta(z_{t-1}) h_{t-1}^\delta, \ \delta, \ \theta, \ \lambda > 0 \tag{1.2}$$

where $g_{\lambda t} = g_\lambda(z_t)$ and $c_{\theta t} = c_\theta(z_t)$ are well-defined functions of z_t. As for $\{g_{\lambda t}\}$ and $\{c_{\theta t}\}$, they are either constants or sequences of independent identically distributed random variables such that they are stochastically independent of h_t^δ. Both $g_{\lambda t}$ and $c_{\theta t}$ contain parameters that affect the moment structure of ε_t. Constrains on the parameters in $g_{\lambda t}$ and $c_{\theta t}$ are necessary to guarantee that $Pr\{h_t^\delta > 0\} = 1$, $\delta > 0$, for all t. I call (1.1) with (1.2) the general GARCH(1,1) model.

Appropriate choices of δ, θ, λ, $g_\lambda(z_t)$ and $c_\theta(z_t)$ yield the following GARCH (1,1) models:

- $g_\lambda(z_t) \equiv \alpha_0$(positive constant):
  - The GARCH(1,1) (Bollerslev, 1986) model for $\delta = 2$ and $c_{2,t-1} = \beta + \alpha_1 z_{t-1}^2$.
  - The Absolute Value GARCH (AVGARCH(1,1)) model (Taylor, 1986, and Schwert, 1989) for $\delta = 1$ and $c_{1,t-1} = \beta + \alpha_1 |z_{t-1}|$.
  - The Asymmetric Power ARCH(A-PARCH(1,1,δ); Ding, Granger and Engle, 1993) model for $\delta = \theta$ and $c_{\delta,t-1} = \beta + \alpha_1\{(1 - \eta \text{sgn}(z_{t-1}))\,|z_{t-1}|\}^\delta$ where $I(z_{t-1}) = 1$ if $z_{t-1} < 0$, and $I(z_{t-1}) = 0$ otherwise.
  - The Parametric Family of GARCH(1,1) (PF-GARCH(1,1, δ, θ); Hentschel, 1995) model for $c_{\theta,t-1} = \beta + \alpha_1 \delta\{|z_{t-1} - b| - c(z_{t-1} - b)\}^\theta$.

- The Generalized Threshold GARCH (1,1) (Bollerslev, Engle and Nelson, 1994) model for $\delta = \theta$, a positive integer, and $c_{\delta,t-1} = \beta + \{\alpha^+(1 - I(z_{t-1})) + \alpha^- I(z_{t-1})\} |z_{t-1}|^\delta$. When $\delta = 1$, the model is the Threshold GARCH (TARCH (1,1)) of Zakoïan (1994), and when $\delta = 2$ one has the GJR -GARCH (1,1) (Glosten, Jagannathan and Runkle, 1993) model.
- The Exponential GARCH(1,1) (Nelson, 1991) model is the limiting case of the PF-GARCH(1,1, δ, 1) model as $\delta \to 0$, $\theta = 1$ and $b = 0$ in the PF-GARCH(1,1, δ, 1) model.
- The log-GARCH (1,1) (Geweke, 1986 and Pantula, 1986) model is the limiting case of the A-PARCH(1,1, δ) model as $\delta \to 0$.

• $g_\lambda(z_t) \neq \alpha_0$:

- The Volatility Switching GARCH(1,1, δ) (VS-GARCH(1,1, δ)) model for $\delta = \theta = \lambda = 1, 2$ and $g_{\delta,t-1} = \alpha_0 + \varphi(|z_{t-1}|^\delta - \xi)\text{sgn}(z_{t-1})$, $c_{\delta,t-1} = \beta + \alpha_1 |z_{t-1}|^\delta + \zeta |z_{t-1}|^\delta \text{ sgn}(z_{t-1})$. When $\delta = 2$, the corresponding model is the VS-GARCH(1,1, 2) model of Fornari and Mele (1996, 1997).
- The Fourth Order Nonlinear Generalized Moving Average Conditional Heteroskedasticity (4NLGMACH(1,1)) model for $\delta = 2$, $\lambda = 4$, $\theta = 1$, and $g_{4,t-1} = \alpha_0 + \alpha_1(z_{t-1} - c)^2 + \alpha_2(z_{t-1} - c)^4$, $c_{1,t} = \beta$. This is a further generalization of the family of Moving Average Conditional Heteroskedasticity (MACH) models that Yang and Bewley (1995) introduced. It has not been applied in practice but is defined in this dissertation for future considerations.
- The Generalized Quadratic ARCH (GQARCH(1,1); Sentana, 1995) model for $\delta = 2$ and $g_{1,t-1} = \alpha_0 + \zeta z_{t-1} h_{t-1}$, $c_{2,t-1} = \beta + \alpha_1 z_{t-1}^2$. Note that $g_{1,t-1}$ in this case depends on z_{t-1} and h_{t-1}, and $\{g_{1,t}\}$ is thus not sequence of iid random variables. Yet it is a sequence of uncorrelated random variables, and this property allows us to place the GQARCH model into this category and consider its statistical properties.

A purpose of this dissertation is to study the moment structure of a large number of special cases of the general GARCH(1,1) model (1.1) and (1.2), and the standard GARCH(p, q) process. Define the unconditional process $\{\varepsilon_t, h_t^\delta\}_{t=-\infty,\infty}$ with (1.1) and the ψ-order unconditional moment $\mu_\psi = \mathsf{E}|\varepsilon_t|^\psi$, $\psi > 0$ as $t \to \infty$. Without further restrictions, there is no guarantee that $\mu_\psi < \infty$. If $\mu_{2\delta}$ exists, define the n-th order of autocorrelation function of $\{|\varepsilon_t|^\delta\}$ as $\rho_n(\delta) = \rho(|\varepsilon_t|^\delta, |\varepsilon_{t-n}|^\delta)$. Then, the following questions arise: For $\psi > 0$, under which conditions does $\mu_\psi < \infty$ hold

as time goes to infinity? If μ_ψ exists, what is the analytic expression for it? If $\mu_\psi < \infty$, what is the analytic expression of $\rho_n(\psi/2)$? In this dissertation, answers to these questions are worked out without the normality assumption.

1.2.1 The general GARCH(1,1) process

The moment structure of a family of first-order GARCH processes are considered in Chapter 2 assuming that $\delta = 1$ or 2 in (1.2). First, a general condition for the existence of any integer moment of the absolute values of the observations is given. Second, a general expression of this absolute moment as a function of lower-order moments is derived. Third, the kurtosis and the autocorrelation function of the squared and absolute-valued observations are derived. It is found that letting $g_\lambda(z_{t-1})$ be stochastic as in Fornari and Mele (1996, 1997), or Sentana (1995), the kurtosis of the corresponding model increases *ceteris paribus* compared to the case where $g_\lambda(z_{t-1}) \equiv \alpha_0$. The results apply to a host of different GARCH parameterizations. Finally, the existence, or the lack of it, of the theoretical counterpart to the so-called Taylor effect in some members of this GARCH family is discussed. Possibilities of extending the results to higher-order GARCH processed are indicated and potential applications of the statistical theory proposed.

1.2.2 The A-PARCH(1,1,δ) process

The statistical properties for the A-PARCH(1,1, δ) model are considered in Chapter 3. This means that $\delta = \theta$ and $g_\lambda(z_{t-1}) \equiv \alpha_0$ in (1.2). The temporal properties of returns, r_t, from speculative assets have recently received considerable attention; see, Granger and Morgenstern (1970), Ding, Granger and Engle (1993), henceforth DGE, Granger and Ding (1995a, b; 1996) and Ding and Granger (1996). DGE found $|r_t|^d$, $d > 0$, is "long-memory" in that the autocorrelations remain high at very long lags and is strongest for $d = 1$ and near it. Consequently, DGE introduced the A-PARCH model with a separate "heteroskedasticity parameter" δ. In this chapter the focus is on the autocorrelation function of $\{|\varepsilon_t|^\delta\}$. The fractional moments and the autocorrelation function, expressed as a function of δ, of the A-PARCH process are derived. A transformation of observations implicit in the A-PARCH specification is introduced in order to discuss the relationship between d of $|r_t|^d$ and the corresponding model parameter δ, where r_t is a return of an asset. It is shown how the heteroskedasticity parameter δ and an asymmetry parameter also present in the A-PARCH model add to the flexibility of the specification compared to the standard GARCH model. Possible advantages of the A-PARCH specification are further demonstrated by reconsidering the empirical example in DGE.

1.2.3 The GARCH(p, q) process

The results of Bollerslev (1986, 1988) make model evaluations possible for the GARCH(1,1) model with normal errors. Chapter 4 generalizes some of those results to GARCH(p, q) process. Assuming a symmetric error distribution with zero mean, a necessary and sufficient condition for the existence of the unconditional fourth moment of the GARCH(p, q) process is given as well as expression for the moment itself. Furthermore, the autocorrelation function of the centred and squared observations of this process is derived. The statistical theory is further illustrated by a few special cases such as the GARCH(2,2) process and the ARCH(q) process.

1.2.4 The GARCH(2,2) process

Nonnegativity constraints on the parameters of the GARCH(p, q) model as defined in Bollerslev (1986) may be relaxed without giving up the requirement of the conditional variance remaining non-negative with probability one. Nelson and Cao (1992) showed that and derived less stringent conditions for the non-negativity. Chapter 5 looks into the consequences of adopting these less severe constraints in the GARCH(2, 2) case and its two second-order special cases, the GARCH(2,1) and GARCH(1, 2) models. It also serves as an illustration of the theory in Chapter 4. In Chapter 5, the autocorrelations function of squared observations under these two sets of constraints are compared with each other. The less severe constraints allow more flexibility in the shape of the autocorrelation function than the constraints restricting the parameters to be nonnegative, and Chapter 5 gives a full characterization of this added flexibility. The statistical theory is further illustrated by an empirical example.

1.3 Conclusions

The discussion of the general GARCH(1,1) model in Chapter 2 is only restricted to the first-order model. Many results concerning the fourth moments may be readily extended to higher-order models by using the statistical theory in Chapter 4. On the other hand, deriving the autocorrelation functions of the squared observations for the EGARCH and log-GARCH processes seems possible by applying the results obtained in Chapters 2 and 3. As mentioned above, results in those chapters are useful in considering how well different GARCH models satisfy the stylized facts found in many financial series. The hope is that practical applications of those tools will be interesting. Furthermore, modifications of the definition of $\rho_n(\delta_1, \delta_2) = \rho(\varepsilon_t^{\delta_1}, \varepsilon_{t-n}^{\delta_2})$ for various δ_1 and δ_2, where δ_1, $\delta_2 > 0$, may also produce useful results. These and related topics will be addressed in further work.

References

Bera, A. and M. Higgins (1993). ARCH models: properties, estimation, and testing. *Journal of Economic Surveys* **7**, 305-366.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings from the American Statistical Association, Business and Economic Statistics Section,* 177-181.

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

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

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.

Diebold, F.X. and J.A. Lopez (1995). Modelling volatility dynamics. In: K.D. Hoover, ed. *Macroeconometrics. Developments, tensions, and prospects*, 427-466. Boston: Kluwer.

Ding, Z. and C.W.J. Granger (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics* **73**, 185-215.

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.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1006.

Engle, R.F., ed. (1995). *ARCH. Selected readings.* Oxford: Oxford University Press.

Fornari, F. and A. Mele (1996). Modeling the changing asymmetry of conditional variances. *Economics Letters* **50**, 197-203.

Fornari, F. and A. Mele (1997). Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics* **12**, 49-65.

Geweke, J. (1986). Modeling the persistence of conditional variances: a comment. *Econometric Review* **5**, 57-61.

Glosten, L.R., R. Jagannathan and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* **48**, 1779-1801.

Gouriéroux, C. (1997). *ARCH models and financial applications.* New York: Springer.

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

Granger, C.W.J. and Z. Ding (1995b). Stylized facts on the temporal and distributional properties of daily data from speculative markets. Department of Economics, University of California, San Diego, Working Paper.

Granger, C.W.J. and Z. Ding (1996). Varieties of long memory models. *Journal of Econometrics* **73**, 61-77.

Guégan, D. (1994). *Séries chronologiques non linéaires à temps discret.* Paris: Economica.

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

Higgins, M.L. and A.K. Bera (1992). A class of nonlinear ARCH models. *International Economic Review* **33**, 137-158.

Kunst, R.M. (1997). Fourth-order moments of augmented ARCH processes. *Communications in Statistics, Theory and Methods* **26**, 1425-1441.

Milhϕj, A. (1985). The moment structure of ARCH processes. *Scandinavian Journal of Statistics* **12**, 281-292.

Nelson, D.B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* **6**, 318-334.

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

Nelson, D.B. and C.Q. Cao (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics* **10**, 229-235.

Palm, F.C. (1996). GARCH models of volatility. In: G.S. Maddala and C.R. Rao, eds. *Handbook of Statistics*, Vol. **14**, 209-240. Amsterdam: Elsevier.

Pantula, S. (1986). Modeling the persistence of conditional variances: a comment. *Econometric Review* **5**, 71-74.

Schwert, G.W. (1989). Why does stock market volatility change over time? *Journal of Finance* **45**, 1129-1155.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies* **62**, 639-661.

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

Taylor, S. (1986). *Modelling financial time series.* New Nork: Wiley.

Yang, M. and R. Bewley (1995). Moving average conditional heteroskedastic processes. *Economics Letters* **49**, 367-372.

Zakoïan, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* **18**, 931-955.

2

Properties of Moments of a Family of GARCH Process

2.1 Introduction

Modelling time series with leptokurtic observations and volatility clusters with GARCH models has become a growth industry, and this has led to a host of modifications and extensions of the original model of Bollerslev (1986) and Taylor (1986, pp. 78-79). For recent overviews of GARCH models see, for example, Bollerslev, Engle and Nelson (1994), Diebold and Lopez (1995), Guégan (1994, ch. 5), Palm (1996), Shephard (1996) and Gouriéroux (1997). Several papers such as Bollerslev (1986, 1988), Fornari and Mele (1996, 1997), Granger and Ding (1996), Nelson (1990, 1991) and Sentana (1995) have paid attention to statistical properties of some of those GARCH models. However, a unified framework for considering statistical properties of various GARCH processes has been lacking in the literature. Hentschel (1995) offered one, but the main purpose of that paper was not to consider statistical properties of GARCH models. The only ones receiving attention were the constraints needed for positivity of the conditional variance and weak stationarity. On the other hand, obtaining existence conditions for moments and analytic expressions for existing moments of various GARCH processes as special cases of a single unifying model would be useful. This would make it possible to investigate what at least seemingly different models actually contribute in practice, i.e., how well they are able to satisfy stylized facts in the observed series. These observational facts include leptokurtosis and nonlinear dependence, of which the latter shows as positive autocorrelations of squared and absolute-valued observations.

This chapter provides a unifying framework for considering statistical properties of many, if not all, GARCH models. The discussion is restricted to the GARCH(1,1) model mainly partly because this is the most commonly applied parameterization in practice. Many results concerning the fourth moments may be readily extended to higher-order models by using the statistical theory in Chapter 4. The results of Chapter 2 are useful in considering how well different GARCH models satisfy the above-mentioned stylized facts found in many financial series. We shall apply these tools for that purpose in future work, although a beginning is made here.

The structure of the present chapter is as follows. The general family of GARCH(1,1) models is defined in Section 2.2. The existence of moments of the sequence of random variables obeying a GARCH(1,1) model that belongs to this family and the expressions for the existing moments are discussed in Section 2.3. The fourth moment and the kurtosis receive special attention. Sections 2.4 to 2.6 deal with the autocorrelation function of the squared and absolute-valued sequences of observations from a GARCH(1,1) process. These results make it possible to discuss the relationship between members of our GARCH family and the so-called Taylor effect (Granger and Ding, 1995a, b), which is done in Section 2.7. Section 2.8 concludes.

2.2 The general GARCH(1,1) model

Consider the following general class of GARCH(1,1) models

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

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

$$h_t^k = g(z_{t-1}) + c(z_{t-1}) h_{t-1}^k, \ k = 1 \text{ or } 2 \tag{2.2}$$

where $g_t = g(z_t)$ and $c_t = c(z_t)$ are well-defined functions of z_t. Finally,

$$Pr\{h_t^k > 0\} = 1 \tag{2.3}$$

for all t. As for $\{g_t\}$ and $\{c_t\}$, these functions are either constants or sequences of independent identically distributed random variables such that they are stochastically independent of h_t^k. Both g_t and c_t contain parameters that affect the moment structure of ε_t. Constraints on the parameters in g_t and c_t are necessary to guarantee that (2.3) holds.

Note that conditions for $\{g_t\}$ can be relaxed in such a way that $\{g_t\}$ is merely a sequence of uncorrelated identically distributed random variables. This case will be discussed later.

Appropriate choices of k, g_t and c_t yield, among others, the following GARCH(1,1) models:

- The GARCH(1,1) (Bollerslev, 1986; Taylor, 1986, pp. 78-79) model for $k = 2$ and $g_{t-1} \equiv \alpha_0$, $c_{t-1} = \beta + \alpha_1 z_{t-1}^2$.

- The Absolute Value GARCH (AVGARCH(1,1)) model (Taylor, 1986, pp. 78-79; Schwert, 1989) for $k = 1$ and $g_{t-1} \equiv \alpha_0$, $c_{t-1} = \beta + \alpha_1 |z_{t-1}|$.

- The GJR-GARCH(1,1) (Glosten, Jagannathan and Runkle, 1993) model for $k = 2$ and $g_{t-1} \equiv \alpha_0$, $c_{t-1} = \beta + (\alpha_1 + \omega I(z_{t-1}))z_{t-1}^2$ where $I(z_{t-1}) = 1$ if $z_{t-1} < 0$, and $I(z_{t-1}) = 0$ otherwise.

- The nonlinear GARCH(1,1) (NLGARCH(1,1,k); Engle, 1990) model. If $k = 2$ then $g_{t-1} \equiv \alpha_0$, $c_{t-1} = \beta + \alpha_1(1 - 2\eta \mathrm{sgn}(z_{t-1}) + \eta^2)z_{t-1}^2$. If $k = 1$ then $g_{t-1} = \alpha_0$, $c_{t-1} = \beta + \alpha_1(1 - \eta \mathrm{sgn}(z_{t-1}))\,|z_{t-1}|$. These two models are special cases of the Asymmetric Power ARCH (A-PARCH) model of Ding, Granger and Engle (1993) obtained by allowing $k > 0$ in (2.2); for a discussion of its statistical properties see also Chapter 3.

- The Volatility Switching GARCH(1,1) (VS–GARCH(1,1); Fornari and Mele, 1996, 1997) model for $k = 2$ and $g_{t-1} = \alpha_0 + \delta(z_{t-1}^2 - \xi)\mathrm{sgn}(z_{t-1})$, $c_{t-1} = \beta + \alpha_1 z_{t-1}^2 + \zeta z_{t-1}^2 \mathrm{sgn}(z_{t-1})$.

- The Threshold GARCH (TGARCH(1,1), Zakoïan, 1994) model for $k = 1$ and $g_{t-1} \equiv \alpha_0$, $c_{t-1} = \beta + (\alpha^+(1 - I(z_{t-1})) + \alpha^- I(z_{t-1}))\,|z_{t-1}|$.

- The Fourth Order Nonlinear Generalized Moving Average Conditional Heteroskedasticity (4NLGMACH(1,1)) model for $k = 2$ and $g_{t-1} = \alpha_0 + \alpha_1(z_{t-1} - c)^2 + \alpha_2(z_{t-1} - c)^4$, $c_{t-1} = \beta$. This is a further generalization of the family of Moving Average Conditional Heteroskedasticity (MACH) models Yang and Bewley (1995) introduced.

- The Generalized Quadratic ARCH (GQARCH(1,1); Sentana, 1995) model for $k = 2$ and $g_{t-1} = \alpha_0 + \zeta z_{t-1} h_{t-1}$, $c_{t-1} = \beta + \alpha_1 z_{t-1}^2$. Note that g_{t-1} in this case depends both on z_{t-1} and h_{t-1}, and $\{g_t\}$ is thus not sequence of iid random variables. Yet it is a sequence of uncorrelated random variables, and this property allows us to obtain results that are similar to those obtained in the pure iid case. It is worth mentioning however, that (2.3) does not necessarily hold for the GQARCH(1,1) model without further parameter restrictions.

The most conspicuous exclusions from the family of GARCH(1,1) models defined by (2.1)-(2.3) are the EGARCH model (Nelson, 1991) and the A-PARCH model. As just mentioned, defining $k > 0$ would allow the latter model to be included. Nonlinear GARCH models such as the smooth transition GARCH (STGARCH; González-Rivera, 1996; Hagerud, 1997) do not fit into this framework either. On the other hand, Hentschel (1995) defined a parametric family of GARCH models that also nests the EGARCH

and A-PARCH models but not, say, VS-GARCH, GQARCH, GMACH or STGARCH models. Following Hentschel (1995), let

$$f(z_t) = |z_t - b| - c(z_t - b).$$

By defining $c_t = \alpha_1 \lambda [f(z_t)]^\delta + \beta$ and $g_t \equiv \lambda \alpha_0 - \beta + 1$ it is seen that a subset of Hentschel's model family defined by $\lambda = \delta = k$ $(= 1, 2)$ is nested in (2.1)-(2.3). As we are interested in analytic expressions of integer moments of the original GARCH process we do not consider the general choice, $k > 0$ in (2.2), here.

2.3 General moment condition and the fourth moment

In this section we consider moment properties of the GARCH(1,1) model (2.1)-(2.3) and begin by introducing notation. Let $\nu_i = \mathsf{E}\,|z_t|^i$, $\gamma_{ci} = \mathsf{E}c_t^i$, $\gamma_{gi} = \mathsf{E}g_t^i$ and $\widetilde{\gamma}_{gi,cj} = \mathsf{E}g_t^i c_t^j$, where i and j are positive integers. In particular, let $\gamma_c = \gamma_{c1}$, $\gamma_g = \gamma_{g1}$ and $\widetilde{\gamma}_{gc} = \widetilde{\gamma}_{g1,c1}$. We require $\gamma_g > 0$, which is not a restrictive assumption. We can now state

Theorem 2.1 *For the general GARCH(1,1) model (2.1)-(2.3), the km-th unconditional moment exists if and only if*

$$\gamma_{cm} = \mathsf{E}c_t^m < 1. \tag{2.4}$$

Under this condition the km-th moment of ε_t can be expressed recursively as

$$\mu_{km} = \mathsf{E}\,|\varepsilon_t|^{km} = \{\nu_{km}/(1-\gamma_{cm})\}\{\sum_{j=1}^{m} \tbinom{m}{j} \widetilde{\gamma}_{gj,c(m-j)} (\mu_{k(m-j)}/\nu_{k(m-j)})\}. \tag{2.5}$$

Proof. See Appendix A.

Condition (2.4) is a necessary and sufficient condition for the existence of the km-th unconditional moment of $\{\varepsilon_t\}$. It appeared in Bollerslev (1986) and Nelson (1990) for the standard GARCH(1,1) model. Note that the existence of μ_{km} does not depend on g_t whereas the value of μ_{km} does.

Since we are interested in the kurtosis of ε_t and the autocorrelation function of $\{\varepsilon_t^2\}$ we first state

Corollary 2.1.1 *For the general GARCH(1,1) model (2.1)-(2.3) with $k = 2$ and $\gamma_{c2} < 1$, the fourth unconditional moment of ε_t is given by*

$$\mu_4 = \mathsf{E}\varepsilon_t^4 = \frac{\nu_4\{\gamma_{g2}(1-\gamma_c) + 2\widetilde{\gamma}_{gc}\gamma_g\}}{(1-\gamma_c)(1-\gamma_{c2})} \tag{2.6}$$

and the kurtosis by

$$\kappa_4 = \frac{\kappa_4(z_t)\{\gamma_{g2}(1-\gamma_c) + 2\widetilde{\gamma}_{gc}\gamma_g\}(1-\gamma_c)}{\gamma_g^2(1-\gamma_{c2})} \tag{2.7}$$

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

In particular, for the case $g_t \equiv \alpha_0$ in (2.2), expression (2.7) simplifies to

$$\kappa_4^0 = \kappa_4(z_t)(1-\gamma_c^2)/(1-\gamma_{c2}). \tag{2.8}$$

This corollary gives the fourth moment and the kurtosis for $\{\varepsilon_t\}$ when the conditional variance is a function of past squared values of the process. Note that if $k = 1$, the squares are replaced by absolute values. Since it is our intention to also consider fourth moments of GARCH processes with $k = 1$ we apply Theorem 2.1 again for obtaining those. We have

Corollary 2.1.2 *For the general GARCH(1,1) model (2.1)-(2.3) with* $k = 1$ *and* $\gamma_{c4} < 1$ *the fourth unconditional moment of* ε_t *is given by*

$$\mu_4^* = \nu_4\Delta_4 \prod_{i=1}^{4}(1-\gamma_{ci})^{-1} \tag{2.9}$$

and the kurtosis

$$\kappa_4^* = \frac{\kappa_4(z)\Delta_4(1-\gamma_c)(1-\gamma_{c2})}{\Delta_2^2(1-\gamma_{c3})(1-\gamma_{c4})} \tag{2.10}$$

where

$$\begin{aligned}
\Delta_2 &= \gamma_{g2}(1-\gamma_c) + 2\widetilde{\gamma}_{gc}\gamma_g && (2.11)\\
\Delta_3 &= (1-\gamma_{c2})[\gamma_{g3}(1-\gamma_c) + 3\widetilde{\gamma}_{g2,c1}\gamma_g] + 3\widetilde{\gamma}_{g1,c2}\Delta_2 && (2.12)\\
\Delta_4 &= (1-\gamma_{c2})(1-\gamma_{c3})[\gamma_{g4}(1-\gamma_c) + 4\widetilde{\gamma}_{g3,c1}\gamma_g] \\
&\quad +6\widetilde{\gamma}_{g2,c2}(1-\gamma_{c3})\Delta_2 + 4\widetilde{\gamma}_{g1,c3}\Delta_3. && (2.13)
\end{aligned}$$

For the large majority of existing absolute value GARCH models $g_t \equiv \alpha_0$ in (2.2). In that case (2.9) and (2.10) reduce to

$$\mu_4^* = \nu_4\alpha_0^4\Delta_4^0 \prod_{i=1}^{4}(1-\gamma_{ci})^{-1} \tag{2.14}$$

and

$$\kappa_4^+ = \frac{\kappa_4(z)\Delta_4^0(1-\gamma_c)(1-\gamma_{c2})}{(1+\gamma_c)^2(1-\gamma_{c3})(1-\gamma_{c4})} \tag{2.15}$$

respectively, where

$$\Delta_4^0 = 1 + 3\gamma_c + 5\gamma_{c2} + 3\gamma_{c3} + 3\gamma_c\gamma_{c2} + 5\gamma_c\gamma_{c3} + 3\gamma_{c2}\gamma_{c3} + \gamma_c\gamma_{c2}\gamma_{c3}. \tag{2.16}$$

Next we consider the case in which $k = 2$ and g_t is also a function of h_t. More precisely, let $g_t = \alpha_0 + f(z_t)h_t$. Defining $f(z_t) = \zeta z_t$ yields the GQARCH model. We have

Corollary 2.1.3 *Assume that* $g_{t-1} = \alpha_0 + f(z_{t-1})h_{t-1}$ *where* $\alpha_0 > 0$, $\gamma_f = \mathsf{E}f(z_{t-1}) = 0$, $\gamma_{f2} = \mathsf{E}f^2(z_{t-1}) > 0$ *and* $\mathsf{E}f(z_{t-1})c_{t-1} = 0$ *in the general GARCH(1,1) model (2.1) with (2.2). The fourth unconditional moment of* ε_t *exists if and only if*

$$\gamma_{c2} = \mathsf{E}c_t^2 < 1. \tag{2.17}$$

Under this condition the fourth moment is given by

$$\mu_4 = \mathsf{E}\varepsilon_t^4 = \frac{\nu_4\{\alpha_0^2(1+\gamma_c) + \alpha_0\gamma_{f2}\}}{(1-\gamma_{c2})(1-\gamma_c)} \tag{2.18}$$

and the kurtosis by

$$\kappa_4 = \frac{\kappa_4(z)\{(1+\gamma_c) + (\gamma_{f2}/\alpha_0)\}(1-\gamma_c)}{(1-\gamma_{c2})}. \tag{2.19}$$

Applying (2.7), (2.8), (2.10), (2.15) and (2.19) to the models discussed in Section 2.2 we obtain the kurtosis of $\{\varepsilon_t\}$ as follows:

(a) $k = 2$:

1. GARCH(1,1) model: $\gamma_c = \beta + \alpha_1\nu_2$ and $\gamma_{c2} = \beta^2 + 2\beta\alpha_1\nu_2 + \alpha_1^2\nu_4$ in (2.8). For $\nu_2 = 1$ and $\nu_4 = 3$ (normality) this result appeared in Bollerslev (1986).

2. GJR-GARCH(1,1) model: $\gamma_c = \beta + (\alpha_1 + \omega/2)\nu_2$ and $\gamma_{c2} = \beta^2 + 2\beta(\alpha_1 + \omega/2)\nu_2 + (\alpha_1 + \omega/2)^2\nu_4$ in (2.8).

3. NLGARCH(1,1,2) model: $\gamma_c = \beta + \alpha_1(1+\eta^2)\nu_2$ and $\gamma_{c2} = \beta^2 + 2\beta\alpha_1(1 + \eta^2)\nu_2 + \alpha_1^2(1 + 6\eta^2 + \eta^4)\nu_4$ in (2.8).

4. VS-GARCH(1,1) model: $\gamma_g \equiv \alpha_0$, $\gamma_{g2} = \alpha_0^2 + \delta^2(\nu_4 - 2\nu_2\xi + \xi^2)$, $\gamma_c = \beta + \alpha_1\nu_2$, $\gamma_{c2} = \beta^2 + 2\beta\alpha_1\nu_2 + (\zeta^2 + \alpha_1^2)\nu_4$ and $\widetilde{\gamma}_{gc} = \alpha_0(\beta + \alpha_1\nu_2) + \delta\zeta(\nu_4 - \xi\nu_2)$ in (2.7). Setting $\nu_2 = 1$ and $\nu_4 = 3$ gives the kurtosis for normal errors; this is the correct version of the corresponding expression in Fornari and Mele (1996, 1997).

5. 4NLGMACH(1,1) model:

$$\kappa_4 = \kappa_4(z)\{(\gamma_{g2}/\gamma_g^2)(1-\beta) + 2\beta\}/(1+\beta)$$

where

$$\gamma_g = \alpha_0 + c^2 + c^4 + (\alpha_1 + 6c^2\alpha_2)\nu_2 + \alpha_2\nu_4$$

and

$$\begin{aligned}\gamma_{g2} &= \alpha_0\{\alpha_0 + 2c^2(\alpha_1 + c^2\alpha_2)\} + (\alpha_1 c^2 + \alpha_2 c^4)^2 \\ &\quad +(2\alpha_0\alpha_1 + 12\alpha_0\alpha_2 c^2 + 6\alpha_1^2 c^2 + 30\alpha_1\alpha_2 c^4 \\ &\quad +28\alpha_2^2 c^6)\nu_2 + (\alpha_1^2 + 2\alpha_0\alpha_2 + 30\alpha_1\alpha_2 c^2 \\ &\quad +70\alpha_2^2 c^4)\nu_4 + (2\alpha_1\alpha_2 + 28\alpha_2^2 c^2)\nu_6 + \alpha_2^2\nu_8.\end{aligned}$$

For normal errors, all moments exist for this model if $0 \leq \beta < 1$ as $\gamma_{cj} = \beta^j$.

6. GQARCH(1,1) model: $\gamma_{f2} = \zeta^2\nu_2$, $\gamma_c = \beta + \alpha_1\nu_2$ and $\gamma_{c2} = \beta^2 + 2\alpha_1\nu_2 + \alpha_1^2\nu_4$ in (2.19). Sentana (1995) derived a fourth moment expression similar to (2.18) for the GQARCH(1,1) model.

(b) $k = 1$:

1. AVGARCH(1,1) model:

$$\begin{aligned}\gamma_c &= \beta + \alpha_1\nu_1 \\ \gamma_{c2} &= \beta^2 + 2\beta\alpha_1\nu_1 + \alpha_1^2\nu_2 \\ \gamma_{c3} &= \beta^3 + 3\beta^2\alpha_1\nu_1 + 3\beta\alpha_1^2\nu_2 + \alpha_1^3\nu_3\end{aligned}$$

and

$$\gamma_{c4} = \beta^4 + 4\beta^3\alpha_1\nu_1 + 6\beta^2\alpha_1^2\nu_2 + 4\beta\alpha_1^3\nu_3 + \alpha_1^4\nu_4 \tag{2.20}$$

in (2.15).

2. NLGARCH(1,1,1) model:

$$\begin{aligned}\gamma_c &= \beta + \alpha_1(1+\eta^2)\nu_2 \\ \gamma_{c2} &= \beta^2 + 2\beta\alpha_1(1+\eta^2)\nu_2 + \alpha_1^2(1+6\eta^2+\eta^4)\nu_4 \\ \gamma_{c3} &= \beta^3 + 3\beta^2\alpha_1\nu_1 + 3\beta\alpha_1^2(1+\eta^2)\nu_2 + \alpha_1^3(1+3\eta^2)\nu_3\end{aligned}$$

and

$$\begin{aligned}\gamma_{c4} &= \beta^4 + 4\beta^3\alpha_1\nu_1 + 6\beta^2\alpha_1^2(1+\eta^2)\nu_2 \\ &\quad +4\beta\alpha_1^3(1+3\eta^2)\nu_3 + \alpha_1^4(1+6\eta^2+\eta^4)\nu_4\end{aligned}$$

in (2.15).

3. TGARCH (1,1) model:

$$\begin{aligned}\gamma_c &= \beta + (1/2)\nu_1(\alpha^+ + \alpha^-) \\ \gamma_{c2} &= \beta^2 + \nu_1\beta(\alpha^+ + \alpha^-) + (1/2)\nu_2((\alpha^+)^2 + (\alpha^-)^2) \\ \gamma_{c3} &= \beta^3 + (3/2)\nu_1\beta^2(\alpha^+ + \alpha^-) + (3/2)\nu_2\beta((\alpha^+)^2 \\ &\quad +(\alpha^-)^2) + (1/2)\nu_3((\alpha^+)^3 + (\alpha^-)^3)\end{aligned}$$

and

$$\begin{aligned}\gamma_{c4} &= \beta^4 + 2\nu_1\beta^3(\alpha^+ + \alpha^-) + 3\nu_2\beta^2((\alpha^+)^2 + (\alpha^-)^2) \\ &\quad +2\nu_3\beta((\alpha^+)^3 + (\alpha^-)^3) + (1/2)\nu_4((\alpha^+)^4 + (\alpha^-)^4)\end{aligned}$$

in (2.15).

