Four Essays on the Econometric Modelling of Volatility and Durations

EFI THE ECONOMIC RESEARCH INSTITUTE

EFI Mission

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

Research Organization

The research activities at the Institute are organized into 20 Research Centres.
Centre Directors are professors at the Stockholm School of Economics.

EFI Research Centre:	*Centre Director:*
Management and Organisation (A)	Sven-Erik Sjöstrand
Centre for Entrepreneurship and Business Creation (E)	Carin Holmquist
Public Management (F)	Nils Brunsson
Information Management (I)	Mats Lundeberg
Centre for People and Organization (PMO)	Andreas Werr
Centre for Innovation and Operations Management (T)	Pär Åhlstrm
Centre for Media and Economic Psychology (P)	Richard Wahlund
Centre for Consumer Marketing (CCM)	Magnus Söderlund
Centre for Information and Communication Research (CIC)	Per Andersson
Marketing, Distribution and Industrial Dynamics (D)	Björn Axelsson
Centre for Strategy and Competitiveness (CSC)	Örjan Sölvell
Accounting and Managerial Finance (B)	Johnny Lind
Centre for Financial Analysis and Managerial Economics in Accounting (BFAC)	Kenth Skogsvik
Finance (FI)	Clas Bergström
Centre for Health Economics (CHE)	Bengt Jönsson
International Economics and Geography (IEG)	Mats Lundahl
Economics (S)	Paul Segerstrom
Economic Statistics (ES)	Anders Westlund
Centre for Business Law (RV)	Johnny Herre
Centre for Tax Law (SR)	Bertil Wiman

Chair of the Board: Professor Carin Holmquist
Director: Associate Professor Filip Wijkström

Address

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

Four Essays on the Econometric Modelling of Volatility and Durations

Cristina Amado

EFI THE ECONOMIC RESEARCH INSTITUTE

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics 2009.

© EFI and the author, 2009
ISBN 978-91-7258-794-6

Keywords:
Financial econometrics; Model specification; Misspecification testing; Nonstationarity; GARCH; Unconditional heteroskedasticity; Long financial time series; Ultra-high-frequency data; Diurnal Variation.

Printed by:
Elanders, Vällingby 2009

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

To my parents, Arminda e José,
for their example

The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say.

The Road goes ever on and on
Out from the door where it began.
Now far ahead the Road has gone,
Let others follow it who can!
Let them a journey new begin,
But I at last with weary feet
Will turn towards the lighted inn,
My evening-rest and sleep to meet.

J.R.R. Tolkien in "The Lord of the Rings"

Contents

Acknowledgements

My adventure in Scandinavian lands and frozen seas is finally coming to an end. Many were the frustrating days, sleepless nights, lonely moments and an uncountable number of lost weekends during my voyage of discovery in the land of Econometric research. The next paragraphs are dedicated to everyone who helped me to see the sunshine in the long dark days during my lonely expedition.

First of all, I would like to express my sincere gratitude to my supervisor, Timo Teräsvirta, for his excellent guidance in the rough hills of Nonlinear Time Series Models. I am particularly thankful to him for being always available to listen to me, for his invaluable comments and for his tolerance during the teaching semesters I spent in Portugal. His exceptional skills and extraordinary intuition have helped me so many times to see the light at the end of the tunnel. It has been an absolute privilege to learn from and to work with him.

Special thanks go to Sune Karlsson who has given me the opportunity to pursue my Ph.D. studies at the former Department of Economic Statistics at Stockholm School of Economics. I also would like to thank Stefan Lundbergh for his guidance and for his extremely useful advices at an early stage of this thesis. Many thanks to my fellow Ph.D. students and to the faculty at SSE for providing a friendly and stimulating working environment. I am also indebted to Pernilla Watson for helping me with administrative problems.

This thesis has benefited greatly from the inspiring research atmosphere at CREATES, University of Aarhus, which I had the opportunity to visit twice during the Spring. Very special thanks go to Niels Haldrup for making my visit possible. Both visits have been financed by the Louis Fraenckels Stiependiefond.

I also express my gratitude to the direction of the Department of Economics at the University of Minho for allowing me, in the past three years, to spend the Spring semesters abroad to work on my doctoral thesis.

Many of my thanks go to my dear friend Jana Eklund for her unconditional help in practical matters surrounding the editing of this thesis. Thank you so much, Jana! I am also extremely grateful to Andrés González. He has taught me how to code my first simulations in Ox programming language when a computer code seemed to me a collection of ancient Egyptian Hieroglyphs.

To all my friends I met in Stockholm, I thank you for making my stay more pleasant and less lonely. My warmest thanks go to Agatha Murgoci, Ana Beleza, Birgit Strikholm, Helinä Laakkonen, Ingvar Strid, Jaewon Kim, Mika Meitz, Milo Bianchi, Mirco Tonin, Rickard Sandberg, Sam Azasu, Tomoaki Nakatani and Zhenfang Zhao. Our social events and nice talks will always be remembered.

Thanks to my friend Ganesh with whom I shared so many happy and not so happy moments in the past years. He has been my office mate, neighbour and my best friend when I was in Stockholm. Thank you for your ever-patience, for the funny talks and for the valuable discussions about the fundamental things in life.

Several colleagues and friends in Portugal also deserve a special mention. Special thanks go to all my dearest friends from the University of Évora who have always believed in me. I really appreciate your loyal friendship and, in my heart, our good *old* times will always be remembered as the best moments of my life. I am also very grateful to Joana Girante and Luís Aguiar-Conraria for being always there for me and for their boosting morale. I am thankful to them and to Miguel Portela for helping me in midterm proctoring and grading when I was away. Carla Monteiro and Henrique Cizeron are the best Vets one can have for their pets. Thank you so much for your friendship and especially for your patience during the exorcism sessions, days before my way back to Stockholm.

Special thanks are dedicated to all my family for their unconditional love and for cheering me on. My dearest aunts, Ivone Silva and Odete Gregório, have never forgot to call me on the important event dates and have always shown me their support throughout these years.

Thanks to my traveller companion, Juno, for making my long working hours during the night a little more sweet and for showing me with his meows and purrs that life is much more than research.

I especially need to thank Filipe Cizeron for the encouraging words and for being on the other side of the phone when I most needed. Without his motivation for doing my doctoral studies abroad it would not be possible to write this dissertation. Thank you for sharing my joys and for smoothing my way through the troubled times.

Special gratitude goes to my parents for giving me the opportunity to learn despite their sacrifice. To my dearest mother, I am deeply thankful for her unconditional love and encouragement in my most challenging endeavors. To my beloved father, for being my source of inspiration, who missed sharing this moment but always has a special place in my heart. To them I dedicate this thesis.

Financial support from the Stockholm School of Economics and the University of Minho is gratefully acknowledged.

Stockholm, May 2009

Cristina Amado

Chapter 1

Introduction

This thesis consists of four research chapters in the area of financial econometrics on topics of the modelling of financial market volatility and the econometrics of ultra-high-frequency data. The aim of the thesis is to develop new econometric methods for modelling and hypothesis testing in these areas. A brief introduction to those research areas and a short description of the specific topics in the chapters follows. For a more detailed overview of the contents of the chapters, the reader is referred to the introductions of the individual chapters.

When making investment decisions, volatility is commonly regarded by market investors as a measure of risk. Modelling volatility is therefore essential in many financial areas such as portfolio diversification, risk management, and derivative asset pricing. The models can be then used for forecasting volatility of stock prices, strike prices or interest rates. The success of the volatility model will depend on how well it predicts and captures the characteristics of financial data. Financial market volatility is also of great importance in financial regulation, monetary policy and economic activity. The central role of risk (or volatility) in financial decision making and the ample evidence that the measures of risk exhibit stochastic behaviour through time have stimulated the development of many sophisticated tools in the field of time series econometrics.

The Autoregressive Conditional Heteroskedastic class of models introduced by Engle (1982)[1] was designed to parameterize time-varying volatility. The Generalized ARCH (GARCH) process defined by Bollerslev (1986) specifies present volatility as a function of past volatilities in addition of past squared returns. The GARCH model is able to capture the temporal dependence in financial time series by allowing the investors to update their risk expectations when new information becomes available on the market. Moreover, the GARCH model also successfully accommodates some special features of financial data such as volatility clustering and excess kurtosis. Since its introduction, richer parameterizations and numerous extensions to the GARCH model have been suggested to increase the flexibility of the original model.

A vast literature focusing on the implications of the assumption of parameter constancy in the GARCH model has been developed in recent years. The occurrence of social, polical or economic events during a long time span may make the structure of volatility to change over time, making the series to become nonstationary. For this reason, as Mikosch and Stărică (2004) documented, the assumption of stationarity (or parameter constancy) may not be very appropriate when the series to be modelled is sufficiently long. In applications it is often found that the sum of the estimated GARCH parameters (excluding the intercept) is close to unity. This so-called 'integrated GARCH effect' may be well explained by occasional level shifts in the intercept of the GARCH model; Diebold (1986) and Lamoureux and Lastrapes (1990). This means that the high persistence (or the observed long-memory) in stock market volatility may not be an inherent feature of the financial data but it can be explained by neglected structural breaks in the variance process.

[1] Robert F. Engle was awarded in 2003 the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel "for methods of analyzing economic time series with time-varying volatility (ARCH)".

Some modelling proposals have been suggested to accommodate deterministic changes in the volatility. One possibility is to use Markov-switching GARCH-type processes for modelling sudden breaks in the volatility at specific points in time. Alternatively, one may consider that volatility is parameterized by a 'smoothly' nonstationary process. One of such models is the spline-GARCH of Engle and Gonzalo Rangel (2008) in which volatility is multiplicatively decomposed into stationary and nonstationary components. More specifically, the nonstationary component is modelled using an exponential spline, and the stationary component is described as a GARCH process.

The chapter *"Modelling Conditional and Unconditional Heteroskedasticity with Smoothly Time-Varying Structure"*[2] introduces a new model, the time-varying GARCH (TV-GARCH) model, in which volatility has a smooth time-varying structure of either additive or multiplicative type. To characterize smooth changes in the (un)conditional variance we assume that the parameters vary smoothly over time according to the logistic transition function. As a result, the parameterizations provide very flexible representations of volatility, and they can describe many types of nonstationary behaviour. These parametric alternatives are particularly useful in applications for modelling long financial data where the non-constancy of parameters becomes an issue. Testing parameter constancy is therefore an important tool for checking the adequacy of the model. For this reason, we provide a modelling framework relying on statistical inference to specify the parametric structure of the TV-GARCH models. We first test the standard GARCH model against these time-varying alternatives and, in case of the rejection of the null hypothesis, determine the structure of the time-varying component is from the data. This is done by testing a sequence of hypotheses by Lagrange multiplier tests presented in the chapter. Misspecification tests are also provided for evaluating the adequacy of the estimated model.

Finite-sample properties of the test statistics and sequential testing are examined by simulation. The Monte Carlo experiments suggest that these procedures have reasonable good properties already in samples of moderate size. The model building strategy is illustrated with an application to the daily S&P 500 index and the spot SPD/USD exchange rate returns. The results show that the tests strongly reject the hypothesis of parameter constancy against the time-varying GARCH alternatives for the two return series. Moreover, our findings suggest that the long-memory type behaviour of the sample autocorrelation functions of the absolute or squared returns can also be explained by deterministic changes in the unconditional variance.

In some applications, the time series used for fitting a GARCH model covers decades of economic activity. For such long series, one may inevitably expect periods of turbulence such as recessions and, possibly, deterministic shifts. In those cases, the assumption of constant unconditional variance of the GARCH model turns out to be too restrictive. Shifts in the unconditional variance then affect the estimation towards an IGARCH model as documented in Lamoureux and Lastrapes (1990).

[2]This is a joint work with Timo Teräsvirta.

Consequently, modelling deterministic changes in the second unconditional moment of the returns is important when the time series covers a long period.

The chapter *"Modelling Changes in the Unconditional Variance of Long Stock Return Series"*[3] addresses the issue of modelling deterministic changes in the unconditional variance over a long return series. For this purpose, we assume that volatility is modelled by a multiplicative decomposition of both conditional and unconditional variance. More specifically, the conditional variance component is parameterized by a GARCH-type model and it describes the short-run dynamics. The unconditional variance component is assumed to be vary slowly over time and it is modelled using a linear combination of logistic transition functions. The structure of the time-varying component is specified using a testing sequence which is similar to the one within the TV-GARCH framework. In order to facilitate the specification, the long series is splitted into non-overlapping subperiods, and parameter estimation in this modelling framework requires special care. The modelling strategy is illustrated with an application to the daily returns of the Dow Jones Industrial Average (DJIA) index from 1920 until 2003. The empirical results sustain the hypothesis that the assumption of constancy of the unconditional variance is not adequate over long return series and indicate that deterministic changes in the unconditional variance may be associated with macroeconomic factors. The observed long-memory property observed in the original series is weakened when the deterministic changes in the unconditional variance are incorporated into the model.

Many financial considerations do not only rely on the behaviour of an individual asset. Instead, standard tools applied by financial analysts typically use information about the covariances or correlations between asset returns. This has motivated the modelling of volatility using multivariate financial time series rather than modelling individual returns separately. However, the growing literature on multivariate GARCH models has so far paid little attention on modelling multivariate financial data with nonstationary volatilities. Recently, Hafner and Linton (2008) proposed a semiparametric generalization of the spline-GARCH model of Engle and Gonzalo Rangel (2008) in which the parametric component is a first-order BEKK model. The authors suggested an estimation procedure for the parametric and nonparametric components and established semiparametric efficiency of their estimators.

In the chapter *"Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations"*[4] we propose an extension of the univariate multiplicative TV-GARCH model to the multivariate Conditional Correlation GARCH (CC-GARCH) framework. The variance equations are parameterized such that they combine the long-run and the short-run dynamic behaviour of the volatilities. In this framework, the long-run behaviour is described by the individual unconditional variances, and it is allowed to vary smoothly over time according to the logistic transition function. Our model differs from the semiparametric model of Hafner and Linton (2008) in the sense that a data-based modelling technique is

[3]This is a joint work with Timo Teräsvirta.

[4]This is a joint work with Timo Teräsvirta.

used for specifying the deterministic time-varying component. It may be of interest to investigate how careful specification of the individual variances affects the correlation structure of several CC-GARCH models. The effects of modelling the nonstationary variance component are examined empirically using pairs of seven daily stock return series from the S&P 500 index. According to the results, the nature and magnitude of the effects on the correlation estimates depend on the correlation structure matrix of the model. The fit of the CC-GARCH models to the data is remarkably improved through taking nonstationarity in variances into account. Another advantage of this framework is that we are able to generalize the news impact surfaces of Kroner and Ng (1998) such that they can vary over time. The so-called time-varying news impact surfaces are now able to distinguish between responses at different levels of turbulence in the market as well as at different correlation levels.

With the increasing availability of intraday databases, new methods in time series econometrics are needed to investigate the recorded information of these more detailed and complex datasets. The available information usually contains the precise time ("time-stamp") at which the order in the stock market has been executed and other associated characteristics with the trade. Transaction data containing recorded financial events at the highest frequency possible are defined as ultra-high-frequency data. An inherent feature of such data is that the events are irregularly spaced. Standard tools of time series econometrics are thus inadequate for such time series as they are based on regularly spaced data. This has contributed to the birth of a new branch of financial econometrics where the so-called high-frequency models have been introduced to account the irregular spacing of the data.

The pioneering work on this area was originated with the class of Autoregressive Conditional Duration (ACD) models of Engle and Russell (1998) and Engle (2000). The focus in this work was on modelling the time elapsed between two market events (or duration). Duration data share some of the stylized facts present in financial data such as duration clustering, high persistence and, among others, fat tails. Another documented feature of the data is their systematic pattern over the day as trading activity tends to be more active near the opening and closing of the market than in the midday. In the chapter "*A Smooth Transition Approach to Modelling Diurnal Variation in Models of Autoregressive Conditional Duration*"[5] we propose a new parameterization for describing the diurnal component. This is done by allowing the durations to change smoothly over the day according to the logistic transition function. For the purpose, we provide a modelling framework for specifying the parameteric structure of the systematic pattern over the trading day. An application to the IBM trade durations suggests that the diurnal pattern may not always have the shape proposed earlier: short durations early and late in the day and lower activity in the middle. For this reason, one should proceed with care when specifying the diurnal component. The estimation of the ACD model should be then preceded by a specification search to determine the structure of the diurnal variation.

[5]This is a joint work with Timo Teräsvirta.

Bibliography

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

DIEBOLD, F. X. (1986): "Comment on "Modelling the Persistence of Conditional Variance", by R. Engle and T. Bollerslev," *Econometric Reviews*, 5, 51–56.

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

ENGLE, R. F. (2000): "The Econometrics of Ultra-High-Frequency Data," *Econometrica*, 68, 1–22.

ENGLE, R. F., AND J. GONZALO RANGEL (2008): "The Spline-GARCH Model for Low-Frequency Volatility and its Global Macroeconomic Causes," *Review of Financial Studies*, 21, 1187–1222.

ENGLE, R. F., AND J. R. RUSSELL (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transition Data," *Econometrica*, 66, 1127–1162.

HAFNER, C., AND O. LINTON (2008): "Efficient Estimation of a Multivariate Multiplicative Model," Unpublished paper.

KRONER, K. F., AND V. K. NG (1998): "Modeling Asymmetric Comovements of Asset Returns," *Review of Financial Studies*, 11, 817–844.

LAMOUREUX, C. G., AND W. D. LASTRAPES (1990): "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business & Economic Statistics*, 8, 225–234.

MIKOSCH, T., AND C. STĂRICĂ (2004): "Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects," *The Review of Economics and Statistics*, 86, 378–390.

Chapter 2

Modelling Conditional and Unconditional Heteroskedasticity with Smoothly Time-Varying Structure

Modelling Conditional and Unconditional Heteroskedasticity with Smoothly Time-Varying Structure [1]

Abstract

In this paper, we propose two parametric alternatives to the standard GARCH model. They allow the conditional variance to have a smooth time-varying structure of either additive or multiplicative type. The suggested parameterizations describe both nonlinearity and structural change in the conditional and unconditional variances where the transition between regimes over time is smooth. A modelling strategy for these new time-varying parameter GARCH models is developed. It relies on a sequence of Lagrange multiplier tests, and the adequacy of the estimated models is investigated by Lagrange multiplier type misspecification tests. Finite-sample properties of these procedures and tests are examined by simulation. An empirical application to daily stock returns and another one to daily exchange rate returns illustrate the functioning and properties of our modelling strategy in practice. The results show that the long memory type behaviour of the sample autocorrelation functions of the absolute returns can also be explained by deterministic changes in the unconditional variance.

---

[1] This paper is a joint work with Timo Teräsvirta.

Acknowledgements: This research has been supported by the Danish National Research Foundation. Material from this paper has been presented at the International Symposium on Econometric Theory and Applications, Xiamen, April 2006; 5th Annual International Conference 'Forecasting Financial Markets and Economic Decision-making', Lodz, May 2006; 13th International Conference on 'Forecasting Financial Markets', Marseille, May-June 2006; 26th International Symposium on Forecasting, Santander, June 2006; Workshop 'Volatility day', Stockholm, November 2006; Nordic Econometric Meeting, Tartu, May 2007; Symposium on "Long Memory", Aarhus, June-July 2007; LACEA-LAMES, Bogotá, October 2007; and at the seminars at Banca d'Italia, Rome, European University Institute, Florence, Humboldt University, Berlin, University of Minho, Braga, Stockholm School of Economics, Leonard N. Stern School of Business at New York University, and University of Vilnius. We would like to thank the participants at these occasions for their comments, and Stefan Lundbergh, Mika Meitz, Anders Rahbek and Esther Ruiz for useful discussions and suggestions. The responsibility for any errors and shortcomings in this article remains ours.

2.1 Introduction

The modelling of time-varying volatility of financial returns has been a flourishing field of research for a quarter of a century following the introduction of the Autoregressive Conditional Heteroskedasticity (ARCH) model by Engle (1982) and the Generalized ARCH (GARCH) model developed by Bollerslev (1986). The increasing popularity of the class of GARCH models has been mainly due to their ability to describe the dynamic structure of volatility clustering of stock return series, specifically over short periods of time. However, one may expect that economic or political events or changes in institutions cause the structure of volatility to change over time. This means that the assumption of stationarity may be inappropriate under the evidence of structural changes in financial return series. Recently, Mikosch and Stărică (2004) argued that stylized facts in financial return series such as the long-range dependence and the 'integrated GARCH effect' can be well explained by unaccounted structural breaks in the unconditional variance (see also Lamoureux and Lastrapes (1990)). Diebold (1986) was the first to suggest that occasional level shifts in the intercept of the GARCH model can bias the estimation towards an integrated GARCH model.

Another line of research has focussed on explaining nonstationary behaviour of volatility by long-memory models, such as the Fractionally Integrated GARCH (FIGARCH) model by Baillie, Bollerslev, and Mikkelsen (1996). The FIGARCH model is not the only way of handling the 'integrated GARCH effect' in return series. Baillie and Morana (2007) generalized the FIGARCH model by allowing a deterministically changing intercept. Hamilton and Susmel (1994) and Cai (1994) suggested a Markov-switching ARCH model for the purpose, and their model has later been generalized by others. One may also assume that the GARCH process contains sudden deterministic switches and try and detect them; see Berkes, Gombay, Horváth, and Kokoszka (2004) who proposed a method of sequential switch or change-point detection.

Yet another way of dealing with high persistence would be to explicitly assume that the volatility process is 'smoothly' nonstationary and model it accordingly. Dahlhaus and Subba Rao (2006) introduced a time-varying ARCH process for modelling nonstationary volatility. Their tvARCH model is asymptotically locally stationary at every point of observation but it is globally nonstationary because of time-varying parameters. Engle and Gonzalo Rangel (2008) assumed that the variance of the process of interest can be decomposed into two components, a stationary and a nonstationary one. The nonstationary component is described by using splines, and the stationary component follows a GARCH process. The parameters of the latter are estimated conditionally on the spline component.

In this paper, we introduce two nonstationary GARCH models for situations in which volatility appears to be nonstationary. First, we propose an additive time-varying parameter model, in which a directly time-dependent component is added to the GARCH specification. In the second alternative, the variance is multiplicatively decomposed into the stationary and nonstationary component as in Engle and Gonzalo Rangel (2008). These two alternatives are quite flexible representations of volatility and can describe many types of nonstationary behaviour. We emphasize the role of model building in this approach. The standard GARCH model is first tested against these time-varying alternatives. If the null hypothesis is rejected, the structure of the time-varying component of the model is determined using the data. This is done by testing a sequence of hypotheses, and these tests are presented in the paper. After parameter estimation, the model is evaluated by misspecification tests, following the ideas in Eitrheim and Teräsvirta (1996) and Lundbergh and Teräsvirta (2002).

The outline of this paper is as follows. In Section 2.2 we present the new Time-Varying (TV-) GARCH model and discuss some of its properties. In Section 2.3 we derive LM parameter constancy tests against an additive and a multiplicative alternative. In Section 2.4 we present a modelling strategy for both specifications. Details regarding the estimation are discussed in Section 2.5, and diagnostic tests for the TV-GARCH model are given in Section 2.6. Section 2.7 contains simulation results on the empirical performance of the tests and the specification strategy. In Section 2.8 we apply our modelling cycle to both stock and exchange rate returns. Finally, Section 2.9 contains concluding remarks.

2.2 The model

Let the model for an asset or index return y_t be

$$y_t = \mu_t + \varepsilon_t$$

where $\{\varepsilon_t\}$ is an innovation sequence with the conditional mean $\mathsf{E}(\varepsilon_t|\mathcal{F}_{t-1}) = 0$ and a potentially time-varying conditional variance $\mathsf{E}(\varepsilon_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$, and $\mathcal{F}_{t-1}$ is the sigma-field generated by the available information until $t-1$. We assume that $\mathsf{E}(y_t|\mathcal{F}_{t-1}) = 0$, because our focus will be on the conditional variance σ_t^2. More precisely, define

$$\varepsilon_t = \zeta_t \sigma_t \tag{2.1}$$

where $\{\zeta_t\}$ is a sequence of independent standard normal variables. Furthermore, assume that σ_t^2 is a time-varying representation measurable with respect to $\mathcal{F}_{t-1}$ with either an additive structure

$$\sigma_t^2 = h_t + g_t \tag{2.2}$$

or a multiplicative one

$$\sigma_t^2 = h_t g_t. \tag{2.3}$$

The function h_t is a component describing conditional heteroskedasticity in the observed process y_t, whereas g_t introduces nonstationarity. Thus, we assume that h_t follows the standard GARCH(p,q) model of Bollerslev (1986):

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{2.4}$$

Then the GARCH(p,q) model is nested in (2.2) when $g_t \equiv 0$ and in (2.3) when $g_t \equiv 1$. More generally, when (2.3) holds, ε_{t-i}^2 is replaced by $\varepsilon_{t-i}^2/g_{t-i}, i = 1, \ldots, q,$

in (2.4). Both parameterizations (2.2) and (2.3) define a time-varying parameter GARCH model.

In order to characterize smooth changes in the conditional variance we assume that the parameters in (2.4) vary smoothly over time. This is done by defining the function g_t in (2.2) as follows:

$$g_t = (\alpha_0^* + \sum_{i=1}^{q} \alpha_i^* \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j^* h_{t-j}) G(t^*; \gamma, \mathbf{c}), \tag{2.5}$$

where $G(t^*; \gamma, \mathbf{c})$ is the so-called transition function which is a continuous and non-negative function bounded between zero and one. Furthermore, $t^* = t/T$, where T is the number of observations. A suitable choice for $G(t^*; \gamma, \mathbf{c})$ is the general logistic smooth transition function defined as follows:

$$G(t^*; \gamma, \mathbf{c}) = \left(1 + \exp\left\{-\gamma \prod_{k=1}^{K} (t^* - c_k)\right\}\right)^{-1}, \ \gamma > 0, \ c_1 \leq c_2 \leq \ldots \leq c_K. \tag{2.6}$$

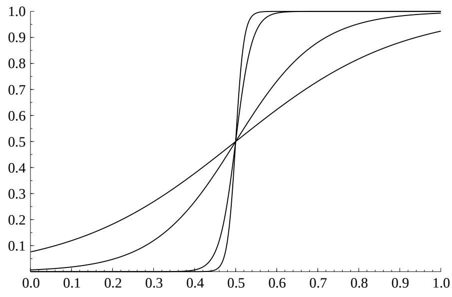
This transition function is such that the parameters of the GARCH model (2.1)-(2.2) fluctuate smoothly over time between (α_i, β_j) and $(\alpha_i + \alpha_i^*, \beta_j + \beta_j^*)$, $i = 0, 1, \ldots, q$, $j = 1, \ldots, p$. The slope parameter γ controls the degree of smoothness of the transition function. When $\gamma \longrightarrow \infty$, the switch from one set of parameters to another in (2.2) is abrupt, that is, the process contains structural breaks at $c_1, c_2, \ldots, c_K$. The order $K \in \mathbb{Z}_+$ determines the shape of the transition function. Typical choices for the transition function in practice are $K = 1$ and $K = 2$. These are illustrated in Figure 2.1 for a set of values for γ, c_1, and c_2. One can observe that large values of γ increase the velocity of transition from 0 to 1 as a function of t^*. When $\gamma \longrightarrow \infty$, a smooth parameter change approaches a structural break because then the process switches instantaneously over time from one regime to another. The TV-GARCH model with $K = 1$ is suitable for describing return processes whose volatility dynamics are different before and after the smooth structural change. When $K = 2$, the parameters first change and eventually move back to their original values.

More generally, one can define an extended version of the additive TV-GARCH model allowing for more than one transition function. A multiple TV-GARCH model can be obtained by adding r transition functions as follows

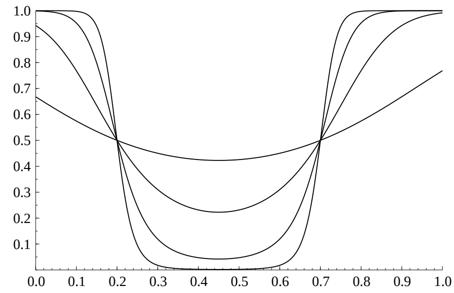
$$g_t = \sum_{l=1}^{r} (\alpha_{0l} + \sum_{i=1}^{q} \alpha_{il} \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{jl} h_{t-j}) G_l(t^*; \gamma_l, \mathbf{c}_l) \tag{2.7}$$

where $G_l(t^*; \gamma_l, \mathbf{c}_l), l = 1, \ldots, r$, are logistic functions as in (2.6) with smoothness parameter γ_l and a threshold parameter vector $\mathbf{c}_l$. The parameters in (2.4) and (2.7) satisfy the restrictions $\alpha_i + \sum_{l=1}^{j} \alpha_{il} > 0, i = 0, \ldots, q; \forall j = 1, \ldots, r$ and $\beta_i + \sum_{l=1}^{j} \beta_{il} \geq 0, i = 1, \ldots, p; \forall j = 1, \ldots, r$. These conditions are sufficient to guarantee strictly positive conditional variances.