In order to see that how g_t affects κ_4 of $\{\varepsilon_t\}$ in (2.7), rewrite (2.7) as

$$\kappa_4 = \kappa_4(z_t)\{\kappa_g(1-\gamma_c) + 2(\widetilde{\gamma}_{gc}/\gamma_g)\}(1-\gamma_c)/(1-\gamma_{c2}) \tag{2.21}$$

where $\kappa_g = \gamma_{g2}/\gamma_g^2$. The case $g_t \equiv \alpha_0$ is given in (2.8): g_t does not affect the kurtosis at all. Next suppose g_t is not a constant. It follows that $\kappa_g > 1$. Assume that $g(z_t)c(z_t)$ is a convex function of z_t. This is not a restrictive assumption in that all models discussed in Section 2.2 satisfy it when their parameters are positive. Thus, by Jensen's inequality $\widetilde{\gamma}_{gc} \geq \gamma_g\gamma_c$. It follows from (2.7) and $\gamma_c < 1$ that $\kappa_g(1-\gamma_c) + 2(\widetilde{\gamma}_{gc}/\gamma_g) > 1 - \gamma_c + 2\gamma_c = 1 + \gamma_c$. From (2.8) and (2.21) we see that

$$\kappa_4 \geq \kappa_4^0. \tag{2.22}$$

We have

Property 2.1.1 *For the general GARCH(1,1) model (2.1)- (2.3), assume that condition $\gamma_{c2} < 1$ holds, κ_g exists and $g(z)c(z)$ is a convex function of z. Then (2.22) holds.*

This property means that assuming g_t stochastic as, say, in the VS-GARCH(1,1) model increases the kurtosis of $\{\varepsilon_t\}$ compared to the case where $g_t \equiv \alpha_0$. In the VS-GARCH(1,1) model setting $\delta = 0$ implies $g_t \equiv \alpha_0$.

A similar property holds between (2.10) and (2.15) for the general GARCH (1,1) model (2.1)-(2.3) when $k = 1$. We have

Property 2.1.2 *For the general GARCH(1,1) model (2.1)- (2.3) with $k = 1$, assume that condition $\gamma_{c4} < 1$ holds, $\gamma_g > 0$, and $g^i(z)c^j(z)$ for $i = 1$, $j = 1, 2, 3$; $i = 2$, $j = 1, 2$; $i = 3$, $j = 1$, are convex functions of z. Then*

$$\kappa_4^* \geq \kappa_4^+. \tag{2.23}$$

Proof. See Appendix A.

A similar result holds for the GQARCH(1,1) model: the kurtosis of $\{\varepsilon_t\}$ following that model is always greater than that of the corresponding standard GARCH(1,1) model. This is illustrated in Figure 2.1.

2.4 Autocorrelations of squared observations, $k = 2$

In this section we consider the autocorrelation function of $\{\varepsilon_t^2\}$ for our general GARCH(1,1) model with $k = 2$. Let $\overline{\gamma}_g = \mathsf{E}(z_t^2 g_t)$ and $\overline{\gamma}_c = \mathsf{E}(z_t^2 c_t)$. The following result defines the n-th order autocorrelation of $\{\varepsilon_t^2\}$, $\rho_n = \rho_n(\varepsilon_t^2, \varepsilon_{t-n}^2)$, $n \geq 1$:

Theorem 2.2 *Assume that $\gamma_{c2} < 1$ in the GARCH(1,1) model (2.1)-(2.3) with $k = 2$. Then the autocorrelation function of $\{\varepsilon_t^2\}$ is defined as follows:*

$$\rho_1 = \frac{\nu_2\{(1-\gamma_c)(\gamma_g\overline{\gamma}_g(1-\gamma_{2c})+\overline{\gamma}_c(\gamma_{g2}(1-\gamma_c)+2\widetilde{\gamma}_{gc}\gamma_g))-\nu_2\gamma_g^2(1-\gamma_{c2})\}}{\nu_4\{\gamma_{g2}(1-\gamma_c)+2\widetilde{\gamma}_{gc}\gamma_g\}(1-\gamma_c)-\nu_2^2\gamma_g^2(1-\gamma_{c2})} \tag{2.24}$$

and $\rho_n = \rho_1\gamma_c^{n-1}$ for $n > 1$.

Proof. See Appendix A.

For $g_t \equiv \alpha_0$, the autocorrelation function of $\{\varepsilon_t^2\}$ has the simplified form

$$\rho_1^0 = \frac{\nu_2\overline{\gamma}_c(1-\gamma_c^2)-\nu_2^2\gamma_c(1-\gamma_{c2})}{\nu_4(1-\gamma_c^2)-\nu_2^2(1-\gamma_{c2})} \tag{2.25}$$

and $\rho_n^0 = \rho_1^0\gamma_c^{n-1}$ for $n > 1$.

Similarly, we can derive the autocorrelation function for the case in which $g_{t-1} = \alpha_0 + f(z_{t-1})h_{t-1}$. We have

Theorem 2.3 *Let $g_{t-1} = \alpha_0 + f(z_{t-1})h_{t-1}$ where $\alpha_0 > 0$, $\gamma_f = \mathsf{E}f(z_{t-1}) = 0$, $\gamma_{f2} = \mathsf{E}f^2(z_{t-1}) > 0$ and $\mathsf{E}f(z_{t-1})c_{t-1} = 0$ in the GARCH(1,1) model (2.1)-(2.2). If condition $\gamma_{c2} < 1$ holds, then the autocorrelation function of $\{\varepsilon_t^2\}$ is given by*

$$\rho_1 = \frac{\nu_2\overline{\gamma}_c(1-\gamma_c)\{\alpha_0(1+\gamma_c)+\gamma_{f2}\}-\alpha_0\nu_2^2\gamma_c(1-\gamma_{c2})}{\nu_4(1-\gamma_c)\{\alpha_0(1+\gamma_c)+\gamma_{f2}\}-\alpha_0\nu_2^2(1-\gamma_{c2})} \tag{2.26}$$

and $\rho_n = \rho_1\gamma_c^{n-1}$ for $n > 1$.

Proof. See Appendix A.

The autocorrelation function of $\{\varepsilon_t^2\}$ given in the above theorems is dominated by an exponential decay from the first autocorrelation. From (2.24)-(2.26) and retaining the previous definitions we obtain the first-order autocorrelation ρ_1 of $\{\varepsilon_t^2\}$ for any member of our GARCH(1,1) subfamily defined by $k = 2$:

1. GARCH(1,1) model: $\overline{\gamma}_c = \beta\nu_2 + \alpha_1\nu_4$ in (2.25). It turns out that (2.25) simplifies to

$$\rho_1 = \alpha_1\nu_2(1-\beta^2-\beta\alpha_1\nu_2)/(1-\beta^2-2\beta\alpha_1\nu_2). \tag{2.27}$$

   Setting $\nu_2 = 1$ and $\nu_4 = 3$ (normality) in (2.27) gives the result in Bollerslev (1988). Note that the existence of the autocorrelation function does depend on the existence of ν_4 although (2.27) is not a function of ν_4.

2. GJR-GARCH(1,1) model: $\overline{\gamma}_c = \beta\nu_2 + (\alpha_1 + 0.5\omega)\nu_4$ in (2.25).

3. NLGARCH(1,1,2) model : $\overline{\gamma}_c = \beta\nu_2 + \alpha_1(1+\eta^2)\nu_4$ in (2.25).

4. VS-GARCH(1,1) model: $\overline{\gamma}_g = \alpha_0\nu_2$ and $\overline{\gamma}_c = \beta\nu_2 + \alpha_1\nu_4$ in (2.24).

5. 4NLGMACH(1,1) model:

$$\rho_1 = \frac{\nu_2\{\gamma_g\overline{\gamma}_g(1-\beta^2) + \nu_2\beta(\gamma_{g2}(1-\beta)+2\beta\gamma_g^2) - \nu_2\gamma_g^2(1+\beta)\}}{\nu_4\{\gamma_{g2}(1-\beta)+2\beta\gamma_g^2\} - \nu_2^2\gamma_g^2(1+\beta)}$$

   where

$$\overline{\gamma}_g = (\alpha_0 + c^2 + c^4)\nu_2 + (\alpha_1 + 6c^2\alpha_2)\nu_4 + \alpha_2\nu_6.$$

6. GQARCH(1,1) model: $\overline{\gamma}_c = \beta\nu_2 + \alpha_1\nu_4$ in (2.26).

Note that (2.26) helps to make more precise a statement in Sentana (1995). The GQARCH and GARCH models do not have the same autocorrelation function for $\{\varepsilon_t^2\}$ as Sentana argued, but the decay of the autocorrelation function is the same in both models.

Through (2.24) and (2.25) we can investigate the dependence of ρ_1 on g_t compared to the special case $g_t \equiv \alpha_0$. From (2.25) we see that when $g_t \equiv \alpha_0$, this term does not affect the autocorrelation function of $\{\varepsilon_t^2\}$ at all. To consider the general case, first rewrite (2.24) as

$$\rho_1 = \frac{(\nu_2/\nu_4)\overline{\gamma}_c - \{1-(\overline{\gamma}_g/(\nu_2\gamma_g))(1-\gamma_c)\}\kappa_4^{-1}}{1-\kappa_4^{-1}}. \tag{2.28}$$

From (2.28) we see that ρ_1 is a function of κ_4 defined on $(\kappa_4(z_t), \infty)$ when the parameters in g_t and c_t are given. It is assumed that ρ_1 is a continuous function of κ_4 on $(\kappa_4(z_t), \infty)$ and is differentiable with respect to κ_4. Under the positive parameter constraints $\rho_1 > 0$ for any $\kappa_4 > \kappa_4(z_t)$. This tells us that $\frac{\partial}{\partial\kappa_4}\rho_1 = (\kappa_4 - 1)^{-2}[\{1-(\overline{\gamma}_g/(\nu_2\gamma_g))(1-\gamma_c)\} - (\nu_2/\nu_4)\overline{\gamma}_c] > 0$ on $(\kappa_4(z_t), \infty)$, *i.e.* ρ_1 is a monotonically increasing function of κ_4. Combining this result and (2.22) we have $\rho_1 \geq \rho_1^0$. Thus when $\kappa_4 = \kappa_4^0$, ρ_1 reaches its lower bound

$$\rho_1^0 = \frac{(\nu_2/\nu_4)\overline{\gamma}_c - \gamma_c(\kappa_4^0)^{-1}}{1-(\kappa_4^0)^{-1}}. \tag{2.29}$$

Assuming that $g(z)c(z)$ is a convex function of z was sufficient for obtaining (2.22). We have

Property 2.2.1 *For the general GARCH(1,1) model (2.1)-(2.3), assume that $g(z)c(z)$ is a convex function of z and that $\gamma_g > 0$. If $\gamma_{c2} < 1$ holds, then*

$$\rho_1 \geq \rho_1^0. \tag{2.30}$$

Similarly, it can be shown starting from (2.26) that Property 2.2.1 holds for the case $g_t = \alpha_0 + f(z_t)h_t$.

2.5 Autocorrelations of squared observations, $k = 1$

In this section we derive the autocorrelation function of $\{\varepsilon_t^2\}$ for our family of absolute value GARCH models. Set $\overline{\gamma}_{c2} = \mathsf{E}(z_t^2 c_t^2)$, $\overline{\gamma}_{g2} = \mathsf{E}(z_t^2 g_t^2)$ and $\overline{\gamma}_{gc} = \mathsf{E}(z_t^2 g_t c_t)$. This notation allows us to formulate the following result:

Theorem 2.4 *If $\gamma_{c4} < 1$ in the GARCH(1,1) model (2.1)-(2.3) with $k = 1$, then the autocorrelation function of $\{\varepsilon_t^2\}$ is defined as follows:*

$$\begin{aligned}
\rho_1^* &= \frac{\nu_2(1-\gamma_c)(1-\gamma_{c2})\{2\overline{\gamma}_{gc}(1-\gamma_{c4})\Delta_3 + \overline{\gamma}_{c2}\Delta_4\}}{\Delta} \\
&\quad - \frac{\nu_2\Delta_2(1-\gamma_{c3})(1-\gamma_{c4})\{\nu_2\Delta_2 - \overline{\gamma}_{g2}(1-\gamma_c)(1-\gamma_{c2})\}}{\Delta} \qquad (2.31) \\
\rho_n^* &= \gamma_{c2}\rho_{n-1}^* + \theta\gamma_c^{n-2}, \; n \geq 2, \qquad (2.32)
\end{aligned}$$

where

$$\Delta = \nu_4\Delta_4(1-\gamma_c)(1-\gamma_{c2}) - \nu_2^2\Delta_2^2(1-\gamma_{c3})(1-\gamma_{c4})$$

and

$$\begin{aligned}
\theta &= \Delta^{-1}\{2\nu_2\widetilde{\gamma}_{gc}(1-\gamma_{c2})(1-\gamma_{c4})[\Delta_3\overline{\gamma}_c(1-\gamma_c) \\
&\quad -\Delta_2(1-\gamma_{c3})(\nu_2\gamma_g - \overline{\gamma}_g(1-\gamma_c))]\}.
\end{aligned}$$

Proof. See Appendix A.

Assume again that $g_t \equiv \alpha_0$. Then (2.31) and (2.32) simplify and are given in

Corollary 2.4.1 *If $\gamma_{c4} < 1$ in the GARCH(1,1) model (2.1)-(2.3) with $k = 1$ and $g_t \equiv \alpha_0$, then the autocorrelation function of $\{\varepsilon_t^2\}$ can be defined as follows:*

$$\rho_1^+ = \frac{\nu_2(1-\gamma_c)(1-\gamma_{c2})\{2\overline{\gamma}_c(1-\gamma_{c4})\Delta_3^0 + \overline{\gamma}_{c2}\Delta_4^0\}}{\Delta^0}$$

$$-\frac{\nu_2^2(1+\gamma_c)(1-\gamma_{c3})(1-\gamma_{c4})\{2\gamma_c+\gamma_{c2}(1-\gamma_c)\}}{\Delta^0} \tag{2.33}$$

$$\rho_n^+ = \gamma_{c2}\rho_{n-1}^+ + \theta^0\gamma_c^{n-1},\ n \geq 2, \tag{2.34}$$

where

$$\begin{aligned}
\Delta^0 &= \nu_4\Delta_4^0(1-\gamma_c)(1-\gamma_{c2}) - \nu_2^2(1+\gamma_c)^2(1-\gamma_{c3})(1-\gamma_{c4}) \\
\theta^0 &= (1/\Delta^0)\{2\nu_2(1-\gamma_{c2})(1-\gamma_{c4})[\Delta_3^0\overline{\gamma}_c(1-\gamma_c) \\
&\quad -\nu_2\gamma_c(1+\gamma_c)(1-\gamma_{c3})]\} \\
\Delta_3^0 &= 1+2\gamma_c+2\gamma_{c2}+\gamma_c\gamma_{c2}
\end{aligned}$$

and

$$\Delta_4^0 = 1+3\gamma_c+5\gamma_{c2}+3\gamma_{c3}+3\gamma_c\gamma_{c2}+5\gamma_c\gamma_{c3}+3\gamma_{c2}\gamma_{c3}+\gamma_c\gamma_{c2}\gamma_{c3}.$$

It follows Theorem 2.4 that the autocorrelation function of $\{\varepsilon_t^2\}$ for the absolute value GARCH(1,1) model ($k=1$) is radically different from that of the GARCH(1,1) models covered by Theorem 2.2. For $k=2$ the autocorrelation function of $\{\varepsilon_t^2\}$ decays exponentially from the first autocorrelation whereas for $k=1$ it does not. It is seen from (2.31) and (2.33) that the rate of decay is slower than exponential. However, as $\gamma_{c4} \to 1$ then the second term on the right-hand side of (2.31) converges to zero. Thus for $\gamma_{c4} \approx 1$ the decay rate is "nearly " exponential with the discount factor γ_{c2}.

Applying Theorem 2.4 or Corollary 2.4.1 to the absolute value GARCH(1,1) models listed above leads to the following expressions for the first-order autocorrelation of $\{\varepsilon_t^2\}$:

1. AVGARCH(1,1) model: $\overline{\gamma}_c = \beta\nu_2 + \alpha_1\nu_3$ and $\overline{\gamma}_{c2} = \beta^2\nu_2 + 2\beta\alpha_1\nu_3 + \alpha_1^2\nu_4$ in (2.33).

2. NLGARCH(1,1,1) model: $\overline{\gamma}_c = \beta\nu_2 + \alpha_1\nu_3$ and $\overline{\gamma}_{c2} = \beta^2\nu_2 + 2\beta\alpha_1\nu_3 + \alpha_1^2(1+\eta^2)\nu_4$ in (2.33).

3. TGARCH(1,1) model: $\overline{\gamma}_c = \beta\nu_2 + (1/2)\nu_3(\alpha^+ + \alpha^-)$ and $\overline{\gamma}_{c2} = \beta^2\nu_2 + \nu_3\beta(\alpha^+ + \alpha^-) + (1/2)\nu_4((\alpha^+)^2 + (\alpha^-)^2)$ in (2.33).

Proceeding as in Section 2.4 we can establish a property similar to Property 2.2.1 for the first-order autocorrelation ρ_1^* of $\{\varepsilon_t^2\}$. Making use of (2.10) we can rewrite (2.31) as a function of κ_4^* as follows:

$$\rho_1^* = \frac{(\nu_2/\nu_4)\{2\overline{\gamma}_{gc}(1-\gamma_{c4})\Delta_3/\Delta_4+\overline{\gamma}_{c2}\} - \{(\nu_2\Delta_2-\overline{\gamma}_{g2}(1-\gamma_c)(1-\gamma_{c2}))/\nu_2\Delta_2\}(\kappa_4^*)^{-1}}{1-(\kappa_4^*)^{-1}}. \tag{2.35}$$

Under the positive parameter constraints $\rho_1^* > 0$. Assume further that ρ_1^* is a continuous function of κ_4^* on $(\kappa_4(z_t), \infty)$ and differentiable with respect to κ_4^*. These are not restrictive assumptions. It is easily seen that $\frac{\partial}{\partial \kappa_4^*}\rho_1^* > 0$ on $(\kappa_4(z_t), \infty)$ so that ρ_1^* is a monotonically increasing function of κ_4^* for any $\kappa_4^* \in (\kappa_4(z_t), \infty)$. For $g_t \equiv \alpha_0$ (2.33) is analogously reduced to

$$\rho_1^+ = \frac{(\nu_2/\nu_4)\{2\overline{\gamma}_c(1-\gamma_{c4})\Delta_3^0/\Delta_4^0 + \overline{\gamma}_{c2}\} - \{(2\gamma_c + \gamma_{c2}(1-\gamma_c))/(1+\gamma_c)\}(\kappa_4^+)^{-1}}{1-(\kappa_4^+)^{-1}}. \tag{2.36}$$

From (2.23) we have

Property 2.2.2 *For the general GARCH(1,1) model (2.1)-(2.3) with $k = 1$, assume that condition $\gamma_{c4} < 1$ holds. If (2.23) also holds then*

$$\rho_1^* \geq \rho_1^+. \tag{2.37}$$

Thus even here, defining g_t to be stochastic increases the degree of autocorrelation in $\{\varepsilon_t^2\}$ compared to the case $g_t \equiv \alpha_0$.

2.6 Autocorrelations of absolute values, $k = 1$

The autocorrelation function of the sequence of absolute values $\{|\varepsilon_t|\}$ for our subfamily of GARCH(1,1) models with $k = 1$ follows as a byproduct of the results in Section 2.4. Let $\overline{\gamma}_{1g} = \mathsf{E}(|z_t|\, g_t)$ and $\overline{\gamma}_{1c} = \mathsf{E}(|z_t|\, c_t)$. A simple substitution of $\overline{\gamma}_{1g}, \overline{\gamma}_{1c}, \nu_1$ and ν_2 for $\overline{\gamma}_g, \overline{\gamma}_c, \nu_2$ and ν_4, respectively, in (2.24), leads to

Theorem 2.5 *If $\gamma_{c2} < 1$ in the GARCH(1,1) model (2.1)-(2.3) with $k = 1$, then the autocorrelation function of $\{|\varepsilon_t|\}$ is defined as follows:*

$$\rho_1(1) = \frac{\nu_1\{(1-\gamma_c)[\gamma_g\overline{\gamma}_{1g}(1-\gamma_{2c}) + \overline{\gamma}_{1c}(\gamma_{g2}(1-\gamma_c) + 2\widetilde{\gamma}_{gc}\gamma_g)] - \nu_1\gamma_g^2(1-\gamma_{c2})\}}{\nu_2\{\gamma_{g2}(1-\gamma_c) + 2\widetilde{\gamma}_{gc}\gamma_g\}(1-\gamma_c) - \nu_1^2\gamma_g^2(1-\gamma_{c2})} \tag{2.38}$$

and $\rho_n(1) = \rho_1(1)\gamma_c^{n-1}$ for $n > 1$.

For $g_t \equiv \alpha_0$, the autocorrelation function of $\{|\varepsilon_t|\}$ has the form

$$\rho_1^0(1) = \frac{\nu_1\overline{\gamma}_{1c}(1-\gamma_c^2) - \nu_1^2\gamma_c(1-\gamma_{c2})}{\nu_2(1-\gamma_c^2) - \nu_1^2(1-\gamma_{c2})} \tag{2.39}$$

and $\rho_n^0(1) = \rho_1^0(1)\gamma_c^{n-1}$ for $n > 1$.

Applying Theorem 2.5 to our absolute value GARCH(1,1) models yields the following autocorrelation functions:

1. AVGARCH(1,1) model: ρ_1 has the form (2.39), where $\overline{\gamma}_{1c} = \beta\nu_1 + \alpha_1\nu_2$.

2. NLGARCH(1,1,1) model: ρ_1 has the form (2.39), where $\overline{\gamma}_{1c} = \beta\nu_1 + \alpha_1\nu_2$.

3. TARCH(1,1) model: ρ_1 has the form (2.39), where $\overline{\gamma}_{1c} = \beta\nu_1 + 0.5\nu_2(\alpha^+ + \alpha^-)$.

A property similar to Property 2.2.1 between $\rho_1(1)$ and $\rho_1^0(1)$ holds. We omit its derivation here.

2.7 The Taylor property

In this section we consider the following question. Define $\rho_n(\delta) = \rho(|\varepsilon_t|^\delta, |\varepsilon_{t-n}|^\delta)$. Granger and Ding (1995a) called the empirical relationship

$$\widehat{\rho}_n(1) > \widehat{\rho}_n(\delta), \text{ for any } \delta \neq 1,\ n \geq 1$$

the Taylor effect. Taylor (1986, pp. 52-55), by studying a variety of speculative return series, found that very often $\widehat{\rho}_1(1) > \widehat{\rho}_1(2) = \widehat{\rho}_1^*$. Granger and Ding (1995a, b) found the Taylor effect in a very large number of high-frequency return series. It was also studied in Ding, Granger and Engle (1993) who considered the long S&P 500 daily stock return series from 3 January 1928 till 30 April 1991 compiled by William Schwert. The results of Sections 2.5 and 2.6 enable us to see if GARCH(1,1) models can generate series with $\widehat{\rho}_n(1) > \widehat{\rho}_n^*$, or, more precisely, whether or not the theoretical relationship $\rho_n(1) > \rho_n^*$ holds for these models. We focus on the first-order autocorrelations and call this theoretical property the Taylor property. For simplicity we restrict ourselves to the AVGARCH(1,1) process with normal errors.

Our starting-point is that for any fixed β, autocorrelations $\rho_1^0(1)$ and ρ_1^+ are functions of α_1 defined on $(0, \alpha)$ where $\alpha > 0$ such $\gamma_{c4} \to 1$ as $\alpha_1 \to \alpha$. We can state the following result:

Theorem 2.6 *Consider the AVGARCH(1,1) model and assume that $z_t \sim$ nid(0,1). Then for any fixed β there exists a subset $\{\alpha_1 : \alpha_1 \in (0, \alpha^*)\}$ such that*

$$\rho_1^0(1) < \rho_1^+. \tag{2.40}$$

Assume $\beta = 0$. Then there exists another subset $\{\alpha_1 : \alpha_1 \in (\alpha^{**}, \alpha)\}$ such that

$$\rho_1^0(1) > \rho_1^+ \tag{2.41}$$

where $\alpha > 0$ such that $\gamma_{c4} = 1$. Furthermore, as $\beta = 0$, $\alpha^* = \alpha^{**}$.

Proof. See Appendix A.

When $\beta > 0$ the limiting value α is a solution of a fourth-order equation in α_1, see (2.20). An analytic solution thus generally does not exist, and we have to resort to numerical considerations. This restricts us to the following conjecture:

Conjecture 2.6.1 *For the AVGARCH(1,1) model with $z_t \sim nid(0,1)$ and any fixed $\beta > 0$ there exists a subset $\{\alpha_1 : \alpha_1 \in (\alpha^{**}, \alpha)\}$ such that*

$$\rho_1^0(1) > \rho_1^+. \tag{2.42}$$

Furthermore, $\alpha^* = \alpha^{**}$.

If (2.41) or (2.42) holds, then there exists a lag $n_0 \geq 1$ such that the Taylor property $\rho_n(1) > \rho_n^*$ is valid for $n \in \{1, ..., n_0\}$.

The situation is illustrated in Figure 2.2 for $\beta = 0$ and $\beta = 0.9$. The two first-order autocorrelations follow each other very closely as a function of α_1; see Panel (a) for $\beta = 0$ and Panel (c) for $\beta = 0.9$. Panels (b) and (d) illustrate the situation after a nonlinear monotonic transformation of α_1: the kurtosis is a monotonically increasing function of this parameter. It is seen how the Taylor property emerges at high values of the kurtosis. It thus seems that the Taylor property is present only for parameterizations that correspond to very strong leptokurtosis and strong, slowly decaying autocorrelation of squared or absolute-valued observations. But then, these are properties that are regularly observed in financial high-frequency series.

We cannot make corresponding comparisons for the standard GARCH(1, 1) process analytically because an analytic expression of the autocorrelation function of $|\varepsilon_t|$ is not available. We have, however, investigated the situation by simulation. The results appear in Tables 2.1 and 2.2. The simulated processes have a high kurtosis because for the AVGARCH model, the Taylor property is found to be present for parameterizations with that property. The results in Table 2.1 do not contain any evidence favouring the existence of the Taylor property. The means of estimated squared autocorrelations based on simulated series with 100000 observations (after deleting 50 to discard the initial effects) and 1000 replications always exceed the corresponding means of the autocorrelations of the absolute-valued series. According to the information in the last two columns of Table 2.1 it is extremely unlikely that the first five autocorrelations of the squared observations would be less than those of the absolute-valued ones. The results are restricted to the standard GARCH(1,1) process with normal errors.

These results appear to contradict the simulation results in Ding, Granger and Engle (1993) who found a standard GARCH(1,1) model capable of generating series with the Taylor effect. But then, the GARCH(1,1) model they simulated does not have a finite fourth moment so that the results are not directly comparable. Furthermore, sufficiently short series generated by the

GARCH(1,1) model may well display the Taylor effect even if the GARCH process does not have the Taylor property. This is seen from Table 2.2 which shows what happens when the experiment of Table 2.1 is repeated with 1000 replications of 1000 instead of 100000 observations each. Chances of generating a series in which the Taylor effect is present are remarkably larger than in the previous experiment as the last two columns of Table 2.2 demonstrate.

There is another detail worth mentioning in Tables 2.1 and 2.2. For the squared observations the theoretical autocorrelation function is known. By comparing its values in Table 2.1 with the estimated ones it is seen that the autocorrelation estimates are heavily downward biased even for 100000 observations. The bias is quite dramatic for sequences with 1000 observations. It is due to the fact that on the average, one has to wait for quite a while to obtain enough information about the (fat) tails of the unconditional marginal distribution to permit accurate estimation. The same is likely to be true, although to a lesser extent, for the autocorrelations of the absolute-valued observations. We do not pursue this issue any further here beyond just noting that the size of the bias and the appearance of the Taylor effect in series generated by the standard GARCH model appear to be related phenomena.

2.8 Conclusions

In this chapter we have derived a general existence condition of any integer moment of absolute-valued observation for our family of GARCH processes as well as the moments themselves. The expressions for the kurtosis and autocorrelations of the squared and absolute-valued observations that follow as special cases are of particular importance. They make it possible to see how well any estimated GARCH(1,1) model reproduces statistical facts observed in the data. Those include high kurtosis and low-starting but persistent autocorrelations of the squared and absolute-valued observations. The investigator may plug estimates of the parameters of a GARCH model into the left-hand side of the existence condition to see what the estimated model implies about the existence of, say, the fourth moment. If the fourth moment condition is satisfied, he or she can do the same for the definitions of the kurtosis and autocorrelations and compare the resulting figures with what is obtained by direct estimation of those statistics from the data. All this can be done for any member of the general GARCH family and for different error distributions. Such comparisons are left for further research.

The only stylized fact investigated to some extent already here is the empirically well-established Taylor effect. The theoretical considerations indicate that some parameterizations of the AVGARCH(1,1) model of Taylor (1986) with normal errors possess a corresponding theoretical property, so

that it is not difficult to generate series with the Taylor effect from these models. On the other hand, our simulation results suggest that for the standard GARCH(1,1) with normal errors the first autocorrelations of the squared observations are greater than those of the absolute-valued observations, which contradicts the empirical Taylor effect. Nevertheless, at least if the sample sizes are not unusually high, it is possible to observe the Taylor effect in series generated by this model as well.

All our results concern first-order GARCH processes. It seems straightforward to generalize those about fourth moments to the GARCH(p, q) case for models with $g_t \equiv \alpha_0$ and $k = 2$. The basic theory is available in Chapter 4. The results obtained for models with $k = 1$ seem much harder to generalize with the present techniques, and we have made no attempt to do that. As first-order GARCH models are the ones most commonly used in empirical work, however, we expect our results to be widely applicable in practice.

References

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

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

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.

Diebold, F.X. and J.A. Lopez (1995). Modelling volatility dynamics. In: K.D. Hoover, ed. *Macroeconometrics. Developments, tensions, and prospects*, 427-466. Boston: Kluwer.

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.

Engle, R.F. (1990). Discussion: Stock market volatility and the crash of 87. *Review of Financial Studies* **3**, 103-106.

Fornari, F. and A. Mele (1996). Modeling the changing asymmetry of conditional variances. *Economics Letters* **50**, 197-203.

Fornari, F. and A. Mele (1997). Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics* **12**, 49-65.

Glosten, L.R., R. Jagannathan and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* **48**, 1779-1801.

González-Rivera, G. (1996). Smooth transition GARCH models. Department of Economics, University of California, Riverside, Working Paper.

Gouriéroux, C. (1997). *ARCH models and financial applications.* New York: Springer.

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

Granger, C.W.J. and Z. Ding (1995b). Stylized facts on the temporal and distributional properties of daily data from speculative markets. Department of Economics, University of California, San Diego, Working Paper.

Granger, C.W.J. and Z. Ding (1996). Varieties of long memory models. *Journal of Econometrics* **73**, 61-77.

Guégan, D. (1994). *Séries chronologiques non linéaires à temps discret.* Paris: Economica.

Hagerud, G.E. (1997). *A new non-linear GARCH model.* Stockholm: EFI Economic Research Institute.

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

Nelson, D.B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* **6**, 318-334.

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. Maddala and C.R. Rao, eds. *Handbook of Statistics*, Vol. **14**, 209-240. Amsterdam: Elsevier.

Schwert, G.W. (1989). Why does stock market volatility change over time? *Journal of Finance* **45**, 1129-1155.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies* **62**, 639-661.

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

Taylor, S. (1986). *Modelling financial time series.* New Nork: Wiley.

Yang, M. and R. Bewley (1995). Moving average conditional heteroskedastic processes. *Economics Letters* **49**, 367-372.

Zakoïan, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* **18**, 931-955.

Appendix A: Proofs

Proof of Theorem 2.1. Raising (2.2) to the power m yields

$$h_t^{km} = c_{t-1}^m h_{t-1}^{km} + \sum_{j=1}^{m} \binom{m}{j} g_{t-1}^j (c_{t-1} h_{t-1}^k)^{m-j}. \tag{A.1}$$

Applying (2.2) to the first term on the right-hand side of (A.1) one obtains

$$\begin{aligned} h_t^{km} &= (c_{t-1}c_{t-2}h_{t-2}^k)^m + \sum_{j=1}^{m} \binom{m}{j} \{g_{t-1}^j (c_{t-1}h_{t-1}^k)^{m-j} \\ &\quad + c_{t-1}^m g_{t-2}^j (c_{t-2}h_{t-2}^k)^{m-j}\}. \end{aligned}$$

Further recursion gives

$$h_t^{km} = (\prod_{i=1}^{n} c_{t-i}^m) h_{t-n}^{km} + \sum_{j=1}^{m} \binom{m}{j} \sum_{i=1}^{n} (\prod_{l=1}^{i-1} c_{t-l}^m) g_{t-i}^j (c_{t-i} h_{t-i}^k)^{m-j}. \tag{A.2}$$

Assuming that the moments μ_{kp} exist up until $p = m-1$, the unconditional expectation of (A.2) is

$$\begin{aligned} \mathsf{E}h_t^{km} &= (\prod_{i=1}^{n} \mathsf{E}c_{t-i}^m) \mathsf{E}h_{t-n}^{km} + \sum_{j=1}^{m} \binom{m}{j} \sum_{i=1}^{n} (\prod_{l=1}^{i-1} \mathsf{E}c_{t-l}^m) \mathsf{E}(g_{t-i}^j c_{t-i}^{m-j}) \mathsf{E}h_{t-i}^{k(m-j)} \\ &= \gamma_{cm}^n \mathsf{E}h_t^{km} + \sum_{j=1}^{m} \binom{m}{j} \sum_{i=1}^{n} (\gamma_{cm}^{i-1} \widetilde{\gamma}_{gj,c(m-j)}) \mathsf{E}h_t^{k(m-j)}. \end{aligned} \tag{A.3}$$

Assume that the process started at some finite value infinitely many periods ago. Then letting $n \to \infty$ in (A.3) gives

$$\begin{aligned} \mu_{km} &= \mathsf{E}z_t^{km} \{\sum_{j=1}^{m} \binom{m}{j} \frac{\widetilde{\gamma}_{gj,c(m-j)}}{1-\gamma_{cm}} \mathsf{E}h_t^{k(m-j)}\} \\ &= (\nu_{km}/(1-\gamma_{cm}))\{\sum_{j=1}^{m} \binom{m}{j} \widetilde{\gamma}_{gj,c(m-j)} (\mu_{k(m-j)}/\nu_{k(m-j)})\} \end{aligned}$$

if and only if $\gamma_{cm} < 1$. ⊣

Proof of Property 2.1.2. Since $\gamma_g > 0$ we can write

$$\frac{\Delta_4}{\Delta_2^2} = \frac{\Delta_4/\gamma_g^4}{\Delta_2^2/\gamma_g^4}.$$

Applying Jensen's inequality to the denominator gives

$$\begin{aligned} (\frac{\Delta_2}{\gamma_g^2})^2 &= [(\gamma_{g2}/\gamma_g^2)(1-\gamma_c) + 2\widetilde{\gamma}_{gc}/\gamma_g]^2 \\ &\geq (1+\gamma_c)^2 \end{aligned} \tag{A.4}$$

On the other hand, applying the same inequality and Corollary 2.1.2 to the numerator yields

$$\begin{aligned}
\Delta_4/\gamma_g^4 &= (1-\gamma_{c2})(1-\gamma_{c3})[(\gamma_{g4}/\gamma_g^4)(1-\gamma_c)+4\widetilde{\gamma}_{g3,c1}/\gamma_g^3] \\
&\quad +6(\widetilde{\gamma}_{g2,c2}/\gamma_g^2)(1-\gamma_{c3})[(\gamma_{g2}/\gamma_g^2)(1-\gamma_c)+2\widetilde{\gamma}_{gc}/\gamma_g] \\
&\quad +4(\widetilde{\gamma}_{g1,c3}/\gamma_g)(1-\gamma_{c2})[(\gamma_{g3}/\gamma_g^3)(1-\gamma_c)+3\widetilde{\gamma}_{g2,c1}/\gamma_g^2] \\
&\quad +12(\widetilde{\gamma}_{g1,c3}/\gamma_g)(\widetilde{\gamma}_{g1,c2}/\gamma_g)[(\gamma_{g2}/\gamma_g^2)(1-\gamma_c)+2\widetilde{\gamma}_{gc}/\gamma_g] \\
&\geq (1-\gamma_{c2})(1-\gamma_{c3})[(1-\gamma_c)+4\gamma_c] \\
&\quad +6\gamma_{c2}(1-\gamma_{c3})[(1-\gamma_c)+2\gamma_c] \\
&\quad +4\gamma_{c3}(1-\gamma_{c2})[(1-\gamma_c)+3\gamma_c] \\
&\quad +12\gamma_{c2}\gamma_{c3}[(1-\gamma_c)+2\gamma_c] \\
&= 1+3\gamma_c+5\gamma_{c2}+3\gamma_{c3}+3\gamma_c\gamma_{c2}+5\gamma_c\gamma_{c3} \\
&\quad +3\gamma_{c2}\gamma_{c3}+\gamma_c\gamma_{c2}\gamma_{c3}=\Delta_4^0. \qquad \text{(A.5)}
\end{aligned}$$

(A.4) and (A.5) together imply that Property 2.1.2 holds. ⊣

Proof of Theorem 2.2. First consider

$$h_t^2 h_{t-1}^2 = (g_{t-1}+c_{t-1}h_{t-1}^2)h_{t-1}^2. \qquad \text{(A.6)}$$

Inserting $h_{t-1}^2 = g_{t-2}+c_{t-2}h_{t-2}^2$ to (A.6) leads to

$$h_t^2 h_{t-2}^2 = (g_{t-1}+c_{t-1}g_{t-2}+c_{t-1}c_{t-2}h_{t-2}^2)h_{t-2}^2.$$

Generally, after $n-1$ recursions,

$$h_t^2 h_{t-n}^2 = \{g_{t-1}+\sum_{j=2}^{n}(g_{t-j}\prod_{i=1}^{j-1}c_{t-i})+(\prod_{i=1}^{n}c_{t-i})h_{t-n}^2\}h_{t-n}^2. \qquad \text{(A.7)}$$

Multiplying both sides of (A.7) by $z_t^2 z_{t-n}^2$ and taking expectations yields

$$\begin{aligned}
\mathsf{E}(\varepsilon_t^2\varepsilon_{t-n}^2) &= (\nu_2^2\gamma_g\sum_{j=0}^{n-1}\gamma_c^j+\nu_2\overline{\gamma}_g\gamma_c^{n-1})\mathsf{E}h_t^2 \\
&\quad +\nu_2\overline{\gamma}_c\gamma_c^{n-1}\mathsf{E}h_t^4 \qquad \text{(A.8)}
\end{aligned}$$

where $\overline{\gamma}_g = \mathsf{E}(z_t^2 g_t)$ and $\overline{\gamma}_c = \mathsf{E}(z_t^2 c_t)$. From (A.3) we have $\mathsf{E}h_t^2 = \gamma_g/(1-\gamma_c)$ and $\mathsf{E}h_t^4 = \{\gamma_{g2}(1-\gamma_c)+2\widetilde{\gamma}_{gc}\gamma_g\}/(1-\gamma_{c2})(1-\gamma_c)$. Inserting these into (A.8) under the assumption $\gamma_{c2}<1$ and applying the definition of $\rho_n = \rho_n(\varepsilon_t^2, \varepsilon_{t-n}^2)$ together with some additional manipulation yield (2.24). ⊣

Proof of Theorem 2.3. Write

$$h_t^2 = \alpha_0 + f_{t-1}h_{t-1} + c_{t-1}h_{t-1}^2 \qquad \text{(A.9)}$$

where $f_{t-1} = f(z_{t-1})$. Recursive substitution as in the proof of Theorem 2.2 leads to

$$\begin{aligned} h_t^2 h_{t-n}^2 &= \{\alpha_0(1+\sum_{j=1}^{n-1}(\prod_{i=1}^{j} c_{t-i})) + \sum_{j=1}^{n}(f_{t-j}h_{t-j}\prod_{i=1}^{j-1} c_{t-i}) \\ &\quad +(\prod_{i=1}^{n} c_{t-i})h_{t-n}^2\}h_{t-n}^2. \end{aligned} \tag{A.10}$$

Multiplying both sides of (A.10) by $z_t^2 z_{t-n}^2$ and taking expectations yields

$$\mathsf{E}(\varepsilon_t^2\varepsilon_{t-n}^2) = (\nu_2^2\alpha_0\sum_{j=0}^{n-1}\gamma_c^j)\mathsf{E}h_t^2 + (\nu_k\overline{\gamma}_c\gamma_c^{n-1})\mathsf{E}h_t^4 \tag{A.11}$$

where $\overline{\gamma}_c = \mathsf{E}(z_t^2 c_t)$. From Theorem 2.1, $\mathsf{E}h_t^4 = \{\alpha_0^2(1+\gamma_c)+\alpha_0\gamma_{f2}\}/\{(1-\gamma_{c2})(1-\gamma_c)\}$ and, besides, $\mathsf{E}h_t^2 = \alpha_0/(1-\gamma_c)$. Inserting these into (A.11) under the condition $\gamma_{c2} < 1$ and some further manipulation yields (2.26). ⊣

Proof of Theorem 2.4. Setting $k = 1$ in (2.2) gives

$$h_t = g_{t-1} + c_{t-1}h_{t-1}. \tag{A.12}$$

Applying (A.12) to h_t in $h_t^2 h_{t-1}^2$ yields

$$h_t^2 h_{t-1}^2 = (g_{t-1}^2 + 2g_{t-1}c_{t-1}h_{t-1} + c_{t-1}^2 h_{t-1}^2)h_{t-1}^2. \tag{A.13}$$

Multiplying both sides of (A.13) by $z_t^2 z_{t-1}^2$ and taking expectations we have

$$\mathsf{E}\varepsilon_t^2\varepsilon_{t-1}^2 = \nu_2\overline{\gamma}_{g2}\mathsf{E}h_t^2 + 2\nu_2\overline{\gamma}_{gc}\mathsf{E}h_t^3 + \nu_2\overline{\gamma}_{c2}\mathsf{E}h_t^4. \tag{A.14}$$

Repeating the above steps for $\mathsf{E}\varepsilon_t^2\varepsilon_{t-2}^2$ one obtains

$$\begin{aligned} \mathsf{E}\varepsilon_t^2\varepsilon_{t-2}^2 &= \nu_2(\nu_2\gamma_{g2} + \gamma_{c2}\overline{\gamma}_{g2} + 2\widetilde{\gamma}_{gc}\overline{\gamma}_g)\mathsf{E}h_t^2 \\ &\quad +2\nu_2(\widetilde{\gamma}_{gc}\overline{\gamma}_c + \overline{\gamma}_{gc}\gamma_{c2})\mathsf{E}h_t^3 + \nu_2\gamma_{c2}\overline{\gamma}_{c2}\mathsf{E}h_t^4. \end{aligned} \tag{A.15}$$

Rewriting the right-hand side of (A.15) by applying (A.14) gives

$$\begin{aligned} \mathsf{E}\varepsilon_t^2\varepsilon_{t-2}^2 &= \gamma_{c2}\mathsf{E}\varepsilon_t^2\varepsilon_{t-1}^2 + \nu_2(\nu_2\gamma_{g2} + 2\widetilde{\gamma}_{gc}\overline{\gamma}_g)\mathsf{E}h_t^2 \\ &\quad +2\nu_2\widetilde{\gamma}_{gc}\overline{\gamma}_c\mathsf{E}h_t^3. \end{aligned}$$

In the same fashion, one can write

$$\begin{aligned} \mathsf{E}\varepsilon_t^2\varepsilon_{t-3}^2 &= \gamma_{c2}\mathsf{E}\varepsilon_t^2\varepsilon_{t-2}^2 + \nu_2(\nu_2\gamma_{g2} + 2\nu_2\widetilde{\gamma}_{gc}\gamma_g \\ &\quad +2\widetilde{\gamma}_{gc}\overline{\gamma}_g\gamma_c)\mathsf{E}h_t^2 + 2\nu_2\widetilde{\gamma}_{gc}\overline{\gamma}_c\gamma_c\mathsf{E}h_t^3. \end{aligned}$$

Generally, recursively applying (A.12) to h_t in $h_t^2 h_{t-n}^2$ gives

$$\begin{aligned} h_t^2 h_{t-n}^2 &= \{g_{t-1}^2 + \sum_{j=2}^{n}(g_{t-j}^2 \prod_{i=1}^{j-1} c_{t-i}^2) \\ &+2\sum_{j=2}^{n-1}[g_{t-j}\prod_{i=1}^{j-1} c_{t-i}^2(\sum_{l=j+1}^{n} g_{t-l}\prod_{k=j}^{l-1} c_{t-k})] \\ &+2g_{t-1}\sum_{j=2}^{n}(g_{t-j}\prod_{i=1}^{j-1} c_{t-i}) + 2g_{t-1}\prod_{i=1}^{n} c_{t-i}h_{t-n} \\ &+2\sum_{j=2}^{n}(g_{t-j}\prod_{i=1}^{j-1} c_{t-i})(\prod_{i=1}^{n} c_{t-i})h_{t-n} \\ &+\prod_{i=1}^{n} c_{t-i}^2 h_{t-n}^2\}h_{t-n}^2. \end{aligned} \tag{A.16}$$

From (A.16) it follows that

$$\begin{aligned} \mathsf{E}\varepsilon_t^2\varepsilon_{t-n}^2 &= \nu_2\{\nu_2\gamma_{g2}\sum_{j=1}^{n-1}\gamma_{c2}^{j-1} + \gamma_{c2}^{n-1}\overline{\gamma}_{g2} \\ &+2\nu_2\widetilde{\gamma}_{gc}\gamma_g(\sum_{j=2}^{n-1}\gamma_{c2}^{j-1}(\sum_{l=j+1}^{n-1}\gamma_c^{l-j-1}) + \sum_{j=2}^{n-1}\gamma_c^{j-2}) \\ &+2\widetilde{\gamma}_{gc}\overline{\gamma}_g(\sum_{j=2}^{n-1}\gamma_{c2}^{j-1}\gamma_c^{n-j-1} + \gamma_c^{n-2})\}\mathsf{E}h_t^2 \\ &+2\nu_2\{\widetilde{\gamma}_{gc}\overline{\gamma}_c(\gamma_c^{n-2} + \sum_{j=2}^{n-1}\gamma_{c2}^{j-1}\gamma_c^{n-j-1}) \\ &+\overline{\gamma}_{gc}\gamma_{c2}^{n-1}\}\mathsf{E}h_t^3 + \nu_2\gamma_{c2}^{n-1}\overline{\gamma}_{c2}\mathsf{E}h_t^4. \end{aligned} \tag{A.17}$$

On the other hand, applying (A.17) to $\mathsf{E}\varepsilon_t^2\varepsilon_{t-(n-1)}^2$ we can rewrite $\mathsf{E}\varepsilon_t^2\varepsilon_{t-n}^2$ as

$$\begin{aligned} \mathsf{E}\varepsilon_t^2\varepsilon_{t-n}^2 &= \gamma_{c2}\mathsf{E}\varepsilon_t^2\varepsilon_{t-(n-1)}^2 + \nu_2(\nu_2\gamma_{g2} + 2\nu_2\widetilde{\gamma}_{gc}\gamma_g\sum_{j=2}^{n-1}\gamma_c^{j-2} \\ &+2\widetilde{\gamma}_{gc}\overline{\gamma}_g\gamma_c^{n-2})\mathsf{E}h_t^2 + 2\nu_2\widetilde{\gamma}_{gc}\overline{\gamma}_c\gamma_c^{n-2}\mathsf{E}h_t^3. \end{aligned} \tag{A.18}$$

Some further manipulations of (A.18) yield (2.31) and (2.32). ⊣

Proof of Theorem 2.6. We assume that $g_t \equiv \alpha_0$ and $c_t = \beta + \alpha_1 |z_t|$ and, furthermore, that $z_t \sim$nid(0,1). For any fixed β, $\rho_1(1)$ and ρ_1^+ are functions of α_1 defined on $(0, \alpha)$ where $\alpha > 0$ such that $\gamma_{c4} = 1$. Set $\rho_1^0(1, \alpha_1) = \rho_1^0(1)$, $\rho_1^+(\alpha_1) = \rho_1^+$ and , furthermore, $\gamma_{c4}(\alpha_1) = \gamma_{c4}$, respectively to stress this fact. We have

$$\frac{\partial}{\partial\alpha_1}\rho_1^0(1, \alpha_1) = \nu_1 + 2\beta\alpha_1\nu_1^2(1 - \beta^2 - \beta\alpha_1\nu_1)/(1 - \beta^2 - 2\beta\alpha_1\nu_1)^2.$$

As α_1 and β are positive, $\frac{\partial}{\partial\alpha_1}\rho_1^0(1,\alpha_1) > 0$. This implies that for a given β, $\rho_1(1,\alpha_1)$ is a monotonically increasing function of α_1 on $(0,\ \alpha)$. Next we show that $\rho_1^+(\alpha_1)$ is also a monotonically increasing function of α_1 on $(0,\ \alpha)$. Choose two arbitrary values $\alpha_{11},\ \alpha_{12} \in (0,\ \alpha)$ such that $\alpha_{11} < \alpha_{12}$. For a fixed β, $\gamma_{c4}(\alpha_{11}) < \gamma_{c4}(\alpha_{12})$. This implies $\rho_1^+(\alpha_{11}) < \rho_1^+(\alpha_{12})$ so that ρ_1^+ is a monotonically increasing function of $\gamma_{c4}(\alpha_1)$.

Second, we show that $\rho_1^0(1,\alpha_1) < \rho_1^+(\alpha_1)$ as $\alpha_1 \to 0$. We have $\frac{\partial}{\partial\alpha_1}\rho_1^0(1,\alpha_1)\Big|_{\alpha_1=0} = \frac{\partial}{\partial\alpha_1}\rho_1^+(\alpha_1)\Big|_{\alpha_1=0} = \nu_1 > 0$. Tedious calculation shows that for normal errors, $\frac{\partial^2}{\partial\alpha_1^2}\rho_1^+(\alpha_1)\Big|_{\alpha_1=0} > \frac{\partial^2}{\partial\alpha_1^2}\rho_1^0(1,\ \alpha_1)\Big|_{\alpha_1=0} > 0$ for any $\beta > 0$. Thus both autocorrelations are convex functions of α_1, and there exists an interval $(0,\ \alpha^*)$ such that $\rho_1^0(1,\alpha_1) < \rho_1^+(\alpha_1)$.

Finally, we show that if $\beta = 0$ then $\rho_1^0(1,\alpha) > \rho_1^+(\alpha)$ as $\alpha_1 \to \alpha$. Note that $\alpha_1 \to \alpha$ implies $\kappa_4^+ \to \infty$ and $\gamma_{c4} \to 1$. It is seen from (2.36) that $\rho_1^+(\alpha_1) \to \nu_2/\nu_4^{1/2}$ as $\alpha_1 \to \alpha$. On the other hand, from (2.39) it follows that $\rho_1^0(1,\alpha_1) \to \nu_1/\nu_4^{1/4}$ as $\alpha_1 \to \alpha$. Under the normality assumption we have $\rho_1^0(1,\alpha) > \rho_1^+(\alpha)$.

Both $\rho_1^+(\alpha_1)$ and $\rho_1^0(1,\alpha_1)$ are monotonically increasing functions of α_1 on $(0,\alpha)$. It is seen that on $(0,\alpha)$, $(\rho_1^0(1,0),\ \rho_1^0(1,\alpha)) \supset (\rho_1^+(0), \rho_1^+(\alpha))$. Then for any $\alpha_1 \in (0,\alpha)$, there exists $\alpha^* \in (0,\alpha)$ such that $\rho_1^0(1,\alpha^*) = \rho_1^+(\alpha^*)$. Therefore for any given α_1 we have $\rho_1^0(1,\alpha_1) < \rho_1^+(\alpha_1)$ if $\alpha_1 < \alpha^*$, otherwise $\rho_1^0(1,\alpha_1) > \rho_1^+(\alpha_1)$. $\dashv$

Tables 2.1-2.2

Table 2.1. The kurtosis and the first-and fifth-order autocorrelations of squared observations for a set of GARCH(1,1) models, the corresponding estimates based on 1000 simulated series of 100000 observations each, the estimates of the autocorrelations of absolute-valued observations from the same simulated series, and the number of times (out of 1000) when the value of the estimated first-order autocorrelation of absolute values exceeded that of the squares

α_1	β	κ_4	ρ_1^+	ρ_5^+	$\widehat{\rho}_1^+$	$\widehat{\rho}_5^+$	$\widehat{\rho}_1^0(1)$	$\widehat{\rho}_5^0(1)$	Score 1*	Score 5**
0.53	0	13.7	0.530	0.042	0.494	0.029	0.421	0.021	36	264
0.575	0	247	0.575	0.063	0.516	0.038	0.457	0.030	88	381
0.23	0.70	13.8	0.427	0.319	0.386	0.280	0.330	0.243	43	86
0.24	0.70	291	0.471	0.368	0.406	0.302	0.362	0.276	96	208
0.09	0.90	16.1	0.350	0.337	0.299	0.283	0.274	0.262	102	126
0.0915	0.90	277	0.389	0.376	0.314	0.299	0.300	0.288	248	299

*Score 1 equals the number of times (out of 1000) in which the simulated value of $\rho_1^0(1)$ exceeds that of ρ_1^+.

**Score 5 equals the number of times (out of 1000) in which the simulated value of $\rho_5^0(1)$ exceeds that of ρ_5^+.

Table 2.2. The kurtosis and the first-and fifth-order autocorrelations of squared observations for a set of GARCH(1,1) models, the corresponding estimates based on 1000 simulated series of 1000 observations each, the estimates of the autocorrelations of absolute-valued observations from the same simulated series, and the number of times (out of 1000) when the value of the estimated first-order autocorrelation of absolute values exceeded that of the squares

α_1	β	κ_4	ρ_1^+	ρ_5^+	$\widehat{\rho}_1^+$	$\widehat{\rho}_5^+$	$\widehat{\rho}_1^0(1)$	$\widehat{\rho}_5^0(1)$	Score 1*	Score 5**
0.53	0	13.7	0.530	0.042	0.432	0.014	0.407	0.015	310	543
0.575	0	247	0.575	0.063	0.452	0.019	0.438	0.021	381	564
0.23	0.70	13.8	0.427	0.319	0.315	0.205	0.302	0.211	399	567
0.24	0.70	291	0.471	0.368	0.329	0.220	0.324	0.234	471	615
0.09	0.90	16.1	0.350	0.337	0.192	0.175	0.197	0.183	554	574
0.0915	0.90	277	0.389	0.376	0.199	0.181	0.207	0.192	573	613

*Score 1 equals the number of times (out of 1000) in which the simulated value of $\rho_1^0(1)$ exceeds that of ρ_1^+.
**Score 5 equals the number of times (out of 1000) in which the simulated value of $\rho_5^0(1)$ exceeds that of ρ_5^+.

Figures 2.1-2.2

Fig. 2.1 The kurtosis of $\{\varepsilon_t\}$ graphed against γ_{c2} from two models: (a) the GQARCH(1,1) model with $\gamma_c = 0.95, \alpha_0 = 0.12, \zeta = 1$, (the upper curve) and (b) the GARCH(1,1) model with $\gamma_c = 0.95$, where $z_t \sim nid(0,1)$

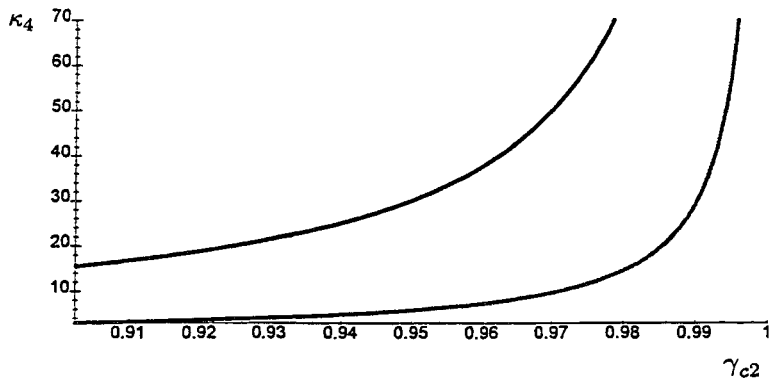
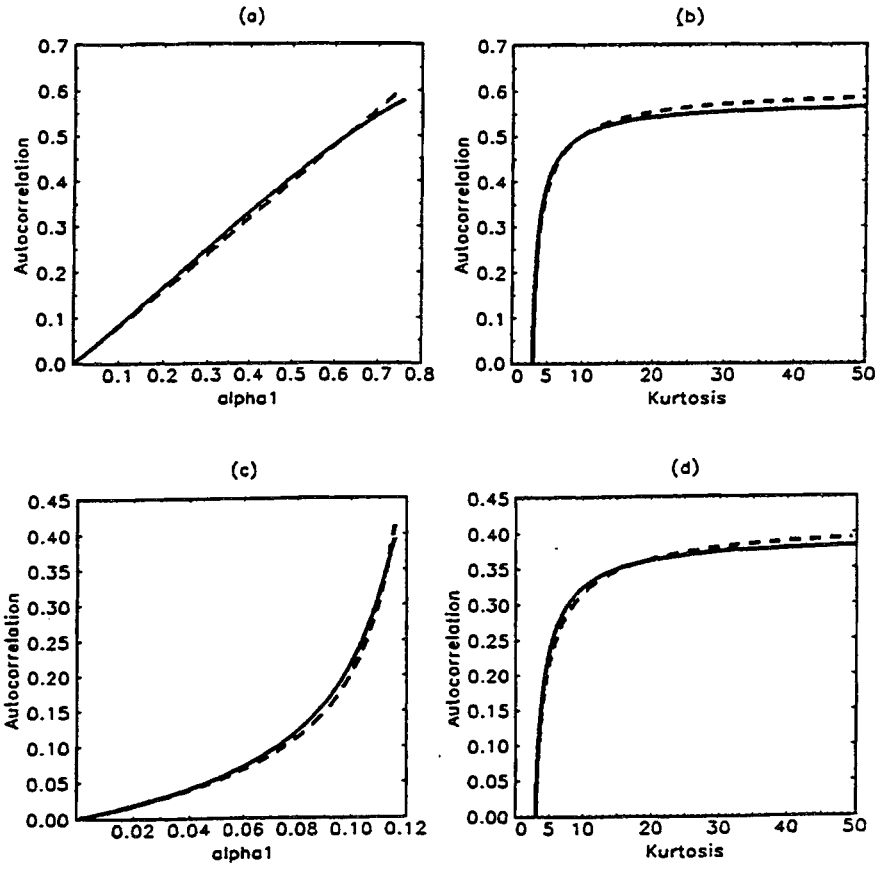


Figure 2.2 First-order autocorrelation of squared observations from the standard GARCH(1,1) model (a) as a function of α_1 with $\beta = 0$, (b) as a function of the kurtosis with $\beta = 0$, (c) as a function of α_1 with $\beta = 0.9$, (d) as a function of the kurtosis with $\beta = 0.9$



3

Statistical Properties of the Asymmetric Power ARCH Process

3.1 Introduction

The temporal properties of power-transformed returns, r_t, from speculative assets have recently received considerable attention. Clive Granger, whose interest in commodity and stock markets can be traced back to Elliott and Granger (1967) and Granger and Morgenstern (1970), has discussed them in a series of papers: see Ding, Granger and Engle (1993), henceforth DGE, Granger and Ding (1995a, b; 1996) and Ding and Granger (1996). In DGE the authors reminded the reader of the fact which Taylor (1986, pp. 52-56) first observed, namely, that $|r_t|$ is positively autocorrelated even at long lags. DGE found that $|r_t|^d$, $d > 0$, is even "long-memory" in that the autocorrelations remain not only positive but rather high at very long lags. Moreover, the authors observed that the property is strongest for $d = 1$ and near it. They noted that this may argue against ARCH specifications that are based on squared returns such as the classical GARCH model of Bollerslev (1986) and Taylor (1986, pp. 78-79). Consequently, DGE introduced a new GARCH model with a separate "heteroskedasticity parameter" δ. This model, the Asymmetric Power ARCH (A-PARCH) model, contained as special cases a host of previous GARCH models. DGE demonstrated its potential by fitting the model with normal errors to the long S&P 500 daily stock return series from 3 January 1928 to 30 April 1991, 17054 observations in all.

In this chapter we take another look at statistical properties of the A-PARCH process. As the motivation for the model lies in observed autocor-

relation functions of $|r_t|^d$, see DGE, we mainly focus on the corresponding autocorrelation function of the A-PARCH process. In particular, the relationship between d and the corresponding model parameter which will be called δ is of interest. It turns out not to be as straightforward as one might think. We demonstrate how the heteroskedasticity parameter and an asymmetry parameter also present in the A-PARCH model add to the flexibility of the specification. In order to proceed we need certain (fractional) moments of the A-PARCH process and having obtained those we have the autocorrelation function necessary for our discussion.

Chapter 3 is organized as follows. The fractional moments and the autocorrelation function of the A-PARCH process are presented in Section 3.2. A transformation of observations implicit in the A-PARCH specification is the topic of Section 3.3. Section 3.4 considers the autocorrelation function of $|r_t|^\delta$ as a function of δ. Section 3.5 discusses possible advantages of the A-PARCH specification in the light of the empirical example in DGE. Section 3.6 concludes.

3.2 Moments and the autocorrelation function

Consider the general A-PARCH(p, q, δ) model

$$\varepsilon_t = z_t h_t \tag{3.1}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with a symmetric density with zero mean and a finite 2δ-th unconditional absolute moment. Furthermore,

$$\begin{aligned} h_t^\delta &= \alpha_0 + \sum_{j=1}^{p} \alpha_j(|\varepsilon_{t-j}| - \phi_j \varepsilon_{t-j})^\delta + \sum_{j=1}^{q} \beta_j h_{t-j}^\delta \\ &= \alpha_0 + \sum_{j=1}^{p} c_{\delta,t-j} h_{t-j}^\delta \end{aligned} \tag{3.2}$$

where $c_{\delta,t-j} = \alpha_j(|z_{t-j}| - \phi_j z_{t-j})^\delta + \beta_j$, $j = 1, ..., p$. Expression (3.2) for $\delta = 2$ and $\phi_j = 0$, $j = 1, ..., p$, appeared in Chapter 4. Following their arguments it is possible to derive the autocorrelation function of $|\varepsilon_t|^\delta$ when $\{\varepsilon_t\}$ is generated by (3.1) and (3.2). The result is not explicitly stated here because we shall concentrate on the first-order case ($p = q = 1$). This implies that (3.2) becomes

$$h_t^\delta = \alpha_0 + c_{\delta,t-1} h_{t-1}^\delta \tag{3.3}$$

where $c_{\delta t} = \alpha_1(|z_t| - \phi z_t)^\delta + \beta_1$ and $\phi = \phi_1$, and $\{c_{\delta t}\}$ is a sequence of iid observations. Letting $\gamma_\delta = \mathrm{E}c_{\delta t}$ and $\gamma_{2\delta} = \mathrm{E}c_{\delta t}^2$ we have

Theorem 3.1 *For the A-PARCH$(1,1,\delta)$ model (3.1) with (3.3), a necessary and sufficient condition for the existence of the 2δ-th absolute moment $\mu_{2\delta} = \mathsf{E}\,|\varepsilon_t|^{2\delta}$ is*

$$\gamma_{2\delta} < 1. \tag{3.4}$$

If (3.4) holds, then

$$\mu_{2\delta} = \alpha_0^2 \nu_{2\delta}(1+\gamma_\delta)/\{(1-\gamma_\delta)(1-\gamma_{2\delta})\} \tag{3.5}$$

and the "moment coefficient"

$$\kappa_{2\delta} = \mu_{2\delta}/\mu_\delta^2 = \kappa_{2\delta}(z_t)(1-\gamma_\delta^2)/(1-\gamma_{2\delta}) \tag{3.6}$$

where $\nu_\psi = \mathsf{E}\,|z_t|^\psi$ and $\kappa_{2\delta}(z_t) = \nu_{2\delta}/\nu_\delta^2$. Furthermore, the autocorrelation function $\rho_n(\delta) = \rho(|\varepsilon_t|^\delta, |\varepsilon_{t-n}|^\delta)$, $n \geq 1$, of $\{|\varepsilon_t|^\delta\}$ has the form

$$\rho_n(\delta) = \frac{\nu_\delta \gamma_\delta^{n-1}[\overline{\gamma}_\delta(1-\gamma_\delta^2) - \nu_\delta\gamma_\delta(1-\gamma_{2\delta})]}{\nu_{2\delta}(1-\gamma_\delta^2) - \nu_\delta^2(1-\gamma_{2\delta})} \tag{3.7}$$

where $\overline{\gamma}_\delta = \mathsf{E}(|z_t|^\delta c_{\delta t})$.

In particular, when $\delta = 2$, the kurtosis of $\{\varepsilon_t^2\}$ equals $\kappa_4 = \kappa_4(z_t)(1-\gamma_2^2)/(1-\gamma_4)$ and $\rho_n(2) = \kappa_4(z_t)\gamma_2^{n-1}(\overline{\gamma}_2 - \gamma_2\kappa_4^{-1})/(1-\kappa_4^{-1})$, $n \geq 1$, where $\kappa_4(z_t)$ is the kurtosis of $\{z_t\}$; see Chapter 2 for a discussion.

The results follow from the general GARCH(p,q) case Chapter 4 considered, but they may also be derived by substituting h_t^δ and $c_{\delta t}$ for their counterparts in the proofs of Theorems 2.1 and 2.4 of Chapter 2. DGE also discussed γ_δ, and Ding and Granger (1996) derived the (conditional) autocorrelation function of $|\varepsilon_t|^\delta$ under the assumption $\nu_\delta = 1$. As DGE pointed out, many well-known GARCH models are special cases of the A-PARCH model. Theorem 3.1 applies to those as well, but they are considered in more detail in Chapter 2.

As stated in the beginning, we shall examine the role of the power parameter δ in the A-PARCH model. The model was introduced to characterize processes $\{\varepsilon_t\}$ in which $\rho(|\varepsilon_t|, |\varepsilon_{t-n}|)$ decays slowly as a function of n. From (3.7) we have

$$\rho_n(\delta) = \rho_1(\delta)\gamma_\delta^{n-1}, \ n \geq 1. \tag{3.8}$$

The decay of the autocorrelation function is exponential, as Ding and Granger (1996) already noted, and γ_δ controls this decay. For a closer look at the maximum value of the first-order autocorrelation (3.7) note that, due to the symmetry of the density of z_t, we can express γ_δ, $\overline{\gamma}_\delta$ and $\gamma_{2\delta}$, respectively, as

$$\gamma_\delta = (\alpha_1/2)\phi_\delta\nu_\delta + \beta_1 \tag{3.9}$$

$$\overline{\gamma}_\delta = (\alpha_1/2)\phi_\delta\nu_{2\delta} + \beta_1\nu_\delta \tag{3.10}$$

$$\gamma_{2\delta} = (\alpha_1^2/2)\phi_{2\delta}\nu_{2\delta} + \alpha_1\beta_1\phi_\delta\nu_\delta + \beta_1^2 \tag{3.11}$$

where $\phi_\delta = (1+\phi)^\delta + (1-\phi)^\delta$ and $\phi_{2\delta} = (1+\phi)^{2\delta} + (1-\phi)^{2\delta}$. Using (3.9), expressions (3.10) and (3.11) may be rewritten as

$$\begin{aligned} \overline{\gamma}_\delta &= \gamma_\delta \nu_\delta + (\alpha_1/2)\phi_\delta(\nu_{2\delta} - \nu_\delta^2) && (3.12) \\ \gamma_{2\delta} &= \gamma_\delta^2 + (\alpha_1^2/2)\phi_{2\delta}\nu_{2\delta} - (\alpha_1^2/4)\phi_\delta^2\nu_\delta^2. && (3.13) \end{aligned}$$

Inserting (3.12) and (3.13) into (3.7) yields

$$\rho_1(\delta) = \frac{(\alpha_1/2)\nu_\delta\phi_\delta(\nu_{2\delta} - \nu_\delta^2)(1-\gamma_\delta^2) + (\alpha_1^2/2)\nu_\delta^2\gamma_\delta(\phi_{2\delta}\nu_{2\delta} - (1/2)\phi_\delta^2\nu_\delta^2)}{(\nu_{2\delta} - \nu_\delta^2)(1-\gamma_\delta^2) + (\alpha_1^2/2)\nu_\delta^2(\phi_{2\delta}\nu_{2\delta} - (1/2)\phi_\delta^2\nu_\delta^2)}. \tag{3.14}$$

From (3.14) it is seen that for given values of α_1, β_1 and ϕ, $\rho_1(\delta)$ is a function of δ expressed in terms of ϕ_δ, $\phi_{2\delta}$, ν_δ and $\nu_{2\delta}$. Assuming $\phi = 0$, $\rho_1(\delta)$ reduces to

$$\rho_1(\delta) = \frac{\alpha_1\nu_\delta(1-\gamma_\delta^2) + \alpha_1^2\gamma_\delta\nu_\delta^2}{(1-\gamma_\delta^2) + \alpha_1^2\nu_\delta^2}. \tag{3.15}$$

On the other hand, setting $\beta_1 = 0$ in (3.14) yields the first-order autocorrelation of squares in the A-PARCH$(1,\delta)$ model:

$$\begin{aligned} \rho_1(\delta) &= \frac{(\alpha_1/2)\nu_\delta\phi_\delta(\nu_{2\delta} - \nu_\delta^2) + (\alpha_1^3/4)\nu_\delta^3\phi_\delta\nu_{2\delta}(\phi_{2\delta} - (1/2)\phi_\delta^2)}{(\nu_{2\delta} - \nu_\delta^2) + (\alpha_1^2/2)\nu_\delta^2\nu_{2\delta}(\phi_{2\delta} - (1/2)\phi_\delta^2)} \\ &= (\alpha_1/2)\phi_\delta\nu_\delta. && (3.16) \end{aligned}$$

Further giving up asymmetry by setting $\phi = 0$ in (3.16) leads to

$$\rho_1(\delta) = \gamma_\delta = \alpha_1\nu_\delta, \tag{3.17}$$

which is analogous to a familiar expression from an ARCH(1) model.