The model (2.2), (2.4) and (2.7) is an additive TV-GARCH model whose intercept, ARCH and GARCH parameters are time-varying. This implies that the model is



(a)



(b)

Figure 2.1 Plots of the logistic transition function (2.6) for: (a) $K = 1$ with location parameter $c_1 = 0.5$; and (b) $K = 2$ with location parameters $c_1 = 0.2$ and $c_2 = 0.7$ for $\gamma = 5, 10, 50$, and 100 where the lowest value of γ corresponds to the smoothest function.

capable of accommodating systematic changes both in the "baseline volatility" (or unconditional variance) and in the amplitude of volatility clusters. Such changes cannot be explained by a constant parameter GARCH model.

Function (2.7) with $r > 1$ is extremely flexible and probably makes the model difficult to estimate in practice. A more applicable but still flexible model is obtained by only letting the "baseline volatility" or the intercept to change smoothly over time. This leads to the following definition for g_t:

$$g_t = \sum_{l=1}^{r} \alpha_{0l} G_l(t^*; \gamma_l, \mathbf{c}_l). \tag{2.8}$$

It may be mentioned that Baillie and Morana (2007) recently proposed a GARCH model which also has a deterministically time-varying intercept. It is modelled using the flexible functional form of Gallant (1984) based on the Fourier decomposition. Their model differs from our time-varying intercept model in the sense that it is in other respects a FIGARCH model, and the authors called it the Adaptive FIGARCH model.

In the GARCH(p, q) model, the unconditional variance of the returns is constant over time, that is, $\mathsf{E}(\varepsilon_t^2) = \alpha_0/(1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j)$ if and only if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$. However, this assumption is not consistent with the behaviour of the volatilities of the stock market returns if the dynamic behaviour of volatility changes in the long run. The additive TV-GARCH model with a time-varying intercept is capable of generating changes in the dynamics of the unconditional variance over time. The model (2.2), (2.4) and (2.8) can be seen as a GARCH(p, q) model with a stochastic time-varying intercept fluctuating smoothly over time between α_0 and $\alpha_0 + \sum_{l=1}^{r} \alpha_{0l} G_l(t^*; \gamma_l, \mathbf{c}_l)$. Therefore, it can generate smooth changes over time in the "baseline volatility". Hence, such parameterization can explain the systematic movements of the conditional variance as in the GARCH model but relaxing the assumption of constancy of the unconditional volatility.

Consider again the model (2.2), (2.4) and (2.7) and assume that $\alpha_{0l} = \alpha_0\delta_l, \alpha_{il} = \alpha_i\delta_l, i = 1, \ldots, q; \beta_{jl} = \beta_j\delta_l, j = 1, \ldots, p$. Furthermore, assume $\delta_l > 0, l = 1, \ldots, r$, if the transition function $G_l(t^*; \gamma_l, \mathbf{c}_l)$ is increasing over time. For the case $G_l(t^*; \gamma_l, \mathbf{c}_l)$ is a decreasing function assume $\sum_{l=1}^{r} \delta_l < 1$ for $l = 1, \ldots, r$. Imposing these restrictions on (2.7) and rewriting (2.2) yields

$$\sigma_t^2 = h_t(1 + \sum_{l=1}^{r} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)). \tag{2.9}$$

Setting $g_t = 1 + \sum_{l=1}^{r} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)$ in (2.9) gives the multiplicative representation (2.3). It is thus seen to be a special case of the additive TV-GARCH model (2.2), (2.4) and (2.7). The multiplicative model has a straightforward interpretation. Writing it in terms of (2.1) as

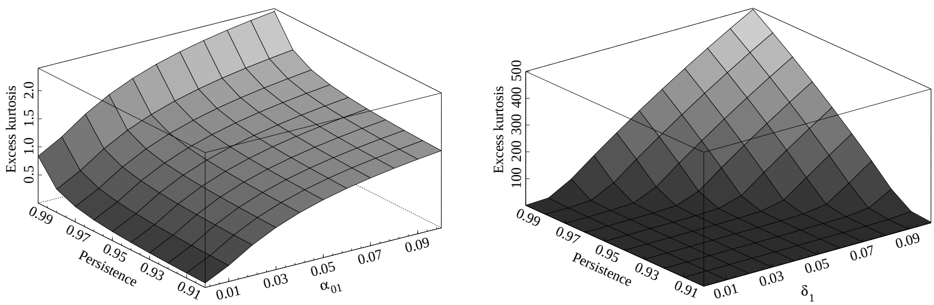
$$\phi_t = \varepsilon_t / g_t^{1/2} = \zeta_t {h_t}^{1/2} \tag{2.10}$$

it is seen that ϕ_t has a constant unconditional variance $\mathsf{E}h_t$ and, moreover, that ϕ_t has a standard stationary GARCH(p, q) representation h_t. Turning (2.10) around, one obtains that $\psi_t = \varepsilon_t / h_t^{1/2}, t = 1, \ldots, T$, form a sequence of independent but not identically distributed observations, as the unconditional variance of ψ_t changes smoothly as a function of time.

We consider properties of both time-varying GARCH specifications by generating 1000 replications with Gaussian errors each with 5000 observations. Figure 2.2 illustrates the relation of the average excess kurtosis of the two models given the persistence and the time-varying constants α_{01} and δ_1. The degree of persistence, measured by the sum $\alpha_1 + \beta_1$, varies between 0.90 and 0.99. The range of parameters α_{01} and δ_1 varies between 0 and 0.1 while $\alpha_0 = 0.01$. Interestingly, simply by assuming normality the proposed models are capable of generating higher kurtosis than the standard GARCH model. Larger values of the time-varying constants generate larger values of the excess kurtosis for both time-varying parameterizations. A high degree of persistence is also able to reproduce heavy-tailed marginal distributions that are often observed in financial return series.

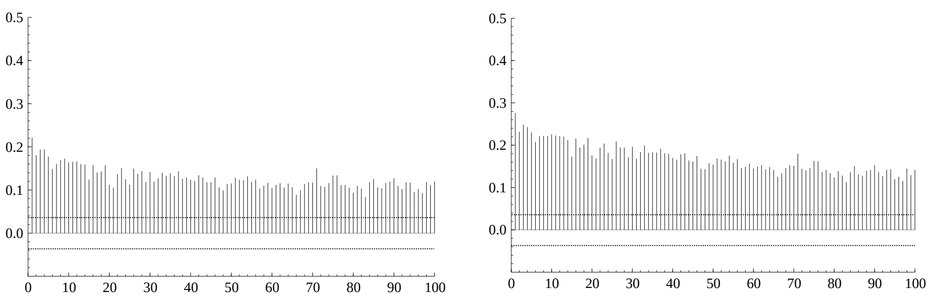
The level of persistence generated by the TV-GARCH models is another property of interest. Figure 2.3 depicts the first 100 autocorrelations of absolute returns of two simulated TV-GARCH processes. The autocorrelations for the additive and multiplicative form are plotted in Figure 2.3(a) and Figure 2.3(b), respectively. The sample length in both cases is 5000 observations. The artificial series are generated with $\alpha_0 = 0.01, \alpha_1 = 0.05, \alpha_{01} = 0.03, \delta_1 = 0.04, \gamma_1 = 10$ and $c_1 = 0.50$. The dotted horizontal lines represent the 95% confidence bounds corresponding to the ACF of an iid Gaussian process. A visual inspection of Figure 2.3 shows that both time-varying specifications can generate long-range dependence looking behaviour.

The dependence structure of each model is also illustrated by the empirical distribution of the GPH estimates of the long-memory parameter d; see Geweke and Porter-Hudak (1983). The results obtained by using absolute values of the returns



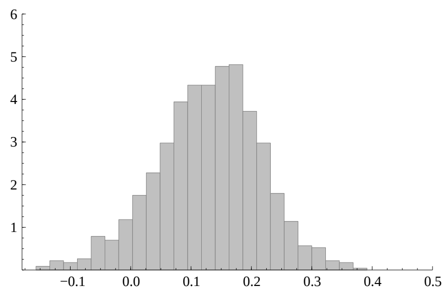
(a) *(b)*

Figure 2.2 Plots of the excess kurtosis, persistence and the constants α_{01} and δ_1 for: (a) an additive TV-GARCH model with a time-varying constant; and (b) a multiplicative TV-GARCH model.

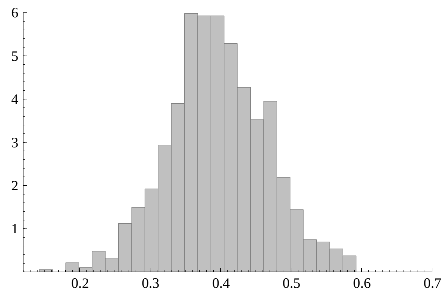


(a) *(b)*

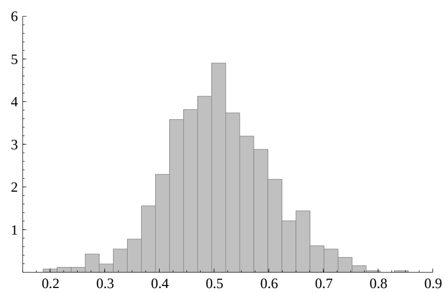
Figure 2.3 Sample autocorrelation functions of absolute returns with the 95% confidence bounds for: (a) an additive TV-GARCH model with a time-varying constant; and (b) a multiplicative TV-GARCH model.



(a)



(b)



(c)

Figure 2.4 Histograms of the GPH long memory parameter estimates for: (a) a GARCH model; (b) an additive TV-GARCH model with a time-varying constant; and (c) a multiplicative TV-GARCH model. The artificial series are generated with $\alpha_0 = 0.01, \alpha_1 = 0.05, \beta_1 = 0.90, \alpha_{01} = 0.03, \delta_1 = 0.04, \gamma_1 = 10$ and $c_1 = 0.50$ for a sample of 5000 observations based on 1000 replications.

are displayed in Figure 2.4. The standard GARCH model is known to have a short memory in the sense that the theoretical autocorrelation function decays to zero at an exponential rate. The exponential decay turns out to be too fast if one wants to adequately describe the high persistence observed in financial data. This may be seen from Figure 2.4(a). If the data are generated by the standard GARCH model, the estimates of the long memory parameter are rather close to zero. However, when the intercept of the GARCH model changes smoothly over time, the degree of the long-memory dependence in the data increases. This is seen from the fact that the empirical distribution for the GPH estimates in Figure 2.4(b) has shifted to the right. As Figure 2.4(c) shows, this effect is even more evident for the TV-GARCH with a multiplicative time-varying structure as more than one half of the probability mass of the empirical distribution of the long-memory parameter is located in the nonstationary area, $d > 0.5$.

2.3 Testing parameter constancy

2.3.1 Testing against an additive alternative

Against the background discussed above, testing parameter constancy is an important tool for checking the adequacy of a GARCH model. If one rejects parameter constancy against a GARCH model with time-varying parameters one may conclude that the structure of the dynamics of volatility is changing over time. Other interpretations cannot be excluded, however, because a rejection of a null hypothesis does not imply that the alternative hypothesis is true. In this section, we propose two parameter constancy tests that allow the parameters to change smoothly over time under the alternative. The first one tests parameter constancy of the GARCH model against an additive TV-GARCH specification. This idea has previously been considered by Lundbergh and Teräsvirta (2002). The second one is a test of constant unconditional variance against the alternative that the variance changes smoothly over time.

We shall first look at the additive alternative where the nonstationary component g_t is defined in (2.5). In order to derive the test statistic rewrite the model as

$$\begin{aligned}
\varepsilon_t &= \zeta_t {h_t}^{1/2}, \qquad \varepsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t) \\
h_t &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \\
&+ (\alpha_{01} + \sum_{i=1}^{q} \alpha_{i1} \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{j1} h_{t-j}) G(t^*; \gamma, \mathbf{c})
\end{aligned} \tag{2.11}$$

where, for simplicity, $r = 1$ and $\mathcal{F}_{t-1}$ is the information set containing all information until $t-1$. The null hypothesis of parameter constancy corresponds to testing $\mathrm{H}_0 : \gamma = 0$ against $\mathrm{H}_1 : \gamma > 0$ in (2.11). Under the null hypothesis, $g_t \equiv 1/2$. One can see that model (2.11) is only identified under the alternative. In particular, when $\gamma = 0$, the parameters $\alpha_{i1}, i = 0, \ldots, q$, and $\beta_{j1}, j = 1, \ldots, p$, as well as $\mathbf{c}$ are not identified. This makes the standard asymptotic inference invalid as the test statistics have a nonstandard asymptotic null distribution. This identification problem was first considered in Davies (1977) and more recently, among others, in Hansen (1996).

In this paper, we circumvent the identification problem following Luukkonen, Saikkonen, and Teräsvirta (1988). Thus we replace the transition function by its first-order Taylor approximation around $\gamma = 0$. Without losing generality, we replace $G(t^*; \gamma, \mathbf{c})$ by $\widetilde{G}(t^*; \gamma, \mathbf{c}) = G(t^*; \gamma, \mathbf{c}) - 1/2$ for notational convenience. From Taylor's theorem one obtains

$$\begin{aligned}\widetilde{G}(t^*;\gamma,\mathbf{c}) &= \widetilde{G}(t^*;0,\mathbf{c}) + \frac{\partial \widetilde{G}(t^*;0,\mathbf{c})}{\partial\gamma}\gamma + R(t^*;\gamma,\mathbf{c}) \\ &= \frac{1}{4}\gamma\prod_{k=1}^{K}(t^*-c_k) + R(t^*;\gamma,\mathbf{c}) \\ &= \sum_{k=0}^{K}\gamma\tilde{c}_k(t^*)^k + R(t^*;\gamma,\mathbf{c}) \end{aligned} \tag{2.12}$$

where $R(t^*;\gamma,\mathbf{c})$ is the remainder term. Replacing $G(t^*;\gamma,\mathbf{c})$ in (2.11) by (2.12) and rearranging terms gives

$$\begin{aligned} h_t &= \alpha_0^* + \sum_{i=1}^{q}\alpha_i^*\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j^* h_{t-j} \\ &+ \sum_{k=1}^{K}(\omega_k(t^*)^k + \sum_{i=1}^{q}\varphi_{ik}(t^*)^k\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\lambda_{jk}(t^*)^k h_{t-j}) + R_1^* \end{aligned} \tag{2.13}$$

where $\alpha_s^* = \alpha_s + \gamma\alpha_{s1}\tilde{c}_0, s = 0,\ldots,q, \beta_j^* = \beta_j + \gamma\beta_{j1}\tilde{c}_0, j = 1,\ldots,p, \omega_k = \gamma\alpha_{01}\tilde{c}_k, \varphi_{ik} = \gamma\alpha_{i1}\tilde{c}_k, i = 1,\ldots,q$, and $\lambda_{jk} = \gamma\beta_{j1}\tilde{c}_k, k = 1,\ldots,K$. The parameters $\tilde{c}_k, k = 0,\ldots,K$, are functions of the original location parameters c_k. In particular, $\tilde{c}_0 = \frac{1}{4}\prod_{k=1}^{K} c_k$ and $\tilde{c}_K = \frac{1}{4}$. Under H_0, the remainder $R_1^* \equiv 0$, so it does not affect the asymptotic null distribution of the test statistic. Using the reparameterization (2.13) it follows that the null hypothesis of parameter constancy becomes

$$\mathrm{H}_0' : \omega_k = \varphi_{ik} = \lambda_{jk} = 0,\ k = 1,\ldots,K,\ i = 1,\ldots,q,\ j = 1,\ldots,p. \tag{2.14}$$

This hypothesis can be tested by a standard LM test. One can also test constancy of a subset of parameters. For example, it may be assumed that $\alpha_{i1} = 0, i = 1,\ldots,q$, and $\beta_{j1} = 0, j = 1,\ldots,p$, which means that only the intercept is time-varying under the alternative. In this case the null hypothesis reduces to $\mathrm{H}_0' : \omega_k = 0, k = 1,\ldots,K$.

In Theorem 1 we present the LM-type statistic for testing parameter constancy against the additive TV-GARCH specification. Under the null hypothesis, the "hats" indicate maximum likelihood estimators and $\hat{h}_t^0$ denotes the conditional variance at time t estimated under H_0.

Theorem 1 *Consider the model (2.13) and let* $\boldsymbol{\theta}_1 = (\alpha_0^*, \alpha_1^*, \ldots, \alpha_q^*, \beta_1^*, \ldots, \beta_p^*)'$ *and* $\boldsymbol{\theta}_2 = (\boldsymbol{\omega}', \boldsymbol{\varphi}_{\boldsymbol{i}}', \boldsymbol{\lambda}_{\boldsymbol{j}}')'$ *where* $\boldsymbol{\omega} = (\omega_1,\ldots,\omega_K)', \boldsymbol{\varphi}_{\boldsymbol{i}} = (\varphi_{i1},\ldots,\varphi_{iK})'$ *and* $\boldsymbol{\lambda}_{\boldsymbol{j}} = (\lambda_{j1},\ldots,\lambda_{jK})'$ *for* $i = 1,\ldots,q$ *and* $j = 1,\ldots,p$. *In addition, denote* $\mathbf{z}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})', \mathbf{Z}_{1t} = [t^{*k}\varepsilon_{t-i}^2](k = 1,\ldots,K, i = 1,\ldots,q)$ *and* $\mathbf{Z}_{2t} = [t^{*k}h_{t-j}]$ $(k = 1,\ldots,K, j = 1,\ldots,p)$. *Furthermore, assume that the maximum likelihood estimator of* $\boldsymbol{\theta}_1$ *is asymptotically normal. Under* $H_0 : \boldsymbol{\theta}_2 = \mathbf{0}$, *the LM type statistic*

$$\xi_{LM} = \frac{1}{2}\sum_{t=1}^{T}\hat{u}_t\hat{\mathbf{x}}_{2t}'\left\{\sum_{t=1}^{T}\hat{\mathbf{x}}_{2t}\hat{\mathbf{x}}_{2t}' - \sum_{t=1}^{T}\hat{\mathbf{x}}_{2t}\hat{\mathbf{x}}_{1t}'\left(\sum_{t=1}^{T}\hat{\mathbf{x}}_{1t}\hat{\mathbf{x}}_{1t}'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_{1t}\hat{\mathbf{x}}_{2t}'\right\}^{-1}\sum_{t=1}^{T}\hat{u}_t\hat{\mathbf{x}}_{2t} \tag{2.15}$$

is asymptotically χ^2*-distributed with* $\dim(\boldsymbol{\theta}_2)$ *degrees of freedom, where* $\hat{u}_t = \hat{\varepsilon}_t^2/\hat{h}_t^0 - 1$,

$$\hat{\mathbf{x}}_{1t} = \frac{1}{\hat{h}_t^0}\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1}\bigg|_{H_0} = (\hat{h}_t^0)^{-1}(\hat{\mathbf{z}}_t + \sum_{j=1}^{p}\hat{\beta}_j^* \frac{\partial \hat{h}_{t-j}}{\partial \boldsymbol{\theta}_1}\bigg|_{H_0}) \tag{2.16}$$

and

$$\hat{\mathbf{x}}_{2t} = \frac{1}{\hat{h}_t^0}\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_2}\bigg|_{H_0} = (\hat{h}_t^0)^{-1}((t^*, \ldots, t^{*K}, (\text{vec }\mathbf{Z}_{1t})', (\text{vec }\mathbf{Z}_{2t})')' + \sum_{j=1}^{p}\hat{\beta}_j^* \frac{\partial \hat{h}_{t-j}}{\partial \boldsymbol{\theta}_2}\bigg|_{H_0}) \tag{2.17}$$

Proof. *See Appendix A.* ■

In practice, the test of Theorem 1 may be carried out in a straightforward way using an auxiliary least squares regression. Thus:

1. Estimate consistently the parameters of the conditional variance under the null hypothesis, and compute $\hat{u}_t = \hat{\varepsilon}_t^2/\hat{h}_t^0 - 1, t = 1, \ldots, T$, and the residual sum of squares, $SSR_0 = \sum_{t=1}^{T}\hat{u}_t^2$.
2. Regress $\hat{u}_t$ on $\hat{\mathbf{x}}_{1t}'$ and $\hat{\mathbf{x}}_{2t}', t = 1, \ldots, T$, and compute the sum of the squared residuals, SSR_1.
3. Compute the χ^2 test statistic as

$$\xi_{LM} = \frac{T(SSR_0 - SSR_1)}{SSR_0}.$$

As a computational detail, note that $\partial \hat{h}_t/\partial \boldsymbol{\theta}_1|_{H_0}$ and $\partial \hat{h}_t/\partial \boldsymbol{\theta}_2|_{H_0}$ in (2.16) and (2.17) are obtained recursively in connection with the parameter estimation, where it is assumed that $\partial \hat{h}_t/\partial \boldsymbol{\theta}_1|_{H_0} = \mathbf{0}$ and $\partial \hat{h}_t/\partial \boldsymbol{\theta}_2|_{H_0} = \mathbf{0}$ for $t = 0, -1, \ldots$. We shall call our LM test statistic LM_K, where K indicates the order of the polynomial in the exponent of the transition function and the tests carried out by means of an auxiliary regression are called LM-type tests.

It should also be mentioned that a robust version of the test statistics (2.15) can be derived when ζ_t are not identically distributed. One can construct a robust version using the procedure by Wooldridge (1990,1991). This test can be carried out as follows:

1. Estimate by quasi maximum likelihood the conditional variance under H_0, compute $\hat{\varepsilon}_t^2/\hat{h}_t^0 - 1, \hat{\mathbf{x}}_{1t}'$ and $\hat{\mathbf{x}}_{2t}', t = 1, \ldots, T$.
2. Regress $\hat{\mathbf{x}}_{2t}$ on $\hat{\mathbf{x}}_{1t}$, and compute the $(\dim \boldsymbol{\theta}_2 \times 1)$ residual vectors $\mathbf{r}_t, t = 1, \ldots, T$.
3. Regress 1 on $\left(\hat{\varepsilon}_t^2/\hat{h}_t^0 - 1\right)\mathbf{r}_t$ and compute the residual sum of squares SSR_0 from this regression. Under the null hypothesis, the test statistic $\xi_{LM_R} = T - SSR_0$ has an asymptotic χ^2 distribution with $\dim \boldsymbol{\theta}_2$ degrees of freedom.

One may extend Theorem 1 to the case where the model has been estimated with $r-1$ transition functions and one wants to test $r-1$ against r transitions. For that purpose, consider the model

$$\begin{aligned} \varepsilon_t &= \zeta_t {h_t}^{1/2}, \qquad \varepsilon_t|\mathcal{F}_{t-1} \sim N(0,h_t) \\ h_t &= (\boldsymbol{\theta}_0 + \sum_{l=1}^{r-1} \boldsymbol{\theta}_{1l} G_l(t^*;\gamma_l,\mathbf{c}_l))'\mathbf{z}_t + \boldsymbol{\theta}'_{1r}\widetilde{G}_r(t^*;\gamma_r,\mathbf{c}_r)\mathbf{z}_t \end{aligned} \tag{2.18}$$

where $\boldsymbol{\theta}_0 = (\alpha_0,\alpha_1,\ldots,\alpha_q,\beta_1,\ldots,\beta_p)'$, $\boldsymbol{\theta}_{1l} = (\alpha_{0l},\alpha_{1l},\ldots,\alpha_{ql},\beta_{1l},\ldots,\beta_{pl})'$, $l=1,\ldots,r-1,r$, and $\mathbf{z}_t = (1,\varepsilon^2_{t-1},\ldots,\varepsilon^2_{t-q},h_{t-1},\ldots,h_{t-p})'$. The null hypothesis is then $\mathrm{H}_0:\gamma_r=0$. Again, model (2.18) is not identified under the null hypothesis. To circumvent the problem we proceed as before and expand the logistic function $G_r(t^*;\gamma_r,\mathbf{c}_r)$ into a first-order Taylor approximation around $\gamma_r=0$. After rearranging terms we have

$$h_t = (\boldsymbol{\eta} + \sum_{l=1}^{r-1} \boldsymbol{\theta}_{1l} G_l(t^*;\gamma_l,\mathbf{c}_l))'\mathbf{z}_t + \sum_{k=1}^{K} \boldsymbol{\mu}'_k (t^*)^k \mathbf{z}_t + R_2^* \tag{2.19}$$

where $\boldsymbol{\eta} = \boldsymbol{\theta}_0 + \gamma_r\boldsymbol{\theta}_{1r}\tilde{c}_0$, $\boldsymbol{\mu}_k = \gamma_r\boldsymbol{\theta}_{1r}\tilde{c}_k$, $k=1,\ldots,K$. The test statistic is based on the following corollary of Theorem 1.

Corollary 2 *Consider the model (2.19) and let* $\boldsymbol{\theta}_1 = (\boldsymbol{\eta}',\boldsymbol{\theta}'_{1l},\gamma_l,\mathbf{c}'_l)'$ *and* $\boldsymbol{\theta}_2 = (\boldsymbol{\mu}'_1,\ldots,\boldsymbol{\mu}'_K)'$. *In addition, denote* $\mathbf{z}_t = (1,\varepsilon^2_{t-1},\ldots,\varepsilon^2_{t-q},h_{t-1},\ldots,h_{t-p})'$, $\mathbf{Z}_{1t} = [t^{*k}\varepsilon^2_{t-i}](k=1,\ldots,K, i=1,\ldots,q)$, $\mathbf{Z}_{2t} = [t^{*k}h_{t-j}](k=1,\ldots,K, j=1,\ldots,p)$ *and* $G_l(t^*) \equiv G_l(t^*;\gamma_l,\mathbf{c}_l)$. *Assume that the maximum likelihood estimator of* $(\boldsymbol{\theta}'_0,\boldsymbol{\theta}'_{11},\ldots,\boldsymbol{\theta}'_{1,r-1},\gamma_1,\ldots,\gamma_{r-1},\mathbf{c}'_1,\ldots,\mathbf{c}'_{r-1})'$ *is asymptotically normal. Under* $H_0:\boldsymbol{\theta}_2=\mathbf{0}$, *the LM type statistic (2.15) with* $\hat{u}_t = \hat{\varepsilon}^2_t/\hat{h}^0_t - 1$,

$$\begin{aligned} \hat{\mathbf{x}}_{1t} &= \frac{1}{\hat{h}^0_t}\frac{\partial\hat{h}_t}{\partial\boldsymbol{\theta}_1}\bigg|_{H_0} \\ &= (\hat{h}^0_t)^{-1}(\hat{\mathbf{z}}_t + \sum_{l=1}^{r-1}\hat{\mathbf{z}}_t\hat{G}_l(t^*) + \sum_{l=1}^{r-1}\hat{\boldsymbol{\theta}}'_{1l}\hat{\mathbf{z}}_t\frac{\partial\hat{G}_l(t^*)}{\partial\boldsymbol{\theta}_1} + \sum_{j=1}^{p}(\hat{\beta}_j + \sum_{l=1}^{r-1}\hat{\beta}^*_{jl}\hat{G}_l(t^*))\left.\frac{\partial\hat{h}_{t-j}}{\partial\boldsymbol{\theta}_1}\right|_{H_0}) \end{aligned}$$

and

$$\begin{aligned} \hat{\mathbf{x}}_{2t} &= \frac{1}{\hat{h}^0_t}\frac{\partial\hat{h}_t}{\partial\boldsymbol{\theta}_2}\bigg|_{H_0} \\ &= (\hat{h}^0_t)^{-1}((t^*,\ldots,t^{*K},(\mathrm{vec}\mathbf{Z}_{1t})',(\mathrm{vec}\mathbf{Z}_{2t})')' + \sum_{j=1}^{p}(\hat{\beta}_j + \sum_{l=1}^{r-1}\hat{\beta}^*_{jl}\hat{G}_l(t^*))\left.\frac{\partial\hat{h}_{t-j}}{\partial\boldsymbol{\theta}_2}\right|_{H_0}) \end{aligned}$$

has an asymptotic χ^2-*distribution with* $\dim(\boldsymbol{\theta}_2)$ *degrees of freedom.*

Remark 3 *The assumption of asymptotic normality in this corollary remains unverified. The existing asymptotic theory of nonlinear GARCH models does not cover the* *case where the transition function is a function of time. Besides, Meitz and Saikkonen (in press) who have worked out asymptotic theory for smooth transition GARCH models, have only obtained results on ergodicity and stationarity. Asymptotic normality of maximum likelihood estimators has not even been proven for 'standard' smooth transition GARCH models in which the transition variable is a stochastic variable. For these reasons, showing asymptotic normality of θ_1 in (2.19) is beyond the scope of this paper. Two things should be emphasized in this context. First, sequential testing to find r is just a model selection device analogous to model selection criteria such as AIC or BIC. The p-values of the tests are simply indicators helping the modeller to choose the number of transitions. Second, our simulation results do not contradict the assumption that the asymptotic null distribution of the test statistic is a χ^2-distribution.*

2.3.2 Testing against a multiplicative alternative

In order to consider the problem of testing parameter constancy in the unconditional variance assume that the error term is parameterized as

$$\varepsilon_t = \zeta_t {h_t}^{1/2}$$

where h_t is a GARCH(p,q) model as in (2.4) and ζ_t is a time-varying random variable satisfying

$$\zeta_t = z_t g_t^{1/2}$$

such that $\{z_t\}$ is a sequence of independent standard normal variables and $g_t = 1 + \sum_{l=1}^{r} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)$. This formulation allows the unconditional variance of ζ_t and thus ε_t to change smoothly over time. As already mentioned, $\{\zeta_t\}$ is a sequence of independent variables. The null hypothesis of constant unconditional variance is then $\mathrm{H}_0 : \delta_l = 0,\ l = 1, \ldots, r$. For the purpose of deriving the test statistic consider $r = 1$ and rewrite the model as follows:

$$\begin{aligned} \varepsilon_t &= z_t (h_t g_t)^{1/2}, \qquad \varepsilon_t | \mathcal{F}_{t-1} \sim N(0, h_t g_t) \\ h_t g_t &= (\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j})(1 + \delta_1 \widetilde{G}(t^*; \gamma, \mathbf{c})). \end{aligned} \tag{2.20}$$

The null hypothesis of constant unconditional variance equals $\mathrm{H}_0 : \gamma = 0$ against $\mathrm{H}_1 : \gamma > 0$. In testing this hypothesis we encounter the same identification problem as the one present in testing parameter constancy against an additive TV-GARCH process. Even here, our solution consists of approximating the transition function with a Taylor expansion around $\gamma = 0$. Proceeding as before, we reparameterize equation (2.20) as follows:

$$h_t g_t = (\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j})(\tilde{\delta}_0 + \sum_{k=1}^{K} \omega_k (t^*)^k + R_3^*) \tag{2.21}$$

where $\tilde{\delta}_0 = 1 + \gamma\delta_1\tilde{c}_0$ and $\omega_k = \gamma\delta_1\tilde{c}_k, k = 1,\ldots,K$. Under the null hypothesis, the remainder $R_3^* \equiv 0$ and does not affect the distribution theory. The null hypothesis of parameter constancy for the multiplicative structure becomes

$$\mathrm{H}_0' : \omega_k = 0,\ k = 1,\ldots,K.$$

The following corollary of Theorem 1 defines the LM-type test statistic for testing parameter constancy in the unconditional variance. The notation $\hat{g}_t^0$ denotes the estimated g_t evaluated under H_0.