3.3 The power parameter

The unconditional moment (3.5) is a fractional moment. Obtaining integer moments for the A-PARCH process analytically seems difficult. Nonetheless, the fractional moments are useful in considering the role of the power parameter δ in the A-PARCH(1,1) model. Transform ε_t as follows:

$$\xi_t = \operatorname{sgn}(\varepsilon_t)\,|\varepsilon_t|^{\delta/2}. \tag{3.18}$$

Assuming $\phi = 0$ in (3.3) so that no asymmetry is present then Theorem 3.1 yields the fourth moment of ξ_t. The kurtosis of ξ_t is

$$\kappa_4(\xi_t) = \mu_{2\delta}/\mu_\delta^2 = \kappa_4^*(z_t)(1-\gamma_\delta^2)/(1-\gamma_{2\delta}) \tag{3.19}$$

where $\kappa_4^*(z_t) = \nu_{2\delta}/\nu_\delta^2$ is the kurtosis of $\operatorname{sgn}(z_t)|z_t|^{\delta/2}$ and $\gamma_\delta = \alpha_1\nu_\delta + \beta_1$. Furthermore, (3.7) is the autocorrelation function of $\{\xi_t^2\}$. In fact, (3.1) with

(3.3) where $\phi = 0$ is the ordinary GARCH(1,1) model for $\{\xi_t\}$. Note, however, that the distribution of z_t also changes although the density retains its symmetry about zero. There does not exist a corresponding transformation for $\phi \neq 0$. Nonetheless, if we reparameterize the A-PARCH(1,1) model by replacing the term $(|\varepsilon_{t-1}| - \phi\varepsilon_{t-1})$, $\phi < 1$, by $(\varepsilon_{t-1} - \phi\,|\varepsilon_{t-1}|)$ we may define

$$\xi_t^* = \operatorname{sgn}(\varepsilon_t - \phi\,|\varepsilon_{t-1}|)\,|\varepsilon_t - \phi\,|\varepsilon_{t-1}||^{\delta/2}. \tag{3.20}$$

When $\{\varepsilon_t\}$ has the modified A-PARCH(1,1) representation the sequence $\{\xi_t^*\}$ obeys a symmetric GARCH(1,1) model. Its kurtosis and the autocorrelation function of $(\xi_t^*)^2$ are given by (3.19) and (3.7) for $\phi \neq 0$.

This may give interesting insight in what the contribution of the extra parameters δ and ϕ is as regards the kurtosis and the autocorrelation function of the squares. To illustrate, consider first real data. We choose the long daily return series of the S&P 500 stock index of 17054 observations from 3 January 1928 to 30 April 1991 compiled by William Schwert and analyzed in DGE, Granger and Ding (1995a, b), and Ding and Granger (1996); see also Rydén, Teräsvirta and Åsbrink (1997). The series is transformed according to (3.18) and (3.20). Figure 3.1 shows that the first-order autocorrelation of the squared observations is a concave function of δ. Value $\delta = 2$ corresponds to the identity transformation. The solid line represents the case $\phi = 0$. The maximum of the autocorrelation function occurs in the vicinity of $\delta = 1$ which agrees with what DGE and Granger and Ding (1995a) observed. The dashed-dotted line is the first-order autocorrelation for the case $\phi = -0.3$. Introducing asymmetry in the transformation seems to reduce the first-order autocorrelation and the value of the autocorrelation-maximizing δ. Figure 3.2 shows the relationship between the kurtosis and the first-order autocorrelation of the transformed variables. The kurtosis is an increasing function of δ. While the kurtosis decreases the autocorrelation as a concave function of the kurtosis first increases and then starts to decline. Although the kurtosis may be reduced to that of the normal distribution the higher-order dependence in the original series does not vanish with the transformation. Some skewness also remains. The asymmetry introduced through ϕ tends to increase kurtosis and, at the same time, reduce the autocorrelation (the dashed-dotted line in Figure 3.2). Given the problem of modelling series displaying high kurtosis and low first-order autocorrelation of squared observations with GARCH models with normal errors (see, for example, Chapter 2) this is an interesting observation. Of course, ϕ also affects the skewness of the transformed observations. A value such as $\phi = -0.1$ already transforms the originally negatively skewed series to be positively skewed for a wide range of values of δ (small values are an exception).

Next we turn to the theoretical A-PARCH model and the role of δ and ϕ there.

3.4 Convexity of the theoretical autocorrelation function

In this section we consider $\rho_1(\delta)$ as a function of δ. We begin by defining $\rho_1(\delta)$ on a set of values of δ such that $\rho_1(\delta)$ exists. Let $I = (0, l)$, $l > 0$, be an open interval such that $\delta_0 \in I$ if and only if $\gamma_{2\delta_0} < 1$. Then there exists $I_\delta = [\delta_a, \delta_b]$, which is a closed interval, such that $I_\delta \subset I$. Obviously, for any $\delta \in I_\delta$ we have $\gamma_{2\delta} < 1$.

We now state our main result:

Property 3.4.1 *Assume that ν_δ is a convex function of δ on I_δ with a finite second derivative for any $\delta \in I_\delta$. Then $\rho_1(\delta)$ given in (3.15) is a convex function of δ on I_δ. Furthermore, if $\frac{\partial \nu_\delta}{\partial \delta}\Big|_{\delta=\delta_0} = 0$ and $\delta_0 \in I_\delta$, then $\frac{\partial \rho_1(\delta)}{\partial \delta}\Big|_{\delta=\delta_0} = 0$.*

Proof. See Appendix B.

Remark 3.1 The convexity assumption for ν_δ is not very restrictive. In fact, what we have assumed for z_t in Section 3.2 is sufficient for ν_δ to be a convex function of δ on I_δ. To see this, suppose that a random variable X has a continuous density $f(x)$ on $(-\infty, +\infty)$ such that $f(x)$ is nonnegative everywhere. Then $\nu''_\delta = \mathsf{E}(|X|^\delta \ln^2 |X|) > 0$ on I_δ so that ν_δ is a convex function of δ on I_δ. Moreover, $\nu'_\delta = \mathsf{E}(|X|^\delta \ln |X|)$ implies that there exists a unique δ_0 such that $\nu'_{\delta_0} = 0$.

Moreover the asymmetry terms ϕ_δ, $\phi_{2\delta}$ and ϕ^2_δ in (3.14) are convex functions of δ on the interval (0,1) and they reach their minimum values on the interval (0, 0.5]. The minimizing value is a function of ϕ, which is defined as an asymmetric parameter in (3.3).

Remark 3.2 If $\phi = 0$, then it follows from Property 3.4.1 that the first-order autocorrelation $\rho_1(\delta)$ reaches its minimum value at $\delta_0 \in I_\delta$, where ν_δ also has its minimum value. For given parameter values β_1 and α_1, $\min\limits_{\delta \in I_\delta} \{\rho_1(\delta)\}$ thus depends on the distribution of z_t.

From (3.15) we see that $\rho_1(\delta)$ only depends on the δ-th absolute moment ν_δ if we do not allow asymmetry. If $\phi \neq 0$ then it follows from (3.14) that $\rho_1(\delta)$ also depends on $\nu_{2\delta}$, ϕ_δ and $\phi_{2\delta}$. This complicates the analysis. Nevertheless, assume as before that ν_δ with a second derivative for any $\delta \in I_\delta$ is a convex function of δ on I_δ. This implies that $\nu_{2\delta}$ is also a convex function of δ on I_δ. Moreover, $\phi''_\delta = (1+\phi)^\delta \ln^2(1+\phi) + (1-\phi)^\delta \ln^2(1-\phi) \geq 0$ indicates that ϕ_δ is a convex function of δ on I_δ. Therefore, there exists a subset $I_{\delta^*} \subseteq I_\delta$ such that ν'_δ, $\nu'_{2\delta}$, ϕ'_δ, $\phi'_{2\delta} < 0$ for any $\delta \in I_{\delta^*}$. Then there exists an unique $\delta_0 \in I_\delta$ such that $u'v - uv' = 0$ at $\delta_0 \in I_\delta$, where $\rho_1(\delta) = u/v$ in (3.14). We thus have $\rho'_1(\delta_0) = 0$. Proving $\rho''_1(\delta_0) > 0$ seems complicated and we content ourselves with

Conjecture 3.4.2 *Assume that ν_δ with a second derivative for any $\delta \in I_\delta$ is a convex function of δ on I_δ. Then, for $\phi \neq 0$, $\rho_1(\delta)$ is a convex function of δ on I_δ. Furthermore, if $\frac{\partial \nu_\delta}{\partial \delta}\Big|_{\delta=\delta_0} = 0$ and $\delta_0 \in I_\delta$, then $\frac{\partial \rho_1(\delta)}{\partial \delta}\Big|_{\delta=\delta_0} \neq 0$.*

Remark 3.3 When $\phi \neq 0$, $\rho_1(\delta)$ is still a convex function of δ on I_δ. However, the minimizing $\delta_0 \in I_\delta$ such that $\nu'_{\delta_0} = 0$ is not the same as the $\delta_* \in I_\delta$ for which $\rho'_1(\delta_*) = 0$. The inequality $\delta_* \neq \delta_0$ is due to the effects of the asymmetric terms ϕ_δ and $\phi_{2\delta}$ in (3.14).

We already noticed that the asymmetry parameter ϕ in (3.14) affects the location of the minimum value of $\rho_1(\delta)$ on the interval I_δ. We are also able to assess the effect of ϕ on the value of the autocorrelation function. Let $\rho_1^*(\delta)$ denote the first order autocorrelation $\rho_1(\delta)$ in (3.14) corresponding to $\phi \neq 0$ whereas $\rho_1^\circ(\delta) = \rho_1(\delta)$ in (3.15) where $\phi = 0$. We have

Property 3.4.3 *Consider the A-PARCH model (3.1) with (3.3) and assume that $\rho_1^*(\delta)$ and $\rho_1^\circ(\delta)$ exist for some δ. Then*

$$\rho_1^*(\delta) > \rho_1^\circ(\delta), \textit{when } \delta \geq 1 \tag{3.21}$$

$$\rho_1^*(\delta) > \rho_1^\circ(\delta), \textit{when } 0 < \delta < 1 \textit{and } \gamma_\delta \textit{is sufficiently large.} \tag{3.22}$$

Proof. See Appendix B.

The situation is illustrated in Figure 3.3. The solid concave curve is the first-order autocorrelation of $|\varepsilon_t|^\delta$ estimated from the S&P 500 series for various values of δ. The solid convex one is the theoretical autocorrelation obtained from (3.15) by setting $\alpha_1 = 0.091$ and $\beta_1 = 0.9$. The dashed-dotted curves are the counterparts of the solid curves after introducing skewness by setting $\phi = -0.3$. It is seen how the nonzero asymmetry parameter has the effect of shifting the autocorrelation-minimizing value of δ to the left. Furthermore, the value of the theoretical autocorrelation function increases with $|\phi|$. The point where the theoretical and the empirical autocorrelations are equal is also shifted to the left. One may expect the estimated value of δ to lie in the neighbourhood of the value at which the two curves intersect.

As Remark 3.2 indicates, the minimizing value of $\rho(|\varepsilon_t|^\delta, |\varepsilon_{t-1}|^\delta)$ may also be shifted away from $\delta_{min} = 0.87$ where it is located under the assumption of normality by assuming another distribution for the error term. It can be shown that if the error term is t-distributed then $\delta_{min} < 0.87$. In fact, the minimizing value is an increasing function of degrees of freedom in the t-distribution. On the other hand, the t-distribution can be generalized further into a generalized t (GT) distribution with the density

$$f(x) = \frac{p}{2\sigma q^{1/p} B(1/p, q)(1 + \frac{|x|^p}{q\sigma^p})^{q+1/p}}, \quad |x| < \infty, \ \sigma, p, q > 0 \tag{3.23}$$

where B is the beta-function, see, for example, McDonald and Newey (1988). Bollerslev, Engle and Nelson (1994) employed the GT distribution as the error distribution when modelling U.S. stock index returns. Suitable choices of the two shape parameters p and q shift the minimizing value to the right. For instance, $p = 1.5$ and $q = 200$ yield $\delta_{min} = 1.07$. For comparison, $p = 2$ for the t-distribution.

3.5 An empirical example

The above theory considers the role of δ and ϕ in the A-PARCH model under the *ceteris paribus* assumption that the other parameters remain constant. If a standard GARCH model is augmented by these parameters, the estimates of the other parameters almost certainly change. DGE considered the contribution of δ and ϕ to the GARCH specification by comparing the values of the estimated log-likelihood for three different models fitted to the above-mentioned S&P 500 daily return series of 17054 observations. The models were the AVGARCH(1,1) model, the standard GARCH(1,1) model, and the A-PARCH(1,1,δ) model. We carry out a similar comparison by studying estimated autocorrelation functions. Estimation results from DGE are reproduced in Table 3.1. The last row of Table 3.1 contains the estimated left-hand side of the existence condition of the absolute moment of order 2δ. It is seen that according to the results from the standard GARCH model, the fourth unconditional moment does not exist for the S&P 500 series. On the other hand, according to the A-PARCH model, the absolute moment of order 2.86 exists. In line with this result, the estimated AVGARCH model signals the existence of the second moment of $\{\varepsilon_t\}$.

Table 3.2 shows the autocorrelation functions of absolute values of $|\varepsilon_t|^\delta$ for $\delta = 2$, 1.43, 1. They are estimated directly from the data (D) or from the estimated models (M) applying (3.14) and (3.15). The autocorrelation function from the model is not available for the standard GARCH model because the fourth moment does not exist. We are thus not able to compute any autocorrelations of squared observations from the model. Yet, the left-hand side of the existence condition for the fourth moment is a monotonically increasing function of both α_1 and β_1, and the autocorrelation function converges to an asymptote as the fourth moment vanishes. In order to obtain an approximation to an asymptotic autocorrelation function of squares we decreased the estimates of these parameters in turn such that the fourth moment just exists ($\gamma_4 = 0.9999$) for the two obtained parameter combinations. This yields two extreme asymptotic autocorrelation functions that, however, lie close to each other; the rest may be obtained by reducing both parameters in size simultaneously until $\gamma_4 < 1$. It is seen from Table 3.2 that the discrepancy between the asymptotic autocorrelation functions and the autocorrelations estimated directly from the data

is large. For the AVGARCH(1,1) model, there is also a large discrepancy between the autocorrelation function obtained from the model and the one estimated from the data.. The autocorrelations from the model are very high compared to those estimated directly from the series. The gap between the two is much smaller for the power ARCH model. Interestingly enough, the model-based autocorrelations seem to decay at a slower rate up to 50 lags than the direct ones although the decay rate of the model-based autocorrelations is just exponential. Nevertheless, the extra parameters in the A-PARCH(1,1) model seem to have considerably improved the correspondence between stylized facts in the data on the one hand and the model on the other.

3.6 Conclusions

The results of this chapter indicate how the power or heteroskedasticity (DGE) parameter adds to the flexibility of the GARCH family when it comes to characterizing stylized facts in the observed series. The considerations are restricted to autocorrelations of powers of absolute values whereas the asymmetric response to news handled by a specific asymmetry parameter is not a topic here. It will be discussed in future work. This chapter merely demonstrates the effect of a nonzero asymmetry parameter on the autocorrelation function of powers of absolute-value transformed observations. It is also shown how the choice of error distribution would further enhance the flexibility of the A-PARCH parametrization. As for the heteroskedasticity parameter, its role turns out to be different from what intuition based on empirical considerations might suggest.

The results on moments and the autocorrelation function generalize to higher-order A-PARCH processes by applying the results in Chapter 4 to the present situation. As such a generalization involves tedious notation and because the (1,1) case is by far the most popular one in practice, we do not discuss higher-order A-PARCH processes in this chapter. They are needed if the first few autocorrelations in the observed autocorrelation function of absolute values or squares do not fit into the general exponential decay pattern of the autocorrelations. This possibility is discussed in Chapter 5 in the context of the standard GARCH process (see also Nelson and Cao, 1992), and their results pertain to the present situation.

References

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

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.

Ding, Z. and C.W.J. Granger (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics* **73**, 185-215.

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.

Elliott, C.M. and C.W.J. Granger (1967). A fresh look at wheat prices and markets in the Eighteenth Century. *Economic History Review*, 357-365.

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

Granger, C.W.J. and Z. Ding (1995b). Stylized facts on the temporal and distributional properties of daily data from speculative markets. Department of Economics, University of California, San Diego, unpublished paper.

Granger, C.W.J. and Z. Ding (1996). Varieties of long memory models. *Journal of Econometrics* **73**, 61-77.

Granger, C.W.J. and O. Morgenstern (1970). *Predictability of stock market prices*. New York: Academic Press.

McDonald, J.B. and W.K. Newey (1988). Partially adaptive estimation of regression models via the generalized t distribution. *Econometric Theory* **4**, 428-457.

Nelson, D.B. and C.Q. Cao (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics* **10**, 229-235.

Rydén, T., T. Teräsvirta and S. Åsbrink (1997). Stylized facts of daily return series and the Hidden Markov model. *Journal of Applied Econometrics* (forthcoming).

Schwert, G.W. (1989). Why does stock market volatility change over time? *Journal of Finance* **45**, 1129-1155.

Taylor, S. (1986). *Modelling financial time series.* New York: Wiley.

Appendix B: Proofs

Proof of Property 3.4.1.

Let $\rho_1(\delta) = u/v$ where $u = \alpha_1(1-\beta_1^2)\nu_\delta - \alpha_1^2\beta_1\nu_\delta^2$ and $v = (1-\beta_1^2) - 2\alpha_1\beta_1\nu_\delta$ from (3.16). It follows that $\rho_1(\delta)$ is continuous and twice differentiable on I_δ. The first derivative $\rho_1'(\delta) = (u'v - uv')/v^2$, where u' and v' are the first order derivatives with respect to δ.

First we show that there exists an unique $\delta_0 \in I_\delta$ such that $\rho_1'(\delta_0) = 0$. Set

$$\begin{aligned} & u'v - uv' \\ = \; & \alpha_1(1-\beta_1^2)^2\nu_\delta' - 2\alpha_1^2\beta_1(1-\beta_1^2)\nu_\delta\nu_\delta' + 2\alpha_1^3\beta_1^2\nu_\delta^2\nu_\delta' = 0 \quad \text{(B.1)} \end{aligned}$$

where ν_δ' is a first derivative of ν_δ with respect to δ. From (B.1) it follows that either (i) $\nu_\delta' = 0$, or (ii) $2\alpha_1^2\beta_1^2\nu_\delta^2 - 2\alpha_1\beta_1(1-\beta_1^2)\nu_\delta - 2(1-\beta_1^2)^2 = 0$. If (ii) holds, then $\nu_\delta = (1-\beta_1^2)(1 \pm \sqrt{3})/2\alpha_1\beta_1$. Consider $\nu_d = (1-\beta_1^2)(1 - \sqrt{3})/2\alpha_1\beta_1$. This implies $\nu_\delta < 0$, which is not possible. Alternatively, the solution $\nu_\delta = (1-\beta_1^2)(1+\sqrt{3})/2\alpha_1\beta_1$. Then $2\alpha_1\beta_1\nu_\delta < 1 - \beta_1^2$, which violates the moment condition $\gamma_{2\delta} < 1$. Thus $\nu_{\delta_0}' = 0$. Assume there exists $\delta_0 \in I_\delta$ such that $\nu_{\delta_0}' = 0$. Then $\rho_1'(\delta_0) = 0$.

Second, we show that the second derivative $\rho_1''(\delta_0) > 0$. Write $\rho_1''(\delta) = [(u''v - uv'') - 2(u'v - uv')v']/v^3$. It suffices to show that $u''v - uv'' > 0$ at $\delta = \delta_0$. From (3.4) we have $1 - \beta_1^2 > \delta_1^2\nu_\delta + 2\alpha_1\beta_1\nu_\delta$ so that

$$\begin{aligned} & u''v - uv''|_{\delta=\delta_0} \\ = \; & \alpha_1\nu_{\delta_0}''[(1-\beta_1^2)^2 + 2\alpha_1\beta_1(1-\beta_1^2)\nu_{\delta_0} - 6\alpha_1^2\beta_1^2\nu_{\delta_0}^2] \\ > \; & \alpha_1\nu_{\delta_0}''[\alpha_1^4\nu_{2\delta_0}^2 + 6\alpha_1^3\beta_1\nu_{\delta_0}\nu_{2\delta_0} + 2\alpha_1^2\beta_1^2\nu_{\delta_0}^2] \\ > \; & 0. \; \dashv \qquad \text{(B.2)} \end{aligned}$$

Proof of Property 3.4.3.

Suppose first that $\delta > 1$ and compare (3.14) with (3.15). In this case, $\phi_\delta > 2$, and $(\phi_{2\delta}\nu_{2\delta} - \frac{1}{2}\phi_\delta^2\nu_\delta^2)/(\nu_{2\delta} - \nu_\delta^2) > 2$ implies that $\rho_1^*(\delta) > \rho_1^\circ(\delta)$. When $\delta = 1$, $\phi_\delta = 2$ and $(\phi_{2\delta}\nu_{2\delta} - \frac{1}{2}\phi_\delta^2\nu_\delta^2)/(\nu_{2\delta} - \nu_\delta^2) > 2$. This proves (3.21). When $0 < \delta < 1$, extra conditions are needed for $\rho_1^*(\delta) > \rho_1^\circ(\delta)$. The moment γ_δ should be sufficiently large so that the numerator on the right side of (3.14) is dominated by γ_δ. This is because $1 < \phi_\delta < 2$ and $(\phi_{2\delta}\nu_{2\delta} - \frac{1}{2}\phi_\delta^2\nu_\delta^2)/(\nu_{2\delta} - \nu_\delta^2) > 2$. (For example, from $\gamma_\delta \geq 0.8$ it already follows that $\rho_1^*(\delta) > \rho_1^\circ(\delta)$. In practice, (3.22) tends to hold for $0 < \delta < 1$ because the estimated $\widehat{\gamma}_\delta$ is usually quite close to unity.) $\dashv$

Tables 3.1-3.2

Table 3.1. GARCH specifications estimated for the S&P 500 daily stock return series, 3 January 1928 to 30 April 1991. Source: DGE (value of the t-statistic in parentheses)

Parameter estimate	GARCH	A-PARCH	AVGARCH
	$\delta = 2$	$\widehat{\delta} = 1.43$	$\delta = 1$
$\widehat{\alpha}_0$	8.0×10^{-7} (12.5)	1.4×10^{-6} (4.5)	9.6×10^{-6} (12.6)
$\widehat{\alpha}_1$	0.091 (50.7)	0.083 (32.4)	0.104 (67)
$\widehat{\beta}_1$	0.906 (43.4)	0.920 (474)	0.913 (517)
$\widehat{\phi}$		−0.373 (−20.7)	
$\widehat{\gamma}^1_{2\delta}$	1.0106	0.99526	0.99591

[1] $\widehat{\gamma}_{2\delta}$ = left-hand side of the 2δ moment condition. A value less than unity indicates the existence of the 2δ-th moment of $\{|\varepsilon_t|\}$.

Table 3.2. Autocorrelation functions of $\{|\varepsilon_t|^\delta\}$, $\delta = 2, 1.43, 1$, *estimated from the linearly filtered S&P 500 daily stock return series, January 3, 1928 to April 30, 1991, and computed from estimated models.*

Model	D/M[1]	Lag									
		1	2	3	4	5	10	20	30	40	50
	D	0.252	0.243	0.182	0.153	0.200	0.115	0.095	0.094	0.067	0.069
GARCH	M(a)	0.390	0.387	0.383	0.380	0.377	0.362	0.332	0.306	0.281	0.258
	M(b)	0.387	0.384	0.381	0.378	0.376	0.362	0.335	0.311	0.288	0.267
A-PARCH	D	0.337	0.328	0.294	0.291	0.287	0.212	0.196	0.181	0.154	0.152
	M	0.421	0.418	0.415	0.412	0.409	0.395	0.368	0.343	0.320	0.298
AVGARCH	D	0.343	0.338	0.330	0.314	0.316	0.258	0.249	0.230	0.207	0.200
	M	0.506	0.504	0.502	0.500	0.498	0.488	0.469	0.450	0.433	0.416

[1]D= Autocorrelations estimated from the S&P 500 daily return series
M= Autocorrelations computed from the estimated model
M(a)=Autocorrelations computed from the estimated model by reducing the estimate of α_1 until the stationarity condition is satisfied ($\widehat{\gamma}_4 = 0.9999$)
M(b)=Autocorrelations computed from the estimated model by reducing the estimate of β_1 until the stationarity condition is satisfied ($\widehat{\gamma}_4 = 0.9999$)

Figures 3.1-3.3

Fig. 3.1. The first-order autocorrelations for the S&P 500 stock returns transformed according to (3.18) (solid curve) and (3.20) with $\phi = -0.3$ (dashed-dotted curve) as a function of the transformation parameter δ

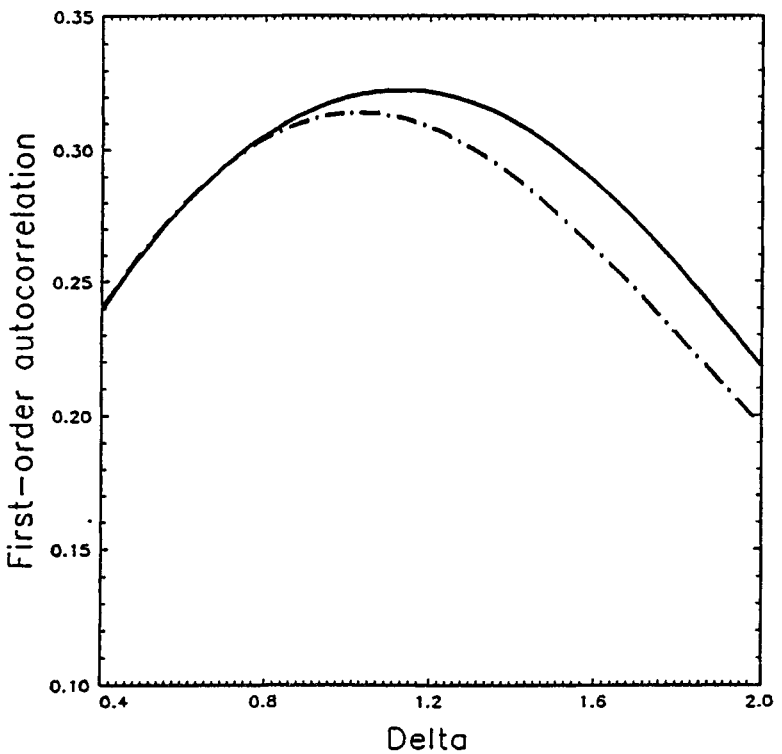


Fig. 3.2. The kurtosis of the S&P 500 stock returns transformed according to (3.18) (solid curve) and (3.20) with $\phi = -0.3$ *(dashed -dotted curve) as a function of the first-order autocorrelation*

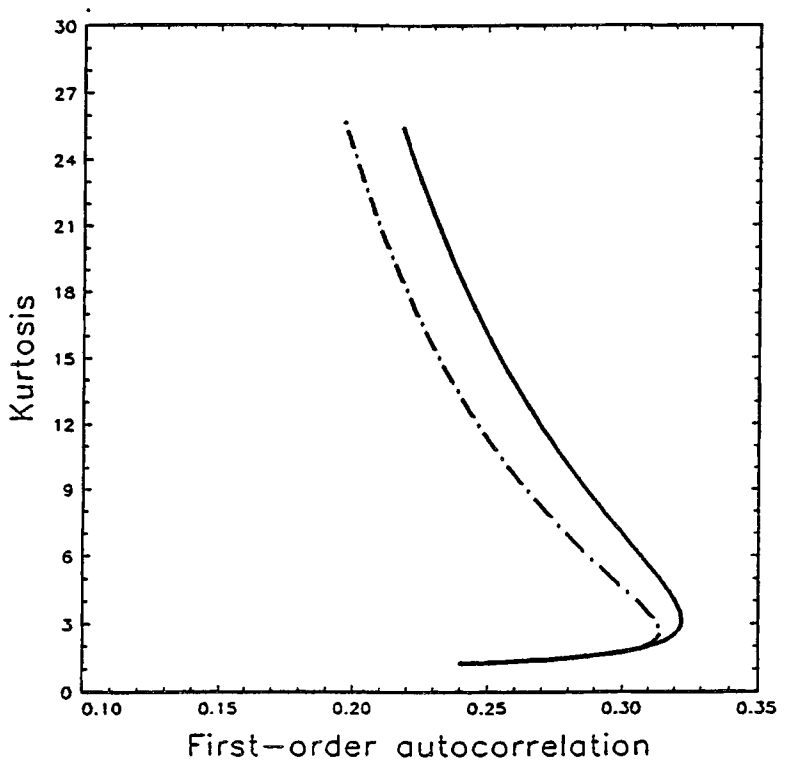
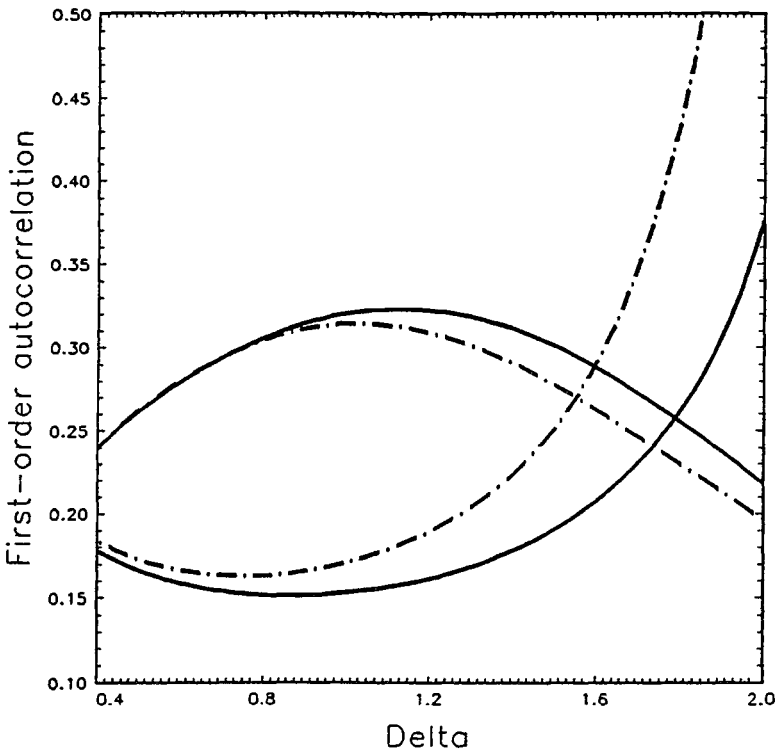


Fig. 3.3. The first-order autocorrelations for the power-transformed S&P 500 series for $\phi = 0$ (solid concave curve) and $\phi = -0.3$ (dashed-dotted concave curve) and the corresponding theoretical autocorrelations from an A-PARCH(1,1) model with $\alpha_1 = 0.091$, $\beta_1 = 0.9$, and $\phi = 0$ (solid convex curve) and $\phi = -0.3$ (dashed-dotted convex curve)



4

Fourth Moment Structure of the GARCH(p, q) Process

4.1 Introduction

The General Autoregressive Conditional Heteroskedasticity (GARCH) model and its many variants are popular in modelling volatility in high-frequency financial series. This is evident from the large number of recent surveys on ARCH and GARCH models and other ways of modelling volatility; see, for example, Bera and Higgins (1993), Bollerslev, Engle and Nelson (1994), Diebold and Lopez (1995), Engle (1995), Guégan (1994, ch. 5), Palm (1996) and Shephard (1996). The statistical properties of the basic GARCH model of Bollerslev (1986) have been discussed in a number of articles. Bollerslev (1986) derived conditions for the existence of unconditional moments of the GARCH(1,1) model under normality. He also derived expressions for these moments as functions of moments of lower order. Teräsvirta (1996) did the same without the normality assumption. Furthermore Bollerslev (1988) found the autocorrelation function for the GARCH(1,1) process with normal errors, and Teräsvirta (1996) generalized the result to symmetric non-normal error distributions. As for higher-order processes, Bollerslev (1986,1988), using Yule-Walker type equations, indicated the shape of the autocorrelation and partial autocorrelation functions of the general GARCH(p, q) process. Milhϕj (1985) already derived the autocorrelation function of an ARCH(q) process which is a special case ($p = 0$) of the GARCH(p, q) model.

Results on the existence of unconditional moments for GARCH models are not only of statistical interest. Practitioners may want to use them to

see what kind of moment implications GARCH models they estimate may have. In particular, the existence of the unconditional fourth moment of stochastic processes generating, say, financial return data has interested researchers. Given both the existence conditions and a suitable expression for the fourth moment, the investigator would be able to check what his estimated model implies about this moment. Obtaining estimates of the kurtosis and the autocorrelations of the centred and squared observations from the model would also be useful. That would enable one to see how well the kurtosis and autocorrelation implied by the estimated model match the estimates obtained directly from the data. Such a comparison would give an indication of how well the model fits the data, because high kurtosis and slowly decaying autocorrelations of the centred and squared observations (or errors) are characteristic features of many financial high-frequency series. Note that the constancy of unconditional moments is a maintained assumption in this chapter; for tests against weak stationarity, see Loretan and Phillips (1994).

The results of Bollerslev (1986,1988) and Teräsvirta (1996) make model evaluations of that type possible for the GARCH(1,1) model. This chapter generalizes some of those results to GARCH(p,q) processes. The focus will be on the fourth moments and the autocorrelation function for the centred and squared observations.