Corollary 4 *Consider the model (2.21) and let* $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ *and* $\boldsymbol{\theta}_2 = (\omega_1, \ldots, \omega_K)'$. *In addition, denote* $\mathbf{z}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})'$ *and* $g_t = 1 + \delta_1 G(t^*; \gamma, \mathbf{c})$. *Under* $H_0 : \boldsymbol{\theta}_2 = \mathbf{0}$, *the LM type statistic (2.15) with* $\hat{u}_t = \hat{\varepsilon}_t^2/\hat{h}_t^0 - 1$,

$$\hat{\mathbf{x}}_{1t} = \frac{1}{\hat{h}_t^0}\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1}\bigg|_{H_0} = (\hat{h}_t^0)^{-1}(\hat{\mathbf{z}}_t + \sum_{j=1}^{p}\hat{\beta}_j^*\frac{\partial \hat{h}_{t-j}}{\partial \boldsymbol{\theta}_1}\bigg|_{H_0})$$

and

$$\hat{\mathbf{x}}_{2t} = \frac{1}{\hat{g}_t^0}\frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2}\bigg|_{H_0} = (t^*, t^{*2}, \ldots, t^{*K})'$$

has an asymptotic χ^2-*distribution with* $\dim(\boldsymbol{\theta}_2)$ *degrees of freedom.*

Once the TV-GARCH model with a single transition has been estimated we may want to investigate the possibility of remaining parameter nonconstancy in the unconditional variance. This is important from the model specification point of view. Thus, similarly to the additive structure, the previous corollary may be extended to the case where we want to test $r = 1$ against $r \geq 2$. To derive the test, consider the model

$$\begin{aligned} \varepsilon_t &= z_t(h_t g_t)^{1/2}, \qquad \varepsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t g_t) \\ h_t g_t &= (\alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j})(1 + \sum\nolimits_{l=1}^{2}\delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)). \end{aligned} \tag{2.22}$$

The null hypothesis is $\mathrm{H}_0 : \gamma_2 = 0$. Again, model (2.22) is only identified under the alternative. The solution to the identification problem consists of replacing the transition function $G_2(t^*; \gamma_2, \mathbf{c}_2)$ by a Taylor approximation around $\gamma_2 = 0$. After a reparameterization, the resulting model is

$$h_t g_t = (\alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j})(\tilde{\delta}_0 + \delta_1 G_1(t^*; \gamma_1, \mathbf{c}_1) + \sum_{k=1}^{K}\omega_k(t^*)^k + R_4^*) \tag{2.23}$$

where $\tilde{\delta}_0 = 1 + \gamma_2\delta_2\tilde{c}_0$ and $\omega_k = \gamma_2\delta_2\tilde{c}_k, k = 1,\ldots,K$. Under the null, the remainder $R_4^* \equiv 0$.

The next corollary to Theorem 1 gives the test statistic.

Corollary 5 *Consider the model (2.23) and let* $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p, \delta_1, \gamma_1, \mathbf{c}_1')'$ *and* $\boldsymbol{\theta}_2 = (\omega_1, \ldots, \omega_K)'$. *In addition, denote* $\mathbf{z}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})'$ *and* $g_t = 1 + \sum_{l=1}^{2} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)$. *Under* $H_0 : \boldsymbol{\theta}_2 = \mathbf{0}$, *the LM type statistic (2.15) with* $\hat{u}_t = \hat{\varepsilon}_t^2 / \hat{h}_t^0 \hat{g}_t^0 - 1$,

$$\hat{\mathbf{x}}_{1t} = \frac{1}{\hat{h}_t^0} \left. \frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1} \right|_{H_0} = (\hat{h}_t^0)^{-1} (\hat{\mathbf{z}}_t \hat{g}_t^0 + \hat{h}_t^0 \frac{\partial \hat{g}_t^0}{\partial \boldsymbol{\theta}_1} + \sum_{j=1}^{p} \hat{\beta}_j \hat{g}_t^0 \left. \frac{\partial \hat{h}_{t-j}}{\partial \boldsymbol{\theta}_1} \right|_{H_0})$$

and

$$\hat{\mathbf{x}}_{2t} = \frac{1}{\hat{g}_t^0} \left. \frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2} \right|_{H_0} = (\hat{g}_t^0)^{-1} (t^*, t^{*2}, \ldots, t^{*K})'$$

has an asymptotic χ^2-*distribution with* $\dim(\boldsymbol{\theta}_2)$ *degrees of freedom.*

Remark 6 *The previous remark is valid even here.*

A special case of this test, in which $h_t \equiv \alpha_0$, will be used in the specification of multiplicative TV-GARCH models in Subsection 2.4.2.

2.4 Model specification

We propose a model-building cycle for TV-GARCH models identical to the specific-to-general strategy for nonlinear models recommended by Granger (1993) or Teräsvirta (1998), among others. The idea is to begin with a parsimonious model and proceed to more complicated ones until the evaluation techniques indicate that an adequate model has been obtained. Adapting this approach to the present situation means determining the number of smooth transitions sequentially by LM-type tests discussed in Section 2.3. These tests can be used to build a GARCH model with time-varying parameters using either the additional or the multiplicative structure. We start off with a restricted specification and gradually increase the number of transition functions as long as the hypothesis of parameter constancy is rejected. The final model is estimated after the first non-rejection of the null hypothesis and evaluated through a sequence of misspecification tests.

2.4.1 Specification of additive TV-GARCH models

In order to describe the specification procedure for TV-GARCH models with an additional time-varying structure, we consider the function g_t defined in (2.7) such that all parameters are changing smoothly over time. However, the strategy may also be applied to a more restrictive functions such as g_t in (2.8). The time-varying conditional variance equals

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} + \sum_{l=1}^{r} (\alpha_{0l} + \sum_{i=1}^{q} \alpha_{il} \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{jl} h_{t-j}) G_l(t^*; \gamma_l, \mathbf{c}_l), \tag{2.24}$$

where the transition function $G_l(t^*; \gamma_l, \mathbf{c}_l)$ is defined in (2.6).

Our specification procedure for building additive TV-GARCH models contains the following stages:

1. Check for the presence of conditional heteroskedasticity by testing the null hypothesis of no ARCH against high-order ARCH. When the order of the ARCH process is sufficiently high, the standard LM test has adequate power against GARCH. If the null hypothesis is rejected, model the conditional variance by a GARCH(1,1) model. Evaluate the estimated GARCH(1,1) model by misspecification tests and, if necessary, expand it to a higher-order model. The squared standardized errors of the selected GARCH model should be free of serial correlation. Neglected autocorrelation may bias tests of parameter constancy.

2. Test the final GARCH model against the alternative of smoothly changing parameters over time using the LM-type statistic described in Theorem 1. If parameter constancy is rejected at a predetermined significance level α, estimate the TV-GARCH model (2.24) with a single transition function. If the null hypothesis of parameter constancy in (2.14) is rejected, the problem of choosing the order of the polynomial of the transition function arises. For the specification of K, we propose a model selection rule based on a sequence of nested tests as in Teräsvirta (1994) and Lin and Teräsvirta (1994). Assume $K = 3$ to ensure a parameterization sufficiently flexible for $G(t^*; \gamma, \mathbf{c})$. If parameter constancy is rejected, test the following sequence of hypotheses:

$$\begin{aligned}
\mathrm{H}_{03} &: \omega_3 = 0, \varphi_{i3} = 0, \lambda_{j3} = 0, \\
\mathrm{H}_{02} &: \omega_2 = 0, \varphi_{i2} = 0, \lambda_{j2} = 0 \mid \omega_3 = 0, \varphi_{i3} = 0, \lambda_{j3} = 0, \\
\mathrm{H}_{01} &: \omega_1 = 0, \varphi_{i1} = 0, \lambda_{j1} = 0 \mid \omega_2 = \omega_3 = 0, \varphi_{i2} = \varphi_{i3} = 0, \lambda_{j2} = \lambda_{j3} = 0,
\end{aligned}$$

where $i = 1, \ldots, q, j = 1, \ldots, p$, in (2.13), by means of LM-type tests. The results of this test sequence may be used as follows. If H_{01} and H_{03} are rejected more strongly, measured by p-values, than H_{02}, then either $K = 1$ or $K = 3$. If testing H_{02} yields the strongest rejection, the choice is $K = 2$. Furthermore, if only H_{01} is rejected at the appropriate significance level or is rejected clearly more strongly than the other two null hypotheses, then the modeller should choose $K = 1$. Visual inspection of the return series is also helpful in making a decision about K. The rules or suggestions based on p-values are based on expressions of the parameters ω_k, φ_{ik} and λ_{jk} in the auxiliary regression as functions of the original parameters at different values of K. The test sequence is analogous to that proposed in Teräsvirta (1994) for specifying the type of the smooth transition autoregressive model, where the choice is between $K = 1$ and $K = 2$.

3. Test the TV-GARCH model with one transition function against the TV-GARCH model with two transition functions at the significance level $\alpha\tau, 0 < \tau < 1$. The significance level is decreased giving a preference for parsimonious models. The overall significance level of the sequence of tests may be approximated by the

Bonferroni upper bound. The user can choose the value for τ. In our simulations we set $\tau = 1/2$. If the null hypothesis is rejected, specify K for the next transition and estimate the TV-GARCH model (2.12) with two transition functions.

4. Proceed sequentially by testing the TV-GARCH model with $r-1$ transition functions against the TV-GARCH model with r transitions at the significance level $\alpha\tau^{r-1}$ until the first non-rejection of the null hypothesis. Evaluate the selected model by misspecification tests and once it passes them accept it as the final model. In the opposite case, modify the specification of the model or try another family of models.

2.4.2 Specification of multiplicative TV-GARCH models

The specific-to-general approach for specifying TV-GARCH models with a multiplicative time-varying component consists in first modelling the unconditional variance as follows:

1. Use the LM-type statistic developed in Section 2.3.2 to test the null hypothesis of constant variance against a time-varying unconditional variance with a single transition function at the significance level α. First, assume $h_t = \alpha_0$ and test $\mathrm{H}_{10} : g_t \equiv 1$ against $\mathrm{H}_{11} : g_t = 1 + \delta_1 G_1(t^*; \gamma_1, \mathbf{c}_1)$. In case of a rejection, test $\mathrm{H}_{20} : g_t = 1 + \delta_1 G_1(t^*; \gamma_1, \mathbf{c}_1)$ against $\mathrm{H}_{21} : g_t = 1 + \sum_{l=1}^{2} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)$ at the significance level $\alpha\tau, 0 < \tau < 1$. Continue until the first non-rejection of the null hypothesis. The significance level is reduced at each step of the testing procedure and converging to zero for reasons previously mentioned.

2. After specifying g_t, test the null hypothesis of no conditional heteroskedasticity in $\{\zeta_t\}$. If it is rejected, model the conditional variance h_t of the standardized variable $\varepsilon_t / g_t^{1/2}$ in the standard fashion, such that

$$h_t = \alpha_0^* + \sum_{i=1}^{q} \alpha_i \left(\frac{\varepsilon_{t-i}^2}{g_{t-i}} \right) + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{2.25}$$

3. The estimated model is evaluated by means of LM-type diagnostic tests proposed in Subsection 2.6.1. If the model passes all the misspecification tests, tentatively accept it. Otherwise, modify it or consider another family of volatility models.

2.5 Estimation of the TV-GARCH model

Suppose that ε_t is generated by a GARCH model with a time-varying structure described in Section 2.2. Let $h_t = h_t(\boldsymbol{\theta}_1)$ and $g_t = g_t(\boldsymbol{\theta}_2)$ where $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ and $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \boldsymbol{\alpha}_1', \ldots, \boldsymbol{\alpha}_r', \boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_r', \gamma_1, \ldots, \gamma_r, \mathbf{c}_1', \ldots, \mathbf{c}_r')'$ with $\boldsymbol{\delta} = (\delta_1, \ldots,$

$\delta_r)', \boldsymbol{\alpha}_i = (\alpha_{1i}, \ldots, \alpha_{qi})'$ and $\boldsymbol{\beta}_i = (\beta_{1i}, \ldots, \beta_{pi})', i = 1, \ldots, r$. For the additive parameterization, $\boldsymbol{\delta} = \mathbf{0}$ and for the multiplicative one, $\boldsymbol{\alpha}_i = \mathbf{0}$ and $\boldsymbol{\beta}_i = \mathbf{0}$. The quasi maximum likelihood (QML) estimator $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\theta}}_1', \widehat{\boldsymbol{\theta}}_2')'$ is obtained maximizing $\sum_{t=1}^{T} \ell_t(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ where the log-likelihood for observation t equals

$$\ell_t(\boldsymbol{\theta}) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\{h_t(\boldsymbol{\theta}_1) + g_t(\boldsymbol{\theta}_2)\} - \frac{1}{2}\frac{\varepsilon_t^2}{h_t(\boldsymbol{\theta}_1) + g_t(\boldsymbol{\theta}_2)} \tag{2.26}$$

for the additive TV-GARCH model or

$$\ell_t(\boldsymbol{\theta}) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\{\ln h_t(\boldsymbol{\theta}_1) + \ln g_t(\boldsymbol{\theta}_2)\} - \frac{1}{2}\frac{\varepsilon_t^2}{h_t(\boldsymbol{\theta}_1) g_t(\boldsymbol{\theta}_2)} \tag{2.27}$$

for the multiplicative TV-GARCH model.

The asymptotic properties of the QML estimators for the GARCH(p, q) process have been studied, among others, by Ling and Li (1997). They showed that the QML estimators are consistent and asymptotic normal provided that $\mathsf{E}\varepsilon_t^4 < \infty$. Ling and McAleer (2003) established consistency for the global maximum of QML estimators under the condition $\mathsf{E}\varepsilon_t^2 < \infty$. Berkes, Horváth, and Kokoszka (2003) obtained consistency of the QML estimators assuming $\mathsf{E}\varepsilon_t^2 < \infty$ and asymptotic normality by assuming $\mathsf{E}\varepsilon_t^4 < \infty$. These results have in common the assumption that the process y_t is stationary and ergodic such that the laws of large numbers apply. More recently, Jensen and Rahbek (2004) relaxed this assumption and allowed the parameters to lie in the region where the process is nonstationary. They showed that for the GARCH(1,1) case, under a finite conditional variance for ζ_t^2, consistency and asymptotic normality still hold independently of whether the process y_t is stationary or not. As already mentioned, asymptotic normality for the parameter estimators of the TV-GARCH models has not yet been proven.

Three remarks are in order regarding numerical aspects of the estimation of TV-GARCH models. The first one concerns the accuracy of the slope estimates when the true parameters γ_l are very large. In order to achieve an accurate estimate for a large γ_l, the number of observations of the transition variable in the neighbourhood of $\mathbf{c}_l$ must be very large. This is due to the fact that even large changes in γ_l only have an effect on the transition function in a small neighbourhood of $\mathbf{c}_l$. But then, for the same reason for large γ_l it is sufficient to obtain an estimate that is large; whether or not it is very accurate is not of utmost importance. Note that if $\hat{\gamma}_l$ is large, an "insignificant" $\hat{\gamma}_l$ is an indication of a large γ_l, not of $\gamma_l \equiv 0$. Besides, because of the identification problem the t-ratio does not have its standard asymptotic distribution when $\gamma_l \equiv 0$. A more serious problem is that large estimates for the smoothness parameter γ_l may lead to numerical problems when carrying out parameter constancy tests. A simple solution, suggested in Eitrheim and Teräsvirta (1996), is to omit those elements of the score that are partial derivatives with respect to the parameters in the transition function. This can be done without significantly affecting the value of the test statistic.

The second comment has to do with the computation of the derivatives of the log-likelihood function. Many of the existing optimization algorithms require the computation of at least the first and, in some cases, also the second derivatives of the log-likelihood function. It is common practice to use numerical derivates that are relatively fast to compute and reliable, and the derivation of exact analytic derivatives is avoided. Fiorentini, Calzolari, and Panattoni (1996), however, encourage the employment of analytic derivatives, because that leads to fewer iterations than optimization with numerical derivatives. Furthermore, the use of analytic derivatives also improves the accuracy of the estimates of the standard errors of the parameter estimates. Consequently, we use analytic first derivatives in all the computations, both in calculating values of the test statistics and in estimating TV-GARCH models.

The third remark is related to the manner in which the parameter estimates are obtained. The parameters in the additive TV-GARCH model are estimated simultaneously by full conditional maximum likelihood. In this context, care is required in the estimation. Since the log-likelihood (2.27) may contain several local maxima, it is advisable to initiate the estimation from different sets of starting-values before settling for the final parameter estimates. Numerical problems in the estimation of the multiplicative TV-GARCH model can be alleviated by concentrating the likelihood iteratively. This considerably reduces the dimensionality problem and is computationally much easier than maximizing the log-likelihood with respect to all parameters simultaneously. The estimation of the TV-GARCH model with multiplicative structure can be simplified since the log-likelihood can be decomposed into two separate sets of parameters: the GARCH and the time-varying parameter vectors. The estimation is divided into two steps which are then repeated one after the other. The iterations start by first estimating $\boldsymbol{\theta}_2$, assuming h_t to be a positive constant, for instance $h_t = \hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\varepsilon_t^2$, and continue by estimating $\boldsymbol{\theta}_1$, given the estimates of $\boldsymbol{\theta}_2$. The estimate of $\boldsymbol{\theta}_1$ will then be used for re-estimating $\boldsymbol{\theta}_2$, and so on. The iterative two-stage estimation procedure is terminated when a local maximum of the log-likelihood has been reached.

2.6 Misspecification testing of TV-GARCH models

The final step of the modelling strategy consists of evaluating the adequacy of the estimated TV-GARCH model by means of a sequence of misspecification tests. We shall assume that the true process of either the additive or the multiplicative time-varying variance is misspecified. The general idea is to construct an augmented version of the TV-GARCH model by introducing a new component $f_t = f(\mathbf{v}_t; \boldsymbol{\theta}_3)$ into the original model. This component is a function that is at least twice continuously differentiable with respect to the elements of $\boldsymbol{\theta}_3$, vector of additional parameters. The vector $\mathbf{v}_t$ is a vector of omitted random variables, and its definition varies from one test to the next.

2.6.1 Misspecification tests for the multiplicative model

The misspecification tests considered here may be divided into three categories. The first two correspond to additive and the third one to multiplicative misspecification.

Let $h_t = h_t(\boldsymbol{\theta}_1)$ and $g_t = g_t(\boldsymbol{\theta}_2)$, such that the parameter vectors $\boldsymbol{\theta}_i, i = 1, 2$, represent the parameters belonging to $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$ and $g_t = 1 + \sum_{l=1}^{r} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l)$. Under $\mathrm{H}_0 : \boldsymbol{\theta}_3 = \mathbf{0}$, the augmented model reduces to the multiplicative TV-GARCH model.

Additive misspecification - case 1

The first category of tests assumes that, under the alternative hypothesis, the original TV-GARCH model may be extended by assuming

$$\varepsilon_t = \zeta_t (h_t + f_t)^{1/2} g_t^{1/2}. \tag{2.28}$$

Under the null hypothesis, $f_t \equiv 0$, which is equivalent to $\boldsymbol{\theta}_3 = \mathbf{0}$. If $g_t \equiv 1$, the test collapses into the additive misspecification test in Lundbergh and Teräsvirta (2002). At least three types of alternative hypotheses can be considered within this family of tests. The test of the GARCH(p, q) component against higher-order alternatives as well as the test against a smooth transition GARCH (ST-GARCH) and, furthermore, the test against an asymmetric component (GJR-GARCH) belong to the additive class (2.28).

The log-likelihood function for observation t of model (2.28) is

$$\ell_t = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \{\ln(h_t + f_t) + \ln g_t\} - \frac{\varepsilon_t^2}{2(h_t + f_t) g_t}. \tag{2.29}$$

When the estimated multiplicative TV-GARCH model is tested against the different types of alternatives, the first component of the score corresponding to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, evaluated under H_0, is equal to

$$\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}}\right|_{H_0} = \frac{1}{2} \left(\frac{\varepsilon_t^2}{h_t g_t} - 1 \right) \mathbf{x}_{1t}$$

where $\mathbf{x}_{1t} = \left(\frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\theta}_1'}, \frac{1}{g_t} \frac{\partial g_t}{\partial \boldsymbol{\theta}_2'} \right)'$ and the parameter vector $\boldsymbol{\theta}$ is partitioned as $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$. The estimated quantities for $\frac{\partial h_t}{\partial \boldsymbol{\theta}_1}|_{H_0}$ and $\frac{\partial g_t}{\partial \boldsymbol{\theta}_2}|_{H_0}$ are defined as

$$\left.\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1}\right|_{H_0} = \hat{\mathbf{z}}_t + \sum_{j=1}^{p} \hat{\beta}_j \left.\frac{\partial \hat{h}_{t-j}}{\partial \boldsymbol{\theta}_1}\right|_{H_0} \tag{2.30}$$

$$\left.\frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2}\right|_{H_0} = \sum_{l=1}^{r} G_l(t^*; \hat{\gamma}_l, \hat{\mathbf{c}}_l) + \sum_{l=1}^{r} \hat{\delta}_l \frac{\partial G_l(t^*; \hat{\gamma}_l, \hat{\mathbf{c}}_l)}{\partial \boldsymbol{\theta}_2}. \tag{2.31}$$

The differences show up in the partial derivatives of (2.29) with respect to $\boldsymbol{\theta}_3$. It follows that the additional block of the score for observation t due to $\boldsymbol{\theta}_3$ has the form

$$\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3} = \frac{1}{2} \left(\frac{\varepsilon_t^2}{(h_t + f_t) g_t} - 1 \right) \frac{1}{h_t} \frac{\partial f_t}{\partial \boldsymbol{\theta}_3}$$

so that, under H_0,

$$\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t g_t} - 1\right)\frac{1}{h_t}\left.\frac{\partial f_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0}$$

where $\frac{\partial f_t}{\partial \boldsymbol{\theta}_3} = \mathbf{v}_t$. The resulting LM test may be easily performed using an auxiliary regression as in Section 2.3. In terms of previous notation, we have

$$\hat{\mathbf{x}}_{1t} = \left(\frac{1}{\hat{h}_t^0}\left.\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1'}\right|_{H_0}, \frac{1}{\hat{g}_t^0}\left.\frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2'}\right|_{H_0}\right)' \tag{2.32}$$

$$\hat{\mathbf{x}}_{2t} = \frac{1}{\hat{h}_t^0}\left.\frac{\partial \hat{f}_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \frac{\hat{\mathbf{v}}_t}{\hat{h}_t^0} \tag{2.33}$$

where $\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1}|_{H_0}$ and $\frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2}|_{H_0}$ are as in (2.30) and (2.31), respectively. We shall now concentrate our attention on tests against higher-order alternatives and a smooth transition GARCH model.

Testing the GARCH(p, q) component against higher-order alternatives

An evident source of misspecification is to select too low an order in the GARCH(p, q) component. A similar testing procedure to the one proposed by Bollerslev (1986) for testing a GARCH(p, q) model against higher-order alternatives is presented. Under the alternative GARCH($p, q + r$), the additional component equals

$$f_t = \sum_{i=q+1}^{q+r} \alpha_i \varepsilon_{t-i}^2 \tag{2.34}$$

or

$$f_t = \sum_{j=p+1}^{p+r} \beta_j h_{t-j} \tag{2.35}$$

if we take the GARCH($p + r, q$) as alternative. The identification problem discussed in Bollerslev (1986) prevents us from considering the alternative GARCH($p + r, q + s$), $r, s > 0$. Under the null hypothesis $\mathrm{H}_0 : \boldsymbol{\theta}_3 = \mathbf{0}$, i.e. $\alpha_{q+1} = ... = \alpha_{q+r} = 0$ for the former case and $\beta_{p+1} = ... = \beta_{p+r} = 0$ for the latter case, the models reduce to the GARCH(p, q) model.

Corollary 7 defines the test statistic for testing $\alpha_{q+1} = ... = \alpha_{q+r} = 0$. A similar result holds for testing $\beta_{p+1} = ... = \beta_{p+r} = 0$ in (2.35) and can be stated by replacing $\boldsymbol{\theta}_3 = (\alpha_{q+1}, ..., \alpha_{q+r})'$ and $\hat{\mathbf{v}}_t = (\varepsilon_{t-(q+1)}^2, \ldots, \varepsilon_{t-(q+r)}^2)'$ in Corollary 7 by $\boldsymbol{\theta}_3 = (\beta_{p+1}, ..., \beta_{p+r})'$ and $\hat{\mathbf{v}}_t = (h_{t-(p+1)}, \ldots, h_{t-(p+r)})'$.

Corollary 7 *Consider the model (2.28) where $\{\zeta_t\}$ is a sequence of independent standard normal variables. Let $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ and $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \gamma_1, \ldots, \gamma_r, \mathbf{c}_1, \ldots, \mathbf{c}_r)'$ with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_r)'$. Furthermore, f_t is defined by (2.34) such that*

$\boldsymbol{\theta}_3 = (\alpha_{q+1}, \ldots, \alpha_{q+r})'$ and $\hat{\mathbf{v}}_t = (\varepsilon^2_{t-(q+1)}, \ldots, \varepsilon^2_{t-(q+r)})'$. Assume that the maximum likelihood estimators of the parameters of (2.28) are asymptotically normal when $H_0 : \boldsymbol{\theta}_3 = \mathbf{0}$ is valid. Thus, under this null hypothesis, the LM statistic (2.15), with $\hat{u}_t = \hat{\varepsilon}^2_t/\hat{h}^0_t \hat{g}^0_t - 1, \hat{\mathbf{x}}_{1t}$ as in (2.32) and $\hat{\mathbf{x}}_{2t}$ as in (2.33) is asymptotically χ^2-distributed with r degrees of freedom.