The plan of Chapter 4 is as follows. In Section 4.2 we derive the necessary and sufficient condition for the existence of the unconditional fourth moment of the GARCH(p,p) process and give an expression for the moment itself. Section 4.3 presents the autocorrelation function for the squared process. In those sections, the GARCH(2,2) model is used as an illustration. Section 4.4 shows how these results are modified for the GARCH(p,q) model with $p \neq q$. Section 4.5 concludes.

4.2 Condition for existence of the fourth moment

We begin with a necessary and sufficient condition for existence of the unconditional fourth moment in the GARCH(p,p) model. The model is defined as

$$\varepsilon_t = z_t\sqrt{h_t} \tag{4.1}$$

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

$$h_t = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{p} \alpha_i^2 \varepsilon_{t-i}^2 \tag{4.2}$$

see Bollerslev (1986). Assume now that this GARCH model has a finite fourth moment and let $\mathrm{E}z_t^j = \nu_j$, $j = 2, 4$. Furthermore, let $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_i \geq 0$ for $i = 1, ..., p$, in (4.2). In particular, $\alpha_p > 0$ and $\beta_p > 0$.

Let $c_{i,t-i} = \beta_i + \alpha_i z_{t-i}^2$ for $i = 1, 2, ..., p$, where $\{c_{it}\}$ is a sequence of i.i.d. random variables such that c_{it} is independent of h_t. This allows us to rewrite (4.2) as

$$h_t = \alpha_0 + \sum_{i=1}^{p} c_{i,t-i} h_{t-i}. \tag{4.3}$$

In this paper we will make heavy use of (4.3). From the assumption that the second moments of $\{\varepsilon_t\}$ exist it follows that $\sum_{i=1}^{p} \gamma_{i1} < 1$ where $\gamma_{i1} = \mathsf{E}c_{it} = \beta_i + \alpha_i \nu_2$, $i = 1, ..., p$. From (4.1) we see that in order to consider the unconditional fourth moment of $\{\varepsilon_t\}$ we have to square (4.3), which gives

$$h_t^2 = \alpha_0^2 + 2\alpha_0 \sum_i c_{i,t-i} h_{t-i} + \sum_i c_{i,t-i}^2 h_{t-i}^2 + 2 \sum_{l<m} c_{l,t-l} c_{m,t-m} h_{t-l} h_{t-m}. \tag{4.4}$$

Taking the unconditional expectations on both sides of (4.4) yields

$$\mathsf{E}h_t^2 = \alpha_0^2 + 2\alpha_0 \gamma_1 \mathsf{E}h_t + \gamma_2 \mathsf{E}h_t^2 + 2 \sum_{l<m} \mathsf{E}(c_{l,t-l} c_{m,t-m} h_{t-l} h_{t-m}) \tag{4.5}$$

where $\mathsf{E}h_t = \alpha_0/(1-\gamma_1)$, $\gamma_i = \sum_{j=1}^{p} \gamma_{ji}$, $i = 1, 2$, and $\gamma_{i2} = \mathsf{E}c_{it}^2 = \beta_i^2 + 2\alpha_i \beta_i \nu_2 + \alpha_i^2 \nu_4$ for $i = 1, ..., p$. Thus, $\mathsf{E}h_t^2$ can be determined through (4.5) if we find a convenient expression for the mixed moments $\mathsf{E}(c_{l,t-l} c_{m,t-m} h_{t-l} h_{t-m})$. We shall tackle this problem by first finding a suitable expression for $h_t h_{t-n}$ and then using it in (4.5). This will be done in the next two subsections.

4.2.1 A useful representation of $h_t h_{t-n}$

First we introduce some notation for the case $1 \leq n \leq p$. Let $\mathbf{e}_r = (1, ..., 1, 0, ..., 0)'$ be a $p \times 1$ vector with the first r components equal to 1. Furthermore, $\mathbf{c}_P = (c_{1,t-1}, ..., c_{p,t-p})'$ is a $p \times 1$ vector with the index set $P = \{1, ..., p\}$, whereas $\mathbf{c}_{P\setminus\{n\}} = (c_{1,t-1}, ..., c_{n-1,t-n+1}, c_{n+1,t-n-1}, ..., c_{p,t-p})'$ is a $(p-1) \times 1$ subvector of $\mathbf{c}_P$ obtained by excluding the element $c_{n,t-n}$ from $\mathbf{c}_P$. Matrix $\mathbf{I}_n$ is an $n \times n$ identity matrix, $\mathbf{0}_{m\times n}$ is an $m \times n$ null matrix and $\mathbf{0}_m$ is an $m \times 1$ null vector. We also define $(p-1) \times (p-1)$ matrices $\mathbf{C}_1 = \mathbf{I}_{p-1}$ and

$$\mathbf{C}_j = \begin{pmatrix} \mathbf{c}'_{(P-1)\setminus\{n-j+1\}} & c_{p,t-j-p+1} \\ \mathbf{I}_{p-2} & \mathbf{0}_{p-2} \end{pmatrix} \tag{4.6}$$

for $j = 2, 3, ..., n$, where $\mathbf{c}_{(P-1)\setminus\{n-j+1\}} = (c_{1,t-j}, ..., c_{n-1,t-j-n+2}, c_{n+1,t-j-n}, ..., c_{p-1,t-j-p+2})'$. Let $\mathbf{G}_{p-j}[c_{j+1,t-(i+j)+1}]$ be a $(p-j) \times (p-j)$ matrix generated by the "seed" element $c_{j+1,t-(i+j)+1}$. It has the form

$$\mathbf{G}_{p-j}[c_{j+1,t-(i+j)+1}] = \begin{pmatrix} \mathbf{c}'_{(P-1)\setminus\{1,...,j\}} & c_{p,t-i-p+2} \\ c_{j,t-i-j+2}\mathbf{I}_{p-j-1} & \mathbf{0}_{p-j-1} \end{pmatrix} \tag{4.7}$$

where $\mathbf{c}_{(P-1)\setminus\{1,\ldots,j\}} = (c_{j+1,t-(i+j-1)}, c_{j+2,t-(i+j)}, \ldots, c_{p-1,t-(i+p-3)})'$ is a $(p-j-1)\times 1$ vector obtained by excluding the first j elements from $\mathbf{c}_{P-1}$. Let $p^* = p(p-1)/2$. Using (4.7), we define a $(p-1)\times p^*$ matrix $\mathbf{C}_{n+1}$ as

$$\mathbf{C}_{n+1} = \begin{pmatrix} \mathbf{G}_{p-1}[c_{2,t-n-2}] & \vdots\, \mathbf{0}'_{p-2} & \vdots \cdots \vdots & \mathbf{0}_{(p-3)\times 2} & \vdots\, \mathbf{0}_{p-2} \\ & \vdots\, \mathbf{G}_{p-2}[c_{3,t-n-3}] & \vdots \cdots \vdots & \mathbf{G}_2[c_{p-1,t-n-p+1}] & \vdots\, \mathbf{G}_1[c_{p,t-n-p}] \end{pmatrix}. \tag{4.8}$$

Note that in (4.8) the number of rows of zero elements in each block of columns grows from zero to $p-2$ while the number of columns decreases from $p-1$ to one. Moreover, for $i > n+1$,

$$\mathbf{C}_i = \begin{pmatrix} \mathbf{G}_{p-1}[c_{2,t-i-1}] & \vdots\, \mathbf{0}'_{p-2} & \vdots \cdots \vdots & \mathbf{0}_{(p-3)\times 2} & \vdots\, \mathbf{0}_{p-2} \\ & \vdots\, \mathbf{G}_{p-2}[c_{3,t-i-2}] & \vdots \cdots \vdots & \mathbf{G}_2[c_{p-1,t-i-p+2}] & \vdots\, \mathbf{G}_1[c_{p,t-i-p+1}] \\ \cdots\cdots\cdots & \cdots\cdots\cdots & \cdots & \cdots\cdots\cdots & \cdots\cdots\cdots \\ \mathbf{I}_{p-2} \vdots \mathbf{0}_{p-2} & \vdots\, \mathbf{0}_{(p-2)\times(p-2)} & \vdots \cdots \vdots & 0 & 0\,0 \\ \mathbf{0}_{(p-3)\times(p-1)} & \vdots\, \mathbf{I}_{p-3} \vdots \mathbf{0}_{p-3} & \vdots \quad \vdots & \vdots & \vdots\,\vdots \\ \vdots & \vdots\,\vdots & \vdots \ddots \vdots & 0 & \vdots\,\vdots \\ \mathbf{0}'_{p-1} & \vdots\, \mathbf{0}'_{p-2} & \vdots \quad \vdots & 1 & 0\,0 \end{pmatrix} \tag{4.9}$$

which is a $p^* \times p^*$ matrix.

When $n > p$, we have to redefine some of the notation as follows. Let $\mathbf{c}^*_{p+1} = (\alpha_0, c_{1,t-1}, \ldots, c_{p,t-p})'$ be a $(p+1)\times 1$ vector. We define the $(p+1)\times(p+1)$ matrices $\mathbf{C}^*_1 = \mathbf{I}_{p+1}$ and

$$\mathbf{C}^*_j = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_0 & c_{1,t-j} & c_{2,t-j-1} & \cdots & c_{p-1,t-j-p+2} & c_{p,t-j-p+1} \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{4.10}$$

for $j = 2, \ldots, n$. Furthermore, using (4.7), we define the $(p+1)\times p^*$ matrix $\mathbf{C}^*_{n+1}$ as

$$\mathbf{C}^*_{n+1} = \begin{pmatrix} \mathbf{0}_{2\times(p-1)} & \vdots\, \mathbf{0}_{3\times(p-2)} & \vdots \cdots \vdots & \mathbf{0}_{(p-1)\times 2} & \vdots\, \mathbf{0}_{p\times 1} \\ \mathbf{G}_{p-1}[c_{2,t-n-2}] & \vdots\, \mathbf{G}_{p-2}[c_{3,t-n-3}] & \vdots \cdots \vdots & \mathbf{G}_2[c_{p-1,t-n-p+1}] & \vdots\, \mathbf{G}_1[c_{p,t-n-p}] \end{pmatrix}.$$

(4.11)

Finally, $\mathbf{C}_i^* = \mathbf{C}_i$ for $i > n+1$.

Now we are ready to consider $h_t h_{t-n}$, $n \geq 1$. We start a recursion by applying (4.3) to h_t. This yields an expression which is dependent on h_{t-1} but no longer on h_t. Applying (4.3) to h_{t-1} completes the next recursion. After the kth recursion, $k \geq n+1$, we have

Lemma 4.1 *Let $n \geq 1$. For $k \geq n+1$, $h_t h_{t-n}$ can be expressed by combinations of the terms of h_{t-i}, h_{t-j}^2 and $h_{t-i}h_{t-j}$ such that*

$$h_t h_{t-n} = S_{10} + \sum_{i=n+1}^{k-1} S_{1i} + S_{20} + \sum_{i=n+1}^{k-1} S_{2i} + S_k \tag{4.12}$$

where, for $1 \leq n \leq p$,

$$\begin{aligned}
S_{10} &= \alpha_0 \left[\left(1 + \mathbf{c}'_{P\setminus\{n\}} \sum_{i=1}^{n-1} \left(\prod_{j=1}^{i} \mathbf{C}_j \right) \mathbf{e}_1 \right) h_{t-n} + \mathbf{c}'_{P\setminus\{n\}} \left(\prod_{i=1}^{n} \mathbf{C}_i \right) \mathbf{h}_{10t} \right] \\
S_{1i} &= \alpha_0 \mathbf{c}'_{P\setminus\{n\}} \left(\prod_{j=1}^{i} \mathbf{C}_j \right) \mathbf{h}_{1it} \\
S_{20} &= \left(c_{n,t-n} + \mathbf{c}'_{P\setminus\{n\}} \sum_{i=1}^{n-1} \left(\prod_{j=1}^{i} \mathbf{C}_j \right) \mathbf{e}_1 c_{n-i,t-n} \right) h_{t-n}^2 \\
&\quad + \mathbf{c}'_{P\setminus\{n\}} \left(\prod_{i=1}^{n} \mathbf{C}_i \right) \mathbf{h}_{20t} \\
S_{2i} &= \mathbf{c}'_{P\setminus\{n\}} \left(\prod_{j=1}^{i} \mathbf{C}_j \right) \mathbf{h}_{2it} \\
S_k &= \mathbf{c}'_{P\setminus\{n\}} \left(\prod_{j=1}^{k} \mathbf{C}_j \right) \mathbf{h}_{kt}
\end{aligned} \tag{4.13}$$

and for $n > p$,

$$\begin{aligned}
S_{10} &= \mathbf{c}^{*\prime}_{p+1} \left(\prod_{i=1}^{n} \mathbf{C}_i^* \right) \mathbf{h}_{10t}^* \\
S_{1i} &= \alpha_0 \mathbf{c}^{*\prime}_{p+1} \left(\prod_{j=1}^{i} \mathbf{C}_j^* \right) \mathbf{h}_{1it} \\
S_{20} &= \mathbf{c}^{*\prime}_{p+1} \left(\prod_{i=1}^{n} \mathbf{C}_i^* \right) \mathbf{h}_{20t}^* \\
S_{2i} &= \mathbf{c}^{*\prime}_{p+1} \left(\prod_{j=1}^{i} \mathbf{C}_j^* \right) \mathbf{h}_{2it} \\
S_k &= \mathbf{c}^{*\prime}_{p+1} \left(\prod_{j=1}^{k} \mathbf{C}_j^* \right) \mathbf{h}_{kt}.
\end{aligned} \tag{4.14}$$

In (4.13), $\mathbf{h}_{10t}$ *and* $\mathbf{h}_{20t}$ *are* $(p-1)\times 1$ *vectors given by*

$$\begin{aligned}\mathbf{h}_{10t} &= (h_{t-(n+1)},...,h_{t-(n+p-1)})' \\ \mathbf{h}_{20t} &= (c_{1,t-(n+1)}h^2_{t-(n+1)},...,c_{p-1,t-(n+p-1)}h^2_{t-(n+p-1)})'\end{aligned} \tag{4.15}$$

whereas in (4.14), $\mathbf{h}^*_{10t}$ *and* $\mathbf{h}^*_{20t}$ *are* $(p+1)\times 1$ *vectors given by*

$$\begin{aligned}\mathbf{h}^*_{10t} &= (h_{t-n},0,\alpha_0 h_{t-(n+1)},...,\alpha_0 h_{t-(n+p-1)})' \\ \mathbf{h}^*_{20t} &= \left(0,h^2_{t-n},c_{1,t-(n+1)}h^2_{t-(n+1)},...,c_{p-1,t-(n+p-1)}h^2_{t-(n+p-1)}\right)'.\end{aligned} \tag{4.16}$$

Finally, $\mathbf{h}_{1it}$, $\mathbf{h}_{2it}$ *and* $\mathbf{h}_{kt}$ *are* $p^*\times 1$ *vectors given by*

$$\begin{aligned}\mathbf{h}_{1it} &= (h_{t-i-1},...,h_{t-(i+p-1)},0,...,0)' \\ \mathbf{h}_{2it} &= (c_{1,t-i-1}h^2_{t-i-1},...,c_{p-1,t-(i+p-1)}h^2_{t-(i+p-1)},0,...,0)' \\ \mathbf{h}_{kt} &= (h_{t-k}h_{t-(k+1)},\cdots,h_{t-k}h_{t-(k+p-1)},h_{t-(k+1)}h_{t-(k+2)}, \\ &\quad \cdots,h_{t-(k+1)}h_{t-(k+p-1)},\cdots,h_{t-(k+p-2)}h_{t-(k+p-1)})'.\end{aligned} \tag{4.17}$$

Proof. See Appendix C1.

Continuing the recursion it follows from Lemma 4.1 that $h_t h_{t-n}$, $1\le n\le p$, is characterized by matrices $\mathbf{C}_1$, $\mathbf{C}_2$, ..., $\mathbf{C}_n$, $\mathbf{C}_{n+1}$ and, finally, $\mathbf{C}_k$ for $k>n+1$. S_{10} and S_{20} are initial terms that are only functions of $\mathbf{C}_1$, $\mathbf{C}_2$, ..., $\mathbf{C}_n$. S_{1i}, S_{2i} and S_k also depend on $\mathbf{C}_k$ when $k\ge n+1$. Moreover, $h_t h_{t-n}$, $n>p$, is characterized by matrices $\mathbf{C}^*_1$, $\mathbf{C}^*_2$, ..., $\mathbf{C}^*_n$, $\mathbf{C}^*_{n+1}$ and $\mathbf{C}^*_k$ for $k>n+1$. These results are what we need for handling the mixed moments in (4.5).

4.2.2 The mixed moment $E(c_{l,t-l}c_{m,t-m}h_{t-l}h_{t-m})$

Let $1\le l<m\le p$ and consider $\mathsf{E}(c_{l,t-l}c_{t-m}h_{t-l}h_{t-m})$, $l<m$. Substituting $t-l$ for t and $m-l$ for n in Lemma 4.1, we obtain a recursion formula for $h_{t-l}h_{t-m}$. Let $\mathsf{E}\mathbf{c}_P=\boldsymbol{\gamma}_P=(\gamma_{11},...,\gamma_{p1})'$, $\mathsf{E}\mathbf{c}_{P\setminus\{m-l\}}=\boldsymbol{\gamma}_{P\setminus\{m-l\}}=(\gamma_{11},...,\gamma_{m-l-1},\gamma_{m-l+1},...,\gamma_{p1})'$, $\mathsf{E}\mathbf{C}_k=\boldsymbol{\Gamma}_k$ for $k=1,2,...,m-l+1$ and $\mathsf{E}\mathbf{C}_k=\boldsymbol{\Gamma}$ for $k>m-l+1$. Thus, $\boldsymbol{\Gamma}_k, k=1,2,...,m-l$, are matrices of order $(p-1)\times(p-1)$ and, in particular, $\boldsymbol{\Gamma}_1=\mathbf{I}_{p-1}$. Furthermore, $\boldsymbol{\Gamma}_{m-l+1}$ is of order $(p-1)\times p^*$, and $\boldsymbol{\Gamma}$ is a $p^*\times p^*$ matrix. The elements of those matrices are functions of $\gamma_{i1}=\mathsf{E}c_{it}$, $i=1,...,p$. In addition, let

$$\lambda(\boldsymbol{\Gamma})=\max\{|\lambda_i|\} \tag{4.18}$$

be the maximum absolute eigenvalue of the matrix $\boldsymbol{\Gamma}$ where λ_i, $i=1,...,p^*$, are the eigenvalues of matrix $\boldsymbol{\Gamma}$.

Applying Lemma 4.1 we have that

$$\begin{aligned}
& \mathsf{E}(c_{l,t-l}c_{t-m}h_{t-l}h_{t-m}) \\
= \; & \mathsf{E}(c_{l,t-l}c_{m,t-m}S_{10}) + \mathsf{E}(c_{l,t-l}c_{m,t-m}S_{20}) \\
& + \sum_{i=m-l+1}^{k-1} \mathsf{E}(c_{l,t-l}c_{m,t-m}S_{1i}) + \sum_{i=m-l+1}^{k-1} \mathsf{E}(c_{l,t-l}c_{m,t-m}S_{2i}) \\
& + \mathsf{E}(c_{l,t-l}c_{m,t-m}S_k). \qquad (4.19)
\end{aligned}$$

Assume that the process started at some finite value infinitely many periods ago. It turns out that the limit of (4.19) exists and is independent of t as $k \to \infty$ if and only if all the eigenvalues of $\mathbf{\Gamma}$ lie inside the unit circle, that is, $\lambda(\mathbf{\Gamma})<1$. We show this and derive the mixed moment $\mathsf{E}(c_{l,t-l}c_{t-m}h_{t-l}h_{t-m})$ as a function of $\mathsf{E}h_t$ and $\mathsf{E}h_t^2$ by proving the next three lemmata.

Lemma 4.2 *For* $1 \le l < m \le p$,

$$\lim_{k\to\infty} \mathsf{E}(c_{l,t-l}c_{m,t-m}S_k) = 0 \qquad (4.20)$$

if and only if $\lambda(\mathbf{\Gamma}) < 1$.

Proof. See Appendix C2.

Lemma 4.3 *For* $1 \le l < m \le p$,

$$\begin{aligned}
& \mathsf{E}\left(c_{l,t-l}c_{m,t-m} \sum_{i=m-l+1}^{\infty} S_{1i}\right) \\
= \; & \alpha_0\gamma_{l1}\gamma_{m1}\gamma'_{P\setminus\{m-l\}} \left(\prod_{i=1}^{m-l+1} \mathbf{\Gamma}_i\right) (\mathbf{I}_{p^*} - \mathbf{\Gamma})^{-1} \mathbf{e}_{p-1}\mathsf{E}h_t \qquad (4.21)
\end{aligned}$$

if and only if $\lambda(\mathbf{\Gamma}) < 1$, *where* $\mathbf{e}_{p-1} = (1, \cdots, 1, 0, \cdots, 0)'$ *is a* $p^* \times 1$ *vector with the first* $p-1$ *elements equal to* 1.

Proof. See Appendix C3 .

Evaluating $\mathsf{E}\left(c_{l,t-l}c_{m,t-m} \sum_{i=m-l+1}^{\infty} S_{2i}\right)$ is not an easy task because $\mathbf{h}_{2it}$ is not stochastically independent of the matrix product $\mathbf{C}_{i-(p-2)} \cdots \mathbf{C}_i$. However, for each nonzero element of $\mathbf{h}_{2it}$, $c_{j,t-(i+j)}$ and $h^2_{t-(i+j)}$ are stochastically independent. For $j \le p-2$ and $i > m-l+2p-1$, this allows us to consider

$$\gamma(i-j, i) = \mathsf{E}(\mathbf{C}_{i-j}\mathbf{C}_{i-(j+1)} \cdots \mathbf{C}_i\mathbf{c}_{2it}) \qquad (4.22)$$

where $\mathbf{c}_{2it} = (c_{1,t-(i+l)}, ..., c_{p-1,t-(i+l+p-2)}, 0, ..., 0)'$ is a $p^* \times 1$ vector whose first $p-1$ elements are nonzero. We see that $\gamma(i-j, i)$ is a $p^* \times 1$ column

vector which is determined by the matrix product $\mathbf{C}_{i-j}\cdots\mathbf{C}_i$ and whose elements consist of $\tilde{\gamma}_{ij} = \mathsf{E}(c_{it}c_{jt})$ for $i<j$. Furthermore, for $j=l<i\leq p$, define

$$\gamma(\mathbf{c}(m-l), i-l-1, i) = \mathsf{E}(\mathbf{c}'_{P\setminus\{m-l\}}\mathbf{C}_{i-(l+1)}\cdots\mathbf{C}_i\mathbf{c}_{2it}) \tag{4.23}$$

where $\mathbf{c}(x)$ is shorthand for $\mathbf{c}_{P\setminus\{x\}}$. Using (4.22) and (4.23), we have

Lemma 4.4 *For $1\leq l<m\leq p$,*

$$\begin{aligned}
&\mathsf{E}\left(c_{l,t-l}c_{m,t-m}\sum_{i=m-l+1}^{\infty}S_{2i}\right)\\
=\;&\gamma_{l1}\gamma_{m1}\sum_{i=1}^{p-1}\gamma(\mathbf{c}(m-l),2,m-l+i-1)\mathsf{E}h_t^2\\
&+\gamma_{l1}\gamma_{m1}\gamma'_{P\setminus\{m-l\}}\left(\prod_{i=1}^{m-l+1}\mathbf{\Gamma}_i\right)\\
&\times(\mathbf{I}_{p^*}-\mathbf{\Gamma})^{-1}\gamma(m-l+p+1,m-l+2p-1)\mathsf{E}h_t^2
\end{aligned}\tag{4.24}$$

if and only if $\lambda(\mathbf{\Gamma})<1$.

Proof. See Appendix C4.

It now follows from these three lemmata that if the condition $\lambda(\mathbf{\Gamma})<1$ holds the mixed moment $\mathsf{E}(c_{l,t-l}c_{t-m}h_{t-l}h_{t-m})$ converges to a finite value which is a linear function of $\mathsf{E}h_t$ and $\mathsf{E}h_t^2$ as $k\to\infty$. Combining these results and using (4.19) yields

Theorem 4.1 *Assume that $\lambda(\mathbf{\Gamma})<1$. Under this condition,*

$$\mathsf{E}(c_{l,t-l}c_{m,t-m}h_{t-l}h_{t-m}) = \alpha_0\gamma_{l1}\gamma_{m1}M_1(l,m)\mathsf{E}h_t + \gamma_{l1}M_2(l,m)\mathsf{E}h_t^2 \tag{4.25}$$

where, for $m-l>1$,

$$\begin{aligned}
&M_1(l,m)\\
=\;&1+\gamma'_{P\setminus\{m-l\}}\Big[\sum_{i=1}^{m-l-1}\Big(\prod_{j=1}^{i}\mathbf{\Gamma}_j\Big)\mathbf{e}_1\\
&+\prod_{i=1}^{m-l}\mathbf{\Gamma}_i(\mathbf{j}_{p-1}+\mathbf{\Gamma}_{m-l+1}(\mathbf{I}_{p^*}-\mathbf{\Gamma})^{-1}\mathbf{e}_{p-1})\Big]
\end{aligned}\tag{4.26}$$

and, in particular,

$$M_1(m-1,m) = 1+\gamma'_{P\setminus\{1\}}\left[\mathbf{j}_{p-1}+\mathbf{\Gamma}_2(\mathbf{I}_{p^*}-\mathbf{\Gamma})^{-1}\mathbf{e}_{p-1}\right]$$

where $\mathbf{j}_{p-1}=(1,1,...,1)'$ is a $(p-1)\times 1$ vector. Furthermore,

$$M_2(l,m) = M_{21}(l,m)+\gamma_{m1}\sum_{i=2}^{4}M_{2i}(l,m) \tag{4.27}$$

such that for $m-l>1$

$$M_{21}(l,m) = \widetilde{\gamma}_{m-l,m}+\boldsymbol{\gamma}'_{P\setminus\{m-l\}}\left[\sum_{i=1}^{m-l-1}\left(\prod_{j=1}^{i}\boldsymbol{\Gamma}_j\right)\mathbf{e}_1\widetilde{\gamma}_{m-l-i,m}\right] \quad (4.28)$$

$$M_{22}(l,m) = \sum_{i=1}^{m-l-1}\gamma_{i1}M_{22}(m-l-i)+\sum_{j=m-l+1}^{p}\widetilde{\gamma}_{j-m+l,j} \quad (4.29)$$

and, for $m-l=1$, $M_{21}(m-1,m)=\widetilde{\gamma}_{1m}$ *and* $M_{22}(m-1,m)=\sum_{j=2}^{p}\widetilde{\gamma}_{j-1,j}$. *In addition, in (4.27), for any* $l<m$,

$$M_{23}(l,m) = \sum_{i=m-l+1}^{m-l+p-1}\gamma(\mathbf{c}(m-l),2,i-1) \quad (4.30)$$

$$M_{24}(l,m) = \boldsymbol{\gamma}'_{P\setminus\{m-l\}}\left(\prod_{j=1}^{m-l+1}\boldsymbol{\Gamma}_j\right)(\mathbf{I}_{p^*}-\boldsymbol{\Gamma})^{-1}\boldsymbol{\gamma}(m-l+p+1,m-l+2p-1). \quad (4.31)$$

Proof. See Appendix C5.

4.2.3 A necessary and sufficient condition for existence of the fourth moment

Theorem 4.1 allows us to express the expectation (4.5) as a function of $\mathsf{E}h_t$ and $\mathsf{E}h_t^2$. We can write

$$\mathsf{E}h_t^2 = \alpha_0^2+2\alpha_0\left(\gamma_1+\sum_{l<m}\gamma_{l1}\gamma_{m1}M_1(l,m)\right)\mathsf{E}h_t + \left(\gamma_2+2\sum_{l<m}\gamma_{l1}M_2(l,m)\right)\mathsf{E}h_t^2. \quad (4.32)$$

From (4.32) and the assumption $\lambda(\boldsymbol{\Gamma})<1$ we have

Theorem 4.2 *A necessary and sufficient condition for the existence of the fourth unconditional moment of* $\{\varepsilon_t\}$ *in the GARCH(p,p) model (4.1) with (4.2) is*

$$\gamma_2+2\sum_{l<m}\gamma_{l1}M_2(l,m)<1. \quad (4.33)$$

The fourth moment is

$$\mathsf{E}(\varepsilon_t^4)=\frac{\alpha_0^2\nu_4[1+\gamma_1+2\sum_{l<m}\gamma_{l1}\gamma_{m1}M_1(l,m)]}{(1-\gamma_1)[1-\gamma_2-2\sum_{l<m}\gamma_{l1}M_2(l,m)]}. \quad (4.34)$$

4.2.4 Special case: GARCH(2,2)

To illustrate the above general theory we consider the GARCH(2,2) process. We have

Corollary 4.1 *For the GARCH$(2,2)$ model,*

$$\mathsf{E}\left(c_{1,t-1}c_{2,t-2}h_{t-1}h_{t-2}\right) = \alpha_0\gamma_{11}\gamma_{21}M_1\left(1,2\right)\mathsf{E}h_t + \gamma_{11}M_2(1,2)\mathsf{E}h_t^2 \quad (4.35)$$

where $M_1(1,2) = (1-\gamma_{21})^{-1}$ and $M_2\left(1,2\right) = \widetilde{\gamma}_{12}(1-\gamma_{21})^{-1}$ if and only if $\gamma_{21} < 1$.

Proof. See Appendix C6.

Corollary 4.2 *For the GARCH$(2,2)$ model a necessary and sufficient condition for the existence of the unconditional fourth moment of $\{\varepsilon_t\}$ is that*

$$\left(\gamma_{12}+\gamma_{22}\right)\left(1-\gamma_{21}\right) + 2\gamma_{11}\widetilde{\gamma}_{12} + \gamma_{21} < 1. \quad (4.36)$$

When (4.36) holds,

$$\mathsf{E}(\varepsilon_t^4) = \frac{\alpha_0^2\nu_4\left[\left(1+\gamma_{11}+\gamma_{21}\right)\left(1-\gamma_{21}\right)+2\gamma_{11}\gamma_{21}\right]}{\left(1-\gamma_{11}-\gamma_{21}\right)\left[\left(1-\gamma_{21}\right)\left(1-\gamma_{12}-\gamma_{22}\right)-2\gamma_{11}\widetilde{\gamma}_{12}\right]}. \quad (4.37)$$

If $\beta_2 = \alpha_2 = 0$, we have the GARCH(1,1) model, and (4.36) is simply

$$\gamma_{12} = \beta_1^2 + 2\beta_1\alpha_1\nu_2 + \alpha_1^2\nu_4 < 1. \quad (4.38)$$

This is the existence condition in Teräsvirta (1996). Setting $\nu_2 = 1$ and $\nu_4 = 3$ in (4.38) yields the condition which Bollerslev (1986) obtained for normal errors.

4.3 The autocorrelation function for the squared process

Milhϕj (1985) proved that the squared process $\{\varepsilon_t^2\}$ has an autocorrelation function similar to that of a standard autoregressive process of order p when $\{\varepsilon_t\}$ follows an ARCH(p) process with normal errors. He also defined the first p autocorrelations of $\{\varepsilon_t^2\}$ as

$$\boldsymbol{\rho} = \left(\mathbf{I}_p - \boldsymbol{\Psi}\right)^{-1}\boldsymbol{\alpha} \quad (4.39)$$

where $\boldsymbol{\rho} = (\rho_1, ..., \rho_p)'$, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$ and the $p \times p$ matrix $\boldsymbol{\Psi} = (\psi_{ij})$ is defined by $\psi_{ij} = \alpha_{i+j} + \alpha_{i-j}$ making use of the fact that $\alpha_k = 0$ for $k \leq 0$ and $k > p$. For $n > p$,

$$\rho_n = \sum_{i=1}^{p} \alpha_i\rho_{n-i}. \quad (4.40)$$

Bollerslev (1986,1988) applied this approach to the GARCH(p,p) case and obtained the analogue of the Yule-Walker equations for the autocorrelation function of $\{\varepsilon_t^2\}$. He was able to show that for any $n > p$, the autocorrelations have the form

$$\rho_n = \sum_{i=1}^{p} (\beta_i + \alpha_i)\, \rho_{n-i}. \tag{4.41}$$

However, he did not yet provide the first p autocorrelations $\rho_1, ..., \rho_p$. We shall do that here.

To fix notation, write the nth order autocorrelation of $\{\varepsilon_t^2\}$ as

$$\rho_n = \rho(\varepsilon_t^2, \varepsilon_{t-n}^2) = \frac{\mathsf{E}(\varepsilon_t^2 \varepsilon_{t-n}^2) - (\mathsf{E}\varepsilon_t^2)^2}{\mathsf{E}(\varepsilon_t^4) - (\mathsf{E}\varepsilon_t^2)^2} \tag{4.42}$$

for $n \geq 1$. In order to obtain ρ_n, we must find an expression for $\mathsf{E}(\varepsilon_t^2 \varepsilon_{t-n}^2)$ as a function of $\mathsf{E}\varepsilon_t^4$ and $\mathsf{E}\varepsilon_t^2$. This can be done by applying results in Section 4.2.

4.3.1 The mixed moment $\mathsf{E}(\varepsilon_t^2 \varepsilon_{t-n}^2)$

It is obvious that combining (4.1) and Lemma 1 will give us the mixed moment $\mathsf{E}(\varepsilon_t^2 \varepsilon_{t-n}^2)$, $n \geq 1$. Consider first the case $1 \leq n \leq p$. Let $\overline{\gamma}_{i1} = \mathsf{E}(z_t^2 c_{it})$ for $i = 1, ..., p$. Setting $M_1(n) = M_1(0,n)$ and $M_2(n) = M_2(0,n)$, it follows directly from Theorem 4.1 that

$$\mathsf{E}(\varepsilon_t^2 \varepsilon_{t-n}^2) = \alpha_0 \nu_2^2 M_1(n)\, \mathsf{E}h_t + \nu_2 M_2(n)\, \mathsf{E}h_t^2. \tag{4.43}$$

As in (4.27), let

$$M_2(n) = M_{21}(n) + \nu_2 \sum_{i=2}^{4} M_{2i}(n) \tag{4.44}$$

such that for $n = 2, ..., p$,

$$M_{21}(n) = \overline{\gamma}_{n1} + \gamma'_{P\setminus\{n\}} \left[\sum_{i=1}^{n-1} \left(\prod_{j=1}^{i} \mathbf{\Gamma}_j \right) \mathbf{e}_1 \overline{\gamma}_{n-i,1} \right] \tag{4.45}$$

$$M_{22}(n) = \sum_{i=1}^{n-1} \gamma_{i1} M_{22}(n-i) + \sum_{j=n+1}^{p} \widetilde{\gamma}_{j-n,j}. \tag{4.46}$$

Furthermore, for $n = 1$, $M_{21}(1) = \overline{\gamma}_{11}$ and $M_{22}(1) = \sum_{j=2}^{p} \widetilde{\gamma}_{j-1,j}$. In addition, for $n = 1, ..., p$,

$$M_{23}(n) = \sum_{i=n+1}^{n+p-1} \gamma(\mathbf{c}(n), 2, i-1) \tag{4.47}$$

and

$$M_{24}(n) = \gamma'_{P\backslash\{n\}} \left[\left(\prod_{j=1}^{n+1} \mathbf{\Gamma}_j \right) (\mathbf{I}_{p*} - \mathbf{\Gamma})^{-1} \right] \gamma (n+p+1, n+2p-1). \tag{4.48}$$

Next, let $n > p$. Under this assumption it can be shown that (4.43) still holds if we define the coefficients of $\mathsf{E}h_t$ and $\mathsf{E}h_t^2$ as functions of n in a proper way. In order to do that, first note an analogy to the situation in Section 4.2. Thus, while $c_{i,t-(n+j)}$ and $h^2_{t-(n+j)}$ are stochastically independent, the product $\mathbf{C}^*_{n-(p-1)}\mathbf{C}^*_{n-(p-2)} \cdots \mathbf{C}^*_n$ is not stochastically independent of $\mathbf{h}^*_{20t}$. This is because there is an element such as $c_{i,t-n}$ which is a function of z^2_{t-n} for some n on the second row of each matrix, say, $\mathbf{C}^*_{n-j}$, $j = 0, 1, ..., p-1$. As in Section 4.2, define the $(p+1) \times 1$ vector $\mathbf{c}^*_{20t} = \left(0, 1, c_{1,t-(n+1)}, ..., c_{p-1,t-(n+p-1)}\right)'$. We now have to consider the expectation $\mathsf{E}\left(z^2_{t-n}\mathbf{C}^*_{n-(p-1)}\mathbf{C}^*_{n-(p-2)} \cdots \mathbf{C}^*_n \mathbf{c}^*_{20t}\right)$, $n > p$, which is given in the following lemma.

Lemma 4.5 *For $n > p$,*

$$\begin{aligned} & \mathsf{E}\left(z^2_{t-n}\mathbf{C}^*_{n-(p-1)}\mathbf{C}^*_{n-(p-2)} \cdots \mathbf{C}^*_n \mathbf{c}^*_{20t}\right) \\ = \; & (0, M_{21}(p) + \nu_2 M_{22}(p), M_{21}(p-1) \\ & + \nu_2 M_{22}(p-1), \cdots, M_{21}(1) + \nu_2 M_{22}(1))'. \end{aligned} \tag{4.49}$$

Proof. See Appendix C7.

Lemma 4.5 implies that for $n > p$, $\mathsf{E}\left(z^2_{t-n}\mathbf{C}^*_{n-(p-1)}\mathbf{C}^*_{n-(p-2)} \cdots \mathbf{C}^*_n \mathbf{h}^*_{20t}\right)$ does not depend on n.

Let $\mathsf{E}\mathbf{c}^*_{p+1} = \gamma_{p+1} = (\alpha_0, \gamma_{11}, ..., \gamma_{p1})'$ be a $(p+1) \times 1$ vector, $\mathsf{E}\mathbf{C}^*_i = \mathbf{\Gamma}_*$ for $i = 2, ..., n$, $\mathsf{E}\mathbf{C}^*_{n+1} = \mathbf{\Gamma}^*_{n+1}$ and $\mathsf{E}\mathbf{C}^*_i = \mathbf{\Gamma}$ for $i > n+1$. We observe that here $\mathsf{E}\mathbf{C}^*_i$ $(i = 2, ..., n+1)$ are quite different from $\mathsf{E}\mathbf{C}_j$ $(j = 2, ..., m-l+1)$ defined in Section 4.2. However, when $i > n+1$ and $j > m-l+1$, $\mathsf{E}\mathbf{C}^*_i = \mathsf{E}\mathbf{C}_j = \mathbf{\Gamma}$. Thus the condition $\lambda(\mathbf{\Gamma}) < 1$ is required for expressing $\mathsf{E}(\varepsilon^2_t \varepsilon^2_{t-n})$ as a function of $\mathsf{E}\varepsilon^4_t$ and $\mathsf{E}\varepsilon^2_t$.

Moreover, results of Lemma 4.5 make it possible to prove that $M_2(n)$ can be expressed as a function of $M_2(1), ..., M_2(p)$ for $n \geq p+1$. We have

Lemma 4.6 *Define the mixed moment $\mathsf{E}\left(u^2_i u^2_{t-n}\right)$ as in (4.43) and let $n > p$. Then, for $n \geq p+1$,*

$$M_2(n) = \gamma'_{p+1}\mathbf{\Gamma}^{n-(p+1)}_* \mathbf{m}_2 \tag{4.50}$$

where $\mathbf{m}_2 = (0, M_2(p), ..., M_2(1))'$ is a $(p+1) \times 1$ vector and $M_2(i)$, $i = 1, ..., p$, are given by (4.44)-(4.48).

Proof. See Appendix C8.

It now follows from Lemma 4.1 that

$$\begin{aligned}\mathsf{E}(\varepsilon_t^2\varepsilon_{t-n}^2) &= \mathsf{E}\left(z_t^2 z_{t-n}^2 S_{10}\right)+\mathsf{E}\left(z_t^2 z_{t-n}^2 S_{20}\right)+\sum_{i=n+1}^{k-1}\mathsf{E}\left(z_t^2 z_{t-n}^2 S_{1i}\right)\\ &\quad+\sum_{i=n+1}^{k-1}\mathsf{E}\left(z_t^2 z_{t-n}^2 S_{2i}\right)+\mathsf{E}\left(z_t^2 z_{t-n}^2 S_k\right). \qquad (4.51)\end{aligned}$$

Assume again that the process started infinitely far in the past at some finite value. Then, given the previous results, the limit of (4.51) as $k \to \infty$ exists and is independent of t if and only if all the eigenvalues of Γ lie inside the unit circle. This is obvious because it was seen that $\mathsf{E}\varepsilon_t^4 < \infty$ requires $\lambda(\Gamma) < 1$. As $k \to \infty$, the last three terms on the right-hand side of (4.51) can be evaluated by Lemmata 4.2 to 4.4. As for the first two terms, when $n \leq p$ we may apply Theorem 4.1, otherwise we apply Lemmata 4.5 and 4.6. The next theorem gives the moments.

Theorem 4.3 *The mixed moment* $\mathsf{E}(\varepsilon_t^2\varepsilon_{t-n}^2)$ *has the form*

$$\mathsf{E}(\varepsilon_t^2\varepsilon_{t-n}^2) = \alpha_0\nu_2^2 M_1(n)\,\mathsf{E}h_t + \nu_2 M_2(n)\,\mathsf{E}h_t^2 \qquad (4.52)$$

where for $n \geq 1$,

$$M_1(n) = \gamma_{p+1}'\Gamma_*^{n-1}\left[\mathbf{e}_{\alpha_0} + \Gamma_{n+1}^*\left(\mathbf{I}_{p^*}-\Gamma\right)^{-1}\mathbf{e}_{p-1}\right] \qquad (4.53)$$

with $\mathbf{e}_{\alpha_0} = \left(\alpha_0^{-1}, 0, 1, ..., 1\right)'$ *is a* $(p+1)\times 1$ *vector.* $M_2(n)$ *in (4.52) is defined by (4.44)-(4.48) for* $n \leq p$*; otherwise* $M_2(n)$ *is given by (4.50).*

Proof. See Appendix C9.

4.3.2 The autocorrelation function for the squared process

Next we derive the autocorrelation function of $\{\varepsilon_t^2\}$ and begin by introducing some notation. Let $\gamma_{M_1}(l,m) = (1-\Delta)M_1(l,m)$ and $\gamma_{M_2}(l,m) = (1-\Delta)M_2(l,m)$, where $1-\Delta = |\mathbf{I}_{p^*}-\Gamma|$ is the determinant of $(\mathbf{I}_{p^*}-\Gamma)$. Furthermore, let

$$\begin{aligned}\gamma_{S_1} &= (1+\gamma_1)(1-\Delta) + 2\sum_{l<m}\gamma_{l1}\gamma_{m1}\gamma_{M_1}(l,m) \qquad (4.54)\\ \gamma_{S_2} &= (1-\gamma_2)(1-\Delta) - 2\sum_{l<m}\gamma_{l1}\gamma_{M_2}(l,m). \qquad (4.55)\end{aligned}$$

A straightforward calculation shows that condition (4.33) is equivalent to

$$0 < \gamma_{S_2} < 1. \qquad (4.56)$$

Applying Theorem 4.2 and Theorem 4.3 to equation (4.42) gives

Theorem 4.4 *Assume that condition (4.56) holds. For the GARCH(p,p) model (4.1) with (4.3), the autocorrelation function for $\{\varepsilon_t^2\}$ is, for any $n \geq 1$,*

$$\rho_n = \frac{\nu_2 \gamma_{S_1} (1-\gamma_1) M_2(n) - \nu_2^2 \gamma_{S_2} [1-(1-\gamma_1) M_1(n)]}{\nu_4 \gamma_{S_1} (1-\gamma_1) - \nu_2^2 \gamma_{S_2}} \tag{4.57}$$

where $M_1(n)$ and $M_2(n)$ are defined in Theorem 4.3 and $M_1(l,m)$ and $M_2(l,m)$ as in Theorem 4.1.

Properties of the autocorrelation function $\{\rho_n\}$, $n = 1, 2, \cdots$, can be established through (4.57). Some of them are listed below:

1. The first p autocorrelations are positive if the parameter restrictions $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_i \geq 0$, $i = 1, \cdots, p$, hold.

2. The autocorrelation function satisfies the difference equation $\rho_n = \sum_{i=1}^{p} \gamma_{i1} \rho_{n-i}$ at lags $n > p$, and $\rho_n > 0$.

3. The autocorrelation function is dominated by an exponential decay and $\lim_{n\to\infty} \rho_n = 0$.

4. When γ_{S_2} is sufficiently close to zero the autocorrelation function is persistent; otherwise the autocorrelations decay quickly with increasing n.

4.3.3 Special case: GARCH(2, 2)

To illustrate the general result we again consider the GARCH(2,2) process. In the GARCH(2,2) model,

$$\begin{aligned} \gamma_i &= \gamma_{1i} + \gamma_{2i}, \ i = 1, 2 \\ \gamma_{S_1} &= (1+\gamma_1)(1-\gamma_{21}) + 2\gamma_{11}\gamma_{21} \\ \gamma_{S_2} &= (1-\gamma_2)(1-\gamma_{21}) - 2\gamma_{11}\widetilde{\gamma}_{12}. \end{aligned}$$

By Theorem 4.4 we have

Corollary 4.3 *Assume that condition (4.36) holds. For the GARCH$(2,2)$ process, the autocorrelations of $\{\varepsilon_t^2\}$ equal*

$$\rho_n = \frac{\nu_2 \gamma_{S_1} (1-\gamma_1) M_2(n) - \nu_2^2 \gamma_{S_2} [1-(1-\gamma_1) M_1(n)]}{\nu_4 \gamma_{S_1} (1-\gamma_1) - \nu_2^2 \gamma_{S_2}}, \ n \geq 1. \tag{4.58}$$

In (4.58), for $n \geq 1$

$$M_1(n) = \boldsymbol{\gamma}_3' \boldsymbol{\Gamma}_*^{n-1} \begin{pmatrix} \alpha_0^{-1} \\ 0 \\ (1-\gamma_{21})^{-1} \end{pmatrix} \tag{4.59}$$

and for $n \geq 3$

$$M_2(n) = \boldsymbol{\gamma}_3' \boldsymbol{\Gamma}_*^{n-3} \begin{pmatrix} 0 \\ M_2(2) \\ M_2(1) \end{pmatrix} \tag{4.60}$$

with $\boldsymbol{\gamma}_3 = (\ \alpha_0 \quad \gamma_{11} \quad \gamma_{21}\ )'$ *and* $\boldsymbol{\Gamma}_* = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_0 & \gamma_{11} & \gamma_{21} \\ 0 & 1 & 0 \end{pmatrix}$, *whereas*

$$M_2(1) = \frac{1}{1-\gamma_{21}} \left[\overline{\gamma}_{11}(1-\gamma_{21}) + \nu_2 \widetilde{\gamma}_{12}\right], \tag{4.61}$$

$$M_2(2) = \frac{1}{1-\gamma_{21}} \left[(\gamma_{11}\overline{\gamma}_{11} + \overline{\gamma}_{21})(1-\gamma_{21}) + \nu_2 \gamma_{11} \widetilde{\gamma}_{12}\right]. \tag{4.62}$$

Proof. See Appendix C10.

From Corollary 4.3 it follows that the first two autocorrelations are

$$\rho_1 = \frac{\nu_2 \gamma_{S_1} \left[\overline{\gamma}_{11}(1-\gamma_{21}) + \nu_2 \widetilde{\gamma}_{12}\right](1-\gamma_1) - \nu_2^2 \gamma_{11} \gamma_{S_2}}{(1-\gamma_{21})\left[\nu_4 \gamma_{S_1}(1-\gamma_1) - \nu_2^2 \gamma_{S_2}\right]} \tag{4.63}$$

and

$$\rho_2 = \frac{\nu_2 \gamma_{S_1} \left[(\gamma_{11}\overline{\gamma}_{11} + \overline{\gamma}_{21})(1-\gamma_{21}) + \nu_2 \gamma_{11} \widetilde{\gamma}_{12}\right](1-\gamma_1) - \nu_2^2 \gamma_{S_2}\left(\gamma_{11}^2 + \gamma_{21} - \gamma_{21}^2\right)}{(1-\gamma_{21})\left[\nu_4 \gamma_{S_1}(1-\gamma_1) - \nu_2^2 \gamma_{S_2}\right]}. \tag{4.64}$$

For higher orders, ρ_n is determined by the powers of matrix $\boldsymbol{\Gamma}_*$ as well as $M_2(1)$ and $M_2(2)$ through equation (4.58). In other words, for $n \geq 3$, ρ_n depends on the first two autocorrelations, which is also clear from the corresponding Yule-Walker equations.

The corresponding results for GARCH(1,1) model are obtained by setting $\gamma_{21} = \widetilde{\gamma}_{12} = \overline{\gamma}_{21} = 0$ in (4.58). The autocorrelation function thus has the form

$$\rho_n = \frac{\nu_2\left(1-\gamma_{11}^2\right) M_2(n) - \nu_2^2(1-\gamma_{12})\left[1-(1-\gamma_{11}) M_1(n)\right]}{\nu_4\left(1-\gamma_{11}^2\right) - \nu_2^2(1-\gamma_{12})}, n \geq 1. \tag{4.65}$$

$M_1(n)$ and $M_2(n)$ can be determined as follows. We have $\mathbf{\Gamma}_* = \begin{pmatrix} 1 & 0 \\ \alpha_0 & \gamma_{11} \end{pmatrix}$. Thus, for $n \geq 1$,

$$\begin{aligned} M_1(n) &= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \mathbf{\Gamma}_*^{n-1} \begin{pmatrix} \alpha_0^{-1} \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \begin{pmatrix} \alpha_0^{-1} \\ \sum_{i=0}^{n-2} \left(\prod_{j=0}^{i} \gamma_{11}^{j} \right) \end{pmatrix} \\ &= \frac{1-\gamma_{11}^n}{1-\gamma_{11}}. \end{aligned} \tag{4.66}$$

Furthermore, for $n \geq 3$,

$$\begin{aligned} M_2(n) &= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \mathbf{\Gamma}_*^{n-3} \begin{pmatrix} 0 \\ \gamma_{11}\overline{\gamma}_{11} \end{pmatrix} \\ &= \begin{pmatrix} \alpha_0 & \gamma_{11} \end{pmatrix} \begin{pmatrix} 0 \\ \gamma_{11}^{n-2}\overline{\gamma}_{11} \end{pmatrix} \\ &= \gamma_{11}^{n-1}\overline{\gamma}_{11} \end{aligned} \tag{4.67}$$

with $M_2(1) = \overline{\gamma}_{11}$ and $M_2(2) = \gamma_{11}\overline{\gamma}_{11}$. In fact, $M_2(n) = \gamma_{11}^{n-1}\overline{\gamma}_{11}$ for $n \geq 1$. Inserting (4.66) and (4.67) into (4.65) yields

$$\begin{aligned} \rho_n &= \frac{\nu_2\gamma_{11}^{n-1}[\overline{\gamma}_{11}(1-\gamma_{11}^2) - \nu_2\gamma_{11}(1-\gamma_{12})]}{\nu_4(1-\gamma_{11}^2) - \nu_2^2(1-\gamma_{12})} \\ &= \frac{\alpha_1\nu_2\gamma_{11}^{n-1}(1-\beta_1^2-\alpha_1\beta_1\nu_2)}{1-\beta_1^2-2\alpha_1\beta_1\nu_2} \end{aligned} \tag{4.68}$$

see Teräsvirta (1996). Setting $\nu_2 = 1$ and $\nu_4 = 3$ in (4.68) yields the corresponding result for normal errors given in Bollerslev (1988).

4.3.4 The ARCH(p) process

As a further illustration, we reproduce the autocorrelation function of $\{\varepsilon_t^2\}$ in an ARCH(p) model and compare our result with that of Milhφj (1985); see (4.39) and (4.40). In an ARCH(p) model, $\beta_i = 0$ $i = 1, ..., p$. The autocorrelations thus have the form

$$\rho_n = \frac{\nu_2\gamma_{S_1}\left(1-\nu_2\sum_{i=1}^{p}\alpha_i\right)M_2(n) - \nu_2^2\gamma_{S_2}\left[1-\left(1-\nu_2\sum_{i=1}^{p}\alpha_i\right)M_1(n)\right]}{\nu_4\gamma_{S_1}\left(1-\nu_2\sum_{i=1}^{p}\alpha_i\right) - \nu_2^2\gamma_{S_2}} \tag{4.69}$$

for $n \geq 1$. Further simplification of (4.69) gives

Corollary 4.4 *For the ARCH(p) model, the autocorrelation function of $\{\varepsilon_t^2\}$ has the form*

$$\rho_n = 1 - \left(1 - \nu_2 \sum_{i=1}^{p} \alpha_i\right) M_1(n), \; n \geq 1 \tag{4.70}$$

where

$$M_1(n) = \gamma_{p+1}' \Gamma_*^{n-1} \left[\mathbf{e}_{\alpha_0} + \Gamma_{n+1}^* \left(\mathbf{I}_{p^*} - \Gamma\right)^{-1} \mathbf{e}_{p-1}\right]. \tag{4.71}$$

Proof. See Appendix C11.

If the errors are assumed normal, then (4.70) gives the corresponding result which is the one Milhϕj (1985) obtained. As an example, for the ARCH(2) model it follows from (4.70), or (4.39) and (4.40), that $\rho_1 = \alpha_1(1-\alpha_2)^{-1}$, $\rho_2 = \alpha_2 + \alpha_1^2(1-\alpha_2)^{-1}$, and $\rho_n = \alpha_1\rho_{n-1} + \alpha_2\rho_{n-2}$ for $n \geq 2$. Alternatively, a general expression for ρ_n, $n \geq 2$, is obtained from Corollary 4.4.

Milhϕj (1985) also found a necessary and sufficient condition for the existence of the unconditional fourth moment of the ARCH(p) process under normality. It has the form

$$3\boldsymbol{\alpha}' \left(\mathbf{I}_p - \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\alpha} < 1 \tag{4.72}$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$ and the $p \times p$ matrix $\boldsymbol{\Psi} = (\psi_{ij})$ is defined by $\psi_{ij} = \alpha_{i+j} + \alpha_{i-j}$ with $\alpha_k = 0$ for $k \leq 0$ and $k > p$. Of course, (4.33) under normality and (4.72) are equivalent, although this may not be easy to see immediately. As an illustration, we demonstrate this equivalence in the ARCH(2) case. Setting $\gamma_{i1} = \alpha_i, \gamma_{i2} = 3\alpha_i^2$, $i = 1, 2$, and $\widetilde{\gamma}_{ij} = 3\alpha_1\alpha_2$ in (4.36) yields

$$3\alpha_1^2 + 3\alpha_2^2 + 3\alpha_1^2\alpha_2 - 3\alpha_2^3 + \alpha_2 < 1. \tag{4.73}$$

On the other hand, setting $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ and $\boldsymbol{\Psi} = \begin{pmatrix} \alpha_2 & 0 \\ \alpha_1 & 0 \end{pmatrix}$ in (4.72) also yields (4.73).

4.4 The GARCH(p,q) model

The results in Theorems 4.1 to 4.4 apply directly to the GARCH(p,q) model with $p \neq q$. This is because we may, without any loss of generality, nest that model in a model with $p = q$ by assuming certain parameters in a GARCH(p,p) or a GARCH(q,q) model equal to zero.

Theorem 4.5.
1. *Let $p > q$ in the GARCH(p,q) model. Then the necessary and sufficient condition for the existence of the fourth moment $\mathsf{E}\varepsilon_t^4$ and the autocorrelation function for $\{\varepsilon_t^2\}$ are given in Theorems 4.2 and 4.4 for the GARCH(p,p) model. In that case, $\gamma_{i1}, \gamma_{i2}, \overline{\gamma}_{i1}$ and $\widetilde{\gamma}_{ij}$ simplify to*

$$\gamma_{i1} = \begin{cases} \gamma_{i1} & for \;\; i = 1, ..., q \\ \beta_i & for \;\; i = q+1, ..., p \end{cases}, \quad \gamma_{i2} = \begin{cases} \gamma_{i2} & for \;\; i = 1, ..., q \\ \beta_i^2 & for \;\; i = q+1, ..., p \end{cases}$$

$$\overline{\gamma}_{i1} = \begin{cases} \overline{\gamma}_{i1} & for \;\; i = 1, ..., q \\ \beta_i \nu_2 & for \;\; i = q+1, ..., p \end{cases}, \quad \widetilde{\gamma}_{ij} = \begin{cases} \widetilde{\gamma}_{ij} & for \;\; i < j \leq q \\ \beta_j \gamma_{i1} & for \;\; i \leq q < j \\ \beta_i \beta_j & for \;\; q < i < j. \end{cases}$$

2. *Let $p < q$ in the GARCH(p,q) model. Then the necessary and sufficient condition for the existence of the fourth moment $\mathsf{E}\varepsilon_t^4$ and the autocorrelation function ρ_n for $\{\varepsilon_t^2\}$ are given in Theorems 4.2 and 4.4 for the GARCH(q,q) model. In that case, $\gamma_{i1}, \gamma_{i2}, \overline{\gamma}_{i1}$ and $\widetilde{\gamma}_{ij}$ simplify to*

$$\gamma_{i1} = \begin{cases} \gamma_{i1} & if \;\; i = 1, ..., p \\ \alpha_i \nu_2 & if \;\; p+1, ..., q \end{cases}, \quad \gamma_{i2} = \begin{cases} \gamma_{i2} & if \;\; i = 1, ..., p \\ \alpha_i^2 \nu_4 & if \;\; p+1, ..., q \end{cases}$$

$$\overline{\gamma}_{i1} = \begin{cases} \overline{\gamma}_{i1} & if \;\; i = 1, ..., p \\ \alpha_i \nu_4 & if \;\; p+1, ..., q \end{cases}, \quad \widetilde{\gamma}_{ij} = \begin{cases} \widetilde{\gamma}_{ij} & if \;\; i < j \leq p \\ \alpha_j \overline{\gamma}_{i1} & if \;\; i \leq p < j \\ \alpha_i \alpha_j \nu_4 & if \;\; p < i < j \leq q. \end{cases}$$

Examples of the use of this theorem can be found in Chapter 5.

4.5 Conclusions

We have obtained a complete characterization of the fourth moment structure of a general GARCH(p,q) process. With our results, an investigator can see what an estimated GARCH model implies about the second and the fourth moments, kurtosis, and the autocorrelation function of the centred and squared observations. Such considerations have previously been possible in the GARCH(1,1) case. These results can be extended to other GARCH processes which are generalizations of the original GARCH(p,q) process. For example, some GARCH processes allowing for asymmetric effects to shocks belong to this category. Those generalizations are a topic of further work.

We have not considered moments of higher than fourth order. Deriving those using the present techniques would no doubt be tedious. On the other hand, as underlined above, the fourth moments are probably more interesting in practice than any higher order ones.

References

Bera, A. and M. Higgins (1993). ARCH models: properties, estimation, and testing. *Journal of Economic Surveys* **7**, 305-366.

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

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

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.

Diebold, F.X. and J.A. Lopez (1995). Modelling volatility dynamics. In: K.D. Hoover, ed.: *Macroeconometrics. Developments, tensions, and prospects*, 427-466. Boston: Kluwer.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987-1008.

Engle, R.F., ed. (1995). *ARCH. Selected readings.* Oxford: Oxford University Press.

Guégan, D. (1994). *Séries chronologiques non linéaires à temps discret.* Paris: Economica.

Loretan, M. and P.C.B. Phillips (1994). Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets. *Journal of Empirical Finance* **1**, 211-248.

Milhϕj, A. (1985). The moment structure of ARCH processes. *Scandinavian Journal of Statistics* **12**, 281-292.

Palm, F.C. (1996). GARCH models of volatility. In: G.S. Maddala and C.R. Rao, eds. *Handbook of Statistics*, Vol. **14**, 209-240. Amsterdam: Elsevier.

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

Teräsvirta, T. (1996). Two stylized facts and the GARCH(1,1) model. Stockholm School of Economics, Working Paper Series in Economics and Finance, No. 96.

Appendix C: Proofs

Appendix C1. Proof of Lemma 4.1.

Let $1 \leq n \leq p$. Applying (4.3) to h_t in $h_t h_{t-n}$ yields

$$h_t h_{t-n} = \alpha_0 h_{t-n} + c_{n,t-n} h_{t-n}^2 + \mathbf{c}'_{P\setminus\{n\}} \mathbf{h}_{1t} h_{t-n}, \tag{C.1}$$

where $\mathbf{h}_{1t} = (h_{t-1}, \cdots, h_{t-n+1}, h_{t-n-1}, \cdots, h_{t-p})'$. Applying (4.3) to h_{t-1} on the right-hand side of (C.1) and continuing the iteration until the appearance of the matrix $\mathbf{C}_{n+1}$ defined by (4.8) gives

$$\begin{aligned} h_t h_{t-n} &= \alpha_0 h_{t-n} + c_{n,t-n} h_{t-n}^2 + \alpha_0 \mathbf{c}'_{P\setminus\{n\}} \sum_{i=1}^{n-1} \left(\prod_{j=1}^{i} \mathbf{C}_j \right) \mathbf{e}_1 h_{t-n} \\ &\quad + \mathbf{c}'_{P\setminus\{n\}} \sum_{i=1}^{n-1} \left(\prod_{j=1}^{i} \mathbf{C}_j \right) \mathbf{e}_1 c_{n-i,t-n} h_{t-n}^2 \\ &\quad + \alpha_0 \mathbf{c}'_{P\setminus\{n\}} \prod_{i=1}^{n} \mathbf{C}_i \mathbf{h}_{10t} + \mathbf{c}'_{P\setminus\{n\}} \prod_{i=1}^{n} \mathbf{C}_i \mathbf{h}_{20t} \\ &\quad + \mathbf{c}'_{P\setminus\{n\}} \prod_{i=1}^{n+1} \mathbf{C}_i \mathbf{h}_{n+1,t}, \end{aligned} \tag{C.2}$$

where $\mathbf{C}_1, ..., \mathbf{C}_{n+1}$ are given by (4.6)-(4.8) and $\mathbf{h}_{n+1,t}$ has the form

$$\begin{aligned} \mathbf{h}_{n+1,t} &= (h_{t-n-1}h_{t-n-2}, \cdots, h_{t-n-1}h_{t-n-p}, \\ &\quad h_{t-n-2}h_{t-n-3}, \cdots, h_{t-n-2}h_{t-n-p}, \cdots, h_{t-n-p+1}h_{t-n-p})'. \end{aligned} \tag{C.3}$$

In particular, $\mathbf{C}_{n+1}$ is a $(p-1) \times p^*$ matrix corresponding to the p^*-component column vector given in (C.3). When $i > n+1$, applying (4.3) to h_{t-i} in vector $\mathbf{h}_{it}$ on the left-hand side of (C.4) yields

$$\mathbf{h}_{it} = \alpha_0 \mathbf{h}_{1,i+1,t} + \mathbf{h}_{2,i+1,t} + \mathbf{C}_{i+1} \mathbf{h}_{i+1,t} \tag{C.4}$$

where $\mathbf{h}_{it}$, $\mathbf{h}_{i+1,t}$, $\mathbf{h}_{1,i+1,t}$ and $\mathbf{h}_{2,i+1,t}$ are defined in (4.17). Further recursions of (C.1) by applying (C.4) lead to equation (4.12).

Similarly, it can be shown that (4.12) and (4.14) hold when $n > p$. $\dashv$

Appendix C2. Proof of Lemma 4.2.

Substituting $t-l$ for t and $m-l$ for n in Lemma 4.1 we obtain a recursion formula for $h_{t-l}h_{t-m}$, $l < m$. This technique will also be used in Appendices C3-C5. Let $k \geq p+m-l$. From (4.13),

$$\begin{aligned} &\mathsf{E}\left(c_{l,t-l} c_{m,t-m} S_k\right) \\ = \; &\gamma_{l1}\gamma_{m1} \mathsf{E}(\mathbf{c}'_{P\setminus\{m-l\}} \mathbf{C}_1 \cdots \mathbf{C}_{m-l+1}) \\ &\times \mathsf{E}(\mathbf{C}_{m-l+2} \cdots \mathbf{C}_{k-p+1}) \mathsf{E}(\mathbf{C}_{k-p+2} \cdots \mathbf{C}_k \mathbf{h}_{kt}). \end{aligned} \tag{C.5}$$

First, when $k \geq p+m-l$, the product $\mathbf{C}_{k-p+2} \cdots \mathbf{C}_k \mathbf{h}_{kt}$ is not a function of k, so that the same is true for its mean. Second, for $i, j > m-l+1$, $i \neq j$, we have $\mathsf{E}\mathbf{C}_i\mathbf{C}_j = \mathsf{E}\mathbf{C}_i\mathsf{E}\mathbf{C}_j$. Thus, for any $k \geq p+m-l$,

$$\mathsf{E}\left(\mathbf{C}_{m-l+2} \cdots \mathbf{C}_{k-p+1}\right) = \mathbf{\Gamma}^{k-(p+m-l)}. \tag{C.6}$$

Furthermore,

$$\lim_{k\to\infty} \mathbf{\Gamma}^{k-(p+m-l)} = \mathbf{0} \tag{C.7}$$

if and only if $\lambda(\mathbf{\Gamma}) < 1$. (C.5) and (C.7) together imply that

$$\mathsf{E}\left(c_{l,t-l}c_{m,t-m}S_k\right) \to 0$$

as $k \to \infty$, if and only if $\lambda(\mathbf{\Gamma}) < 1$. ⊣

Appendix C3. Proof of Lemma 4.3.

By (4.13),

$$\begin{aligned}
&\mathsf{E}\left(c_{l,t-l}c_{m,t-m}\sum_{i=m-l+1}^{k-1} S_{1i}\right) \\
= \;& \gamma_{l1}\gamma_{m1}\sum_{i=m-l+1}^{k-1}\mathsf{E}\left(S_{1i}\right) \\
= \;& \alpha_0\gamma_{l1}\gamma_{m1}\gamma'_{P\setminus\{m-l\}}\mathsf{E}\left(\mathbf{C}_1\cdots\mathbf{C}_{m-l+1}\right) \\
& \times\left[\sum_{i=m-l+1}^{k-1}\mathsf{E}\left(\mathbf{C}_{m-l+2}\cdots\mathbf{C}_i\right)\right]\mathsf{E}\left(\mathbf{h}_{1it}\right) \\
= \;& \alpha_0\gamma_{l1}\gamma_{m1}\gamma'_{P\setminus\{m-l\}}\mathbf{\Gamma}_1\cdots\mathbf{\Gamma}_{m-l+1}\left(\sum_{i=m-l+1}^{k-1}\mathbf{\Gamma}^{i-(m-l+1)}\right)\mathbf{e}_{p-1}\mathsf{E}h_t.
\end{aligned} \tag{C.8}$$

Note that

$$\sum_{i=m-l+1}^{k-1}\mathbf{\Gamma}^{i-(m-l+1)} \to \left(\mathbf{I}_{p^*}-\mathbf{\Gamma}\right)^{-1}$$

as $k \to \infty$, if and only if $\lambda(\mathbf{\Gamma}) < 1$. Thus (4.21) is valid. ⊣

Appendix C4. Proof of Lemma 4.4.

Suppose that $k > m-l+p$. By (4.13),

$$\begin{aligned}
&\mathsf{E}\left(c_{l,t-l}c_{m,t-m}\sum_{i=m-l+1}^{k-1} S_{2i}\right) \\
= \;& \gamma_{l1}\gamma_{m1}\sum_{i=m-l+1}^{m-l+p-1}\mathsf{E}\left(S_{2i}\right) + \gamma_{l1}\gamma_{m1}\sum_{i=m-l+p}^{k-1}\mathsf{E}\left(S_{2i}\right)
\end{aligned}$$

$$= \gamma_{l1}\gamma_{m1}\sum_{i=m-l+1}^{m-l+p-1}\mathsf{E}\left(S_{2i}\right)+\gamma_{l1}\gamma_{m1}\gamma'_{P\setminus\{m-l\}}\prod_{j=1}^{m-l+1}\mathbf{\Gamma}_j$$
$$\times\left[\sum_{i=m-l+p}^{k-1}\mathsf{E}\left(\mathbf{C}_{m-l+2}\cdots\mathbf{C}_{i-p+1}\right)\mathsf{E}\left(\mathbf{C}_{i-p+2}\cdots\mathbf{C}_i\mathbf{h}_{2it}\right)\right]. \quad \text{(C.9)}$$

Arguing as in Appendix C2, $\mathsf{E}\left(\mathbf{C}_{i-p+2}\cdots\mathbf{C}_i\mathbf{h}_{2it}\right)$ does not depend on i for $i > m-l+2p-2$. Setting $i = m-l+p$ in (4.22) gives

$$\mathsf{E}\left(\mathbf{C}_{m-l+p+1}\cdots\mathbf{C}_{m-l+2p-1}\mathbf{c}_{2,m-l+p,t}\right) = \gamma\left(m-l+p+1, m-l+2p-1\right).$$

Then, for any $i > m-l+2p-2$,

$$\mathsf{E}\left(\mathbf{C}_{i-p+2}\cdots\mathbf{C}_i\mathbf{c}_{2it}\right) = \gamma\left(m-l+p+1, m-l+2p-1\right). \quad \text{(C.10)}$$

On the other hand, by (C.6),

$$\sum_{i=m-l+p}^{k-1}\mathsf{E}\left(\mathbf{C}_{m-l+2}\cdots\mathbf{C}_{i-p+1}\right) = \sum_{i=m-l+p}^{k-1}\mathbf{\Gamma}^{i-(m-l+p)}$$

which implies

$$\lim_{k\to\infty}\sum_{i=m-l+p}^{k-1}\mathbf{\Gamma}^{i-(m-l+p)} = \left(\mathbf{I}_{p_*}-\mathbf{\Gamma}\right)^{-1} \quad \text{(C.11)}$$

if and only if $\lambda\left(\mathbf{\Gamma}\right) < 1$. Under this condition and applying (C.10), (C.11) and (4.23) to (C.9) we see that equation (4.24) holds. ⊣

Appendix C5. Proof of Theorem 4.1.