Remark 8 *Note that the result stated in Corollary 7 depend on an assumption of asymptotic normality which so far remains unproven. Asymptotic normality has, however, been proven in the special case $\boldsymbol{\theta}_2 = \mathbf{0}$ when the null model (2.28) is a standard GARCH(p,q) model. A similar remark will hold for Corollaries 9, 10, 11 and 12.*

Testing the GARCH(p, q) component against a nonlinear specification

It is possible that responses of volatility in financial series to negative and positive shocks are not symmetric around zero (or some other value). The GARCH literature offers a variety of parameterizations for describing asymmetric effects of shocks on the conditional variance. The ST-GARCH model, discussed in Hagerud (1997), González-Rivera (1998) and Anderson, Nam, and Vahid (1999), is one of them. Symmetry of the estimated TV-GARCH can be tested against asymmetry or, more generally, against nonlinearity, using these models as alternatives. To this end, let

$$f_t = \sum_{i=1}^{q} (\alpha^*_{1i} + \alpha^*_{2i}\varepsilon^2_{t-i})G(\varepsilon_{t-i}; \gamma, \mathbf{c}) \tag{2.36}$$

where $G(\varepsilon_{t-i}; \gamma, \mathbf{c})$ is the transition function given in (2.6) with ε_{t-i} as the transition variable. With the purpose of simplifying the derivation of the test we replace $G(\varepsilon_{t-i}; \gamma, \mathbf{c})$ by $\widetilde{G}(\varepsilon_{t-i}; \gamma, \mathbf{c}) = G(\varepsilon_{t-i}; \gamma, \mathbf{c}) - 1/2$. The null hypothesis of linearity is $\mathrm{H}_0 : \gamma = 0$ under which $G(\varepsilon_{t-i}; \gamma, \mathbf{c}) \equiv 1/2$. However, the remaining parameters in (2.36) are not identified under the null hypothesis. Again the identification problem may be circumvented using a Taylor series approximation of the transition function around $\gamma = 0$. After rearranging terms, one obtains

$$h_t + f_t = \alpha^*_0 + \sum_{i=1}^{q} \alpha^*_i \varepsilon^2_{t-i} + \sum_{j=1}^{p} \beta_j h_{t-j} + \sum_{i=1}^{q}\sum_{k=1}^{K} (\varpi_{ik}\varepsilon^k_{t-i} + \pi_{ik}\varepsilon^{k+2}_{t-i}) + R^*_5 \tag{2.37}$$

where $\alpha^*_0 = \alpha_0 + \sum_{i=1}^{q} \gamma\alpha^*_{1i}\tilde{c}_0, \alpha^*_i = \alpha_i + \gamma\alpha^*_{2i}\tilde{c}_0, \varpi_{ik} = \gamma\alpha^*_{1i}\tilde{c}_k$ and $\pi_{ik} = \gamma\alpha^*_{2i}\tilde{c}_k$. The component given in (2.36) can be rewritten as

$$f_t = \sum_{i=1}^{q}\sum_{k=1}^{K} (\varpi_{ik}\varepsilon^k_{t-i} + \pi_{ik}\varepsilon^{k+2}_{t-i}) + R^*_5 \tag{2.38}$$

When the null hypothesis holds, the remainder R^*_5 vanishes, and so does not affect the distributional properties of the test. Using this notation, the hypothesis of no additional nonlinear structure becomes $\mathrm{H}'_0 : \varpi_{ik} = \pi_{ik} = \mathbf{0}, i = 1, \ldots, q, k = 1, \ldots, K$. The next corollary gives the test statistic.

Corollary 9 *Consider the model (2.28) where* $\{\zeta_t\}$ *is a sequence of independent standard normal variables. Let* $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ *and* $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \gamma_1, \ldots, \gamma_r, \mathbf{c}_1, \ldots, \mathbf{c}_r)'$ *with* $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_r)'$. *Furthermore,* f_t *is defined by (2.38) such that* $\boldsymbol{\theta}_3 = (\varpi_i', \pi_i')'$, *where* $\varpi_i = (\varpi_{i1}, \ldots, \varpi_{iK})'$ *and* $\pi_i = (\pi_{i1}, \ldots, \pi_{iK})', i = 1, \ldots, q$. *In addition, let* $\hat{\mathbf{v}}_t = (\hat{\mathbf{v}}_{1,t}', \ldots, \hat{\mathbf{v}}_{K+2,t}')'$ *with* $\mathbf{v}_{it} = (\varepsilon_{t-1}^i, \ldots, \varepsilon_{t-q}^i)', i = 1, \ldots, K+2$. *Assume that the maximum likelihood estimators of the parameters of (2.28) are asymptotically normal when* $H_0 : \boldsymbol{\theta}_3 = \mathbf{0}$ *is valid. Thus, under this null hypothesis, the LM statistic (2.15), with* $\hat{u}_t = \hat{\varepsilon}_t^2/\hat{h}_t^0 \hat{g}_t^0 - 1, \hat{\mathbf{x}}_{1t}$ *as in (2.32) and* $\hat{\mathbf{x}}_{2t}$ *as in (2.33) is asymptotically* χ^2-*distributed with* $\dim(\boldsymbol{\theta}_3)$ *degrees of freedom.*

Additive misspecification - case 2

We shall now consider the case in which the true model has the following form:

$$\varepsilon_t = \zeta_t h_t^{1/2} (g_t + f_t)^{1/2}. \tag{2.39}$$

Under the null hypothesis, $f_t \equiv 0$, which is again equivalent to $\boldsymbol{\theta}_3 = \mathbf{0}$. The model again reduces to (2.1) and (2.3). The log-likelihood for the observation t equals

$$\ell_t = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\{\ln h_t + \ln(g_t + f_t)\} - \frac{\varepsilon_t^2}{2h_t(g_t + f_t)}.$$

The block of the score containing the first partial derivatives with respect to $\boldsymbol{\theta}_3$ is

$$\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t(g_t + f_t)} - 1\right)\frac{1}{g_t}\frac{\partial f_t}{\partial \boldsymbol{\theta}_3}$$

which, under H_0, is equal to

$$\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t g_t} - 1\right)\frac{1}{g_t}\left.\frac{\partial f_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0}.$$

For this alternative, the quantity $\hat{\mathbf{x}}_{1t}$ is defined as in (2.32) and

$$\hat{\mathbf{x}}_{2t} = \frac{1}{\hat{g}_t^0}\left.\frac{\partial \hat{f}_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \frac{\hat{\mathbf{v}}_t}{\hat{g}_t^0}. \tag{2.40}$$

Testing the hypothesis of no additional transitions

Once the TV-GARCH model has been estimated, one may use this set-up, for example, to re-check the need for another transition function in g_t. Taking the multiplicative TV-GARCH model with $r + s$ transitions as the alternative, it follows that

$$f_t = \sum_{l=r+1}^{r+s} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l) \tag{2.41}$$

The hypothesis of no additional transitions is $\mathrm{H}_0 : \gamma_{r+1} = ... = \gamma_{r+s} = 0$. Under this hypothesis, the parameters $(\delta_l, \mathbf{c}_l')'$ are not identified. To circumvent this problem, we replace the transition function $G_l(t^*; \gamma_l, \mathbf{c}_l)$ by its first-order Taylor expansion around $\gamma_l = 0, l = r+1, \ldots, r+s$. After merging terms, we obtain

$$\begin{aligned} g_t + f_t &= 1 + \sum_{l=1}^{r} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l) + \sum_{l=r+1}^{r+s} \delta_l (\gamma_l \tilde{c}_0 + \sum_{k=1}^{K} \gamma_l \tilde{c}_k (t^*)^k) + R_6^* \\ &= \delta_l^* + \sum_{l=1}^{r} \delta_l G_l(t^*; \gamma_l, \mathbf{c}_l) + \sum_{l=r+1}^{r+s} \sum_{k=1}^{K} \psi_{lk} (t^*)^k + R_6^* \qquad (2.42) \end{aligned}$$

where $\delta_l^* = 1 + \sum_{l=r+1}^{r+s} \gamma_l \delta_l \tilde{c}_0$ and $\psi_{lk} = \gamma_l \delta_l \tilde{c}_k, l = r+1, \ldots, r+s, k = 1, \ldots, K$. It is convenient to reparameterize (2.41) as follows:

$$f_t = \sum_{l=r+1}^{r+s} \sum_{k=1}^{K} \psi_{lk} (t^*)^k + R_6^* \qquad (2.43)$$

Under the null hypothesis, the remainder R_6^* vanishes. It seems that the coefficients $\psi_{lk}, l = r+1, \ldots, s$, for a fixed k, are not identified because they are all related to the same variable $(t^*)^k$. They have to be merged, which leads to

$$f_t = \sum_{k=1}^{K} \psi_k^* (t^*)^k + R_6^*.$$

In other words, the test statistic is the same, independent of whether we would be testing against including G_{r+1} or including $G_{r+1}, \ldots, G_{r+s}, s \geq 2$. Compare this with Corollary 5, which is a special case. In fact, Corollary 5 contains another example of a misspecification test of the multiplicative model in which the misspecification is of the type $h_t(g_t + f_t)$.

Multiplicative misspecification

Under multiplicative misspecification, the parametric alternative to the TV-GARCH model is formulated as

$$\varepsilon_t = \zeta_t (h_t g_t f_t)^{1/2}. \qquad (2.44)$$

In this framework, $\mathrm{H}_0 : f_t \equiv 1$, which is equivalent to $\boldsymbol{\theta}_3 = \mathbf{0}$. Under the null hypothesis, the model reduces to the multiplicative TV-GARCH model. For this specification, the log-likelihood function for observation t may be written

$$\ell_t = -\frac{1}{2} \ln 2\pi - \frac{1}{2} (\ln h_t + \ln g_t + \ln f_t) - \frac{\varepsilon_t^2}{2 h_t g_t f_t}.$$

The additional block of the score has the form

$$\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3} = \frac{1}{2} \left(\frac{\varepsilon_t^2}{h_t g_t f_t} - 1 \right) \frac{\partial f_t}{\partial \boldsymbol{\theta}_3}$$

which, under H_0, reduces to

$$\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t g_t} - 1\right)\left.\frac{\partial f_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0}.$$

Taking (2.44) as the alternative, the vector $\hat{\mathbf{x}}_{1t}$ is given in (2.32) and

$$\hat{\mathbf{x}}_{2t} = \left.\frac{\partial \hat{f}_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \hat{\mathbf{v}}_t. \tag{2.45}$$

This category includes general misspecification tests of adequacy of the estimated specification. After the estimation of the TV-GARCH model, one may want to check whether the estimated standardized errors still contain some structure. In the GARCH context, Lundbergh and Teräsvirta (2002) proposed a Lagrange multiplier statistic for testing the hypothesis of no remaining ARCH which is asymptotically equivalent to the portmanteau statistic introduced by Li and Mak (1994). A similar test statistic can be obtained for the multiplicative TV-GARCH model.

Testing the hypothesis of no remaining ARCH

An important misspecification test for the multiplicative TV-GARCH specification is the so-called 'ARCH-in-GARCH' test. The original model

$$\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}, \qquad \zeta_t \sim \mathrm{nid}(0,1)$$

is extended by assuming that, under the alternative, $\zeta_t = \xi_t f_t^{1/2}$, where $\xi_t \sim \mathrm{nid}(0,1)$, and

$$f_t = 1 + \sum_{j=1}^{s} \phi_j \zeta_{t-j}^2. \tag{2.46}$$

The hypothesis of interest is $\mathrm{H}_0 : \phi_1 = ... = \phi_s = 0$ and $\frac{\partial \hat{f}_t}{\partial \boldsymbol{\theta}_3}|_{H_0} = (\widehat{\zeta}_1^2, \ldots, \widehat{\zeta}_s^2)'$. Some special cases may be mentioned. If $g_t \equiv 1$, the test collapses into the test of 'no ARCH-in-GARCH' in Lundbergh and Teräsvirta (2002). If $h_t \equiv 1$ as well, the test coincides with the Engle's test of no ARCH. Setting only $h_t \equiv 1$, it reduces to the test of no ARCH in $\varepsilon_t/\widehat{g}_t^{1/2}$. The test is presented in the next corollary.

Corollary 10 *Consider the model (2.44) where $\{\zeta_t\}$ is a sequence of independent standard normal variables. Let $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ and $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \gamma_1, \ldots, \gamma_r, \mathbf{c}_1, \ldots, \mathbf{c}_r)'$ with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_r)'$. Furthermore, f_t is defined by (2.46) such that $\boldsymbol{\theta}_3 = (\phi_1, \ldots, \phi_s)'$ and $\hat{\mathbf{v}}_t = (\widehat{\zeta}_1^2, \ldots, \widehat{\zeta}_s^2)'$. Assume that the maximum likelihood estimators of the parameters of (2.44) are asymptotically normal when $H_0 : \boldsymbol{\theta}_3 = \mathbf{0}$ is valid. Thus, under this null hypothesis, the LM statistic (2.15), with $\hat{u}_t = \hat{\varepsilon}_t^2/\hat{h}_t^0 \hat{g}_t^0 - 1, \hat{\mathbf{x}}_{1t}$ as in (2.32) and $\hat{\mathbf{x}}_{2t} = \hat{\mathbf{v}}_t$ is asymptotically χ^2-distributed with s degrees of freedom.*

2.6.2 Misspecification tests for the additive model

In this section we shall consider the additive TV-GARCH model and assume that it is either additively or multiplicatively misspecified. The former possibility may include, for example, tests against remaining nonlinearity and additional transitions, whereas the test of the adequacy of the estimated model belongs to the latter one. To this end, let $h_t = h_t(\boldsymbol{\theta}_1)$ and $g_t = g_t(\boldsymbol{\theta}_2)$, such that $\boldsymbol{\theta}_i, i = 1, 2$, represent the parameters belonging to $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$ and $g_t = \sum_{l=1}^{r}(\alpha_{0l} + \sum_{i=1}^{q} \alpha_{il}\varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{jl} h_{t-j}) G_l(t^*; \gamma_l, \mathbf{c}_l)$. Under the null hypothesis of no misspecification, the extended model reduces to the additive TV-GARCH parameterization.

Additive misspecification

In order to define the set of alternative models for this class, consider a general alternative written as

$$\varepsilon_t = \zeta_t (h_t + g_t + f_t)^{1/2}. \tag{2.47}$$

Under the null hypothesis, $f_t \equiv 0$. If $g_t \equiv 0$, the test coincides to the additive test developed in Lundbergh and Teräsvirta (2002). In the case of the additive parameterization, the diagnostic tests mentioned in Subsections 2.6.1 and 2.6.1 belong to the class (2.47). Such tests can be easily adapted into the present context, where the quantities $\hat{u}_t, \hat{\mathbf{x}}_{it}, i = 1, 2$, and $\hat{\mathbf{v}}_t$ have to be modified accordingly. We shall therefore be concerned with a general alternative hypothesis rather than describing individual situations.

The log-likelihood function for observation t is

$$\ell_t = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\{\ln(h_t + g_t + f_t)\} - \frac{\varepsilon_t^2}{2(h_t + g_t + f_t)}$$

and the vector of the first partial derivatives with respect to $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$ under H_0 equals

$$\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}}\right|_{H_0} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t + g_t} - 1\right)\mathbf{x}_{1t}$$

where $\mathbf{x}_{1t} = \left(\frac{1}{h_t+g_t}\frac{\partial h_t}{\partial \boldsymbol{\theta}_1'}, \frac{1}{h_t+g_t}\frac{\partial g_t}{\partial \boldsymbol{\theta}_2'}\right)'$. The appropriate estimates of $\frac{\partial h_t}{\partial \boldsymbol{\theta}_1}|_{H_0}$ and $\frac{\partial g_t}{\partial \boldsymbol{\theta}_2}|_{H_0}$ are

$$\left.\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1}\right|_{H_0} = \hat{\mathbf{z}}_t + \sum_{j=1}^{p}\hat{\beta}_j \left.\frac{\partial \hat{h}_{t-j}}{\partial \boldsymbol{\theta}_1}\right|_{H_0} \tag{2.48}$$

$$\left.\frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2}\right|_{H_0} = \sum_{l=1}^{r}\hat{\mathbf{z}}_t \hat{G}_l(t^*) + \sum_{l=1}^{r}\hat{\boldsymbol{\theta}}_{2l}'\hat{\mathbf{z}}_t \frac{\partial \hat{G}_l(t^*)}{\partial \boldsymbol{\theta}_2} + \sum_{j=1}^{p}\sum_{l=1}^{r}\hat{\beta}_{jl}\hat{G}_l(t^*) \left.\frac{\partial \hat{g}_{t-j}}{\partial \boldsymbol{\theta}_2}\right|_{H_0} \tag{2.49}$$

where $\mathbf{z}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})', \boldsymbol{\theta}_{2l} = (\alpha_{0l}, \alpha_{1l}, \ldots, \alpha_{ql}, \beta_{1l}, \ldots, \beta_{pl})', l = 1, \ldots, r$, and $G_l(t^*) \equiv G_l(t^*; \gamma_l, \mathbf{c}_l)$. The additional block of the score for observation t, under H_0, equals

$$\left.\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t + g_t} - 1\right)\frac{1}{h_t + g_t}\left.\frac{\partial f_t}{\partial \boldsymbol{\theta}_3}\right|_{H_0}$$

where $\frac{\partial f_t}{\partial \boldsymbol{\theta}_3} = \mathbf{v}_t$. To define the LM statistic, set

$$\hat{\mathbf{x}}_{1t} = \left(\frac{1}{\hat{h}_t^0 + \hat{g}_t^0} \frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1'}\bigg|_{H_0}, \frac{1}{\hat{h}_t^0 + \hat{g}_t^0} \frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2'}\bigg|_{H_0} \right)' \tag{2.50}$$

$$\hat{\mathbf{x}}_{2t} = \frac{1}{\hat{h}_t^0 + \hat{g}_t^0} \frac{\partial \hat{f}_t}{\partial \boldsymbol{\theta}_3}\bigg|_{H_0} = \frac{\hat{\mathbf{v}}_t}{\hat{h}_t^0 + \hat{g}_t^0} \tag{2.51}$$

where $\frac{\partial \hat{h}_t}{\partial \boldsymbol{\theta}_1}|_{H_0}$ and $\frac{\partial \hat{g}_t}{\partial \boldsymbol{\theta}_2}|_{H_0}$ are given in (2.48) and (2.49), respectively. These results apply to the test against remaining nonlinearity. The test will be presented in the following corollary.

Corollary 11 *Consider the model (2.47) where $\{\zeta_t\}$ is a sequence of independent standard normal variables. Let $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$, $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \gamma_1, \ldots, \gamma_r,$ $\mathbf{c}_1, \ldots, \mathbf{c}_r)'$ with $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_r)'$ and $\boldsymbol{\theta}_3 = (\varpi_i', \pi_i')'$, where $\varpi_i = (\varpi_{i1}, \ldots, \varpi_{iK})'$ and $\pi_i = (\pi_{i1}, \ldots, \pi_{iK})', i = 1, \ldots, q$. Assume that the maximum likelihood estimators of the parameters of (2.47) are asymptotically normal when $H_0 : \boldsymbol{\theta}_3 = \mathbf{0}$ is valid. Thus, under this null hypothesis, the LM statistic (2.15), with $\hat{u}_t = \hat{\varepsilon}_t^2/(\hat{h}_t^0 + \hat{g}_t^0) - 1, \hat{\mathbf{x}}_{1t}$ as in (2.50) and $\hat{\mathbf{x}}_{2t}$ as in (2.51) with $\hat{\mathbf{v}}_t = (\hat{\mathbf{v}}_{1,t}', \ldots, \hat{\mathbf{v}}_{K+2,t}')'$ where $\mathbf{v}_{it} = (\varepsilon_{t-1}^i, \ldots, \varepsilon_{t-q}^i)', i = 1, \ldots, K+2$, is asymptotically χ^2-distributed with* $\dim(\boldsymbol{\theta}_3)$ *degrees of freedom.*

Multiplicative misspecification

Consider the following extended TV-GARCH model

$$\varepsilon_t = \zeta_t (h_t + g_t)^{1/2} f_t^{1/2}. \tag{2.52}$$

Under the null hypothesis, $f_t \equiv 1$. This category entails the test for assessing the adequacy of the functional form of the estimated model. This test was already discussed when the TV-GARCH model was in the multiplicative form and the same considerations apply here.

The log-likelihood function for a single observation on (2.52) is

$$\ell_t = -\frac{1}{2} \ln 2\pi - \frac{1}{2}\{\ln(h_t + g_t) + \ln f_t)\} - \frac{\varepsilon_t^2}{2(h_t + g_t) f_t}$$

and the relevant block of the score due to $\boldsymbol{\theta}_3$, under H_0, has the form

$$\frac{\partial \ell_t}{\partial \boldsymbol{\theta}_3}\bigg|_{H_0} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t + g_t} - 1\right) \frac{\partial f_t}{\partial \boldsymbol{\theta}_3}\bigg|_{H_0}.$$

The hypothesis of interest is that the squared standardized error sequence is iid. Under the alternative, f_t is defined in (2.46). In this framework, the vector $\hat{\mathbf{x}}_{1t}$ is given as in (2.50) and $\hat{\mathbf{x}}_{2t} = \frac{\partial \hat{f}_t}{\partial \boldsymbol{\theta}_3}|_{H_0} = \hat{\mathbf{v}}_t$. The following Corollary defines the test statistic.

Corollary 12 *Consider the model (2.52) where* $\{\zeta_t\}$ *is a sequence of independent standard normal variables. Let* $\boldsymbol{\theta}_1 = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ *and* $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \gamma_1, \ldots, \gamma_r, \mathbf{c}_1, \ldots, \mathbf{c}_r)'$ *with* $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_r)'$. *Furthermore,* f_t *is defined by (2.46) such that* $\boldsymbol{\theta}_3 = (\phi_1, \ldots, \phi_s)'$ *and* $\hat{\mathbf{v}}_t = (\widehat{\zeta_1^2}, \ldots, \widehat{\zeta_s^2})'$. *Assume that the maximum likelihood estimators of the parameters of (2.52) are asymptotically normal when* $H_0 : \boldsymbol{\theta}_3 = \mathbf{0}$ *is valid. Thus, under this null hypothesis, the LM statistic (2.15), with* $\hat{u}_t = \hat{\varepsilon}_t^2/(\hat{h}_t^0 + \hat{g}_t^0) - 1$, $\hat{\mathbf{x}}_{1t}$ *as in (2.50) and* $\hat{\mathbf{x}}_{2t} = \hat{\mathbf{v}}_t$ *is asymptotically* χ^2-*distributed with* s *degrees of freedom.*

2.7 Simulation study

2.7.1 Monte Carlo design

In this section, we conduct a small simulation experiment to evaluate the finite-sample properties of the proposed parameter constancy tests. These are the tests against an additive and a multiplicative TV-GARCH specifications. Specifically, we shall investigate the size and power properties of the LM-type tests involved in the modelling strategies as well as the success rate of the specification procedures. Sample lengths of 1000, 2500 and 5000 observations have been used in all simulations. For each design, the total number of replications equals 2000. To avoid the initialization effects, the first 1000 observations have been discarded before generating the actual series. All the computations have been carried out using Ox, version 3.30 (see Doornik (2002)). The behaviour of the test statistics is examined for several data generating processes (DGP's) that can be nested in the following TV-GARCH specification:

$$\begin{aligned} y_t &= \varepsilon_t, \qquad \varepsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t) \\ h_t &= \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} + (\alpha_{01} + \alpha_{11}\varepsilon_{t-1}^2 + \beta_{11}h_{t-1})G_1(t^*; \gamma_1, c_1). \end{aligned} \tag{2.53}$$

The data generating processes are as following:

DGP (i):

$h_t = 0.10 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$

$\alpha_1 = \{0.05, 0.09, 0.10\}$ and $\beta_1 = \{0.80, 0.85, 0.90\}$

DGP (ii):

$h_t = 0.10 + \alpha_{01}G_1(t^*; \gamma_1, c_1) + 0.10\varepsilon_{t-1}^2 + 0.80h_{t-1}$

$\alpha_{01} = \{0.10, 0.30\}$

DGP (iii):

$h_t = 0.10 + (0.10 + \alpha_{11}G_1(t^*; \gamma_1, c_1))\varepsilon_{t-1}^2 + 0.80h_{t-1}$

$\alpha_{11} = \{0.05, 0.09\}$

DGP (iv):

$$h_t = (0.10 + \alpha_{01}G_1(t^*;\gamma_1,c_1)) + (0.10 + \alpha_{11}G_1(t^*;\gamma_1,c_1))\varepsilon_{t-1}^2 + 0.80h_{t-1}$$
$$\alpha_{01} = \{0.10, 0.30\} \text{ and } \alpha_{11} = \{0.05, 0.09\}$$

DGP (v):

$$h_t = 0.10 + 0.10\varepsilon_{t-1}^2 + (0.80 + \beta_{11}G_1(t^*;\gamma_1,c_1))h_{t-1}$$
$$\beta_{11} = \{0.05, 0.09\}$$

DGP (vi):

$$h_t = 0.10 + \alpha_{01}G_1(t^*;\gamma_1,c_1) + 0.10\varepsilon_{t-1}^2 + (0.80 + \beta_{11}G_1(t^*;\gamma_1,c_1))h_{t-1}$$
$$\alpha_{01} = \{0.10, 0.30\} \text{ and } \beta_{11} = \{0.05, 0.09\}$$

DGP (vii):

$$h_t = (0.10 + 0.10\varepsilon_{t-1}^2 + 0.85h_{t-1})(1 + \delta_1 G_1(t^*;\gamma_1,c_1))$$
$$\delta_1 = \{0.05, 0.08\}$$

The first six designs concern the additive TV-GARCH model, whereas the remaining one relates to the multiplicative model. In all these seven experiments, the midpoint of the change in volatility is at $c_1 = 0.5$, whereas the slope parameter γ_1 varies in the interval $\gamma_1 = \{5, 10\}$. Following the suggestion in Bollerslev (1986), recursive computation of h_t is initialized by using the estimated unconditional variance for the pre-sample values $t \leq 0$.

2.7.2 Finite sample properties

In this section we shall look at the small-sample properties of the modelling strategy for the TV-GARCH model. We first report results on the size and power properties of our parameter constancy tests. Then we turn to the specification of TV-GARCH models.

Size and power simulations

The size and the power results of the tests are presented in graphs following the recommendation by Davidson and MacKinnon (1998). Both the ordinary and the robustified versions of each test are computed using auxiliary regressions. Results of the size simulations appear in the form of p-value discrepancy plots in Figure 2.5. In these graphs, the difference between the empirical size and the nominal size is plotted against the nominal size. The upper panel of Figure 2.5 presents the results for the size simulations for the test against an additive alternative, whereas the bottom panel shows the empirical size results of the test against a multiplicative alternative. For each test we calculate the actual rejection frequencies for the three sample sizes at the following nominal levels: 0.1%, 0.3%, 0.5%, 0.7%, 0.9%, 1%, ..., 10%. The series are

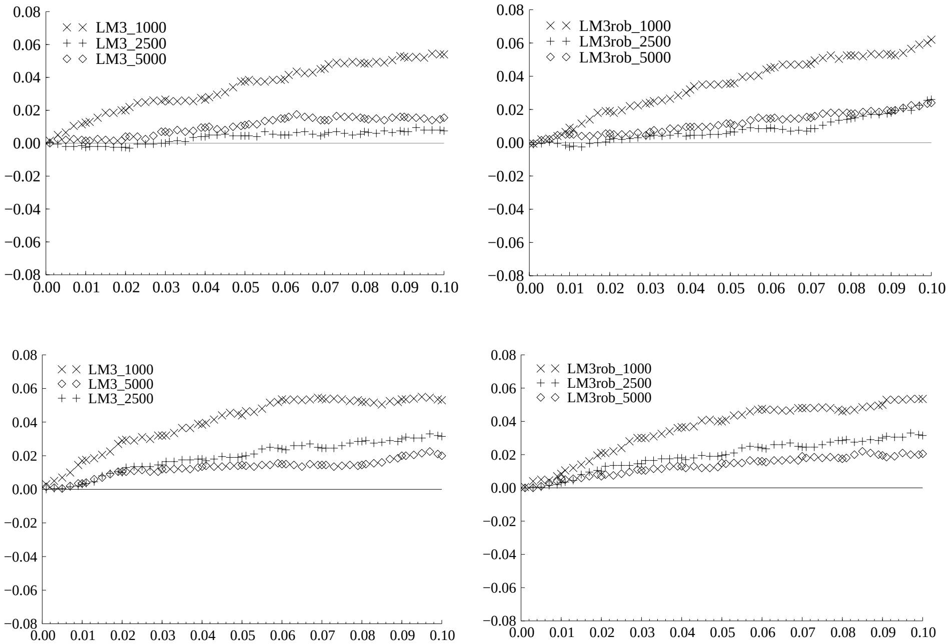


Figure 2.5 Size discrepancy plots of the additive (upper panel) and multiplicative (lower panel) parameter constancy tests. Both the ordinary (left) and the robust (right) versions of the tests are plotted.

generated from the GARCH model given by the DGP (i) where $\alpha_0 = 0.10, \alpha_1 = 0.10$ and $\beta_1 = 0.85$.

Both tests are somewhat size-distorted at the sample size $T = 1000$, but the results become more accurate as the sample size increases. For sample sizes typically used for modelling volatility clustering, such as $T = 2500$ and $T = 5000$, the tests are reasonably well-sized. Furthermore, the size distortions in the robust version of the tests do not differ too much from those in the non-robust test. Our main conclusion is that both the non-robust and robust versions of the test statistics are rather good approximations to the finite-sample distributions for $T \geq 2500$. Employing a robust test even when the errors are normal does not seem to lead to a large loss of power.

Although there exist several parameter constancy tests in the GARCH literature, none of them can be considered a direct benchmark for our parameter constancy tests. Because of this, in Figure 2.6 we only report power results for our tests. In these graphs the rejection frequencies are plotted against the nominal significance levels 0.1%, 0.3%, 0.5%, 0.7%, 0.9%, 1%, …, 10%. Instead of the size-adjusted power-size curves suggested by Davidson and MacKinnon (1998), we simply report power curves as the tests have good size properties.

The power results in Figure 2.6 have been obtained by generating artificial data

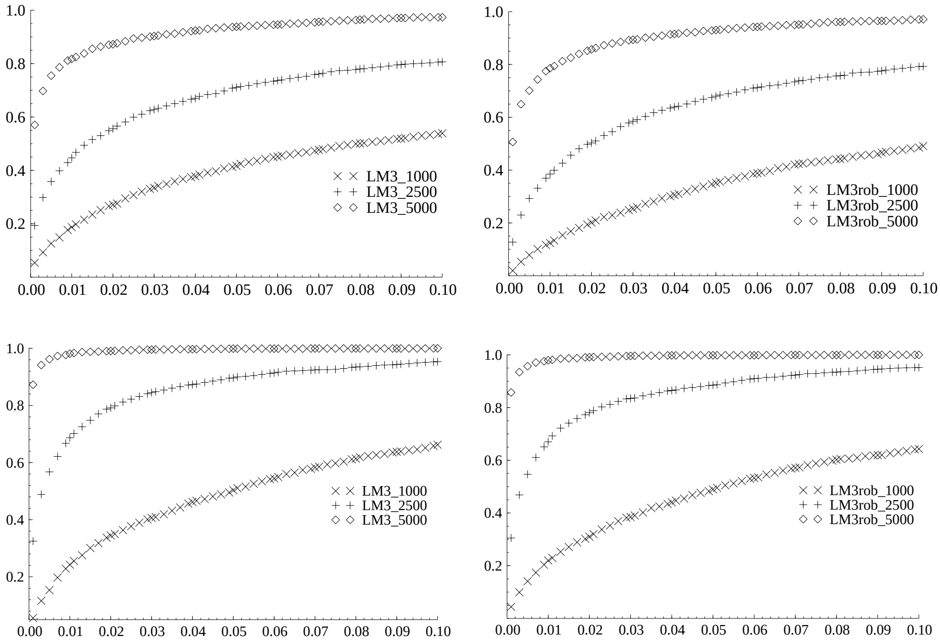


Figure 2.6 Power curves of the additive (upper panel) and multiplicative (lower panel) parameter constancy tests. Both the ordinary (left) and robust (right) versions of the tests are plotted.

from the DGP (ii) where the coefficient $\alpha_{01} = 0.10$, the slope parameter $\gamma_1 = 5$ and the location parameter $c_1 = 0.5$. The rejection frequencies of the additive LM test statistics shown in the top panel are moderate when $T = 1000$ and increase with the sample size. The pattern of the power results for the robustified version of the test is very similar to the non-robust one.

Rejection frequencies for the LM-type test against a multiplicative alternative are shown in the lower panel of Figure 2.6. The results refer to power simulations when the data generating process is a multiplicative TV-GARCH model (DGP vii). The coefficient $\delta_1 = 0.05$ and $\gamma_1 = 5$ as before. As expected, the rejection frequencies are an increasing function of the sample size and of the parameter δ_1 (as well as of the parameter α_{01} in the additive case). Moreover, the LM-type test statistic turns out to be very powerful even for short time series. Again, the behaviour of the robust version of the test in the power simulations is quite similar to that of the non-robust version.

Simulating the model selection strategy

In this section we consider the performance of the specific-to-general specification strategy for TV-GARCH models with an additive time-varying structure. This is done by studying the selection frequencies of various models. The specification procedure has been discussed in Section 2.4.1. A total of 2000 replications are carried out for each DGP and all three sample sizes. The first 1000 observations of each generated series are discarded to avoid the initialization effects. Throughout, we set $\alpha = 0.05$ for both the LM_1 and LM_3 versions of the test. The maximum number of transitions considered equals two. Furthermore, $\tau = 1/2$, which means that we halve the significance level of the test at each stage of the sequence.

Results for DGP (i) are reported in Table 2.1 (see Appendix B). The frequencies of the correct number of transitions are shown in boldface. The column labelled 'choice' refers to the number of transition functions selected. In general, the statistic LM_1 has better size properties than LM_3. However, in most cases, the test based on the third-order Taylor expansion also has an empirical size very close to the nominal size except when the sum $\alpha_1 + \beta_1$ is close to one and the sample size is less than 2500 observations.

Results for series generated from a model with a single transition function can be found in Table 2.2. We report separately an additive time-varying structure in each parameter of the GARCH model when $c_1 = 0.50$. This corresponds to the DGP's (ii), (iii) and (v). For all the cases, the parameters of the linear GARCH are $\alpha_0 = 0.10, \alpha_1 = 0.10$ and $\beta_1 = 0.80$. Clearly, the constant-parameter GARCH model is chosen too often for parameterizations with smoothest changes and shortest series. For large sample sizes, the selection frequencies of the true model become quite high even for very smooth changes. Again, the LM_1-test has higher power than LM_3. As expected, the correct model is selected more frequently for high than for low values of α_{01}, α_{11} or β_{11}. Moreover, the correct model is selected slightly more often when the change only occurs either in the constant α_0 or in the GARCH parameter β_1 than when it does in the ARCH parameter α_1.

The model selection frequencies when the series are generated from DGP (iv) are given in Table 2.3. The correct model is chosen more frequently when the change in α_{01} and α_{11} becomes large. It also becomes easier to identify a single transition when the slope parameter γ increases. Again, the results concern the case when the change occurs in the middle of the sample. Finally, Table 2.4 contains the frequencies of the selected models for the DGP (vi). In this case, the power of our procedure turns out to be very similar to that shown in Table 2.3. This may be explained by the fact that either changes in α_{01} and α_{11} or the ones in α_{01} and β_{11} simultaneously change the amplitude of clusters as well as the unconditional variance. We also carried out simulations for the DGP (vii) which are not reported in the paper. The results are almost identical to what is reported for the additive TV-GARCH model. Overall, the sequential procedure seems to work relatively well for all combinations of parameters considered and for sample sizes $T \geq 1000$.

2.8 Applications

In this section we shall present two empirical examples involving two financial time series, a stock index and an exchange rate return series. The former is the Standard

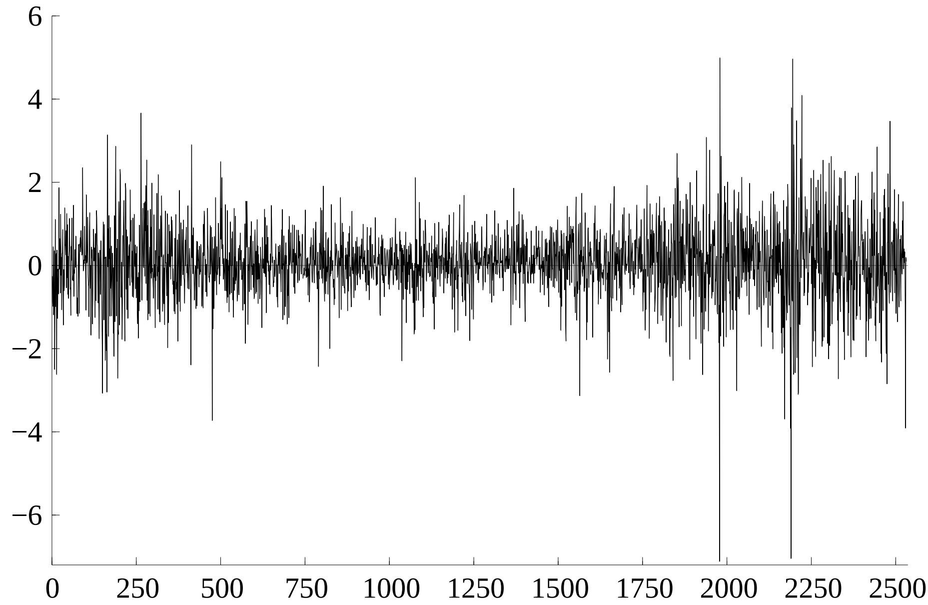


Figure 2.7 Daily returns of the S&P 500 composite index from January 2, 1990 until December 31, 1999 (2531 observations).

and Poor 500 composite index (S&P 500) and the latter the spot exchange rate of the Singapore dollar versus the U.S. dollar (SPD/USD). Both series are observed at a daily frequency and transformed into the continuously compounded rates of return.

2.8.1 Stock index returns

The daily S&P 500 return series was provided by the Yahoo-Quotes database. The sample extends from January 2, 1990, to December 31, 1999, which amounts to 2531 observations. The series is plotted in Figure 2.7. It contains periods of large volatility both in the beginning and at the end of the sample period, whereas the average volatility in the middle of the sample is somewhat lower than in both ends.

Summary statistics for the series can be found in the second column of Table 2.5. It is seen that there is both negative skewness and excess kurtosis in the series. Normality of the marginal distribution of the S&P 500 returns is strongly rejected. Robust skewness and kurtosis estimates (see Kim and White (2004) and Teräsvirta and Zhao (2007)) are also provided. The robust skewness measure is positive but very close to zero, which suggests that the asymmetry of the empirical distribution of the returns is due to a small number of outliers. The robust centred kurtosis that has value zero for the normal distribution indicates some excess kurtosis but much less than the conventional measure. This is in line with the robust skewness estimate. As expected, the null hypothesis of no ARCH is strongly rejected.

We first estimate a standard GARCH(1,1) model to this series. In order to save space, the results are not shown here. Results of the parameter constancy test against an additive time-varying structure are reported in Table 2.7. The test of parameter constancy against an additive TV-GARCH model, when several parameters are assumed to change under the alternative, rejects the null hypothesis. The tests against alternatives in which some parameters remain constant, suggest that the the intercept may be the main source of nonconstancy.

Instead of specifying and estimating an additive TV-GARCH model with a time-varying intercept, we test the iid hypothesis of our stochastic sequence $\{\varepsilon_t\}$ against deterministic change. This is Step 1 in the specification of multiplicative TV-GARCH models outlined in Subsection 2.4.2. The results can be found in Table 2.8. The null hypothesis is rejected very strongly as the p-value of the test equals 3×10^{-23}. The test sequence for specifying the structure of the deterministic function g_t points towards $K = 2$. Fitting the TV-GARCH model with a single transition function and $K = 2$ to the series and testing for another transition still leads to rejecting the null hypothesis. The p-value, however, is now considerably larger, equalling 0.0028, and the specification test sequence now clearly suggests $K = 1$. Accepting this outcome, fitting the corresponding TV-GARCH model to the series and testing for yet another transition yields the p-value 0.0623. If the null hypothesis is tested directly against a standard logistic transition function, the p-value equals 0.0197. Given the relatively large number of observations, this is not a small value, and the model with two transitions is tentatively accepted as the final model.

In this model, the estimate of g_t has the following form:

$$\widehat{g}_t = \{1 + \underset{(0.4265)}{1.7041} G_1(t^*; \widehat{\gamma}_1, \widehat{\mathbf{c}}_1) + \underset{(0.5455)}{1.7335} G_2(t^*; \widehat{\gamma}_2, \widehat{c}_2)\} \tag{2.54}$$

with

$$G_1(t^*; \widehat{\gamma}_1, \widehat{\mathbf{c}}_1) = (1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0100)}{0.1643})(t^* - \underset{(0.0831)}{0.6950})\})^{-1} \tag{2.55}$$

and

$$G_2(t^*; \widehat{\gamma}_2, \widehat{c}_2) = (1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0043)}{0.8534})\})^{-1}. \tag{2.56}$$

The graph of the deterministic component $\widehat{g}_t$ is depicted in Figure 2.9. The two transitions are clearly visible and illustrate how volatility first decreases and then increases over time. A GARCH model is fitted to the standardized residuals $\varepsilon_t/\widehat{g}_t^{1/2}$, and the estimated model is subjected to misspecification tests described in Subsection 2.6.1. Table 2.9 contains the test results. The hypothesis of 'no ARCH in GARCH' is not rejected for any lag length considered. As may be expected, the hypothesis of no additional transitions is not rejected either. There is, however, some indication of nonlinearity in the conditional variance as the GARCH(1,1) component is strongly rejected against a STGARCH(1,1) one for $K = 1$. In order to remedy this problem, we specify a GJR-GARCH(1,1) model for h_t.

The parameter estimates of the GJR-GARCH model can be found in Table 2.6. It is seen that the persistence factor equals $\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\gamma}_1/2 = 0.993$, so that the estimated model is practically an integrated GJR-GARCH model. For illustration, Table 2.6 also contains the parameter estimates at the point where the parameters in h_t have been estimated for the first time. It is seen that there is already a large change in the value of the log-likelihood compared to the maximum found for the GJR-GARCH(1,1) model. The persistence, however, has not yet decreased very much. Figure 2.10 contains the autocorrelations of $|\varepsilon_t|$ (Panel (a)) and those of $|\varepsilon_t|/\widehat{g}_t^{1/2}$ after a single iteration (Panel (b)). It is seen that the increase in the log-likelihood is mainly due to a decrease in the general level of the autocorrelations. At the same time, the autocorrelations retain the 'long-memory property', the very slow decay as a function of the lag, that is obvious in the autocorrelations of $|\varepsilon_t|$.

The log-likelihood considerably increases with further iterations, and the final persistence indicator has the remarkably low value $\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\gamma}_1/2 = 0.918$. A clear trade-off is observed here. When it is assumed that the process is stationary there is only one level (unconditional variance) to which the conditional variance converges when it is assumed that $z_t = 0$ for $t > t_0$. This convergence then takes a very long time ($\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\gamma}_1/2 = 0.993$ is very close to unity). In the TV-GJR-GARCH model this level is time-varying, and the rate of convergence to a particular level can thus be much more rapid than it is in the standard GJR-GARCH model. Panel (c) of Figure 2.10 now shows that the autocorrelations of $|\varepsilon_t|/\widehat{g}_t^{1/2}$ have decreased even further, and only few of them exceed two standard deviations of $|\varepsilon_t|$ under the iid normality assumption, marked by the straight line in the figure. A major part of the variation in the daily S&P 500 return series can thus be attributed to the slow-moving component g_t, and surprisingly little remains to be explained by the traditional GJR-GARCH component.

Table 2.10 contains the misspecification test results for this model. Even if the GJR-GARCH model is a rather crude representation of asymmetry compared to the smooth transition GARCH specification, it manages to capture most of the asymmetry. The p-value of the test of no additional nonlinearity, when applied to the TV-GJR-GARCH model, equals 0.024, which is much larger than 1×10^{-10} obtained when the test was applied to the estimated TV-GARCH(1,1) model. Applying the 1% significance level, the other misspecification tests do not reject the model either, and the TV-GJR-GARCH model is thus accepted to be our final model.

Figure 2.11 that contains the estimated conditional standard deviations $h_t^{1/2}$ of $\{\varepsilon_t\}$ for the GJR-GARCH(1,1) model and the ones of $\{\varepsilon_t/\widehat{g}_t^{1/2}\}$ illustrates the situation as well. For the GJR-GARCH model, see Panel (a), the graph looks rather 'nonstationary'. Some nonstationarity remains after a single iteration, as the autocorrelations of $\{\varepsilon_t/\widehat{g}_t^{1/2}\}$ in Panel (b) also demonstrate. From the graph in Panel (c) (the final model) it is seen that volatility is still changing over time, but there no longer seem to be persistent level changes. They have been absorbed by the deterministic component.

Column 4 in Table 2.5 contains the skewness and kurtosis estimates for $\varepsilon_t/\widehat{g}_t^{1/2}$. The negative skewness remains but, as can be expected from the other results, the excess kurtosis of the final $\varepsilon_t/\widehat{g}_t^{1/2}$ series is considerably less (2.8) than the original number (5.3). This is another illustration of the fact that volatility to be modelled by

h_t in the TV-GJR-GARCH model is much smaller than it is in the GJR-GARCH(1,1) model without the nonstationary component. Even the robust kurtosis estimate in Table 2.5 shows some decrease, but because its nonrobust value was already small, the decrease has remained rather modest.

In Figure 2.12, the estimated news impact curve of the standard GJR-GARCH(1,1) model is compared with corresponding curves of the TV-GJR-GARCH(1,1) model. The news impact curve of the TV-GJR-GARCH model is time-varying because it depends on g_{t-1}. The news impact curve of the GJR-GARCH model is time-invariant, and from the figure it is seen how the curve can vary over time in the TV-GJR-GARCH model. This curve is completely flat for $\varepsilon_{t-1} > 0$ because $\alpha_1 = 0$ in the model. Its estimate was originally slightly negative but statistically insignificant, and the model was re-estimated after restricting α_1 to zero. The curves based on the TV-GJR-GARCH model clearly show the obvious fact that when there is plenty of turbulence in the market, the news impact of a particular negative shock is smaller than it is when calm prevails. In the latter case, even a minor piece of 'bad news' (a negative shock) can be 'news', whereas in the former case, even a relatively large negative shock can have a rather small news component. This distinction cannot be made in the standard GJR-GARCH model. According to our TV-GJR-GARCH model, 'good news' (positive shocks) have no impact on volatility in this application.

2.8.2 Exchange rate data

The data of this section consist of daily returns of the spot SPD/USD exchange rate provided by the Federal Reserve Bank of New York. The time series is shown in Figure 2.8. It covers the period from May 1, 1997 until July 11, 2005, yielding a total of 2060 observations. At first sight, it appears that one can distinguish two different regimes in the series. A period of high volatility occurs during the East Asian financial crisis due to the significant depreciation of the Singapore dollar relative to the U.S. dollar. After the crisis, the volatility of the currency returns descends to a low level.

Descriptive statistics for the SPD/USD exchange rate returns are reported in Table 2.5. There is plenty of excess kurtosis, and the estimated skewness is strongly negative. These values are due to a limited number of large negative returns early in the series during the so-called Asian crisis. Naturally, the marginal distribution of the returns is far from normal. The robust measure of skewness indicates that there is in fact little skewness and the robust centred kurtosis is substantially smaller than its standard measure. The hypothesis of no ARCH is strongly rejected, as can be expected. The GARCH(1,1) model fitted to this exchange rate return series again shows high persistence of volatility. The estimate of α_1 is larger and that of β_1 smaller than in the S&P 500 model, which is a consequence of the fact that the kurtosis is larger in the exchange rate series than it is in the S&P 500 returns.

Parameter constancy of the GARCH(1,1) model is rejected against an additive TV-GARCH model. These test results are presented in Table 2.7. In this case, however, the rejection is not due to the intercept but rather to the other two parameters. As in the previous application, we shall not fit any additive TV-GARCH models to our return series but choose to work with the multiplicative model. The test of

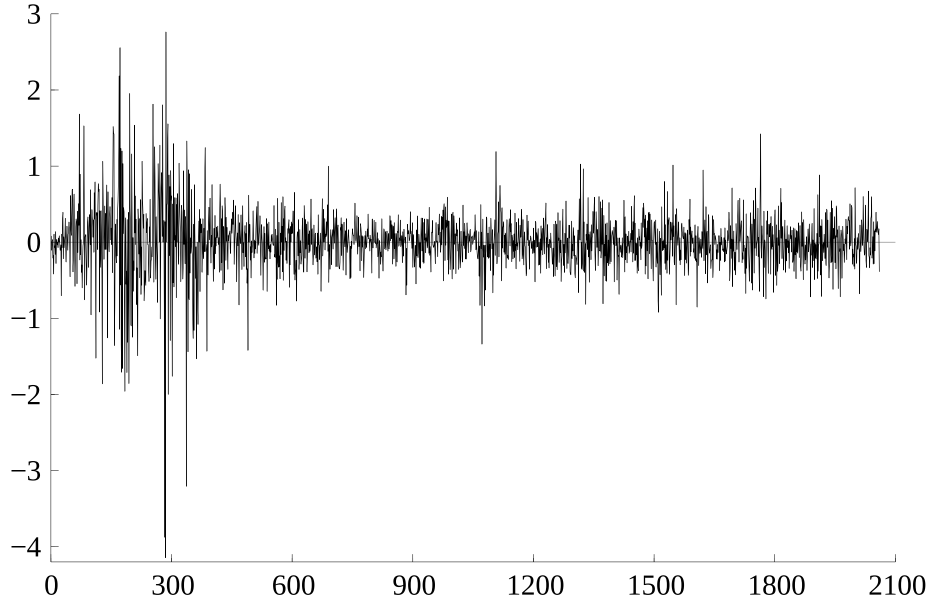


Figure 2.8 Daily returns of the Singapore Dollar versus US dollar exchange rate from May 1, 1997 until July 11, 2005 (2060 observations).

constant unconditional variance against a time-varying one has the p-value equal to 1×10^{-20}. Table 2.8 contains the outcomes of the sequence of specification tests. The results indicate that one should choose $K = 1$, that is, have a monotonically increasing transition function. A multiplicative TV-GARCH model with a single transition appears adequate in the sense that the test for another transition has $p = 0.14$. The diagnostic tests of this model in Table 2.9 do not reject the model. There is no remaining ARCH in the standardized errors, no evidence of higher-order structure in the GARCH component, and nothing suggests the existence of additional transitions. Finally, the linearity test against the smooth transition GARCH does not indicate remaining nonlinearity. Judging from these statistics, the model seems to be adequately specified. It is thus tentatively accepted as our final model for the SPD/USD daily return series.

The final estimates for the function g_t are as follows:

$$\widehat{g}_t = \{1 - \underset{(0.0074)}{0.7890} G_1(t^*; \widehat{\gamma}_1, \widehat{c}_1)\}, \tag{2.57}$$

where

$$G_1(t^*; \widehat{\gamma}_1, \widehat{c}_1) = (1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0014)}{0.2101})\})^{-1} \tag{2.58}$$

The graph of the transition function can be found in Figure 2.13. Figure 2.8 already shows that the volatility is high in the beginning and settles down to a lower level after about 500 observations (two years). From Table 2.5 it is seen that the excess kurtosis has decreased substantially from its value for $\{\varepsilon_t\}$ and, furthermore, that the skewness has been reduced from -0.9 to less than -0.3. This large reduction can be ascribed to the fact that the original skewness was due to a couple of very large negative returns during the Asian crisis. Their significance has subsequently been reduced in $\{\varepsilon_t/\widehat{g}_t^{1/2}\}$ where the conditional heteroskedasticity component has been standardized by the underlying nonstationary volatility component. Besides, according to the robust estimates the skewness has not been affected, which is in line with this conclusion as well.

The parameter estimates of the model appear in Table 2.6. It can be seen that even for the exchange rate series, the first iteration already has a large effect on the value of the log-likelihood. Figure 2.14 shows that at that stage, the autocorrelations of $|\varepsilon_t|/\widehat{g}_t^{1/2}$ are considerably lower than those of $|\varepsilon_t|$, although their decay as a function of the lag length is still slow. The final estimates indicate more persistence than there is in the S&P 500 case, but the decrease is still large compared to the GARCH(1,1) model. The decay rate of the autocorrelations of $|\varepsilon_t|/\widehat{g}_t^{1/2}$ in Figure 2.14 is quite rapid and looks more or less exponential. The first-order autocorrelation that was about 0.304 for $|\varepsilon_t|$ equals 0.121 for $|\varepsilon_t|/\widehat{g}_t^{1/2}$. The graph of the conditional variance h_t in Panel (a) of Figure 2.15 clearly shows the period of high volatility, which is the cause of the high persistence suggested by the GARCH(1,1) model. Panel (c) shows that in the final model this high-volatility period is explained by the deterministic component g_t, and that the graph of h_t does not show signs of nonstationarity. This is precisely what one would expect after a look at the parameter estimates in Table 2.6.

Figure 2.16 contains the estimated news impact curves of the traditional GARCH(1,1) model and the ones of the TV-GARCH(1,1) model for three regimes. It is seen that symmetry in the response of volatility to news is preserved in the latter model. This is obviously because of certain 'symmetry' of the exchange rates: good news for the US dollar may be bad news for the SPD, and vice versa. An additional result, similarly to the previous application, is the ability of the time-varying news impact curves to distinguish different reaction levels of volatility to news in calm and turbulent times. In general, the impact of news on volatility tends to be high in expansions and low in recessions.

2.9 Concluding remarks

In this paper we introduce two new nonstationary GARCH models whose parameters are allowed to have a smoothly time-varying structure. Time-variation of the (un)conditional variance is incorporated in the model either in an additive or a multiplicative form. This approach is appealing since most daily financial return series cover a long time period and non-constancy of parameters in models describing them therefore appears quite likely. We also develop a modelling strategy for our TV-GARCH specifications. In order to determine the appropriate number of transitions we propose a procedure consisting of a sequence of Lagrange multiplier tests. The test statistics can be robustified against deviations from the iid assumption. Our simulation experiments suggest that the parameter constancy tests have reasonable good properties already in samples of moderate size. The modelling strategy appears to work quite well for the data-generating processes that we simulate.

We put our TV-GARCH models to test by applying the modelling strategy to daily stock index and exchange rate returns. We find that parameter constancy against an additive and a multiplicative structure is strongly rejected for both return series. Fitting a traditional GARCH model to these series yields results that are quite different from the ones obtained by our approach and suggest the presence of long memory in volatility. Our results show that the long-memory type behaviour of the sample autocorrelation functions of the absolute returns may also be induced by changes in the unconditional variance. Once the model accounts for the time-variation in the baseline volatility or unconditional variance, the evidence for long memory is considerably weakened or even vanishes altogether.

An extension to multivariate GARCH models appears possible. The so-called Constant Conditional Correlation (CCC-) GARCH model by Bollerslev (1990) and its extensions typically make use of a standard GARCH(1,1) specification for conditional variances. These GARCH equations could be generalized to account for time-variation in parameters. An interesting question to investigate with our TV-GARCH specifications is how such a generalization would affect estimates of time-varying correlations in a situation in which there are changes in the unconditional variance of the return series. This and other extensions to multivariate models will be left for future work.