We shall show that (4.25)-(4.31) hold. From (4.13) we obtain that

$$\mathsf{E}\left(c_{l,t-l}c_{m,t-m}S_{10}\right)$$
$$= \alpha_0\gamma_{l1}\gamma_{m1}\left[1+\gamma'_{P\setminus\{m-l\}}\sum_{i=1}^{m-l-1}\left(\prod_{j=1}^{i}\mathbf{\Gamma}_j\right)\mathbf{e}_1+\prod_{i=1}^{m-l}\mathbf{\Gamma}_i\mathbf{j}_{p-1}\right]\mathsf{E}h_t \quad \text{(C.12)}$$

provided $\mathbf{C}_{m-l}$ is not an identity matrix. Following Lemma 4.3 and equation (C.12) we may define the coefficients of $\mathsf{E}h_t$ for the sum of $\mathsf{E}\left(c_{l,t-l}c_{m,t-m}S_{10}\right)$ and $\mathsf{E}\left(c_{l,t-l}c_{m,t-m}\sum_{i=m-l+1}^{\infty}S_{1i}\right)$ as follows. When $m-l > 1$,

$$M_1\left(l,m\right)$$
$$= 1+\gamma'_{P\setminus\{m-l\}}\Big[\sum_{i=1}^{m-l-1}\left(\prod_{j=1}^{i}\mathbf{\Gamma}_j\right)\mathbf{e}_1$$
$$+\prod_{i=1}^{m-l}\mathbf{\Gamma}_i\left(\mathbf{j}_{p-1}+\mathbf{\Gamma}_{m-l+1}\left(\mathbf{I}_{p^*}-\mathbf{\Gamma}\right)^{-1}\mathbf{e}_{p-1}\right)\Big] \quad \text{(C.13)}$$

and when $m-l=1$, (C.13) simplifies to

$$M_1(m-1,m) = 1 + \gamma'_{P\setminus\{1\}}\left[\mathbf{j}_{p-1} + \mathbf{\Gamma}_2\left(\mathbf{I}_{p^*} - \mathbf{\Gamma}\right)^{-1}\mathbf{e}_{p-1}\right] \tag{C.14}$$

since $\mathbf{\Gamma}_1 = \mathbf{I}_{p-1}$. Equation (4.26) is thus valid.

To evaluate $\mathsf{E}\left(c_{l,t-l}c_{m,t-m}S_{20}\right)$, consider first

$$\mathsf{E}\left(c_{l,t-l}c_{m,t-m}c_{m-l,t-m}h_{t-m}^2\right) = \gamma_{l1}\widetilde{\gamma}_{m-l,m}\mathsf{E}h_t^2. \tag{C.15}$$

Second, when $m-l>1$,

$$\begin{aligned}
&\mathsf{E}\left[c_{l,t-l}c_{m,t-m}\mathbf{c}'_{P\setminus\{m-l\}}\sum_{i=1}^{m-l-1}\left(\prod_{j=1}^{i}\mathbf{C}_j\right)\mathbf{e}_1 c_{m-l-i,t-m}h_{t-m}^2\right] \\
= &\left[\gamma_{l1}\gamma'_{P\setminus\{m-l\}}\sum_{i=2}^{m-l}\left(\prod_{j=2}^{i}\mathbf{\Gamma}_j\right)\mathbf{e}_1\widetilde{\gamma}_{m-l-i,m}\right]\mathsf{E}h_t^2.
\end{aligned} \tag{C.16}$$

Combining (C.15) and (C.16) and observing (4.25) yields (4.28). Finally, if $m-l=1$, then

$$\mathsf{E}\left(c_{l,t-l}c_{m,t-m}\mathbf{c}'_{P\setminus\{1\}}\mathbf{h}_{20t}\right) = \gamma_{l1}\gamma_{m1}\left(\sum_{j=2}^{p}\widetilde{\gamma}_{j-1,j}\right)\mathsf{E}h_t^2. \tag{C.17}$$

For $m-l=2$,

$$\begin{aligned}
&\mathsf{E}\left(c_{l,t-l}c_{m,t-m}\mathbf{c}'_{P\setminus\{2\}}\mathbf{C}_1\mathbf{C}_2\mathbf{h}_{20t}\right) \\
= &\gamma_{l1}\gamma_{m1}\left(\gamma_{11}\sum_{j=2}^{p}\widetilde{\gamma}_{j-1,j} + \sum_{j=3}^{p}\widetilde{\gamma}_{j-2,j}\right)\mathsf{E}h_t^2.
\end{aligned} \tag{C.18}$$

Let $M_{22}(1) = M_{22}(m-1,m) = \sum_{j=2}^{p}\widetilde{\gamma}_{j-1,j}$. We may then write the coefficients of $\mathsf{E}h_t^2$ in equation (C.18) as

$$M_{22}(2) = M_{22}(m-2,m) = \gamma_{11}M_{22}(1) + \sum_{j=3}^{p}\widetilde{\gamma}_{j-2,j}. \tag{C.19}$$

Continuing the recursion for $m-l=3,...,p-1$ proves (4.29).

Next, (4.30) and (4.31) follow from Lemma 4.4. Letting $k\to\infty$ in (4.19) gives

$$\begin{aligned}
&\mathsf{E}\left(c_{l,t-l}c_{m,t-m}h_{t-l}h_{t-m}\right) \\
= &\alpha_0\gamma_{l1}\gamma_{m1}M_1(l,m)\mathsf{E}h_t \\
&+\gamma_{l1}\left[M_{21}(l,m) + \gamma_{m1}\sum_{i=2}^{4}M_{2i}(l,m)\right]\mathsf{E}h_t^2
\end{aligned} \tag{C.20}$$

which, given definition (4.27), equals (4.25). This completes the proof. ⊣

Appendix C6. Proof of Corollary 4.1.

In the GARCH(2,2) model, $\mathbf{\Gamma} = \gamma_{21}$. Thus (4.35) holds if and only if $\gamma_{21} < 1$. Note that $m - l = 1$ and $\mathbf{\Gamma}_2 = \gamma_{21}$. By (4.26),

$$M_1(1,2) = 1 + \gamma_{21}\left[1 + \gamma_{21}(1-\gamma_{21})^{-1}\right] = (1-\gamma_{21})^{-1}.$$

Furthermore, by (4.28)-(4.31),

$$\begin{aligned} M_{21}(1,2) &= M_{22}(1,2) = \widetilde{\gamma}_{12} \\ M_{23}(1,2) &= \gamma(\mathbf{c}(1),2) = \mathsf{E}(c_{2,t-3}c_{2,t-4}c_{1,t-4}) = \gamma_{21}\widetilde{\gamma}_{12} \\ M_{24}(1,2) &= \gamma_{21}\mathbf{\Gamma}_1\mathbf{\Gamma}_2(\mathbf{I}_1 - \mathbf{\Gamma})^{-1}\widetilde{\gamma}_{12} = \gamma_{21}^2\widetilde{\gamma}_{12}(1-\gamma_{21})^{-1}. \end{aligned}$$

Finally, from (4.27),

$$M_2(1,2) = \widetilde{\gamma}_{12}(1-\gamma_{21})^{-1}. \qquad \dashv$$

Appendix C7. Proof of Lemma 4.5.

Define the $(p+1)\times 1$ vector

$$\mathbf{h}(n) = \left(0, \mathsf{E}\left(z_{t-n}^2 h_2(n)\right), \cdots, \mathsf{E}\left(z_{t-n}^2 h_{p+1}(n)\right)\right).$$

First, we show that for $n > p$,

$$\mathsf{E}\left(z_{t-n}^2 \mathbf{C}_{n-(p-1)}^* \mathbf{C}_{n-(p-2)}^* \cdots \mathbf{C}_n^* \mathbf{c}_{20t}^*\right) = \mathbf{h}(n) \tag{C.21}$$

where, for $i = 2, ..., p+1$,

$$h_i(n) = \sum_{j=2}^{p+1} \phi_{ij}(n)\, c_{j-2,t-(n+j-2)}. \tag{C.22}$$

In (C.22), $c_{0,t-n} = 1$, and $\phi_{ij}(n)$ are determined recursively with respect to integers $m = n-(p-1), n-(p-2), ..., n$, such that

$$\phi_{ij}(m) = \begin{cases} \phi_{i2}(m-1)\, c_{j-1,t-(m-j+2)} + \phi_{i,j+1}(m-1) & \text{for } j = 2, ..., p \\ \phi_{i2}(m-1)\, c_{p,t-(m+p-1)} & \text{for } j = p+1 \end{cases} \tag{C.23}$$

with the initial values $\phi_{ij}(n-(p-1)) = c_{ij}$ for any i and j, where c_{ij} is the (i,j)th element of matrix $\mathbf{C}_{n-(p-1)}^*$.

Let $\prod_{m=n-(p-1)}^{n} \mathbf{C}_m^* = \mathbf{\Phi}_n$, where $\mathbf{C}_m^*$ are given by (4.10). Since each $\mathbf{C}_m^*$ has the same first row $(1, 0, ..., 0)$, we can define

$$\mathbf{\Phi}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \phi_{21}(n) & \phi_{22}(n) & \cdots & \phi_{2,p+1}(n) \\ \vdots & \vdots & & \vdots \\ \phi_{p+1,1}(n) & \phi_{p+1,2}(n) & \cdots & \phi_{p+1,p+1}(n) \end{pmatrix} \tag{C.24}$$

and rewrite (C.24) as

$$\mathbf{\Phi}_n = \mathbf{\Phi}_{n-1}\mathbf{C}_n^*. \tag{C.25}$$

Equalities between the elements of the matrices on both sides of (C.25) appear in (C.23). We see that $\mathsf{E}(z_{t-n}^2\mathbf{\Phi}_n\mathbf{c}_{20t}^*)$ can be expressed in terms of $h_i(n)$ defined by (C.22) and (C.23).

Next, let $n = p+1$. Then by (C.21),

$$\mathsf{E}\left(z_{t-(p+1)}^2\mathbf{C}_2^*\mathbf{C}_3^*\cdots\mathbf{C}_{p+1}^*\mathbf{c}_{20t}^*\right) = \mathbf{h}(p+1) \tag{C.26}$$

where $\mathbf{c}_{20t}^* = (0,1,c_{1,t-(p+2)},\cdots,c_{p-1,t-2p})'$. For any $n \geq p+1$, it follows from (C.21) with (C.22) and (C.23) that

$$\mathsf{E}\left(z_{t-n}^2\mathbf{C}_{n-(p-1)}^*\mathbf{C}_{n-(p-2)}^*\cdots\mathbf{C}_n^*\mathbf{c}_{20t}^*\right) = \mathbf{h}(p+1). \tag{C.27}$$

Finally, without loss of generality, we merely show that (4.49) is true for $p=2$. From (C.26),

$$\mathsf{E}\left(z_{t-3}^2\mathbf{C}_2^*\mathbf{C}_3^*\begin{pmatrix}0\\1\\c_{1,t-4}\end{pmatrix}\right) = \begin{pmatrix}0\\ \gamma_{11}(\overline{\gamma}_{11}+\nu_2\widetilde{\gamma}_{12})+\overline{\gamma}_{21}\\ \overline{\gamma}_{11}+\nu_2\widetilde{\gamma}_{12}\end{pmatrix} \tag{C.28}$$

where $\overline{\gamma}_{11} = \mathsf{E}(z_t^2c_{1t}), \overline{\gamma}_{21} = \mathsf{E}(z_t^2c_{2t})$ and $\widetilde{\gamma}_{12} = \mathsf{E}(c_{1t}c_{2t})$. On the other hand, by equations (4.45) and (4.46),

$$\begin{aligned} M_{21}(1)+\nu_2M_{22}(1) &= \overline{\gamma}_{11}+\nu_2\widetilde{\gamma}_{12}, \\ M_{21}(2)+\nu_2M_{22}(2) &= \overline{\gamma}_{21}+\gamma_{11}(\overline{\gamma}_{11}+\nu_2\widetilde{\gamma}_{12}). \quad \dashv \end{aligned} \tag{C.29}$$

Appendix C8. Proof of Lemma 4.6.

First, for any $n > p$, we show that the coefficient of $\mathsf{E}h_t^2$ in (4.43) is a function of n. By (4.14),

$$\begin{aligned} &\sum_{i=n+1}^{\infty}\mathsf{E}\left(z_t^2z_{t-n}^2S_{2i}\right) \\ =\ & \nu_2^2\sum_{i=n+1}^{n+p-1}\gamma(\mathbf{c}_{p+1},2,i)\,\mathsf{E}h_t^2 \\ &+\nu_2^2\gamma_{p+1}'\mathbf{\Gamma}_*^{n-1}\mathbf{\Gamma}_{n+1}^*\left(\mathbf{I}_{p^*}-\mathbf{\Gamma}\right)^{-1}\gamma(n+p,n+2p-2)\,\mathsf{E}h_t^2 \end{aligned} \tag{C.30}$$

Set $n = p+1$. Then (C.30) simplifies to

$$\begin{aligned} &\sum_{i=p+2}^{\infty}\mathsf{E}\left(z_t^2z_{t-(p+1)}^2S_{2i}\right) \\ =\ & \nu_2^2\gamma_{p+1}'\mathbf{\Gamma}_*^2\sum_{i=1}^{p-1}\mathbf{\Gamma}_*^{i-1}\gamma(3+i,p+1+i)\,\mathsf{E}h_t^2 \\ &+\nu_2^2\gamma_{p+1}'\mathbf{\Gamma}_*^{n-1}\mathbf{\Gamma}_{n+1}^*\left(\mathbf{I}_{p^*}-\mathbf{\Gamma}\right)^{-1}\gamma(2p+1,3p-1)\,\mathsf{E}h_t^2. \end{aligned} \tag{C.31}$$

For any $n > p$, $\gamma(2p+1, 3p-1)$ in (C.31) remains unchanged. Thus

$$\begin{aligned}
&\sum_{i=n+1}^{\infty} \mathsf{E}\left(z_t^2 z_{t-n}^2 S_{2i}\right) \\
= \;& \nu_2^2 \gamma'_{p+1} \mathbf{\Gamma}_*^{n-(p-1)} \sum_{i=1}^{p-1} \mathbf{\Gamma}_*^{i-1} \gamma(3+i, p+1+i)\, \mathsf{E}h_t^2 \\
& + \nu_2^2 \gamma'_{p+1} \mathbf{\Gamma}_*^{n-1} \mathbf{\Gamma}_{n+1}^* \left(\mathbf{I}_{p^*} - \mathbf{\Gamma}\right)^{-1} \gamma(2p+1, 3p-1)\, \mathsf{E}h_t^2 \\
= \;& \nu_2^2 M_{23}(n) + \nu_2^2 M_{24}(n), \; n > p. \qquad \text{(C.32)}
\end{aligned}$$

Furthermore, by (4.14) and Lemma 4.5 we can write, for any $n > p$,

$$\begin{aligned}
\mathsf{E}(z_{t-n}^2 S_{20}) &= (\gamma'_{p+1} \mathbf{\Gamma}_*^{n-(p+1)} \mathbf{m}_1) \mathsf{E}h_t^2 \\
&= (M_{21}(n) + \nu_2 M_{22}(n)) \mathsf{E}h_t^2 \qquad \text{(C.33)}
\end{aligned}$$

where $\mathbf{m}_1 = (0, M_{21}(p) + \nu_2 M_{22}(p), \cdots, M_{21}(1) + \nu_2 M_{22}(1))'$. It follows from (C.32) and (C.33) that $M_2(n) = M_{21}(n) + \nu_2 \sum_{i=2}^{4} M_{2i}(n)$ holds for any $n > p$.

It remains to show that (4.50) holds for any $n > p$. Without loss of generality, we prove that it holds for $p = 2$. For proof, see Appendix C10. ⊣

Appendix C9. Proof of Theorem 4.3.

Note that only the second column of $\mathbf{\Phi}_n = \prod_{i=n-(p-1)}^{n} \mathbf{C}_i^*$ depends on variables z_{t-n}^2. Thus,

$$\mathsf{E}\left(z_t^2 z_{t-n}^2 \mathbf{c}_{p+1}^{*\prime} \mathbf{C}_1^* \cdots \mathbf{C}_n^* \mathbf{h}_{10t}^*\right) = \left(\alpha_0 \nu_2^2 \gamma'_{p+1} \mathbf{\Gamma}_*^{n-1} \mathbf{e}_{\alpha_0}\right) \mathsf{E}h_t \qquad \text{(C.34)}$$

and

$$\sum_{i=n+1}^{\infty} \mathsf{E}\left(z_t^2 z_{t-n}^2 S_{1i}\right) = \left(\alpha_0 \nu_2^2 \gamma'_{p+1} \mathbf{\Gamma}_*^{n-1} \mathbf{\Gamma}_{n+1}^* \left(\mathbf{I}_{p^*} - \mathbf{\Gamma}\right)^{-1} \mathbf{e}_{p-1}\right) \mathsf{E}h_t. \qquad \text{(C.35)}$$

From (C.34) and (C.35) we see that (4.53) holds. ⊣

Appendix C10. Proof of Corollary 4.3.

Here we only illustrate how to calculate $M_1(n)$ and $M_2(n)$ for the GARCH(2,2) model. Write

$$\mathbf{\Gamma}_* = \mathsf{E}\left(\mathbf{C}_i^*\right) = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_0 & \gamma_{11} & \gamma_{12} \\ 0 & 1 & 0 \end{pmatrix}$$

for $i = 2, ..., n$, where $\mathbf{C}_i^*$ are defined by (4.10). By (4.53),

$$M_1(n) = \gamma'_3 \mathbf{\Gamma}_*^{n-1} \left[\begin{pmatrix} \alpha_0^{-1} \\ 0 \\ 1 \end{pmatrix} + \frac{1}{1-\gamma_{21}} \begin{pmatrix} 0 \\ 0 \\ \gamma_{21} \end{pmatrix} \right]$$

$$= \gamma_3' \mathbf{\Gamma}_*^{n-1} \begin{pmatrix} \alpha_0^{-1} \\ 0 \\ (1-\gamma_{21})^{-1} \end{pmatrix} \tag{C.36}$$

for $n = 1, 2, \cdots$. By (4.45)-(4.48),

$$\begin{aligned}
M_{21}(1) &= \overline{\gamma}_{11} \\
M_{22}(1) &= \widetilde{\gamma}_{12} \\
M_{23}(1) &= \gamma(\mathbf{c}(1), 2) = \mathsf{E}(c_{2,t-3}c_{2,t-4}c_{1,t-4}) = \gamma_{21}\widetilde{\gamma}_{12} \\
M_{24}(1) &= \gamma_{21}\mathbf{\Gamma}_2 \frac{1}{1-\gamma_{21}} \widetilde{\gamma}_{12} = \gamma_{21}^2 \widetilde{\gamma}_{12}(1-\gamma_{21})^{-1}.
\end{aligned}$$

Thus, from (4.44),

$$M_2(1) = \frac{\overline{\gamma}_{11}(1-\gamma_{21}) + \nu_2 \widetilde{\gamma}_{12}}{1-\gamma_{21}}. \tag{C.37}$$

From (C.36) it follows that $M_1(1) = (1-\gamma_{21})^{-1}$. Similarly,

$$\begin{aligned}
M_{21}(2) &= \gamma_{11}\overline{\gamma}_{11} + \overline{\gamma}_{21} \\
M_{22}(2) &= \gamma_{11}\widetilde{\gamma}_{12} \\
M_{23}(2) &= \gamma(\mathbf{c}(2), 2, 3) = \mathsf{E}(c_{1,t-1}c_{2,t-3}c_{2,t-4}c_{1,t-4}) = \gamma_{11}\gamma_{21}\widetilde{\gamma}_{12} \\
M_{24}(2) &= \gamma_{11}\mathbf{\Gamma}_2\mathbf{\Gamma}_3 \frac{1}{1-\gamma_{21}} \widetilde{\gamma}_{12} = \gamma_{11}\gamma_{21}^2 \widetilde{\gamma}_{12}(1-\gamma_{21})^{-1}.
\end{aligned}$$

Thus, by (4.44),

$$M_2(2) = \frac{(\gamma_{11}\overline{\gamma}_{11} + \overline{\gamma}_{21})(1-\gamma_{21}) + \nu_2\gamma_{11}\widetilde{\gamma}_{12}}{1-\gamma_{21}} \tag{C.38}$$

and, from (C.36),

$$M_1(2) = (1 + \gamma_{11} - \gamma_{21})(1-\gamma_{21})^{-1}.$$

For any $n \geq 3$, by Lemma 4.5,

$$M_{21}(n) + \nu_2 M_{22}(n) = \gamma_3' \mathbf{\Gamma}_*^{n-3} \begin{pmatrix} 0 \\ M_{21}(2) + \nu_2 M_{22}(2) \\ M_{21}(1) + \nu_2 M_{22}(1) \end{pmatrix}$$

and by (C.32),

$$\begin{aligned}
M_{23}(n) &= \gamma_3' \mathbf{\Gamma}_*^{n-1} \gamma(4) = \gamma_3' \mathbf{\Gamma}_*^{n-1} \mathsf{E} \begin{pmatrix} 0 \\ 0 \\ c_{2,t-5} \end{pmatrix} (c_{1,t-5}) \\
&= \gamma_3' \mathbf{\Gamma}_*^{n-1} \begin{pmatrix} 0, & 0, & \widetilde{\gamma}_{12} \end{pmatrix}' \\
M_{24}(n) &= \gamma_3' \mathbf{\Gamma}_*^{n-1} \begin{pmatrix} 0, & 0, & \gamma_{21} \end{pmatrix}' \left(\frac{\widetilde{\gamma}_{12}}{1-\gamma_{21}}\right).
\end{aligned}$$

Therefore, for $n \geq 3$, applying (C.32) and (C.33) yields

$$\begin{aligned} M_2(n) &= \gamma_3' \Gamma_*^{n-3} \left[\begin{pmatrix} 0 \\ M_{21}(2) + \nu_2 M_{22}(2) \\ M_{21}(1) + \nu_2 M_{22}(1) \end{pmatrix} + \nu_2 \Gamma_*^2 \begin{pmatrix} 0 \\ 0 \\ \frac{\tilde{\gamma}_{12}}{1-\gamma_{21}} \end{pmatrix} \right] \\ &= \gamma_3' \Gamma_*^{n-3} \begin{pmatrix} 0 \\ \frac{\gamma_{11}[\overline{\gamma}_{11}(1-\gamma_{21}) + \nu_2 \tilde{\gamma}_{12}] + \overline{\gamma}_{21}(1-\gamma_{21})}{1-\gamma_{21}} \\ \frac{\overline{\gamma}_{11}(1-\gamma_{21}) + \nu_2 \tilde{\gamma}_{12}}{1-\gamma_{21}} \end{pmatrix}. \end{aligned} \tag{C.39}$$

Finally, by (C.37) and (C.38),

$$M_2(n) = \gamma_3' \Gamma_*^{n-3} \begin{pmatrix} 0, & M_2(2), & M_2(1) \end{pmatrix}', \; n \geq 3. \dashv \tag{C.40}$$

Appendix C11. Proof of Corollary 4.4.

By Theorem 4.4, the autocorrelation function for the squared process has the form

$$\rho_n = \frac{\nu_2 \gamma_{S_1} \left(1 - \nu_2 \sum_{i=1}^{p} \alpha_i\right) M_2(n) - \nu_2^2 \gamma_{S_2} \left[1 - \left(1 - \nu_2 \sum_{i=1}^{p} \alpha_i\right) M_1(n)\right]}{\nu_4 \gamma_{S_1} \left(1 - \nu_2 \sum_{i=1}^{p} \alpha_i\right) - \nu_2^2 \gamma_{S_2}} \tag{C.41}$$

for $n \geq 1$. Showing that Corollary 4.4 is true is equivalent to demonstrating that

$$M_2(n) = (\nu_4/\nu_2) \left(1 - \left(1 - \nu_2 \sum_{i=1}^{p} \alpha_i\right) M_1(n)\right)$$

holds. Then we can show the validity of (4.70) by (C.41).

Without loss of generality, we only investigate the case $p = 2$. It follows from Corollary 4.3 that

$$M_2(i) = (\nu_4/\nu_2)(1 - (1 - \nu_2\alpha_1 - \nu_2\alpha_2) M_1(i)) \tag{C.42}$$

for $i = 1, 2$. We have to prove that (C.42) holds for any $n \geq 3$. From (4.59) and (4.60) we see that this is equivalent to showing that

$$\begin{pmatrix} \alpha_0 & \nu_2\alpha_1 & \nu_2\alpha_2 \end{pmatrix} \Gamma_*^k \boldsymbol{\alpha}_* = 1 \tag{C.43}$$

for any k, where $\boldsymbol{\alpha}_* = (\alpha_0^{-1}(1 - \nu_2\alpha_1 - \nu_2\alpha_2), 1, 1)'$. That (C.43) holds follows from the fact that

$$(\nu_4/\nu_2) \begin{pmatrix} 0 \\ M_2(2) \\ M_2(1) \end{pmatrix} + \Gamma_*^2 \begin{pmatrix} \alpha_0^{-1}(1 - \nu_2\alpha_1 - \nu_2\alpha_2) \\ 0 \\ (1 - \nu_2\alpha_1 - \nu_2\alpha_2)(1 - \nu_2\alpha_2)^{-1} \end{pmatrix} = \boldsymbol{\alpha}_*.$$

Note that $\Gamma_*^k \boldsymbol{\alpha}_* = \boldsymbol{\alpha}_*$ for any k and $\begin{pmatrix} \alpha_0 & \nu_2\alpha_1 & \nu_2\alpha_2 \end{pmatrix} \boldsymbol{\alpha}_* = 1$. This completes the proof. $\dashv$

5

Properties of the Autocorrelation Function of Squared Observations for Second Order GARCH Processes under Two Sets of Parameter Constraints

5.1 Introduction

In GARCH models for parameterizing nonconstant conditional variance, the parameters have to satisfy conditions that restrict the conditional variance to be positive with probability one. In a recent paper, Nelson and Cao (1992) showed that the conditions on the coefficients given in Bollerslev (1986) are sufficient but not necessary for the positivity restriction to hold. These conditions require that the parameters of the standard GARCH(p, q) model be nonnegative. Nelson and Cao (1992) derived a set of less severe sufficient conditions which allow some parameters to be negative in higher order models than the classical GARCH(1,1) one. They pointed out that the estimated higher-order GARCH models sometimes do contain negative parameter estimates. Their conclusion was that empirical violations of Bollerslev's conditions may be the result neither of sampling error nor model misspecification.

It thus appears that GARCH models with negative parameters may characterize some features of the series which are not adequately captured by a GARCH model with nonnegative coefficients. It may be argued that the great popularity of GARCH models is due to the fact that they adequately describe two stylized facts in the data: high kurtosis and a slowly decaying autocorrelation function for the squared observations. In this chapter we assume that processes modelled by GARCH models have finite fourth moments.. This being the case, the autocorrelation function of the squared observations exists. The question we ask is whether replacing the Bollerslev

conditions by those of Nelson and Cao leads to increased flexibility in the shape of the autocorrelation function.

In order to consider this problem we make use of the necessary and sufficient condition for the existence of the unconditional fourth moment in the GARCH(p,q) process Chapter 4 derived. Chapter 4 also provided the autocorrelation function needed here. We restrict our considerations to the GARCH(2,2) model and its two special cases. This is because the most commonly used higher-order GARCH models are of order two and because the algebra becomes tedious if we go beyond the second-order case. The results show that relaxing Bollerslev's conditions does allow more flexibility in the shape of the autocorrelation function. The details will be discussed later.

The plan of this chapter is as follows. Section 5.2 contains the assumptions. Section 5.3 discusses properties of the autocorrelation function of the squared observations. In Section 5.4, we present the four types of autocorrelation functions for the GARCH(2,2) model, and consider their relation to the two sets of conditions for nonnegativity of the conditional variance. Section 5.5 contains an example and Section 5.6 concludes.

5.2 Assumptions

Since GARCH(p,q) models with $p,q \geq 2$ are rare in practice, we restrict the discussion to the case $p,q \leq 2$. Thus, we begin with the GARCH(2,2) model

$$\varepsilon_t = z_t\sqrt{h_t} \tag{5.1}$$

where $\{z_t\}$ is a sequence of independent identically distributed random variables with zero mean, a symmetric density, and a finite fourth moment, and

$$h_t = \alpha_0 + \beta_1 h_{t-1} + \beta_2 h_{t-2} + \alpha_1^2\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2. \tag{5.2}$$

In particular, setting $\beta_2 = 0$ in (5.2) yields a GARCH(1,2) model, whereas assuming $\alpha_2 = 0$ leads to the GARCH(2,1) model. To simplify the discussion, we assume that $\{z_t\} \sim$nid$(0,1)$. Consider the following set of assumptions:

(A.1) *Bollerslev's conditions*: Assume that $\alpha_0 > 0, \beta_i \geq 0$ and $\alpha_i \geq 0$ for $i = 1,2$, in (5.2). Furthermore, if $\alpha_1 = 0$ then $\alpha_2 > 0$ and vice versa to guarantee identification.

(A.2) *Nelson & Cao's conditions*: Relax (A.1) by assuming that some parameters in (5.2) may be negative but make sure that the conditional variances are nonnegative almost surely for all t. More precisely,

(A.21) In the GARCH(1,2) model, (a) $\alpha_0 > 0$, (b) $0 \leq \beta_1 < 1$, (c) $\beta_1\alpha_1 + \alpha_2 \geq 0$ and (d) $\alpha_1 \geq 0$.

(A.22) In the GARCH(2,1) model, (a) $\alpha_0 > 0, \beta_1, \alpha_1 \geq 0$, (b) $\beta_1+\beta_2 < 1$ and (c) $\beta_1^2 + 4\beta_2 \geq 0$.

(A.23) In the GARCH(2,2) model, (a) the roots of $(1-\sum_{i=1}^{2}\beta_i z^i)$ lie outside the unit circle and the polynomials $(1-\sum_{i=1}^{2}\beta_i z^i)$ and $\sum_{i=1}^{2}\alpha_i z^{i-1}$ have no common roots, (b) λ_1 and λ_2 are real numbers and $\lambda_1 > 0$, where λ_1 and λ_2 are the roots of $(1-\beta_1 z^{-1}-\beta_2 z^{-2})$, (c) $\alpha_0/(1-\lambda_1-\lambda_2+\lambda_1\lambda_2) > 0$, (d) $\alpha_1\lambda_1+\alpha_2 > 0$, (e) $\alpha_1 \geq 0$, (f) $\beta_1\alpha_1+\alpha_2 \geq 0$ and (g) $\beta_1(\beta_1\alpha_1+\alpha_2)+\beta_2\alpha_1 \geq 0$.

(A.3) In order to discuss (A.1) and (A.2) we use the condition in Chapter 4 for the existence of the unconditional fourth moment of (5.1) with (5.2). It has the form

$$(\gamma_{12}+\gamma_{22})(1-\gamma_{21})+2\gamma_{11}\widetilde{\gamma}_{12}+\gamma_{21} < 1 \tag{5.3}$$

where $\gamma_{i1} = \beta_i+\alpha_i$, $\gamma_{i2} = \beta_i^2+2\beta_i\alpha_i+3\alpha_i^2$ for $i = 1,2$, and $\widetilde{\gamma}_{12} = \beta_1\beta_2+\beta_1\alpha_2+\beta_2\alpha_1+3\alpha_1\alpha_2$, with $|\gamma_{21}| < 1$. In particular,

(A.31) For the GARCH(1,2) process, (5.3) becomes

$$(\beta_1^2+2\beta_1\alpha_1+3\alpha_1^2+3\alpha_2^2)(1-\alpha_2)+2(\beta_1+\alpha_1)(\beta_1\alpha_2+3\alpha_1\alpha_2)+\alpha_2 < 1. \tag{5.3.1}$$

(A.32) For the GARCH(2,1) process, we have from (5.3)

$$(\beta_1^2+\beta_2^2+2\beta_1\alpha_1+3\alpha_1^2)(1-\beta_2)+2(\beta_1+\alpha_1)(\beta_1\beta_2+\beta_2\alpha_1)+\beta_2 < 1. \tag{5.3.2}$$

Furthermore, in the GARCH(2,2) case (5.3) is called condition (A.33).

The set (A.2) contains conditions that are weaker than those in (A.1). Parameter α_2 in the GARCH(1,2) model is allowed to be negative under some conditions and the same is true for β_2 in the GARCH(2,1) model. Moreover, in the GARCH(2,2) model both α_2 and β_2 may be negative under certain conditions. Condition (A.3) is necessary and sufficient for the existence of the fourth unconditional moment of ε_t and thus guarantees the existence of the autocorrelation function of the squared process $\{\varepsilon_t^2\}$.