Appendix A: Proof of Theorem 1

Proof of Theorem 1. Assuming the independent innovations to be normally distributed, it follows that for model (2.11), the conditional log-likelihood function is given by

$$L_T(\boldsymbol{\theta}) = -\frac{T}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln h_t - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{h_t}.$$

Let $\boldsymbol{\theta}$ be a parameter vector partitioned as $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$. The null hypothesis is $\boldsymbol{\theta}_2 = \mathbf{0}$. The corresponding partition of the average score vector $\mathbf{q}_{(T)}(\boldsymbol{\theta})$ is $\mathbf{q}_{(T)}(\boldsymbol{\theta}) = (\mathbf{q}_{1(T)}(\boldsymbol{\theta}_1)', \mathbf{q}_{2(T)}(\boldsymbol{\theta}_2)')'$. Let h_t^0 denote the conditional variance under the null hypothesis and let the true parameter vector under H_0 be $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_1^{0\prime}, \mathbf{0}')'$. The Lagrange multiplier statistic is defined as follows:

$$\xi_{LM} = T\mathbf{q}_{(T)}(\hat{\boldsymbol{\theta}})'\mathbf{I}(\hat{\boldsymbol{\theta}})^{-1}\mathbf{q}_{(T)}(\hat{\boldsymbol{\theta}})$$

where T is the sample size, $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}_1', \mathbf{0}')'$ is the constrained maximum likelihood estimator of $\boldsymbol{\theta}$,

$$\mathbf{q}_{(T)}(\hat{\boldsymbol{\theta}}) = (\mathbf{0}', \mathbf{q}_{2(T)}(\mathbf{0})')' = (\mathbf{0}', \frac{1}{T}\sum_{t=1}^{T}\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2'}|_{\mathrm{H}_0})'$$

is the average score vector and $\mathbf{I}(\hat{\boldsymbol{\theta}})$ the information matrix, both evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. In this case, the partial derivatives with respect to $\boldsymbol{\theta}$ have the form

$$\frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\theta}} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\mathbf{x}_t$$

where $\mathbf{x}_t = (\mathbf{x}_{1t}', \mathbf{x}_{2t}')'$, with $\mathbf{x}_{1t} = \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\theta}_1}$ and $\mathbf{x}_{2t} = \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\theta}_2}$. Accordingly,

$$\mathbf{q}_{(T)}(\hat{\boldsymbol{\theta}}) = \frac{1}{2T}\sum_{t=1}^{T}\left(\frac{\varepsilon_t^2}{\hat{h}_t^0} - 1\right)\hat{\mathbf{x}}_t = \left(\mathbf{0}', \frac{1}{2T}\sum_{t=1}^{T}\left(\frac{\varepsilon_t^2}{\hat{h}_t^0} - 1\right)\hat{\mathbf{x}}_{2t}'\right)'$$

where $\hat{h}_t^0$ and $\hat{\mathbf{x}}_t = (\hat{\mathbf{x}}_{1t}', \hat{\mathbf{x}}_{2t}')'$ denote h_t^0 and $\mathbf{x}_t$, respectively, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Under normality, the population information matrix equals the negative expected value of the average Hessian matrix:

$$\mathbf{I}(\boldsymbol{\theta}) = -\mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right].$$

The Hessian of the log-likelihood equals

$$\sum_{t=1}^{T}\frac{\partial^2 \ell_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} = -\frac{1}{2}\sum_{t=1}^{T}\left[\frac{\varepsilon_t^2}{h_t^3}\frac{\partial h_t}{\partial \boldsymbol{\theta}}\frac{\partial h_t}{\partial \boldsymbol{\theta}'} + \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\left(\frac{\partial^2 h_t}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'} - \frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\theta}}\frac{\partial h_t}{\partial \boldsymbol{\theta}'}\right)\right]$$

so the information matrix becomes

$$\mathbf{I}(\boldsymbol{\theta}) = \frac{1}{2T}\sum_{t=1}^{T}\mathsf{E}\left[\frac{\varepsilon_t^2}{h_t^3}\frac{\partial h_t}{\partial\boldsymbol{\theta}}\frac{\partial h_t}{\partial\boldsymbol{\theta}'}\right] = \frac{1}{2T}\sum_{t=1}^{T}\mathsf{E}\mathbf{x}_t\mathbf{x}_t'.$$

As the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ is consistent for $\boldsymbol{\theta}_0$,

$$\mathbf{I}(\hat{\boldsymbol{\theta}}) = \frac{1}{2T}\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'$$

is consistent for $\mathbf{I}(\boldsymbol{\theta})$. Then the Lagrange multiplier type test statistic for testing parameter constancy has the standard form:

$$\begin{aligned}\xi_{LM} &= \frac{1}{2}\sum_{t=1}^{T}\hat{u}_t\hat{\mathbf{x}}_t'\left(\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{u}_t \\ &= \frac{1}{2}\sum_{t=1}^{T}\hat{u}_t\hat{\mathbf{x}}_{2t}'\left\{\sum_{t=1}^{T}\hat{\mathbf{x}}_{2t}\hat{\mathbf{x}}_{2t}' - \sum_{t=1}^{T}\hat{\mathbf{x}}_{2t}\hat{\mathbf{x}}_{1t}'\left(\sum_{t=1}^{T}\hat{\mathbf{x}}_{1t}\hat{\mathbf{x}}_{1t}'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_{1t}\hat{\mathbf{x}}_{2t}'\right\}^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_{2t}\hat{u}_t.\end{aligned}$$

where $\hat{u}_t = \varepsilon_t^2/\hat{h}_t^0 - 1$. Under H_0 and standard regularity conditions, the statistic ξ_{LM} has an asymptotic χ^2-distribution with $\dim(\boldsymbol{\theta}_2)$ degrees of freedom. ■

Appendix B: Tables and Figures

Table 2.1 Model selection frequencies based on the additive sequential procedure

α_1	β_1	Number of transitions	$T = 1000$ LM$_1$	$T = 1000$ LM$_3$	$T = 2500$ LM$_1$	$T = 2500$ LM$_3$	$T = 5000$ LM$_1$	$T = 5000$ LM$_3$
DGP (i): GARCH model with $\alpha_0 = 0.10$								
0.10	0.80	$r = 0$	**95.45**	**94.15**	**95.00**	**95.35**	**95.50**	**94.35**
		$r = 1$	3.36	4.10	3.90	3.05	3.30	3.90
		$r \geq 2$	1.20	1.75	1.10	1.60	1.20	1.75
0.10	0.85	$r = 0$	**94.70**	**91.90**	**94.30**	**94.10**	**95.10**	**93.70**
		$r = 1$	3.65	5.20	4.60	3.90	3.40	4.45
		$r \geq 2$	1.65	2.90	1.05	2.00	1.50	1.85
0.05	0.90	$r = 0$	**94.45**	**90.45**	**94.20**	**93.75**	**94.40**	**93.25**
		$r = 1$	4.00	6.35	4.50	4.20	4.00	4.40
		$r \geq 2$	1.55	3.20	1.30	2.05	1.60	2.35
0.09	0.90	$r = 0$	**90.45**	**76.95**	**92.80**	**88.10**	**94.30**	**90.85**
		$r = 1$	8.20	14.10	4.95	7.10	4.10	5.50
		$r \geq 2$	1.35	8.95	2.25	4.80	1.60	3.65

Notes: Selection frequencies in percentage of the standard LM parameter constancy test based on 2000 replications. The initial nominal significance level equals 5%. The columns 'LM$_1$' and 'LM$_3$' correspond to the test procedure based on the first-order and third-order Taylor expansions, respectively.

Table 2.2 Model selection frequencies based on the additive sequential procedure

Parameters		Number of transitions	$T = 1000$ LM$_1$	$T = 1000$ LM$_3$	$T = 2500$ LM$_1$	$T = 2500$ LM$_3$	$T = 5000$ LM$_1$	$T = 5000$ LM$_3$
DGP (ii): Change only in the constant								
α_{01}=0.10	γ_1=5	$r = 0$	64.75	72.65	31.75	50.55	6.70	19.25
		$r = 1$	**33.25**	**24.00**	**65.80**	**46.95**	**91.45**	**78.10**
		$r \geq 2$	2.00	3.35	2.45	2.50	1.85	3.65
	γ_1=10	$r = 0$	46.50	58.10	9.45	21.90	0.20	1.20
		$r = 1$	**51.35**	**37.35**	**88.35**	**75.75**	**97.45**	**96.65**
		$r \geq 2$	2.15	4.55	2.20	2.35	2.35	2.15
α_{01}=0.30	γ_1=5	$r = 0$	17.90	33.55	0.25	2.85	0.00	0.00
		$r = 1$	**78.35**	**60.90**	**97.15**	**93.75**	**97.35**	**96.75**
		$r \geq 2$	3.75	5.55	2.60	3.40	2.65	3.25
	γ_1=10	$r = 0$	5.75	11.15	0.00	0.00	0.00	0.00
		$r = 1$	**90.45**	**82.00**	**98.05**	**96.70**	**97.15**	**96.25**
		$r \geq 2$	3.80	6.85	1.95	3.30	2.85	3.75
DGP (iii): Change only in the ARCH component								
α_{11}=0.05	γ_1=5	$r = 0$	80.25	84.55	55.80	69.65	22.65	41.30
		$r = 1$	**18.70**	**13.40**	**42.70**	**28.95**	**75.15**	**56.45**
		$r \geq 2$	1.05	2.05	1.50	1.40	2.20	2.25
	γ_1=10	$r = 0$	68.10	76.35	29.05	47.00	3.65	10.80
		$r = 1$	**30.20**	**21.60**	**68.85**	**51.00**	**93.55**	**86.95**
		$r \geq 2$	1.70	2.05	2.10	2.00	2.80	2.25
α_{11}=0.09	γ_1=5	$r = 0$	50.30	62.05	7.75	22.20	0.00	1.10
		$r = 1$	**46.95**	**34.40**	**89.20**	**74.95**	**96.35**	**95.80**
		$r \geq 2$	2.75	3.55	3.05	2.85	3.65	3.10
	γ_1=10	$r = 0$	27.80	38.65	0.95	3.05	0.00	0.00
		$r = 1$	**68.95**	**56.90**	**95.70**	**94.70**	**96.85**	**96.80**
		$r \geq 2$	3.35	4.45	3.35	2.25	3.15	3.20
DGP (v): Change only in the GARCH component								
β_{11}=0.05	γ_1=5	$r = 0$	68.40	76.50	30.60	50.80	5.00	14.90
		$r = 1$	**30.00**	**21.10**	**67.40**	**47.05**	**93.30**	**83.05**
		$r \geq 2$	1.60	2.40	2.00	2.15	1.70	2.05
	γ_1=10	$r = 0$	50.15	62.30	8.35	21.20	0.10	0.70
		$r = 1$	**47.90**	**34.90**	**89.55**	**76.50**	**97.35**	**97.25**
		$r \geq 2$	1.95	2.80	2.30	2.30	2.55	2.05
β_{11}=0.09	γ_1=5	$r = 0$	18.75	30.10	0.00	1.80	0.00	0.00
		$r = 1$	**77.75**	**64.90**	**95.95**	**94.85**	**96.35**	**96.65**
		$r \geq 2$	3.50	5.00	4.05	3.35	3.65	3.35
	γ_1=10	$r = 0$	9.80	11.25	0.00	0.00	0.00	0.00
		$r = 1$	**85.90**	**81.70**	**96.60**	**96.20**	**97.05**	**96.00**
		$r \geq 2$	4.30	7.05	3.40	3.80	1.95	4.00

Notes: Selection frequencies in percentage of the standard LM parameter constancy test based on 2000 replications. The initial nominal significance level equals 5%. The columns 'LM$_1$' and 'LM$_3$' correspond to the test procedure based on the first-order and third-order Taylor expansions, respectively.

Table 2.3 Model selection frequencies based on the additive sequential procedure

α_{01}	α_{11}	γ_1	Number of transitions	$T = 1000$ LM$_1$	$T = 1000$ LM$_3$	$T = 2500$ LM$_1$	$T = 2500$ LM$_3$	$T = 5000$ LM$_1$	$T = 5000$ LM$_3$
DGP (iv): Change in the intercept and ARCH component									
0.10	0.05	5	$r = 0$	24.55	41.40	0.75	3.80	0.00	0.00
			$r = 1$	**72.75**	**55.05**	**96.70**	**93.85**	**97.65**	**97.50**
			$r \geq 2$	2.70	3.55	2.55	2.35	2.35	2.50
		10	$r = 0$	7.80	17.05	0.00	0.05	0.00	0.00
			$r = 1$	**88.80**	**79.45**	**97.80**	**97.65**	**97.35**	**97.40**
			$r \geq 2$	3.40	3.50	2.20	2.30	2.65	2.60
0.10	0.09	5	$r = 0$	11.80	25.50	0.00	0.80	0.00	0.00
			$r = 1$	**84.35**	**69.45**	**96.90**	**95.85**	**96.35**	**96.20**
			$r \geq 2$	3.85	5.05	3.10	3.35	3.65	3.80
		10	$r = 0$	3.25	7.45	0.00	0.00	0.00	0.00
			$r = 1$	**91.95**	**87.25**	**96.90**	**96.90**	**96.40**	**97.10**
			$r \geq 2$	4.80	5.30	3.10	3.10	3.60	2.90
0.30	0.05	5	$r = 0$	2.30	8.80	0.00	0.00	0.00	0.00
			$r = 1$	**93.20**	**86.35**	**97.35**	**97.40**	**97.05**	**96.90**
			$r \geq 2$	4.50	4.85	2.65	2.60	2.95	3.10
		10	$r = 0$	0.95	0.90	0.00	0.00	0.00	0.00
			$r = 1$	**94.95**	**93.65**	**96.95**	**97.15**	**97.15**	**97.20**
			$r \geq 2$	4.10	5.45	3.05	0.50	2.85	2.80
0.30	0.09	5	$r = 0$	1.65	5.80	0.00	0.00	0.00	0.00
			$r = 1$	**92.40**	**86.65**	**96.60**	**96.25**	**96.60**	**96.60**
			$r \geq 2$	5.95	7.55	3.40	3.75	3.40	3.40
		10	$r = 0$	0.55	0.45	0.00	0.00	0.00	0.00
			$r = 1$	**92.40**	**91.30**	**96.05**	**95.50**	**95.80**	**95.70**
			$r \geq 2$	7.05	8.25	3.95	4.50	4.20	4.30

Notes: Selection frequencies in percentage of the standard LM parameter constancy test based on 2000 replications. The initial nominal significance level equals 5%. The columns 'LM$_1$' and 'LM$_3$' correspond to the test procedure based on the first-order and third-order Taylor expansions, respectively.

Table 2.4 Model selection frequencies based on the additive sequential procedure

			Number of	$T = 1000$		$T = 2500$		$T = 5000$	
α_{01}	β_{11}	γ_1	transitions	LM_1	LM_3	LM_1	LM_3	LM_1	LM_3
DGP (vi): Change in the intercept and GARCH component									
0.10	0.05	5	$r = 0$	12.45	27.20	0.00	0.70	0.00	0.00
			$r = 1$	**84.85**	**69.55**	**97.30**	**96.60**	**97.60**	**97.25**
			$r \geq 2$	2.70	3.25	2.70	2.70	2.40	2.75
		10	$r = 0$	4.50	8.45	0.00	0.00	0.00	0.00
			$r = 1$	**92.50**	**87.85**	**97.75**	**97.50**	**97.05**	**97.30**
			$r \geq 2$	3.00	3.70	2.25	2.50	2.95	2.70
0.10	0.09	5	$r = 0$	2.95	7.45	0.00	0.00	0.00	0.00
			$r = 1$	**91.70**	**86.35**	**95.65**	**95.55**	**96.50**	**96.50**
			$r \geq 2$	5.35	6.20	4.35	4.45	3.50	3.50
		10	$r = 0$	3.20	1.55	0.00	0.00	0.00	0.00
			$r = 1$	**90.90**	**89.85**	**95.40**	**93.90**	**96.15**	**95.15**
			$r \geq 2$	5.90	8.60	4.60	6.10	3.85	4.85
0.30	0.05	5	$r = 0$	1.25	4.35	0.00	0.00	0.00	0.00
			$r = 1$	**94.65**	**90.00**	**96.70**	**96.50**	**97.25**	**97.15**
			$r \geq 2$	4.10	5.65	3.30	3.50	2.75	2.85
		10	$r = 0$	1.15	0.55	0.00	0.00	0.00	0.00
			$r = 1$	**95.05**	**95.00**	**96.80**	**96.20**	**97.20**	**96.50**
			$r \geq 2$	3.80	4.45	3.20	3.80	2.80	3.50
0.30	0.09	5	$r = 0$	0.60	1.70	0.00	0.00	0.00	0.00
			$r = 1$	**91.70**	**89.10**	**95.35**	**94.15**	**96.10**	**95.95**
			$r \geq 2$	7.70	9.20	4.65	5.85	3.90	4.05
		10	$r = 0$	1.25	0.20	0.00	0.00	0.00	0.00
			$r = 1$	**91.60**	**88.00**	**94.60**	**91.35**	**94.30**	**92.70**
			$r \geq 2$	7.15	11.80	5.40	8.65	5.70	7.30

Notes: Selection frequencies in percentage of the standard LM parameter constancy test based on 2000 replications. The initial nominal significance level equals 5%. The columns 'LM_1' and 'LM_3' correspond to the test procedure based on the first-order and third-order Taylor expansions, respectively.

Table 2.5 Descriptive statistics and diagnostics for the daily returns

	S&P 500 returns			SPD/USD returns		
	ε_t	$\varepsilon_t/\hat{g}_t^{1/2}$	$\varepsilon_t/(\hat{h}_t\hat{g}_t)^{1/2}$	ε_t	$\varepsilon_t/\hat{g}_t^{1/2}$	$\varepsilon_t/(\hat{h}_t\hat{g}_t)^{1/2}$
Minimum	−7.1127	−4.3309	−6.3139	−4.1444	−1.9042	−6.0724
Maximum	4.9887	3.0374	4.0498	2.7618	1.4231	4.0671
Skewness	−0.3678	−0.3361	−0.3898	−0.9045	−0.2839	−0.2424
Robust SK	0.0325	0.0318	0.0229	−0.0045	−0.0165	−0.0217
Ex.kurtosis	5.2867	2.7996	2.1736	14.593	3.2055	2.1941
Robust KR	0.2541	0.1737	0.1503	0.1662	0.1120	0.1030
Std. dev.	0.8912	0.6120	0.9980	0.4150	0.2887	0.9971
Mean	0.0538	0.0407	0.0621	0.0077	0.0035	0.0142
LJB	3004.5	874.21	562.31	18558.6	909.62	433.4
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ARCH(4)	154.19	55.34	4.056	339.69	108.07	5.111
	(3×10^{-32})	(3×10^{-11})	(0.3478)	(3×10^{-72})	(2×10^{-22})	(0.2761)
T	2531	2531	2531	2060	2060	2060

Notes: LJB denotes the Lomnicki-Jarque-Bera test. ARCH(4) is the fourth-order ARCH LM test statistic described in Engle (1982). Robust SK denotes the robust measure for skewness based on quantiles proposed by Bowley (see Kim and White (2004)) and the robust KR denotes the robust centred coefficient for kurtosis proposed by Moors (see Kim and White (2004)). The numbers in parentheses are p-values.

Table 2.6 Estimation results for the GJR-GARCH and GARCH models in the two applications

GJR-GARCH model						
			$\varepsilon_t/\hat{g}_{t_{S\&P500}}^{1/2}$			
	S&P 500		First iteration		Last iteration	
	Estimate	(Std. error)	Estimate	(Std. error)	Estimate	(Std. error)
$\hat{\alpha}_0$	0.009	(0.003)	0.024	(0.007)	0.033	(0.008)
$\hat{\alpha}_1$	0.014	(0.007)	–	(–)	–	(–)
$\hat{\beta}_1$	0.939	(0.012)	0.901	(0.019)	0.855	(0.031)
$\hat{\gamma}_1$	0.079	(0.017)	0.123	(0.024)	0.125	(0.023)
$\hat{\alpha}_1 + \hat{\beta}_1 + \frac{\hat{\gamma}_1}{2}$	0.993		0.963		0.918	
Log-Lik	−3034.98		−2760.91		−2280.98	

GARCH model						
			$\varepsilon_t/\hat{g}_{t_{SPD/USD}}^{1/2}$			
	SPD/USD		First iteration		Last iteration	
	Estimate	(Std. error)	Estimate	(Std. error)	Estimate	(Std. error)
$\hat{\alpha}_0$	9×10^{-4}	(3×10^{-4})	0.002	(0.001)	0.003	(0.001)
$\hat{\alpha}_1$	0.057	(0.017)	0.058	(0.019)	0.065	(0.021)
$\hat{\beta}_1$	0.937	(0.020)	0.929	(0.027)	0.901	(0.031)
$\hat{\alpha}_1 + \hat{\beta}_1$	0.994		0.987		0.966	
Log-Lik	−635.66		−455.38		−283.45	

Table 2.7 Test of parameter constancy of the GARCH(1,1) model against a time-varying GARCH model with additive structure for several combinations of parameters

S&P 500 returns

	Parameter constancy test		Decision rule for selecting K					
			H_{03}		H_{02}		H_{01}	
	LM	p-value	LM_3	p-value	LM_2	p-value	LM_1	p-value
α_0	16.695	8×10^{-4}	0.087	0.7680	15.713	7×10^{-5}	0.901	0.3425
α_1	9.477	0.0236	3.364	0.0667	4.652	0.031	1.473	0.2249
β_1	8.779	0.0324	1.886	0.1697	6.083	0.014	0.817	0.3660
α_0 and α_1	22.829	7×10^{-4}	4.293	0.1169	17.071	2×10^{-4}	1.507	0.4708
α_0 and β_1	19.974	0.0028	2.181	0.3360	16.900	2×10^{-4}	0.914	0.6331
α_0, α_1 and β_1	26.415	0.0017	6.694	0.0823	16.518	9×10^{-4}	3.277	0.3509

SPD/USD returns

	Parameter constancy test		Decision rule for selecting K					
			H_{03}		H_{02}		H_{01}	
	LM	p-value	LM_3	p-value	LM_2	p-value	LM_1	p-value
α_0	5.169	0.1598	1.879	0.1705	3.011	0.0827	0.283	0.5946
α_1	15.291	0.0016	3.841	0.0500	7.598	0.0058	3.888	0.0486
β_1	10.967	0.0119	1.944	0.1633	6.721	0.0095	2.318	0.1279
α_0 and α_1	17.435	0.0078	3.750	0.1533	4.901	0.0863	8.830	0.0121
α_0 and β_1	13.523	0.0354	1.679	0.4319	4.209	0.1219	7.660	0.0217
α_0, α_1 and β_1	20.816	0.0135	6.886	0.0756	4.795	0.1874	9.203	0.0267

Table 2.8 Results of the sequence of tests of constant unconditional variance against a time-varying GARCH model with multiplicative structure

Transitions in the alternative model	Parameter constancy test		Decision rule for selecting K					
			H_{03}		H_{02}		H_{01}	
	LM	p-value	LM_3	p-value	LM_2	p-value	LM_1	p-value
Single transition								
S&P 500	107.79	3×10^{-23}	5.96	0.0146	66.99	3×10^{-16}	36.03	2×10^{-9}
SPD/USD	95.74	1×10^{-20}	0.10	0.7528	32.73	1×10^{-8}	63.93	1×10^{-15}
Double transition								
S&P 500	14.10	0.0028	1.31	0.2523	0.06	0.8006	12.74	0.0004
SPD/USD	5.51	0.1380	3.09	0.0788	0.71	0.3993	1.72	0.1903
Triple transition								
S&P 500	7.32	0.0623	1.86	0.1730	0.03	0.8611	5.44	0.0197
SPD/USD	–	–	–	–	–	–	–	–

Table 2.9 Misspecification tests for the GARCH models in the two applications

	No ARCH-in-GARCH			GARCH(1,1) vs GARCH(1,2)	GARCH(1,1) vs GARCH(2,1)	No additional transition	No STGARCH with $K=1$
	$r=1$	$r=5$	$r=10$				
S&P 500							
LM test	0.841	6.167	7.223	0.001	0.303	9.993	20.088
p-value	0.359	0.290	0.704	0.971	0.582	0.019	4×10^{-5}
$\varepsilon_t/\hat{g}_{t_{S\&P500}}^{1/2}$							
LM test	0.155	5.115	6.824	0.558	0.738	0.677	45.387
p-value	0.694	0.402	0.742	0.455	0.390	0.879	1×10^{-10}
	No ARCH-in-GARCH			GARCH(1,1) vs GARCH(1,2)	GARCH(1,1) vs GARCH(2,1)	No additional transition	No STGARCH with $K=1$
	$r=1$	$r=5$	$r=10$				
SPD/USD							
LM test	0.343	12.83	15.94	0.666	2.772	14.93	2.741
p-value	0.558	0.025	0.101	0.415	0.096	0.002	0.254
$\varepsilon_t/\hat{g}_{t_{SPD/USD}}^{1/2}$							
LM test	0.014	6.988	10.34	0.064	2.607	2.133	1.167
p-value	0.905	0.222	0.411	0.780	0.106	0.545	0.558

Notes: The tests are those against remaining ARCH in the standardized residuals, GARCH(1,2) and GARCH(2,1) models, additional transition in the function g_t, and STGARCH(1,1) model of order 1.

Table 2.10 Misspecification tests for the GJR-GARCH model in the stock returns application

	No ARCH-in-GARCH			GARCH(1,1) vs	GARCH(1,1) vs	No additional	No STGARCH
	$r = 1$	$r = 5$	$r = 10$	GARCH(1,2)	GARCH(2,1)	transition	with $K = 1$
S&P 500							
LM test	0.518	2.094	3.060	21.874	8.317	17.176	16.804
p-value	0.472	0.836	0.980	2×10^{-5}	0.004	7×10^{-4}	0.002
$\varepsilon_t/\hat{g}_{t_{S\&P500}}^{1/2}$							
LM test	2.450	4.963	7.821	8.113	3.293	0.184	11.235
p-value	0.118	0.420	0.646	0.017	0.070	0.980	0.024

Notes: The tests are those against remaining ARCH in the standardized residuals, GARCH(1,2) and GARCH(2,1) models, additional transition in the function g_t, and STGARCH(1,1) model of order 1.

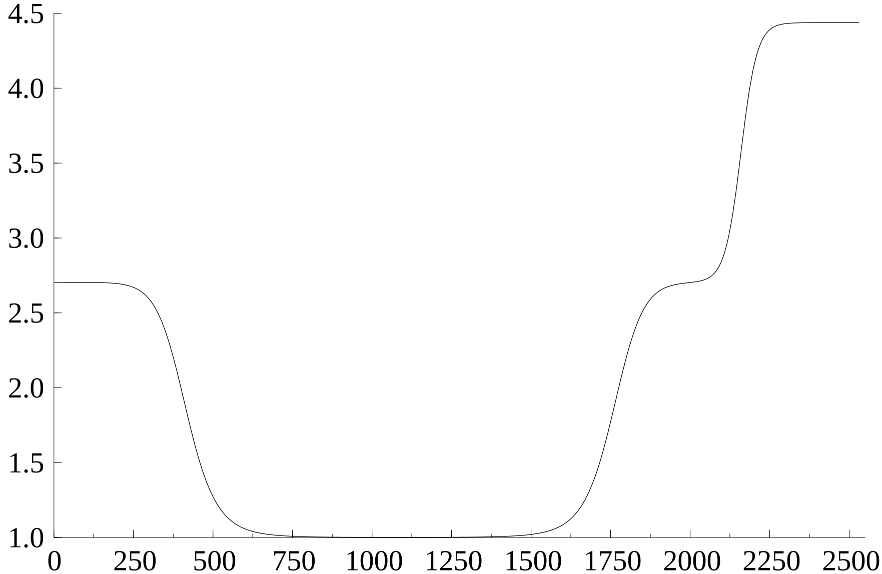
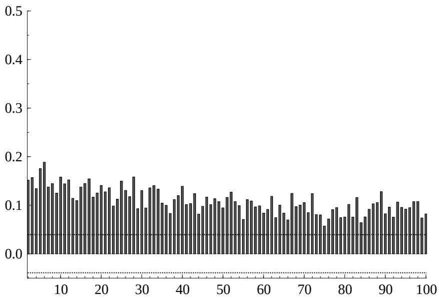
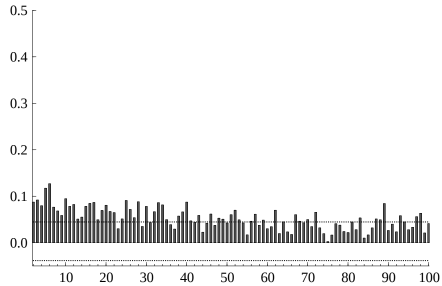


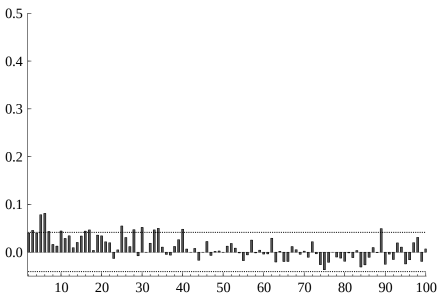
Figure 2.9 Graph of the final estimated function g_t for the S&P 500 returns model as a smooth function of the rescaled time variable t^* as given in (2.54)-(2.56)



(a) S&P 500 returns

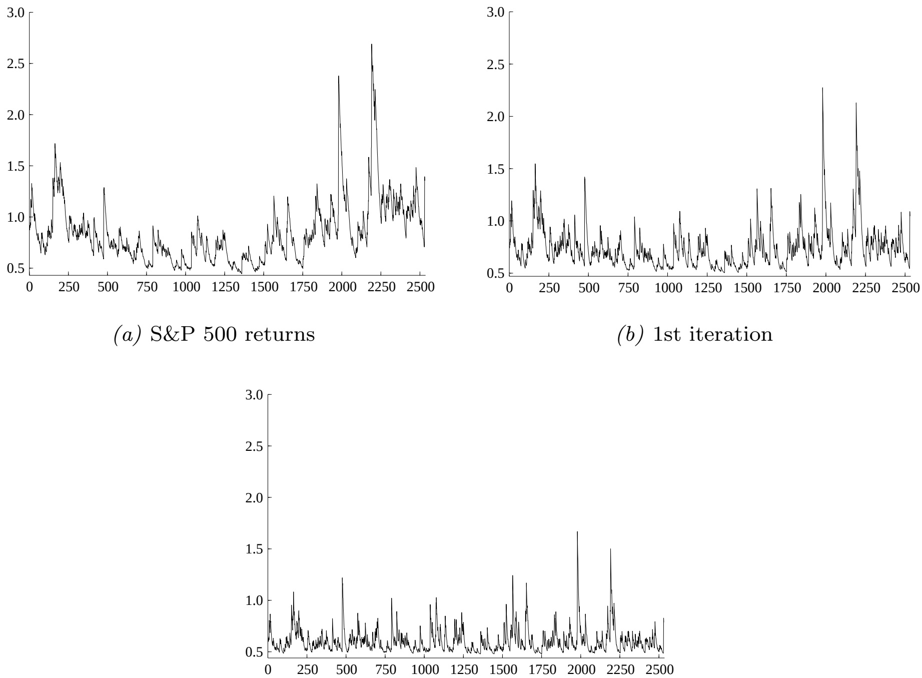


(b) 1st iteration



(c) 23rd iteration

Figure 2.10 Sample autocorrelations of absolute log returns of the S&P 500 and the standardized variable $|\varepsilon_t|/\hat{g}_{t_{S\&P500}}^{1/2}$ for the first and the final iterations with the 95% confidence bounds.



(a) S&P 500 returns

(b) 1st iteration

(c) 23rd iteration

Figure 2.11 Conditional standard deviation of the GJR-GARCH(1,1) model for the S&P 500 returns and the standardized variable $\varepsilon_t/\hat{g}_{t_{S\&P500}}^{1/2}$ for the first and the final iterations.

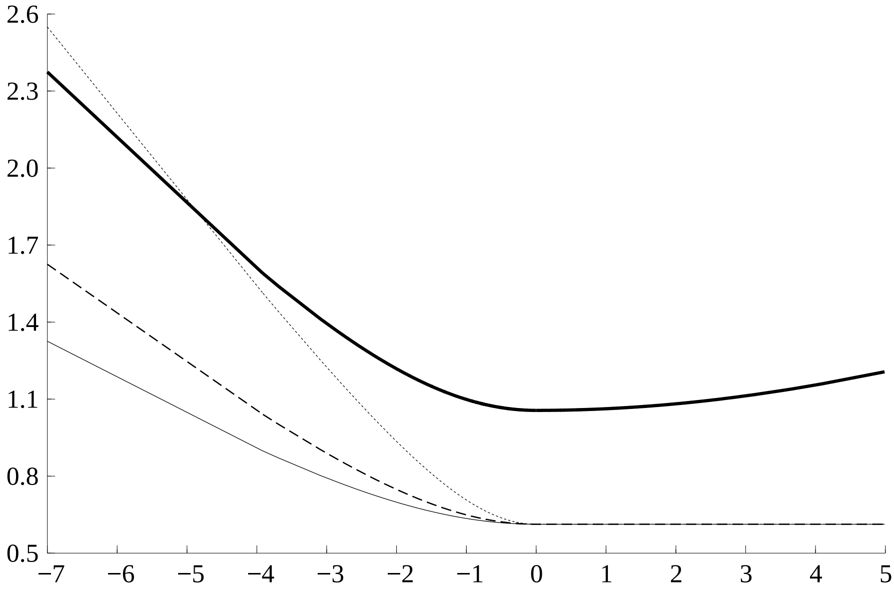


Figure 2.12 News impact curves of the GJR-GARCH(1,1) (solid line in boldface) and the TV-GJR-GARCH(1,1) models for several regimes. The time-varying news impact curves are plotted for the lower regime, i.e. $G_1(t^*) = G_2(t^*) = 0$ (dotted line), for an intermediate regime, i.e. $G_1(t^*) = 1$ and $G_2(t^*) = 0$ (dashed line) and for the higher regime, i.e. $G_1(t^*) = G_2(t^*) = 1$ (solid line).

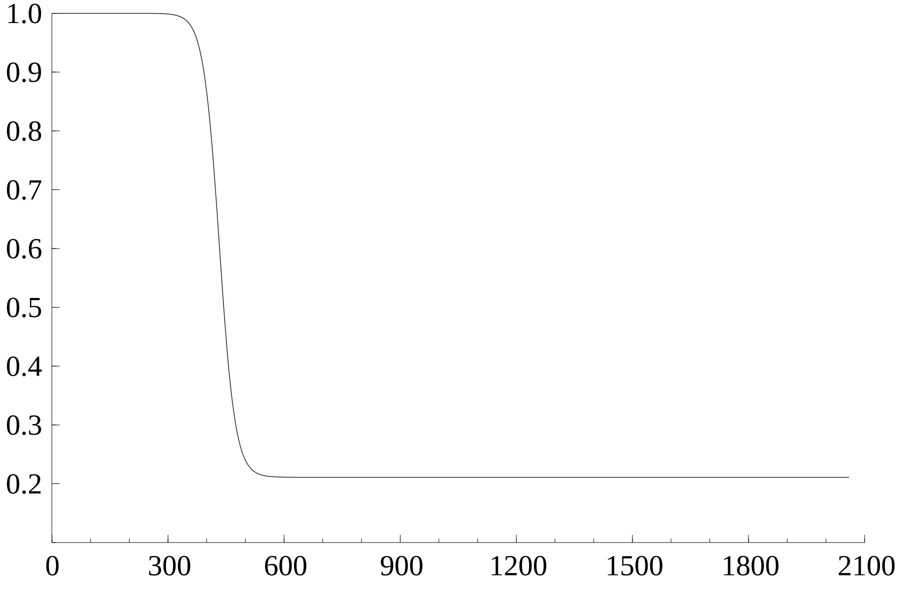
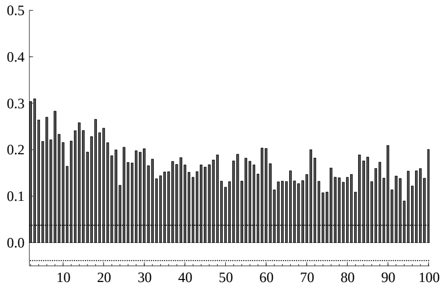
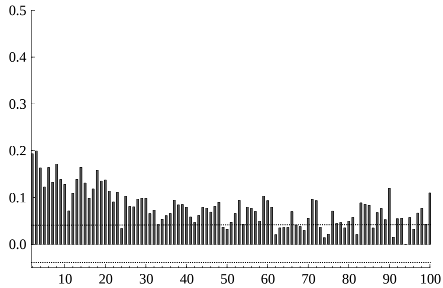


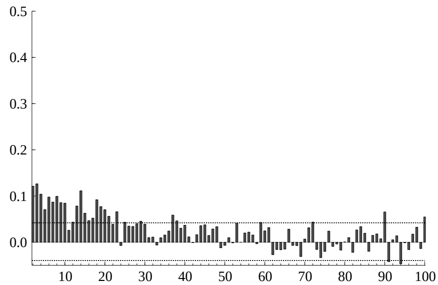
Figure 2.13 Graph of the final estimated function g_t for the SPD/USD returns model as a smooth function of the rescaled time variable t^* as given in (2.57)-(2.58).



(a) SPD/USD returns

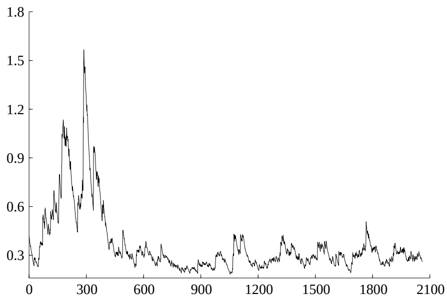


(b) 1st iteration

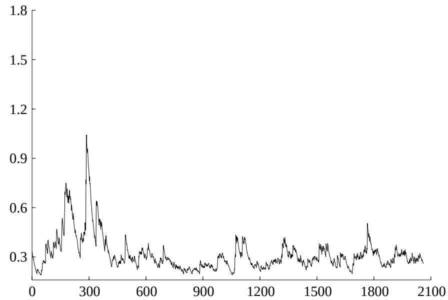


(c) 41st iteration

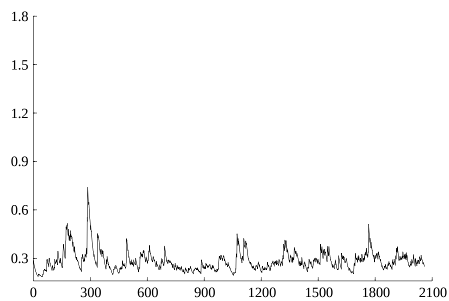
Figure 2.14 Sample autocorrelations of absolute log returns of the SPD/USD and for the standardized variable $|\varepsilon_t|/\hat{g}^{1/2}_{t_{SPD/USD}}$ for the first and the final iterations with the 95% confidence bounds.



(a) SPD/USD returns



(b) 1st iteration



(c) 41st iteration

Figure 2.15 Conditional standard deviation of the GARCH(1,1) model for the SPD/USD returns and for the standardized variable $\varepsilon_t/\hat{g}^{1/2}_{t_{SPD/USD}}$ for the first and the final iterations.

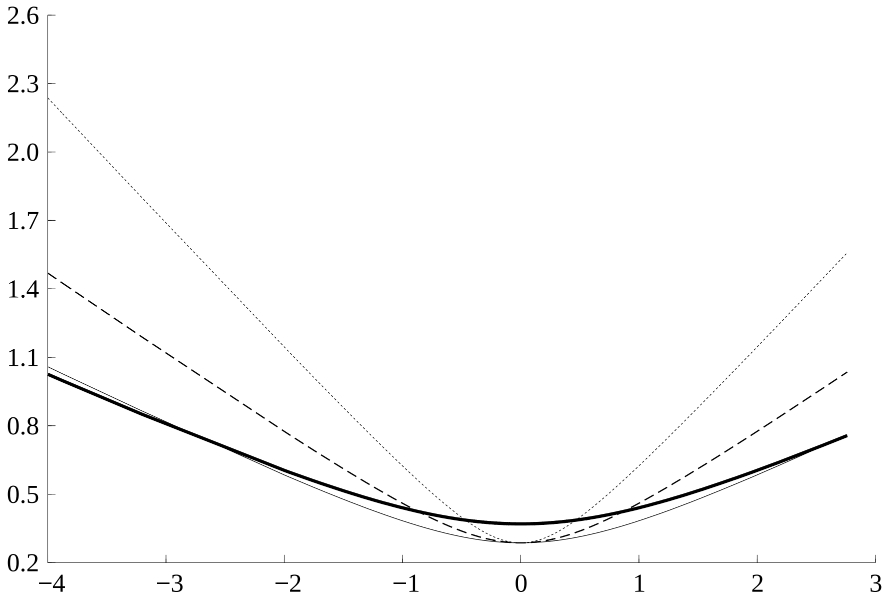


Figure 2.16 News impact curves of the GARCH(1,1) (solid line in boldface) and the TV-GARCH(1,1) models for several regimes. The time-varying news impact curves are plotted for the lower regime, i.e. $G_1(t^*) = 0$ (dotted line), for an intermediate regime, i.e. $G_1(t^*) = 0.5$ (dashed line) and for the higher regime, i.e. $G_1(t^*) = 1$ (solid line)

Bibliography

Anderson, H. M., K. Nam, and F. Vahid (1999): "Asymmetric Nonlinear Smooth Transition GARCH Models," in *Nonlinear Time Series Analysis of Economic and Financial Data*, ed. by P. Rothman, pp. 191–207. Kluwer Academic Publishers, Boston.

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

Baillie, R. T., and C. Morana (2007): "Modeling Long Memory and Structural Breaks in Conditional Variances: An Adaptive FIGARCH Approach," Working Paper No. 593, Queen Mary University of London, Department of Economics.

Berkes, I., E. Gombay, L. Horváth, and P. Kokoszka (2004): "Sequential Change-Point Detection in GARCH(p,q) Models," *Econometric Theory*, 20, 1140–1167.

Berkes, I., L. Horváth, and P. Kokoszka (2003): "GARCH Processes: Structure and Estimation," *Bernoulli*, 9, 201–207.

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

Bollerslev, T. (1990): "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *The Review of Economic and Statistics*, 72, 498–505.

Cai, J. (1994): "A Markov Model of Unconditional Variance in ARCH," *Journal of Business & Economic Statistics*, 12, 309–316.

Dahlhaus, R., and S. Subba Rao (2006): "Statistical Inference for Time-Varying ARCH Processes," *Annals of Statistics*, 34, 1075–1114.

Davidson, R., and J. G. MacKinnon (1998): "Graphical Methods for Investigating the Size and Power of Hypothesis Tests," *The Manchester School*, 66, 1–26.

Davies, R. B. (1977): "Hypothesis Testing When a Nuisance Parameter is Present Only under the Alternative," *Biometrika*, 64, 247–254.

Diebold, F. X. (1986): "Comment on "Modelling the Persistence of Conditional Variance", by R. Engle and T. Bollerslev," *Econometric Reviews*, 5, 51–56.

Doornik, J. A. (2002): *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press, see also www.nuff.ox.ac.uk/Users/Doornik, 3rd edn.

Eitrheim, Ø., and T. Teräsvirta (1996): "Testing the Adequacy of Smooth Transition Autoregressive Models," *Journal of Econometrics*, 74, 59–75.

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

Engle, R. F., and J. Gonzalo Rangel (2008): "The Spline-GARCH Model for Low-Frequency Volatility and its Global Macroeconomic Causes," *Review of Financial Studies*, 21, 1187–1222.

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

Gallant, R. (1984): "The Fourier Flexible Form," *American Journal of Agricultural Economics*, 66, 204–208.

Geweke, J., and S. Porter-Hudak (1983): "The Estimation and Application of Long-Memory Time-Series Models," *Journal of Time Series Analysis*, 4, 221–238.

González-Rivera, G. (1998): "Smooth Transition GARCH Models," *Studies in Nonlinear Dynamics and Econometrics*, 3, 61–78.

Granger, C. W. J. (1993): "Strategies for Modelling Nonlinear Time-Series Relationships," *Economic Record*, 69, 233–239.

Hagerud, G. E. (1997): "A New Non-Linear GARCH Model," Ph.D. thesis, EFI, Stockholm School of Economics.

Hamilton, J. D., and R. Susmel (1994): "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307–333.

Hansen, B. E. (1996): "Inference When a Nuisance Parameter is Not Identified under the Null Hypothesis," *Econometrica*, 64, 413–430.

Jensen, S. T., and A. Rahbek (2004): "Asymptotic Inference for Nonstationary GARCH," *Econometric Theory*, 20, 1203–1226.

Kim, T.-H., and H. White (2004): "On More Robust Estimation of Skewness and Kurtosis," *Finance Research Letters*, 1, 56–73.

Lamoureux, C. G., and W. D. Lastrapes (1990): "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business & Economic Statistics*, 8, 225–234.

Li, W. K., and T. K. Mak (1994): "On the Squared Residual Autocorrelations in Non-Linear Time Series With Conditional Heteroskedasticity," *Journal of Time Series Analysis*, 15, 627–636.

Lin, C.-F. J., and T. Teräsvirta (1994): "Testing the Constancy of Regression Parameters Against Continuous Structural Change," *Journal of Econometrics*, 62, 211–228.

Ling, S., and W. Li (1997): "Diagnostic Checking of Nonlinear Multivariate Time Series with Multivariate ARCH Errors," *Journal of Time Series Analysis*, 18, 447–464.

Ling, S., and M. McAleer (2003): "Asymptotic Theory for a New Vector ARMA-GARCH Model," *Econometric Theory*, 19, 280–310.

Lundbergh, S., and T. Teräsvirta (2002): "Evaluating GARCH Models," *Journal of Econometrics*, 110, 417–435.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

Meitz, M., and P. Saikkonen (in press): "Ergodicity, Mixing, and Existence of Moments of a Class of Markov Models with Applications to GARCH and ACD Models," *Econometric Theory*.

Mikosch, T., and C. Stărică (2004): "Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects," *The Review of Economics and Statistics*, 86, 378–390.

Teräsvirta, T. (1994): "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89, 208–218.

——— (1998): "Modeling Economic Relationships with Smooth Transition Regressions," in *Handbook of Applied Economic Statistics*, ed. by A. Ullah, and D. Giles, pp. 507–552. New York: Dekker.

Teräsvirta, T., and Z. Zhao (2007): "Stylized Facts of Return Series, Robust Estimates, and Three Popular Models of Volatility," SSE/EFI Working Paper Series in Economics and Finance No. 662.

Wooldridge, J. M. (1990): "A Unified Approach to Robust, Regression-Based Specification Tests," *Econometric Theory*, 6, 17–43.

——— (1991): "On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances," *Journal of Econometrics*, 47, 5–46.

Chapter 3

Modelling Changes in the Unconditional Variance of Long Stock Return Series

Modelling Changes in the Unconditional Variance of Long Stock Return Series[1]

Abstract

In this paper we develop a testing and modelling procedure for describing the long-term movements over very long return series. For the purpose, we assume that volatility is multiplicatively decomposed into a conditional and an unconditional component as in Amado and Teräsvirta (2008). The latter component is modelled by incorporating smooth changes so that the unconditional variance is allowed to evolve slowly over time. Statistical inference is used for specifying the parameterization of the time-varying component by applying a sequence of Lagrange multiplier tests. The model building procedure is illustrated with an application to the daily returns of the DJIA index covering a period of eighty three years of financial market history. Two major conclusions are as follows. First, the LM tests strongly reject the assumption of constancy of the unconditional variance. Second, the results show that the long-memory property in volatility may be explained by ignored changes in the unconditional variance of the long series.

[1] This paper is a joint work with Timo Teräsvirta.

Acknowledgements: The first author would like to acknowledge financial support from the Louis Fraenckels Stipendiefond. Part of this research was done while the first author was visiting CREATES, University of Aarhus, whose kind hospitality is gratefully acknowledged. The Center for Research in Econometric Analysis of Time Series, CREATES, is funded by the Danish National Research Foundation. The responsibility for any errors and shortcomings in this paper remains ours.

3.1 Introduction

The observation that deterministic shifts in long return series can generate long-memory behaviour has received much attention in recent years. Most of the work in this topic is related with the study of the behaviour of standard statistical tools and model misspecification under nonstationarity. Early studies include Diebold (1986) and Lamoureux and Lastrapes (1990) who suggested that occasional level shifts in the intercept of the first-order GARCH model can bias the estimation towards an integrated GARCH model. More recently, Mikosch and Stărică (2004) argued that the so-called 'integrated GARCH effect' is caused by the nonstationary behaviour of very long return series. They show how the long-range dependence in volatility and the IGARCH effect may be explained by neglected deterministic changes in the unconditional variance of the stochastic process. Moreover, Granger and Hyung (2004) claimed that occasional breaks in a long time series of absolute stock returns can also explain the observed slow decay of the autocorrelation functions of absolute returns in long return series.

It is well documented that shocks to the conditional variance of the standard GARCH model of Bollerslev (1986) decay at an exponential rate. This has motivated the development of more flexible models to describe the observed dependence structure in financial market volatility. One of these models is the Fractionally Integrated GARCH model of Baillie, Bollerslev, and Mikkelsen (1996) which belongs to the class of long-memory models. In these processes, shocks to the conditional variance decay at a slow hyperbolic rate which is more strongly supported by financial data than the GARCH model. A generalization of the FIGARCH model was recently proposed by Baillie and Morana (2007) in which they allow the intercept to change deterministically according to the flexible functional form of Gallant (1984).

The question of explicitly modelling nonstationarity in stock market volatility has, however, received somewhat less attention. There have been some attempts to incorporate nonstationarity directly into the model. Stărică and Granger (2005) introduced a nonstationary approach in which the returns are modelled as nonstationary sequence of independent random variables with time-varying unconditional variance but their model does not allow for volatility clustering. More recently, Engle and Gonzalo Rangel (2008) proposed modelling the volatility process by a multiplicative decomposition into a nonstationary and a stationary component. The nonstationary component (or the unconditional variance) is described by an exponential spline, and the stationary component (or the short-run dynamics of volatility) follows a first-order GARCH process.

This paper addresses the issue of modelling deterministic changes in the unconditional variance of long return series. It is assumed that volatility is modelled by decomposing the variance into a conditional and an unconditional component as in Amado and Teräsvirta (2008). The conditional variance follows a GARCH process, and describes the short-run dynamics of volatility. The nonstationary component of volatility describes the long-volatility dynamics, and it is represented by a linear combination of logistic transition functions. Statistical inference is used for specifying the parametric structure of the time-varying component by applying a sequence of

Lagrange multiplier tests. Our modelling strategy is applied to describe the long-run properties of the long daily Dow Jones Industrial Average (DJIA) return series from 1920 to 2003. One may expect that the longer the observation period, the more likely the occurrence of structural changes or shifts in the second unconditional moment of returns. The test results strongly support the time-variation of the unconditional variance in the period under study. The estimation results indicate that the strongest deterministic changes in the unconditional variance are associated with the largest economic recessions. This in turn suggests that the unconditional variance behaviour may be related to the evolution of the deterministic conditions in the economy. Our findings also suggest that the observed long-memory property in volatility may well be due to deterministic changes in the unconditional variance of the return series.

The paper is organized as follows. The TV-GARCH model and the modelling strategy are presented in Section 3.2. Details regarding the estimation of the model are discussed in Section 3.3. Section 3.4 contains the application. In Section 3.5 we show by a small Monte Carlo simulation how ignored deterministic changes in the unconditional variance affect the estimation of a misspecified model. Finally, Section 3.6 concludes.

3.2 A model for the long-term volatility component

3.2.1 The time-varying GARCH framework

In this paper the tool for modelling an asset return series over a long period is a GARCH-type model in which the unconditional variance is assumed to evolve smoothly over time. To motivate the introduction of our model we shall begin by focusing on the long-run properties of the GJR-GARCH(p, q) model of Glosten, Jagannathan, and Runkle (1993). Let $\mathfrak{F}_{t-1}$ be the information set containing the historical information of the series of interest available at time $t-1$ and write the asset returns $\{y_t\}$ as

$$y_t = \mathsf{E}(y_t|\mathfrak{F}_{t-1}) + \varepsilon_t \tag{3.1}$$

$$\varepsilon_t = \zeta_t h_t^{1/2} \tag{3.2}$$

where $\{\zeta_t\}$ is a sequence of independent standard normal variables. Under this assumption the conditional distribution of the innovation sequence $\{\varepsilon_t\}$ is $\varepsilon_t|\mathfrak{F}_{t-1} \sim N(0, h_t)$. For simplicity, the conditional mean of the asset returns is set equal to zero, i.e. $\mathsf{E}(\mathbf{y}_t|\mathfrak{F}_{t-1}) = 0$. The component h_t describes the dynamics of the conditional variance of the asset returns. To allow positive and negative shocks to have an asymmetric effect on the stock market volatility we choose the GJR-GARCH(p, q) model for h_t. It has the form

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \kappa_i \varepsilon_{t-i}^2 I_{t-i}(\varepsilon_{t-i} < 0) + \sum_{j=1}^{p} \beta_j h_{t-j}. \tag{3.3}$$

where the set of conditions for positivity and stationarity are imposed and $I_{t-i}(\varepsilon_{t-i} < 0)$ is an indicator function that equals 1 when $\varepsilon_{t-i} < 0, i = 1, \ldots, q$, and 0 otherwise.

Re-writing the dynamic structure of (3.3) in terms of the unconditional variance σ^2 one obtains

$$h_t = \sigma^2+\sum_{i=1}^{q}\alpha_i(\varepsilon_{t-i}^2-\sigma^2)+\sum_{i=1}^{q}\kappa_i\left(\varepsilon_{t-i}^2 I_{t-i}\left(\varepsilon_{t-i}<0\right)-\sigma^2\right)+\sum_{j=1}^{p}\beta_j(h_{t-j}-\sigma^2) \quad (3.4)$$

where $\sigma^2 \equiv \mathsf{E}(\varepsilon_t^2) = \omega/(1-\sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{q}\kappa_i/2 - \sum_{j=1}^{p}\beta_j)$. When the persistence rate $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{q}\kappa_i/2 + \sum_{j=1}^{p}\beta_j < 1$ then the conditional variance mean reverts to σ^2 at the geometric rate $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{q}\kappa_i/2 + \sum_{j=1}^{p}\beta_j$.

The assumption that the volatility process reverts to a constant level is very restrictive especially when modelling asset returns over long periods. In order to account for changes in the long-run volatility we shall consider a more flexible specification in which the unconditional variance σ^2 can be time-varying. We incorporate smooth changes in the unconditional variance of returns so that the variance evolves slowly over time. The variance is thus modelled using a multiplicative decomposition of the variance as follows:

$$\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}, \qquad \varepsilon_t|\mathfrak{F}_{t-1} \sim N(0, h_t g_t). \quad (3.5)$$

In equation (3.5) the short-run (or the stationary) component h_t is modelled as the GJR-GARCH process as in (3.3) with the exception that $\varepsilon_t^* = \varepsilon_t/g_t^{1/2}$:

$$h_t = \omega + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^{*2} + \sum_{i=1}^{q}\kappa_i\varepsilon_{t-i}^{*2}I_{t-i}(\varepsilon_{t-i}^* < 0) + \sum_{j=1}^{p}\beta_j h_{t-j}. \quad (3.6)$$

The long-run (or the nonstationary) component g_t is a slowly time-varying trend that functions as a proxy for all factors that affect the unconditional variance. More specifically, we follow Amado and Teräsvirta (2008) and let the time-varying unconditional variance component be a linear combination of logistic transition functions:

$$g_t = \delta_0 + \sum_{l=1}^{r}\delta_l G_l(t/T; \gamma_l, \mathbf{c}_l) \quad (3.7)$$

where $\delta_l, l = 0, \ldots, r$, are parameters. Furthermore, $G_l(t/T; \gamma_l, \mathbf{c}_l), l = 1, \ldots, r$, are generalized logistic transition functions:

$$G_l(t/T; \gamma_l, \mathbf{c}_l) = \left(1 + \exp\left\{-\gamma_l \prod_{j=1}^{k}(t/T - c_{lj})\right\}\right)^{-1} \quad (3.8)$$

satisfying the identification restrictions $\gamma_l > 0, l = 0, \ldots, r$, and $c_{l1} \leq c_{l2} \leq \ldots \leq c_{lk}$. The transition functions $G_l(t/T; \gamma_l, \mathbf{c}_l)$ allow the unconditional variance to change smoothly as a function of the calendar time t/T. The parameters, c_{lj} and γ_l, determine the location and the speed of the transition between different regimes. Equations (3.5)−(3.8) define the time-varying GARCH (TV-GARCH) model. The unconditional variance in this model is time-varying and equals $\mathsf{E}_t(\varepsilon_t^2) = \mathsf{E}(\zeta_t^2 h_t g_t) = g_t \mathsf{E} h_t$. This approach of introducing nonstationarity in the long run volatility component has been discussed in detail by Amado and Teräsvirta (2008).

Some special cases of the TV-GARCH model are of interest. Under $\delta_1 = \ldots = \delta_r = 0$, the unconditional variance $\mathsf{E}_t(\varepsilon_t^2)$ becomes constant. When $r = 1$ and $k = 1, g_t$ increases (decreases) monotonically over time from δ_0 to $\delta_0 + \delta_1$ when $\delta_1 > 0 (\delta_1 < 0)$, with the location centred at $t = c_1 T$. The slope parameter γ_1 in (3.8) controls the degree of smoothness of the transition: the larger γ_1, the faster the transition is between the extreme regimes. When $\gamma_1 \to \infty, g_t$ collapses into a step function. For small values of γ_1, the transition between regimes is approximately linear around c_1. When $\delta_l \neq 0$, for values $r > 1$ and $k > 1$, (3.7)$-$(3.8) form a very flexible parameterization capable of describing nonmonotonic deterministic changes in the unconditional variance.

3.2.2 Model specification

Since the nonlinear model in (3.5)$-$(3.8) is our most general parameterization, a systematic modelling strategy is required when a TV-GARCH model is fitted to the data. The strategy for building TV-GARCH models is based on statistical inference and it consists of the specification, estimation and evaluation of the model. At the specification stage, one first specifies the structure of g_t and, once that has been done, models the dynamics of the short-run component h_t. In practice, the parametric structure of the unconditional variance component has to be determined from the data, which involves two sets of decision problems. First, the number of transitions r in (3.7) has to be determined. Second, when $r \geq 1$, the integer k for each transition function has to be selected. This specification procedure is sequential and based on statistical inference. We shall apply the procedure of Amado and Teräsvirta (2008) for selecting r and k.

An important feature of the modelling strategy in this paper is that, since we are modelling very long return series, we shall divide the observation period into a number of subperiods. To introduce notation, let r be the total number of transitions in the whole period and $r_i, i = 1, \ldots, N$, be the number of transitions in the subperiod i, so $r = \sum_{i=1}^N r_i$. Define h_{it} as the conditional variance and $g_{it} = 1 + \sum_{l=1}^{r_i} \delta_{il} G_{il}(t/T; \gamma_l, \mathbf{c}_l), i = 1, \ldots, N$, for each subperiod.

The sequence of LM tests for specifying a TV-GARCH model is as follows:

1. Split the original time series into N non-overlapping subsamples. To facilitate specification the splits should preferably be located in tranquil periods.

2. For each $i = 1, \ldots, N$, specify g_{it} under the assumption that the conditional variance is constant, i.e. $h_{it} \equiv \omega_i > 0$. This is done as follows. First, test the hypothesis of constant unconditional variance $\mathrm{H}_{01} : \gamma_{i1} = 0$ against $\mathrm{H}_{11} : \gamma_{i1} > 0$ in

$$g_{it} = \omega_i^{-1} \left\{1 + \delta_{i1} G_{i1}\left(t/T; \gamma_{i1}, \mathbf{c}_{i1}\right)\right\} = \omega_i^{-1} + \delta_{i1}^* G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1}) \qquad (3.9)$$

where $\delta_{i1}^* = \omega_i^{-1}\delta_{i1}$, at the significance level $\alpha^{(1)}$. Under the null hypothesis, the function contains unidentified nuisance parameters δ_{i1} and $\mathbf{c}_{i1}$. To circumvent this identification problem we follow Luukkonen, Saikkonen, and Teräsvirta (1988) and approximate $G_{i1}(t/T;\gamma_{i1},\mathbf{c}_{i1})$ by its third-order Taylor expansion around $\gamma_{i1} = 0$. After reparameterizing, we obtain

$$g_{it} = \omega_i^* + \sum_{j=1}^{3} \phi_{ij}(t/T)^j + R_3(t/T;\gamma_{i1},\mathbf{c}_{i1}) \tag{3.10}$$

where $\phi_{ij} = \gamma_{i1}\widetilde{\delta}_{ij}^*, i = 1,\ldots,N$, and $R_3(t/T;\gamma_{i1},\mathbf{c}_{i1})$ is the remainder. Furthermore, $R_3(t/T;\gamma_{i1},\mathbf{c}_{i1}) \equiv 0$ under H_{01}, so the remainder of the Taylor expansion does not affect the asymptotic distribution theory. The null hypothesis of constant unconditional variance becomes $\mathrm{H}_{01}^{'} : \phi_{i1} = \phi_{i2} = \phi_{i3} = 0$. Under $\mathrm{H}_{01}^{'}$, the standard LM statistic has an asymptotic χ^2-distribution with three degrees of freedom. See Amado and Teräsvirta (2008) for details on how to compute the test statistic.