5.3 The autocorrelation function of the squares and some of its properties

The results in Chapter 4 applied to the GARCH(2,2) model give the autocorrelation function of $\{\varepsilon_t^2\}$ as

$$\rho_1 = \frac{\gamma_{S_1}\left[\overline{\gamma}_{11}(1-\gamma_{21})+\widetilde{\gamma}_{12}\right](1-\gamma_{11}-\gamma_{21})-\gamma_{11}\gamma_{S_2}}{(1-\gamma_{21})\left[3\gamma_{S_1}(1-\gamma_{11}-\gamma_{21})-\gamma_{S_2}\right]} \tag{5.4}$$

$$\rho_2 = \frac{\gamma_{S_1}\left[(\gamma_{11}\overline{\gamma}_{11}+\overline{\gamma}_{21})(1-\gamma_{21})+\gamma_{11}\widetilde{\gamma}_{12}\right](1-\gamma_{11}-\gamma_{21})-\gamma_{S_2}\left(\gamma_{11}^2+\gamma_{21}-\gamma_{21}^2\right)}{(1-\gamma_{21})\left[3\gamma_{S_1}(1-\gamma_{11}-\gamma_{21})-\gamma_{S_2}\right]} \tag{5.5}$$

and for $n \geq 3$,

$$\rho_n = \frac{\gamma_{S_1}(1-\gamma_{11}-\gamma_{21})M_2(n) - \gamma_{S_2}[1-(1-\gamma_{11}-\gamma_{21})M_1(n)]}{3\gamma_{S_1}(1-\gamma_{11}-\gamma_{21}) - \gamma_{S_2}} \quad (5.6)$$

where

$$\gamma_{S_1} = (1+\gamma_{11}+\gamma_{21})(1-\gamma_{21}) + 2\gamma_{11}\gamma_{21} \quad (5.7)$$

$$\gamma_{S_2} = (1-\gamma_{12}-\gamma_{22})(1-\gamma_{21}) - 2\gamma_{11}\widetilde{\gamma}_{12} \quad (5.8)$$

$$M_1(n) = \begin{pmatrix} \alpha_0 & \gamma_{11} & \gamma_{21} \end{pmatrix} \mathbf{\Gamma}^{n-1} \begin{pmatrix} \alpha_0^{-1} \\ 0 \\ (1-\gamma_{21})^{-1} \end{pmatrix}, \; n = 3, 4, \ldots \quad (5.9)$$

$$M_2(n) = \begin{pmatrix} \alpha_0 & \gamma_{11} & \gamma_{21} \end{pmatrix} \mathbf{\Gamma}^{n-3} \begin{pmatrix} 0 \\ M_2(2) \\ M_2(1) \end{pmatrix}, \; n = 3, 4, \ldots \quad (5.10)$$

In (5.9) and (5.10),

$$\mathbf{\Gamma} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_0 & \gamma_{11} & \gamma_{21} \\ 0 & 1 & 0 \end{pmatrix}$$

and, furthermore,

$$M_2(1) = \frac{1}{1-\gamma_{21}}\left[\overline{\gamma}_{11}(1-\gamma_{21}) + \widetilde{\gamma}_{12}\right] \quad (5.11)$$

$$M_2(2) = \frac{1}{1-\gamma_{21}}\left[(\gamma_{11}\overline{\gamma}_{11}+\overline{\gamma}_{21})(1-\gamma_{21}) + \gamma_{11}\widetilde{\gamma}_{12}\right] \quad (5.12)$$

where $\overline{\gamma}_{i1} = \beta_i + 3\alpha_i$ for $i = 1, 2$.

We consider some properties for the autocorrelation function of the GARCH (2,2) process which hold both under (A.1) and (A.2).

Property 5.3.1 *Assume that model (5.1) with (5.2) satisfies either (A.1) and (A.33) or (A.23) and (A.33). Then the first and second-order autocorrelations for the squared process $\{\varepsilon_t^2\}$ are positive.*

From those conditions, it follows that

$$\gamma_{S_1}(1-\gamma_{11}-\gamma_{21}) > \gamma_{S_2} \quad (5.13)$$

$$\overline{\gamma}_{11}(1-\gamma_{21}) + \widetilde{\gamma}_{12} > \gamma_{11} \quad (5.14)$$

$$(\gamma_{11}\overline{\gamma}_{11}+\overline{\gamma}_{21})(1-\gamma_{21}) + \gamma_{11}\widetilde{\gamma}_{12} > \gamma_{11}^2 + \gamma_{21} - \gamma_{21}^2. \quad (5.15)$$

Inequalities (5.13)-(5.15) are sufficient to guarantee positivity of the numerators of ρ_1 and ρ_2 on the right-hand side of equations (5.4) and (5.5).

From Bollerslev (1986, 1988) and Chapter 4 we know that for $n > 2$, the autocorrelations ρ_n obey the Yule-Walker equation

$$\rho_n = \sum_{i=1}^{2} (\alpha_i + \beta_i)\, \rho_{n-i}, \quad n > 2. \tag{5.16}$$

Thus we have

Property 5.3.2 *The autocorrelations of (5.6) satisfy the difference equation (5.16) for $n > 2$.*

Properties 5.3.1 and 5.3.2 together imply that under (A.3) and (A.1), or (A.3) and (A.2), the autocorrelations for the squared process are positive.

Property 5.3.3 *The autocorrelations decay exponentially as $n \to \infty$.*

5.4 Types of autocorrelation function of the squares

In this section we see what difference, if any, the Nelson-Cao conditions make concerning the behaviour of the autocorrelation function of $\{\varepsilon_t^2\}$ generated from a GARCH(2,2) process compared to the more severe conditions of Bollerslev (1986). We first classify the autocorrelation functions according to the behaviour of the first two autocorrelations and the shape of the function thereafter. In general, the autocorrelation function for the squared process behaves like a mixture of exponentials. As in ARMA(2,2) models (Box and Jenkins, 1976), it is possible to distinguish between four different types of autocorrelation functions. First, the autocorrelation function may decay smoothly from the first value ρ_1 and the decay is monotonic from ρ_2 onwards. This we call Type 1. In Type 2, the autocorrelation function also decays monotonically starting from ρ_2, but the autocorrelation function reaches its maximum at $n = 2$. It may also decay exponentially from ρ_2 with an oscillation while $\rho_1 > \rho_2$; this is Type 3. Finally, there is Type 4, the case where the autocorrelation function decays exponentially from ρ_2 with an oscillation, but $\rho_1 < \rho_2$. Those four types of autocorrelation function are illustrated in Figure 5.1.

In order to study the behaviour of ρ_1 and ρ_2 under different parameter restrictions we express ρ_2 as a function of ρ_1:

$$\rho_2 = \gamma_{11}\rho_1 + \alpha_2 + \theta\beta_2 \tag{5.17}$$

where $\theta = \{\gamma_{S_1}(1 - \gamma_{11} - \gamma_{21}) - \gamma_{S_2}\}/\{3\gamma_{S_1}(1 - \gamma_{11} - \gamma_{21}) - \gamma_{S_2}\}$ and $0 < \theta < 1/3$ when (A.3) holds. Obviously, from equation (5.17), we may have $\rho_1 > \rho_2$ although $\alpha_2 + \theta\beta_2 < 0$. On the other hand, $\rho_1 < \rho_2$ when

$\alpha_2 + \theta\beta_2$ is sufficiently large. Therefore, instead of the exponential decay from the first lag onwards as in the GARCH(1,1) case, the autocorrelations for the GARCH(2,2) process decay exponentially from the second lag as illustrated in Figure 5.1.

5.4.1 GARCH(2, 2) model

We begin with the general GARCH(2,2) model and state

Property 5.4.11 *Assume that (A.1) and (A.33) are satisfied and that* $M_2(1) < M_2(2)$ and $\gamma_{11} \geq \gamma_{21}$. *Then* $\rho_1 < \rho_2$.

To show that Property 5.4.11 is valid, notice first that, by (5.11) and (5.12),

$$M_2(2) = M_2(1)\gamma_{11} + \overline{\gamma}_{21}.$$

Then, $M_2(1) < M_2(2)$ implies

$$M_2(1)\gamma_{11} + \overline{\gamma}_{21} > M_2(1). \tag{5.18}$$

Moreover, combining (5.4) and (5.11) we may rewrite $M_2(1)$ as a function of ρ_1 such that

$$M_2(1) = \frac{\rho_1(1-\gamma_{21})[3\gamma_{S_1}(1-\gamma_{11}-\gamma_{21}) - \gamma_{S_2}] + \gamma_{11}\gamma_{S_2}}{\gamma_{S_1}(1-\gamma_{11}-\gamma_{21})}. \tag{5.19}$$

Inserting (5.19) into (5.18) we obtain, after some manipulation,

$$\rho_1\gamma_{11} + \frac{\overline{\gamma}_{21}\gamma_{S_1}(1-\gamma_{11}-\gamma_{21})(1-\gamma_{21}) + \gamma_{S_2}(\gamma_{11}^2 - \gamma_{11})}{[3\gamma_{S_1}(1-\gamma_{11}-\gamma_{21}) - \gamma_{S_2}](1-\gamma_{21})} > \rho_1. \tag{5.20}$$

Further algebra leads to

$$\rho_1\gamma_{11} + \alpha_2 + \theta\beta_2 + \frac{\gamma_{S_2}(\gamma_{21} - \gamma_{21}^2 - \gamma_{11} + \gamma_{11}^2)}{3\gamma_{S_1}(1-\gamma_{11}-\gamma_{21}) - \gamma_{S_2}} > \rho_1. \tag{5.21}$$

The sign of the last term on the left-hand side of (5.21) depends on the sign of $\gamma_{21} - \gamma_{21}^2 - \gamma_{11} + \gamma_{11}^2$. Suppose that $\gamma_{11} \geq \gamma_{21}$ and that (A.33) holds. In that case $\gamma_{21} - \gamma_{21}^2 - \gamma_{11} + \gamma_{11}^2 \leq 0$. Finally, applying (5.17) to (5.21) yields $\rho_1 < \rho_2$.

Note that when $M_2(1) > M_2(2)$, we may still have $\rho_1 < \rho_2$. For instance, let $\beta_1 = 0.8$, $\beta_2 = 0.05$, $\alpha_1 = 0.1$, and $\alpha_2 = 0.02$. Then $M_2(1) = 1.172 > M_2(2) = 1.165$, but $\rho_1 = 0.262 < \rho_2 = 0.265$.

It is seen that the existence condition for the fourth moments along with Bollerslev's conditions implies that the autocorrelation function for the GARCH(2,2) process is dominated by Type 2 exponential decay when

$M_2(1) < M_2(2)$ and $\gamma_{11} \geq \gamma_{21}$. Furthermore, this decay may be oscillating. Otherwise the autocorrelation function is characterized by smooth nonoscillating exponential decay starting from ρ_2 with either $\rho_1 > \rho_2$ or $\rho_1 \leq \rho_2$.

Assume next that $\alpha_2 < 0$. It then follows that the autocorrelation function decays exponentially starting from ρ_2, and $\rho_1 > \rho_2$. In other words, we have

Property 5.4.12 *Assume that (A.23) and (A.33) hold and that* $\alpha_2 < 0$. *Then* $\rho_1 > \rho_2$.

To see that Property 5.4.12 is true, we have to show that it holds both for $\alpha_2 + \theta\beta_2 < 0$ and for $\alpha_2 + \theta\beta_2 > 0$. First, let $\alpha_2 + \theta\beta_2 < 0$. If $\gamma_{11} \leq 1$, then from (5.17) we have that $\rho_1 > \rho_2$. If $\gamma_{11} > 1$, let $\delta = \gamma_{11} - 1$. Then $\delta < |\gamma_{21}|$. It follows from (A.33) that, $\rho_2 - \rho_1 = \delta\rho_1 + \alpha_2 + \theta\beta_2 < 0$. Second, consider $\alpha_2 + \theta\beta_2 > 0$. Set $\delta = 1 - \gamma_{11}$. Then $\delta > \gamma_{21} > \alpha_2 + \theta\beta_2$. Since $\rho_1 > \theta$, we see that $\rho_2 - \rho_1 = -\delta\rho_1 + \alpha_2 + \theta\beta_2 < 0$. This completes the proof of Property 4.12.

Next consider the case in which $\alpha_2 > 0$. Conditions (A.23) allow $\beta_2 < 0$. It is seen from (5.17) that ρ_2 depends on the size of α_2. When $\alpha_2 + \theta\beta_2 < 0$ we may have $\rho_1 > \rho_2$, otherwise $\rho_1 < \rho_2$. Assume that $M_2(1) < M_2(2)$ holds as in the definition of Property 4.11. We then find a sufficient condition for $\rho_1 < \rho_2$ under (A.23) and (A.33) which we express as follows:

Property 5.4.13 *Assume that the conditions (A.23) and (A.33) hold and that, furthermore,* $\alpha_2 > 0$, $\beta_2 < 0$, $M_2(1) < M_2(2)$ *and* $\gamma_{11} \geq \gamma_{21}$. *It then follows that* $\rho_1 < \rho_2$.

We may now compare the GARCH(2,2) model satisfying (A.23) and (A.33) with the one fulfilling (A.1) and (A.33). The results appear in Table 1. When $\alpha_2 > 0$, the autocorrelation functions have rather similar properties both under (A.1) and (A.23). Differences occur when $\alpha_2 < 0$, which is allowed by (A.23). In that case, the autocorrelation function is either of Type 1 or Type 3. When $\alpha_2 > 0$ and $\beta_2 < 0$, all four types are possible. Assuming (A.1) excludes Type 3 whereas the remaining alternatives are possible.

5.4.2 GARCH(1,2) model

Next we consider the effects of relaxing (A.1) in the GARCH(1,2) model. Setting $\beta_2 = 0$ in (5.17) yields

$$\rho_2 = \gamma_{11}\rho_1 + \alpha_2. \tag{5.22}$$

Following the same argument in the GARCH(2,2) case, we can show that the following two properties are valid:

Property 5.4.21 *Assume that (A.1) and (A.31) are satisfied and that* $M_2(1) < M_2(2)$ and $\gamma_{11} \geq \alpha_2$. *Then* $\rho_1 < \rho_2$.

Property 5.4.22 *Suppose that (A.21) and (A.31) hold. If* $\alpha_2 < 0$, *then* $\rho_1 > \rho_2$.

These properties correspond to Properties 5.4.11 and 5.4.12 for the GARCH(2,2) model. Note that $\alpha_2 < 0$ restricts the autocorrelation function to be of Type 1; see Table 1. Oscillation of Type 3 is not possible if the model is a GARCH(1,2) model.

5.4.3 *GARCH*(2, 1) *model*

Finally, we take a look at the GARCH(2,1) model. Setting $\alpha_2 = 0$ in (5.17) we obtain

$$\rho_2 = \gamma_{11}\rho_1 + \theta\beta_2. \tag{5.23}$$

Setting $\delta = 1 - \gamma_{11}$, under the conditions of (A.1) and (A.33), $\rho_2 - \rho_1 = -\delta\rho_1 + \theta\beta_2 < 0$ holds, because $\delta > \beta_2$ and $\rho_1 > \theta$. The autocorrelations function decays exponentially starting from the first lag with or without an oscillation. It is dominated by a mixture of exponential decays consisting of Type 1 and Type 3. It is impossible to have $\rho_2 > \rho_1$ in the GARCH(2,1) case and we have

Property 5.4.31 *If (A.1) and (A.32) hold, then* $\rho_1 > \rho_2$.

Conditions (A.22) and (A.32) allow $\beta_2 < 0$. It is seen that in that case $\rho_1 > \rho_2$ and the autocorrelations decay smoothly from the first lag so that only a Type 1 autocorrelation function is possible. In other words,

Property 5.4.32 *If (A.22) and (A.32) hold, then* $\rho_1 > \rho_2$.

5.4.4 *Summary of findings*

The situation as a whole is summed up in Table 5.1. We see that conditions (A.1) in principle already allow all four types of autocorrelation functions to exist. The oscillating autocorrelation function of Type 3, however, is only possible under (A.1) if $\alpha_2 = 0$, that is, the model is a GARCH(2,1) model. Thus, allowing negative parameters leads to increased flexibility in characterizing squared processes with a Type 3 autocorrelation function. On the other hand, if $\alpha_2 = 0$ while (A.1) holds then only autocorrelation functions with $\rho_1 > \rho_2$ are possible. In GARCH(2,1) and GARCH(1,2) models a negative parameter, β_2 or α_2, respectively, indicates a Type 1

autocorrelation function. In that case, the negativity allows more scope for the shape within this class of autocorrelation functions.

5.5 An empirical example

In this section we shall illustrate some of the above theory. There do not seem to exist many empirical examples of GARCH models with negative coefficient estimates in the literature. One paper containing such models is French, Schwert and Stambaugh (1987); see also Nelson and Cao (1992). Their Table 2 shows results on MA(1)-GARCH(1,2) models estimated for the daily S&P 500 return series. The estimates of α_2 are negative for all three models for periods January 1928 to December 1953, January 1953 to December 1984 and January 1928 to December 1984, respectively. To investigate the situation further we estimated an AR(5)-GARCH(1,2) model for each of the three periods assuming normal errors as French et al. (1987) did. The AR(5) part was intended to filter out all linear dependence in the original series. The maximum likelihood estimation of parameters was carried out using the Newton algorithm with analytical first and second derivatives (the programme we used is written by Stefan Lundbergh). After estimation we filtered out the linear dependence in the original return series by our AR(5) filter and computed the autocorrelation functions of the squares of the filtered series. The estimation results appear in Table 5.2 and the autocorrelation functions in Table 5.3. Our GARCH parameter estimates are, not surprisingly, very close to those of French et al. (1987) and we also have $\hat{\alpha}_2 < 0$. Two of the three autocorrelation functions start decaying roughly exponentially from the third lag, which suggests a GARCH model with an order higher than two for the filtered series. In these functions, the second autocorrelation is the highest one. For the period January 1953 to December 1984, however, the autocorrelation function starts decaying slowly from the second lag (Type 1), which agrees with a GARCH(1,2) model with $\alpha_2 < 0$ (Table 5.1). We estimated GARCH(1,3) models for the other two periods, but the results in Table 5.2 indicate that this extension does not contribute much to the explanation.

To investigate the situation further we computed the value of the left-hand side of the fourth moment existence condition (5.3.1) for each of the three GARCH(1,2) models. The same was done for the GARCH(1,3) models. Table 5.2 shows that the fourth moment of $\{\varepsilon_t^2\}$ only exists for the GARCH(1,2) model of the second subperiod, January 1953 to December 1984. This allows us to estimate the autocorrelation function of $\{\varepsilon_t^2\}$ from the model, applying (5.4) and (5.22). The result can be found in Table 5.3. The autocorrelation function is clearly of Type 1 as the theory prescribes but it starts much higher than the corresponding autocorrelation function in Table 5.3. Thus the correspondence between the estimated

GARCH model and our "stylized fact", the autocorrelation function estimated from the linearly filtered series, is not satisfactory, although the two autocorrelation functions have the same general shape.

For an illustration of the last remark consider a GARCH(1,2) model with $\beta_1 = 0.9246$ and let $\alpha_1 + \alpha_2 = 0.07$. These values correspond to the estimates of the second model in Table 5.2. The linear restriction bounds α_2 between -0.07 (bound for the existence of the fourth moment) and 0.07 ($\alpha_1 \geq 0$ is necessary for positive conditional variance). The left-hand panel of Figure 5.2 shows what happens to ρ_1 and ρ_2 when α_2 changes. The step from ρ_1 to ρ_2 increases with decreasing α_2 but is never very large. For most positive values of α_2 we have $\rho_1 < \rho_2$ but both autocorrelations remain high. The right-hand panel depicts another case where $\beta_1 = 0.87$ whereas α_1 and α_2 are bound together as before. In this case the lower bound for α_2 equals -0.47 and is due to the condition (A.21) (c). The upper bound still equals 0.07 for the same reason as before. It is seen that the difference between ρ_1 and ρ_2 can now be large even for negative values of α_2 close to zero. At the same time, ρ_1 is already rather low. However, in this case the decay of the autocorrelation function is much faster than in the preceding situation due to the relatively low value of β_1. This is of course just an example, but it suggests that it may be quite difficult to characterize a low-starting and slowly decaying autocorrelation function of the squared observations with a low-order GARCH model.

5.6 Summary and conclusions

We have considered the effects of relaxing the nonnegativity restriction of parameters of the second-order GARCH models and replacing it by the less severe restriction of Nelson and Cao (1992). The latter restrictions still guarantee the positivity of the conditional variance with probability one. This has been done by studying the autocorrelation function of the centred and squared observations. All four different types of autocorrelation function in principle already exist in the GARCH(2,2) model with nonnegative coefficients and finite fourth moments, but one of them (Type 3) actually requires either that $\alpha_2 = 0$ (which makes the model GARCH(1,2)) or that $\alpha_2 < 0$. On the other hand, if the second-order coefficient of a GARCH(1,2) or a GARCH(2,1) model is negative then we know the type of the autocorrelation function (Type 1). This may help in the specification of the model. Our numerical example shows that the negative coefficients may in some cases considerably increase the possibilities of reproducing an observed autocorrelation function with a GARCH model. But then, it also indicates that some types of observed autocorrelation functions may be very hard to represent with the family of (low order) GARCH models.

References

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

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

Box, G.E.P. and G.M. Jenkins (1976). Time series analysis: forecasting and control, 2nd edition. Holden-Day: San Francisco, CA.

French, K.R., G.W. Schwert and R.F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics* **19**, 3-30.

Nelson, D.B. and C.Q. Cao (1992). Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics* **10**, 229-235.

Tables 5.1-5.3

Table 1. Behaviour of the autocorrelation function of the squared observations of a GARCH(2,2) process and its special cases

Model	Conditions	
	(A.3) and (A.1)	(A.3) and (A.2)
GARCH(2,2)	positive Type 1,2,4	positive Type 1,3, if $\alpha_2 < 0$ Type 1,2,3,4, if $\alpha_2 > 0,\ \beta_2 < 0$
GARCH(1,2)	positive Type 1,2,4	positive Type 1, if $\alpha_2 < 0$
GARCH(2,1)	positive Type 1,3	positive Type 1, if $\beta_2 < 0$

Table 5.2. Parameter estimates (standard deviations in parentheses) of the GARCH part of AR(5)-GARCH models for the S&P 500 daily return series, January 1928 to December 1984, and two subperiods

	ARCH model for the error process				
	GARCH(1,2)			GARCH(1,3)	
Parameter estimate	Period				
	Jan 1928 -Dec 1952 (T=7326)	Jan 1953 -Dec 1984 (T=8043)	Jan 1928 -Dec 1984 (T=15369)	Jan 1928 -Dec 1952 (T=7326)	Jan 1928 -Dec 1984 (T=15369)
$\hat{\alpha}_0$	1.43×10^{-6}	4.65×10^{-7}	5.74×10^{-7}	1.51×10^{-6}	5.54×10^{-7}
	(2.28×10^{-7})	$(9{,}71\times10^{-8})$	(8.11×10^{-8})	(2.58×10^{-7})	(9.60×10^{-8})
$\hat{\beta}_1$	0.8964	0.9246	0.9192	0.8921	0.9218
	(0.0137)	(0.0146)	(0.0102)	(0.0137)	(0.0146)
$\hat{\alpha}_1$	0.1089	0.1361	0.1234	0.1092	0.1233
	(0.0152)	(0.0156)	(0.0111)	(0.0173)	(0.0190)
$\hat{\alpha}_2$	−0.0090	−0.0661	−0.0447	−0.0157	−0.0040
	(0.0081)	(0.0068)	(0.0048)	(0.0140)	(0.0144)
$\hat{\alpha}_3$				0.0106	−0.0066
				(0.0102)	(0.0075)
LHS[1]	1.013	0.9999	1.009	1.014	1.008

[1] LHS=Left-hand side of the fourth moment condition. A value less than unity indicates the existence of the fourth moment of $\{\varepsilon_t^2\}$.

Table 5.3. Autocorrelation functions of $\{\varepsilon_t^2\}$ estimated from filtered data and from GARCH(1,2) models

Period	D/M[1]	Lag									
		1	2	3	4	5	10	20	30	40	50
Jan 1928											
-Dec 1952	D	0.264	0.293	0.221	0.206	0.220	0.162	0.131	0.137	0.089	0.095
Jan 1953											
-Dec 1984	D	0.231	0.136	0.128	0.151	0.141	0.133	0.098	0.065	0.063	0.033
	M	0.415	0.374	0.369	0.367	0.365	0.354	0.334	0.315	0.298	0.281
Jan 1928											
-Dec 1984	D	0.279	0.311	0.232	0.221	0.232	0.182	0.153	0.149	0.110	0.111

[1]D=Autocorrelations estimated from the AR(5)-filtered observations
M=Autocorrelations estimated from the GARCH(1,2) model

Figures 5.1-5.2

Figure 5.1. The four types of autocorrelation function of squared observations generated by a GARCH(2,2) process

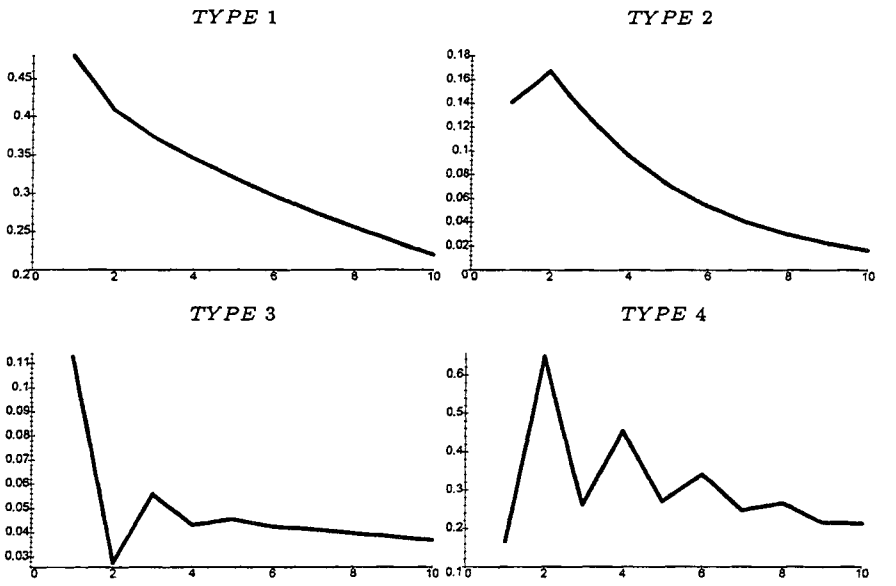
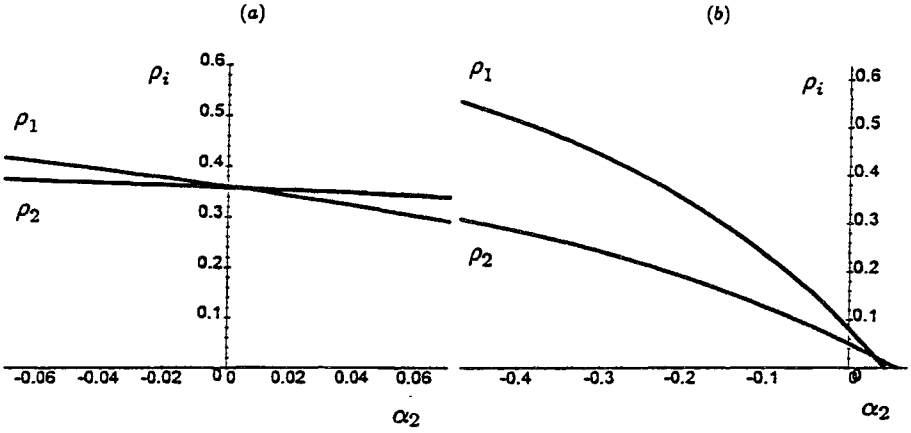


Figure 5.2. First and second-order autocorrelations of $\{\varepsilon_t^2\}$ *in a GARCH(1,2) model represented as a function of parameter* α_2 *in two cases: (a)* $\beta_1 = 0.9246$, $\alpha_1 \geq 0, \alpha_1 + \alpha_2 = 0.07$, *and (b)* $\beta_1 = 0.87, \alpha_1 \geq 0, \alpha_1 + \alpha_2 = 0.07$



EFI
The Economic Research Institute

Reports since 1993

Published in the language indicated by the titel

1993

Ekvall, N., Studies in Complex Financial Instruments and their Valuation.
Flink, T., The Development of High-Speed Trains - An Industrial Dynamics Approach. Research Report.
Gerdtham, U-G., Hertzman, P., Boman, G., Jönsson, B., Impact of Inhaled Corticosteroids on Asthma Hospitalization in Sweden: A Pooled Regression Analysis. Research Report.
Gun-Ho Cha, Reappraisal of Market Efficiency Tests Arising from Nonlinear Dependence, Fractals, and Dynamical Systems Theory.
Hertz, S., The Internationalization Processes of Freight Transport Companies.
Nilsson, K., Förhållanden som påverkar rapporteringstiden - En studie i tre koncerner. Research Report.
Redovisningens roller, red. Samuelson, L.A., Östman, L.
Ruist, E., Sambandsanalys med ordinal-skalevariabler. Vilka statistiska metoder kan användas? Research Report.
Skogsvik, K., Accounting for Pension Obligations. Research Report.
Sjöstrand, S-E., Om företagsformer.
Stein, J., Strategy Formation and Managerial Agency. A Socio-Cognitive Perspective.
Söderlund, M., Omvärldsmodeller hos besluts-fattare i industriföretag - en studie av svenska leverantörer till fordonsindustrin.
Sörensson, T., Swedish Convertible Bonds and their Valuation.
Tyrstrup, M., Företagsledares arbete. En logitudinell studie av arbetet i en företags-ledning.
Uddén-Jondal, E., Wage Formation in a Unionized Economy.
Whitelegg, J., Hultén, S., Flink, T., High Speed Trains. Fast Tracks to the Future.

1994

Andersson, H., Ett industriföretags omvandling. En studie av Hägglunds för-ändringsprocess 1922-81 med bas i företagets produkter, relationer och resurser.
Andersson, H., En produkthistoria. *(separat publicerad bilaga till ovanstående)*
Cowan, R., Hultén, S., Escaping Lock-in: The Case of the Electric Vehicle. Research Report.
Dahlfors, G., Jansson, P., Essays in Financial Guarantees and Risky Debt.
Företag och marknader i förändring - dynamik i nätverk, red. Mattsson, L-G., Hultén, S.
EFI/Nerenius & Santérus Förlag.
Helgesson, C-F., Coordination and Change in Telecommunications. Research Report.
Lundström, T., Wijkström, F., The Swedish
Nonprofit Sector. Research Report.
Nessén, M., Essays on Exchange Rates and International Finance.
Niemeyer, J., Essays on Market Microstructure -Empirical Evidence from Some Nordic Exchanges. **Nilsson, K.**, Rapportering i svenska storföretag 1993. Delrapport 1. Faktiska tider och påverkande faktorer för månatlig resultatrapportering. Research Report.

Nilsson, K., Rapportering i svenska storföretag 1993. Delrapport 2. Rapporteringstider för kvartals-/tertialsrapporter. Research Report.
Nordenlöw, L., Entreprenörskap i industriella nätverk. Research Report.
Norén, N-E., Om Flexibilitetspåverkande beslut. Research Report.
Normark, P., Medlemsägda företag. Organisering av strategiska förändringar.
Nyman, K., Studier av inkomstfördelningar med hjälp av mikrosimuleringsmodeller. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1993. Research Report.
Sjöholm, G., Redovisningsmått vid bolagisering. Utformning och effekter.
Wijkström, F., Det organiserade ideella arbetet i Sverige. Research Report.
Wirsäll, N-E., Holmlöv, PG., Den nyckfulle Svensson. Reflektioner efter ett seminarium om strategier för handeln.
Fonden för Handels- och Distributionsforskning, EFI.

1995

Becker, T., Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption.
Blomberg, J., Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Styrelsearbete i koncerner
Brännström, T., Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.
Ekonomisk politik i omvandling. red. Jonung, L. **Gunnarsson, J., Wahlund, R.**, Hushållens finansiella strategier. En explorativ studie.
Fonden för Handels- och Distributionsforskning, EFI.
Höök, P., Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.
Levin, J., Olsson, P., Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.
Mägi, A., Customer Statisfaction in a Store Performance Framework. Research Report.
Persson, P-G., Modeling the Impact of Sales Promotion on Store Profits.
Roman, L., Institutions in Transition. A Study of Vietnamese Banking.
Sandberg, J., Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.
Sandberg, J., How Do We Justify Knowledge Produced by Interpretative Approaches?Research Report.
Schuster, W., Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering.
Söderberg, K., Farmartjänsten - Ny kooperation inom lantbruket. Research Report.
Söderqvist, T., Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation.
Thorén, B., Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor.

1996

Advancing your Business. People and Information Systems in Concert. red. Lundeberg, M., Sundgren, B.
Att föra verksamheten framåt. Människor och informationssystem i samverkan. red. Lundeberg, M., Sundgren, B.
EFI/Studentlitteratur
Andersson, P., Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.
Andersson, P., The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.
Asplund, M., Essays in Industrial Economics.
Delmar, F., Entrepreneurial Behavior & Business Performance.
Edlund, L., The Marriage Market: How Do You Compare?
Gunnarsson, J., Three Studies of Financial Behavior. Research Report.
Hedborg, A., Studies of Framing, Judgment and Choice.
Holgersson, C., Höök, P., Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.
Holmberg, C., Stores and Consumers - Two Perspectives on Food Purchasing.

Håkansson, P., Wahlund, R., Varumärken. Från teori till praktik.
Fonden för Handels- och Distrubitionsforskning, Ekonomiska Forskningsinstitutet, EFI.
Karlsson, A., The Family Business as an Heirloom. Research Report.
Linghag, S., Man är handelsstudent. Research Report.
Molin, J., Essays on Corporate Finance and Governance.
Mägi, A., The French Food Retailing Industry - A Descriptive Study. The Foundation for Distribution Research.
Mölleryd, B., Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk obiltelefoni. Research Report.
Nielsen, S., Omkostningskalkulation for avancerede produktionsomgivelser - en sammenligning af stokastiske og deterministiske omkostningskalkulationsmodeller.
Normark, P., Danielsson, L., Larsson, A., Lundblad, P., Kooperativa nyckeltal. Research Report.
Sandin, R., Heterogeneity in Oligopoly: Theories and Tests.
Sandén-Håkansson, U., Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.
Stein, J., Söderlund, M., Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildnings-kvalitet. En empirisk studie av civilekonomer. Research Report.
Strömberg, P., Thorburn, K., An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.
Söderlund, M., Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995.
Ulfsdotter, U., Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyoghurt 1982-1994.
Westelius, A., A Study of Patterns of Communication in Management Accounting and Control Projects.
Wijkström, F., Den svenska ideella sektorn och pengarna. Research Report.
Örtendahl, M., Health and Time - A Problem of Discounting.

1997

Alexius, A., Essays on Exchange Rates, Prices and Interest Rates.
Andersson, B., Essays on the Swedish Electricity Market.
Berggren, N., Essays in Constitutional Economics.
Ericsson, J., Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.
Friberg, R., Prices, Profits and Exchange Rates.
Från optionsprissättning till konkurslagstiftning. red. Bergström, C., Björk, T.
Gang, Z., Chinese Rural Enterprises Between Plan and Market.
Hagerud, G.E., A New Non-Linear GARCH Model.
Haksar, A., Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka
Nilsson, H.E., Using Conceptual Data Modelling in Management Accounting: A Case Study.
Löthgren, M., Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.
Sjöberg, L., Ramsberg, J., En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.
Säfvenblad, P., Price Formation in Multi-Asset Securities Markets.
Södergren, B., På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.
Tambour, M., Essays on Performance Measurement in Health Care.
Thorén, B., Berg-Surwee, U., Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.
Åsbrink, S., Nonlinearities and Regime Shifts in Financial Time Series.