3. If $\mathrm{H}_{01}^{'}$ is rejected, for each subperiod select the order $k \leq 3$ in the exponent of $G_{i1}(t/T;\gamma_{i1},\mathbf{c}_{i1})$ using a short sequence of tests within (3.10); for details see Amado and Teräsvirta (2008). Next, estimate g_{it} with a single transition function and test $\mathrm{H}_{02} : g_{it} = \omega_i^{-1} + \delta_{i1}^* G_{i1}(t/T;\gamma_{i1},\mathbf{c}_{i1})$ against $\mathrm{H}_{12} : g_{it} = \omega_i^{-1} + \sum_{l=1}^{2} \delta_{il}^* G_{il}(t/T;\gamma_{il},\mathbf{c}_{il})$ at the significance level $\alpha^{(2)} = \tau\alpha^{(1)}$, where $\tau \in (0,1)$. The significance level is reduced at each stage by a factor τ in order to favour parsimony. In our application we set $\tau = 0.5$. Test the hypothesis of no second transition $\mathrm{H}_{02} : \gamma_{i2} = 0$ in

$$g_{it} = \omega_i^{-1} + \delta_{i1}^* G_{i1}(t/T;\gamma_{i1},\mathbf{c}_{i1}) + \delta_{i2}^* G_{i2}(t/T;\gamma_{i2},\mathbf{c}_{i2}) \tag{3.11}$$

Again, model (3.11) is not identified under the null hypothesis. To circumvent the problem we proceed as before and express the logistic function $G_{i2}(t/T;\gamma_{i2},\mathbf{c}_{i2})$ by a third-order Taylor approximation around $\gamma_{i2} = 0$. After rearranging terms we have

$$g_{it} = \omega_i^* + \delta_{i1}^* G_{i1}(t/T;\gamma_{i1},\mathbf{c}_{i1}) + \sum_{j=1}^{3} \varphi_{ij}(t/T)^j + R_3^*(t/T;\gamma_{i2},\mathbf{c}_{i2}) \tag{3.12}$$

where $\varphi_{ij} = \gamma_{i2}\widetilde{\delta}_{ij}^*, i = 1,\ldots,N$, and $R_3^*(t/T;\gamma_{i2},\mathbf{c}_{i2})$ is the remainder. The new null hypothesis based on this approximation is $\mathrm{H}_{02}^{'} : \varphi_{i1} = \varphi_{i2} = \varphi_{i3} = 0$. Again, this hypothesis can be tested using a LM test. If the null hypothesis is rejected, specify k for the second transition and estimate g_{it} with two transition functions.

4. More generally, when g_{it} has been estimated with $r_i - 1$ transition functions one tests for another transition in g_{it} using the significance level $\alpha^{(r_i)} = \tau\alpha^{(r_i-1)}$. Testing continues until the first non-rejection of the null hypothesis.

In summary, we begin the model specification problem by first modelling the unconditional variance assuming that the conditional variances remain constant. After specifying and estimating g_t, the hypothesis of no conditional heteroskedasticity is tested in $\{\varepsilon_t^*\}$. If the null hypothesis of no ARCH is rejected, the conditional variance component h_t is modelled as in (3.6) with $p = q = 1$. At the evaluation stage the adequacy of the estimated model is tested by means of LM-type misspecification tests (see Amado and Teräsvirta (2008) for further details).

3.3 Estimation of parameters

After the number of transitions and their type in (3.7) have been determined, the parameters of the TV-GARCH model are estimated by quasi-maximum likelihood (QML). For this purpose, let $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'$ be the parameter vector of the model. Let $h_t \equiv h_t(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ and $g_t \equiv g_t(\boldsymbol{\theta}_2)$ where $\boldsymbol{\theta}_1 = (\omega, \alpha_1, \ldots, \alpha_q, \kappa_1, \ldots, \kappa_q, \beta_1, \ldots, \beta_p)'$ and $\boldsymbol{\theta}_2 = (\boldsymbol{\delta}', \gamma_1, \ldots, \gamma_r, \mathbf{c}_1', \ldots, \mathbf{c}_r')'$ with $\boldsymbol{\delta} = (\delta_0, \delta_1, \ldots, \delta_r)'$. The model defined in (3.5)−(3.8) can be now rewritten as follows:

$$\varepsilon_t = \zeta_t \left\{ h_t\left(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\right) g_t\left(\boldsymbol{\theta}_2\right)\right\}^{1/2}. \tag{3.13}$$

Assuming that $\{\zeta_t\}$ is a sequence of independent standard normal variables, the log-likelihood function for observation t equals

$$\ell_t(\boldsymbol{\theta}) = -(1/2)\ln 2\pi - (1/2)\{\ln h_t(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) + \ln g_t(\boldsymbol{\theta}_2)\} - (1/2)\frac{\varepsilon_t^2}{h_t(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) g_t(\boldsymbol{\theta}_2)} \tag{3.14}$$

The unconditional and the conditional variance components are estimated separately using maximization by parts. The iterative algorithm proceeds as follows:

Step 1: Maximize

$$L_T^U(\boldsymbol{\theta}_2) = \sum_{t=1}^{T} \ell_t^U(\boldsymbol{\theta}_2) = -(1/2)\sum_{t=1}^{T}\{\ln g_t(\boldsymbol{\theta}_2) + \tilde{\varepsilon}_t^2/g_t(\boldsymbol{\theta}_2)\}$$

with respect to $\boldsymbol{\theta}_2$, assuming $\tilde{\varepsilon}_t = \varepsilon_t$, that is, setting $h_t(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \equiv 1$. Let the estimator of $\boldsymbol{\theta}_2$ be $\widehat{\boldsymbol{\theta}}_2^{(1)}$. Making use of $\widehat{\boldsymbol{\theta}}_2^{(1)}$, maximize

$$L_T^V(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(1)}) = \sum_{t=1}^{T} \ell_t^V(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(1)}) = -(1/2)\sum_{t=1}^{T}\left\{\ln h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(1)}) + \varepsilon_t^{*2}/h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(1)})\right\}$$

with respect to $\boldsymbol{\theta}_1$, where $\varepsilon_t^* = \varepsilon_t/\{g_t(\widehat{\boldsymbol{\theta}}_2^{(1)})\}^{1/2}$. Denote the estimator as $\widehat{\boldsymbol{\theta}}_2^{(1)}$.

Step 2: Maximize

$$L_T^U(\boldsymbol{\theta}_2) = \sum_{t=1}^{T} \ell_t^U(\boldsymbol{\theta}_2) = -(1/2)\sum_{t=1}^{T}\{\ln g_t(\boldsymbol{\theta}_2) + \tilde{\varepsilon}_t^2/g_t(\boldsymbol{\theta}_2)\}$$

with respect to $\boldsymbol{\theta}_2$, where $\tilde{\varepsilon}_t = \varepsilon_t/\{h_t(\widehat{\boldsymbol{\theta}}_1^{(1)}, \widehat{\boldsymbol{\theta}}_2^{(1)})\}^{1/2}$. Call this estimator $\widehat{\boldsymbol{\theta}}_2^{(2)}$ and maximize

$$L_T^V(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(2)}) = \sum_{t=1}^{T} \ell_t^V(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(2)}) = -(1/2)\sum_{t=1}^{T}\left\{\ln h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(2)}) + \varepsilon_t^{*2}/h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(2)})\right\}$$

with respect to $\boldsymbol{\theta}_1$, where $\varepsilon_t^* = \varepsilon_t/\{g_t(\widehat{\boldsymbol{\theta}}_2^{(2)})\}^{1/2}$. This yields $\widehat{\boldsymbol{\theta}}_1^{(2)}$.

Iterate until convergence.

In the nth iteration, maximization is carried out in the usual way by solving the score equations:

$$\frac{\partial}{\partial \boldsymbol{\theta}_2} L_T^U(\boldsymbol{\theta}_2) = (1/2)\sum_{t=1}^{T}\left(\frac{\tilde{\varepsilon}_t^2}{g_t(\boldsymbol{\theta}_2)} - 1\right)\frac{1}{g_t(\boldsymbol{\theta}_2)}\frac{\partial g_t(\boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2} = 0$$

for $\boldsymbol{\theta}_2$ with $\tilde{\varepsilon}_t = \varepsilon_t/\{h_t(\widehat{\boldsymbol{\theta}}_1^{(n-1)}, \widehat{\boldsymbol{\theta}}_2^{(n-1)})\}^{1/2}$, and

$$\frac{\partial}{\partial \boldsymbol{\theta}_1} L_T^V(\boldsymbol{\theta}_1) = (1/2)\sum_{t=1}^{T}\left(\frac{\varepsilon_t^{*2}}{h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)})} - 1\right)\frac{1}{h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)})}\frac{\partial h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)})}{\partial \boldsymbol{\theta}_1} = 0$$

for $\boldsymbol{\theta}_1$, where $\varepsilon_t^* = \varepsilon_t/\{g_t(\widehat{\boldsymbol{\theta}}_2^{(n)})\}^{1/2}$. Letting $G_{lt} \equiv G_l(t/T; \gamma_l, \mathbf{c}_l)$, $l = 1, \ldots, r$, we have

$$\frac{\partial g_t(\boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2} = (1, G_{1t}, G_{1t}^{(\gamma)}, G_{1t}^{(c)}, \ldots, G_{rt}, G_{rt}^{(\gamma)}, G_{rt}^{(c)})'$$

where, for $k = 1$ in (3.8),

$$\begin{aligned} G_{lt}^{(\gamma)} &= \frac{\partial G_{lt}}{\partial \gamma_l} = \delta_l G_{lt}(1 - G_{lt})(t/T - c_l) \\ G_{lt}^{(c)} &= \frac{\partial G_{lt}}{\partial c_l} = -\gamma_l \delta_l G_{lt}(1 - G_{lt}) \end{aligned}$$

and for $k > 1$

$$\begin{aligned} G_{lt}^{(\gamma)} &= \frac{\partial G_{lt}}{\partial \gamma_l} = \delta_l G_{lt}(1 - G_{lt})\prod\nolimits_{j=1}^{k}(t/T - c_{lj}) \\ G_{ilt}^{(c)} &= \frac{\partial G_{ilt}}{\partial c_l} = -\gamma_l \delta_l G_{lt}(1 - G_{lt})\prod\nolimits_{j=1, j\neq l}^{k}(t/T - c_{lj}) \end{aligned}$$

where c_{lj} denotes the jth element in the parameter vector $\mathbf{c}_l, l = 1, \ldots, r$, and

$$\begin{aligned} \frac{\partial h_t(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)})}{\partial \boldsymbol{\theta}_1} &= (1, \varepsilon_{t-1}^{*2}, \ldots, \varepsilon_{t-q}^{*2}, \varepsilon_{t-1}^{*2} I_{t-1}(\varepsilon_{t-1}^* < 0), \ldots, \varepsilon_{t-q}^{*2} I_{t-q}(\varepsilon_{t-q}^* < 0), \\ &\quad h_{t-1}(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)}), \ldots, h_{t-p}(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)}))' + \sum_{j=1}^{p} \beta_j \frac{\partial h_{t-j}(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(n)}))}{\partial \boldsymbol{\theta}_1} \end{aligned}$$

This algorithm is computationallly attractive for situations in which direct maximization of the log-likelihood function is difficult. Under certain regularity conditions, the resulting estimator coincides with the ML estimator and becomes fully efficient upon convergence; see Song, Fan, and Kalbfleisch (2005) for details. Throughout this paper, we assume that certain regularity conditions are satisfied to ensure consistency and asymptotic normality of the QML estimator. The asymptotic properties of the estimators of the TV-GARCH model are not yet known. Extending the results to the nonstationary TV-GARCH model is not straightforward and is beyond the scope of this paper.

In this work, the long time series requires some modifications to the estimation algorithm. Because the whole series is divided into non-overlapping subperiods, the different data segments can have different "baseline" volatility levels. For this reason, the algorithm iterates from an initial value which is estimated by "chain rule" to accommodate differences in the volatility levels. This proceeds as follows. First, for the first subperiod, estimate the parameters of $g_{1t} = \delta_0 + \sum_{l=1}^{r_1} \delta_{1l} G_{1l}(t/T; \gamma_{1l}, \mathbf{c}_{1l})$ where r_1 is the number of transitions for this period. The estimate $\widehat{g}_{1t}$ serves as the "intercept" in the nonstationary component of the next subperiod. Conditioning on this value, carry out the estimation of the parameters for the next subperiod. More generally, for the ith subperiod, estimate $g_{it} = \widehat{\delta}_0^{(i-1)} + \sum_{l=1}^{r_i} \delta_{il} G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$ by conditioning on $\widehat{\delta}_0^{(i-1)}$, where $\widehat{\delta}_0^{(i-1)} = \widehat{\delta}_0 + \sum_{l=1}^{r_{i-1}} \widehat{\delta}_l G_l(t/T; \widehat{\gamma}_l, \widehat{\mathbf{c}}_l)$ and r_{i-1} is the number of transitions in the $(i-1)$th subperiod. The estimates $\widehat{\gamma}_l$ and $\widehat{\mathbf{c}}_l$ are then used as fixed values in the next iterations. This means that the estimation algorithm is carried out without iterating $\widehat{\gamma}_l$ and $\widehat{\mathbf{c}}_l$, and therefore the parameters $\delta_l, l = 0, \ldots, r$, are estimated conditionally on those estimates.

Another aspect that deserves attention in the estimation of the model is the selection of starting-values of the time-varying parameters. Since the log-likelihood may contain several local maxima, it is advisable to initiate the estimation from different sets of starting-values before settling for the final parameter estimates. In addition, to improve the accuracy of the estimates of the standard errors, we follow Fiorentini, Calzolari, and Panattoni (1996) and use analytic first derivatives both in the estimation of the TV-GARCH models and in the computation of the test statistics. All computations in this paper have been carried out using Ox programming language, version 3.40 (see Doornik (2002)).

3.4 Application to the Dow Jones Industrial Average index

3.4.1 Data description

In this section we illustrate the use of the modelling building procedure of the TV-GARCH model to the daily returns of the Dow Jones Industrial Average (DJIA) index. The entire sample covers the period between January 2, 1920 and December 31, 2003, yielding 21121 observations. The daily returns are defined as the log differences of the closing prices of the index between two consecutive days. The closing

Table 3.1 Summary statistics of the daily DJIA return series: full sample

Series	Min	Max	Mean	S.D.	Skew	Ex.Kr.	Rob.Sk.	Rob.Kr.
ε_t	-25.63	14.27	0.022	1.136	-0.659	25.28	-0.006	0.227
$\varepsilon_t/\widehat{g}_t^{1/2}$	-17.71	6.680	0.018	0.660	-1.185	29.73	-0.006	0.145

Notes: The table contains summary statistics for the DJIA return series. The sample period starts in January 2, 1920 and ends in December 31, 2003 (21121 observations).

prices of the DJIA index have been obtained from the Wharton Research Data Services (WRDS) provided by the Wharton School of the University of Pennsylvania. Descriptive statistics of the return series can be found in Table 3.1. The coefficients of skewness and kurtosis seem to indicate that the stock returns ε_t have a left skewed and a significantly fat-tailed distribution. To check this conclusion, we also provide the robust measures of skewness and kurtosis as recommended by Kim and White (2004) in order to account for outliers. The robust measure for skewness is practically zero whereas the robust kurtosis measure suggests that there is indeed some excess kurtosis in the series. Figure 3.1 graphs the daily returns for the DJIA index for the observation period. The period covers the Great Depression of 1929 and the early 1930's, the Second World War, the 1973 oil crisis, the stock market crash of October 1987 and the recent dot-com bubble. Because of the long observation period it is unlikely that the series is stationary.

We divide the 83 years long series into six non-overlapping subperiods each comprising at least of 2500 observations. In most cases we report the findings for each of the six periods and the full sample. Summary statistics of the subperiods can be found in Table 3.9.

3.4.2 Estimation results

The focus of the empirical analysis lies in the specification of the unconditional variance using the modelling strategy described in Section 3.2.2. We begin by determining the number of transitions for each subperiod separately. This is done using the sequence of specification tests. The initial significance level of the sequence of tests is $\alpha^{(1)} = 0.01$. At each stage of the sequence we halve the significance level of the test, i.e. $\tau = 0.5$. The tests results are presented in the second column of Table 3.2.

We first test the hypothesis of constant unconditional variance against a smoothly time-varying unconditional variance with one transition function. The null hypothesis is rejected for all subperiods with the exception of the subperiod 5 covering the October 1987 crash. The stock market volatility returned to normal levels very quickly after the crash, which suggests that the unconditional variance remained stable during that period. These findings are consistent with the hypothesis of Engle and Lee (1999) that the 1987 crash is more transient than other big shocks. The null hypothesis of constant unconditional variance is, however, rejected very strongly for the subperiods 1, 2, 4 and 6. The first period contains the Great Depression, the second includes the

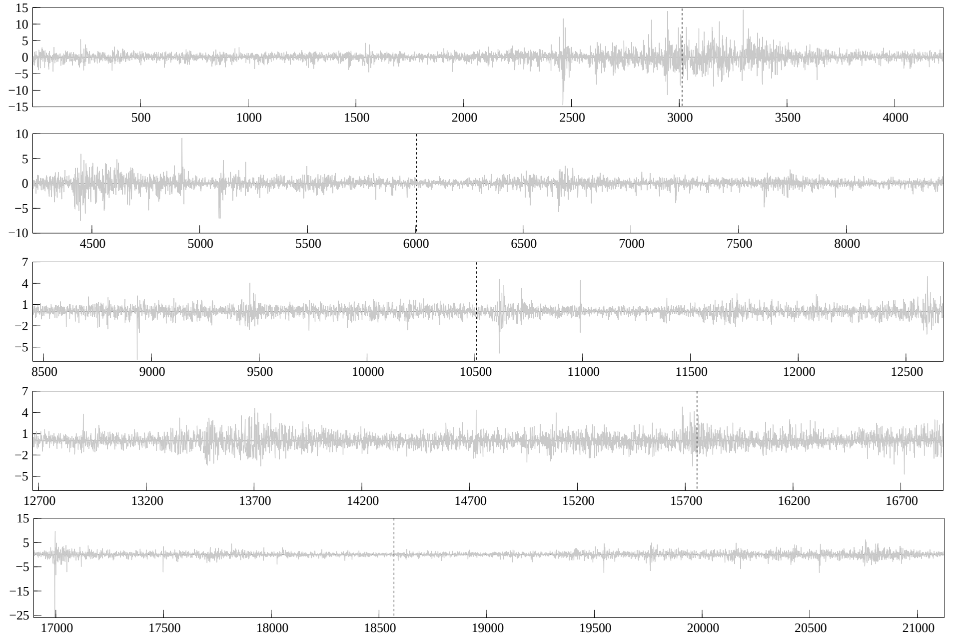


Figure 3.1 The DJIA daily returns from January 2, 1920 until December 31, 2003. The vertical lines represent the split dates.

Second World War, the fourth one the OPEC oil crisis and the most recent one the IT bubble. The results indicate that the strongest deterministic changes in the unconditional variance are associated with the largest economic recessions in the period under study.

The sequence of nested tests based on (3.10) to select k in (3.8) are given in the last three columns of Table 3.2 (see Amado and Teräsvirta (2008) for details). The strongest rejection is when $k = 1$ for all four periods. Note that, for the subperiod 1, the tests H_{01} and H_{02} cannot discriminate between $k = 1$ and $k = 2$ as the p-values are very close to each other. We choose $k = 2$ to minimize the number of transitions to be specified. Misspecification tests to check the validity of this choice can be carried out at the model evaluation stage.

We proceed to first estimate a TV-GARCH model with a single transition and test against a double transition model at $\alpha^{(2)} = 0.005$. We reject the hypothesis in three out of the five cases and select $k = 1$. Fitting the model with two transition functions for the three subsamples and testing against another transition leads to a rejection only for the fourth period. The p-value, however, is 0.0021 which is very close to $\alpha^{(3)} = 0.0025$. Thus, we tentatively accept the model with two transitions as the final parameterization for the first, second and fourth subperiods.

The above results imply that eight transition functions in total are needed to describe the unconditional variance for the whole series. Estimation results for the

Table 3.2 p-values of sequences of Lagrange multiplier tests for the six subperiods

Subperiods	H_0	H_{03}	H_{02}	H_{01}
Subsample 1 (02/01/1920 – 31/12/1931)				
Single transition	5×10^{-44}	7×10^{-4}	3×10^{-22}	3×10^{-24}
Double transition	2×10^{-8}	0.4773	0.0111	2×10^{-8}
Triple transition	0.2632	0.0574	0.7178	0.6213
Subsample 2 (04/01/1932 – 31/12/1943)				
Single transition	5×10^{-77}	3×10^{-19}	3×10^{-20}	2×10^{-46}
Double transition	2×10^{-6}	0.0017	0.0242	1×10^{-4}
Triple transition	0.2818	0.1547	0.3484	0.3388
Subsample 3 (04/01/1944 – 29/12/1961)				
Single transition	0.0079	0.0712	0.9654	0.0034
Double transition	0.0792	0.2789	0.2364	0.0402
Triple transition	–	–	–	–
Subsample 4 (01/01/1962 – 16/11/1982)				
Single transition	1×10^{-17}	2×10^{-4}	0.3328	3×10^{-16}
Double transition	4×10^{-5}	9×10^{-4}	0.8199	6×10^{-4}
Triple transition	0.0021	0.1889	0.0006	0.2988
Subsample 5 (17/11/1982 – 31/12/1993)				
Single transition	0.1018	0.8983	0.0173	0.4688
Double transition	–	–	–	–
Subsample 6 (03/01/1994 – 31/12/2003)				
Single transition	3×10^{-18}	0.0012	0.0017	8×10^{-16}
Double transition	0.0315	0.0361	0.0694	0.2817
Triple transition	–	–	–	–

Notes: The entries are the p-values of the LM-type tests of constant unconditional variance against a time-varying GARCH model for each subperiod of the DJIA stock index returns. The test sequence starts at the significance level $\alpha = 0.01$ and setting $\tau = 0.5$. The order k in (3.8) is chosen from the sequence of nested tests based on (3.10). If H_{0i} is rejected most strongly, measured by the p-value, of the three hypotheses, one selects $k = i$. See Amado and Teräsvirta (2008) for further details.

Table 3.3 Estimation results for the DJIA returns: full sample

Panel (a): parameter estimates of the TV-GJR-GARCH(1,1) model

$$h_t = \underset{(0.0052)}{0.0285} + \underset{(0.00386)}{0.0236}\,\varepsilon^{*2}_{t-1} + \underset{(0.0136)}{0.9011}h_{t-1} + \underset{(0.0166)}{0.0913}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -19112.1 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9704$$

$$\begin{aligned} g_t = \underset{(0.0276)}{0.82506} + \underset{(0.1382)}{1.1680}&(1+\exp\{-\underset{(-)}{52.81}(t/T - \underset{(-)}{0.0119})(t/T - \underset{(-)}{0.1047})\})^{-1} \\ &+\underset{(0.3886)}{6.3472}(1+\exp\{-\underset{(-)}{100}(t/T - \underset{(-)}{0.1154})\})^{-1} \\ &-\underset{(0.37118)}{6.4953}\,(1+\exp\{-\underset{(-)}{96.65}(t/T - \underset{(-)}{0.1643})\})^{-1} \\ &-\underset{(0.0657)}{1.2475}(1+\exp\{-\underset{(-)}{70.03}(t/T - \underset{(-)}{0.2454})\})^{-1} \\ &-\underset{(0.0188)}{0.1522}(1+\exp\{-\underset{(-)}{100}(t/T - \underset{(-)}{0.3697})\})^{-1} \\ &+\underset{(0.1486)}{1.7541}(1+\exp\{-\underset{(-)}{67.17}(t/T - \underset{(-)}{0.6355})\})^{-1} \\ &-\underset{(0.1495)}{1.2656}(1+\exp\{-\underset{(-)}{100}(t/T - \underset{(-)}{0.6544})\})^{-1} \\ &+\underset{(0.0599)}{0.6952}(1+\exp\{-\underset{(-)}{100}(t/T - \underset{(-)}{0.9216})\})^{-1} \end{aligned}$$

Panel (b): parameter estimates of the GJR-GARCH(1,1) model

$$h_t = \underset{(0.0021)}{0.0116} + \underset{(0.0039)}{0.0304}\varepsilon^{*2}_{t-1} + \underset{(0.0085)}{0.9208}h_{t-1} + \underset{(0.0123)}{0.0784}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -27468.5 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9903$$

TV-GARCH model are reported in Table 3.3, Panel (a). The estimation results for each of the subperiods can be found in Table 3.10 in Appendix B.

The estimation is carried out with the sequential quadratic programming optimisation algorithm using analytical derivatives. The numbers in parenthesis below the parameter estimates are the asymptotic standard error estimates and calculated using numerical second derivatives. The standard errors of γ_i and $\mathbf{c}_i, i = 1, \ldots, 8$, are not available because the parameters $\delta_j, j = 0, \ldots, 8$, are estimated conditionally on those parameters. In some subperiods we observe that the transition between the extreme regimes of volatility is quite rapid. For these cases, the maximum value of γ_j is constrained to 100 to avoid convergence problems. This approximation is adequate because the shape of the transition function does not change much beyond values of γ_j larger or equal than 100.

For an idea how the unconditional variance changes over time, the estimated component $g_t^{1/2}$ is plotted in Figure 3.2 (upper panel). The estimated g_t functions for each subperiod are also shown in Figure 3.8 (Appendix A). It is seen that the

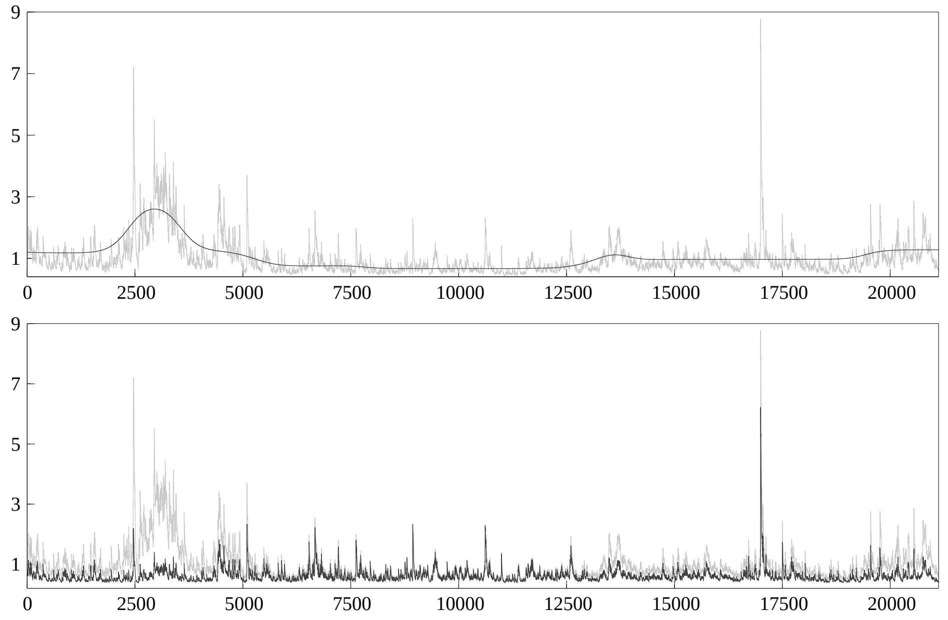


Figure 3.2 First panel, shows the estimated function $g_t^{1/2}$ (black curve) and the conditional standard deviation from the GJR-GARCH(1,1) model (grey curve). Second panel, shows the estimated conditional standard deviations from the GJR-GARCH(1,1) model (grey curve) and from the TV-GJR-GARCH(1,1) (black curve) models.

largest deterministic changes in the unconditional variance occur during the periods of recession in the economy. In particular, the strongest movement in the long-run volatility is observed during the Great Depression. This is in agreement with Mikosch and Stărică (2004) who find that most of the recessions coincide with an increase in the unconditional variance of the series. In their analysis of the S&P 500 returns, they identify the 1973 oil crisis as the major change detected in the unconditional variance, but then they study a time series only covering the period from January 2, 1953, until December 31, 1990.

For comparison, we also report the results of fitting the GJR-GARCH(1,1) model into the complete series. They can be found in Panel (*b*) of Table 3.3. The results for each subperiod appear in Table 3.11 (Appendix B). We find that the subperiods characterized by the largest changes in the unconditional variance have a stronger integrated GJR-GARCH effect. The stationary condition for the full sample model is $\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 < 1$. This model is practically an integrated GJR-GARCH model as the persistence indicator $\widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9903$. The autocorrelation functions of $|\varepsilon_t|$ plotted in Figure 3.3 (upper panel) lead to the same conclusion. The graph clearly displays the long-memory property: relatively rapid decay at short lags fol-

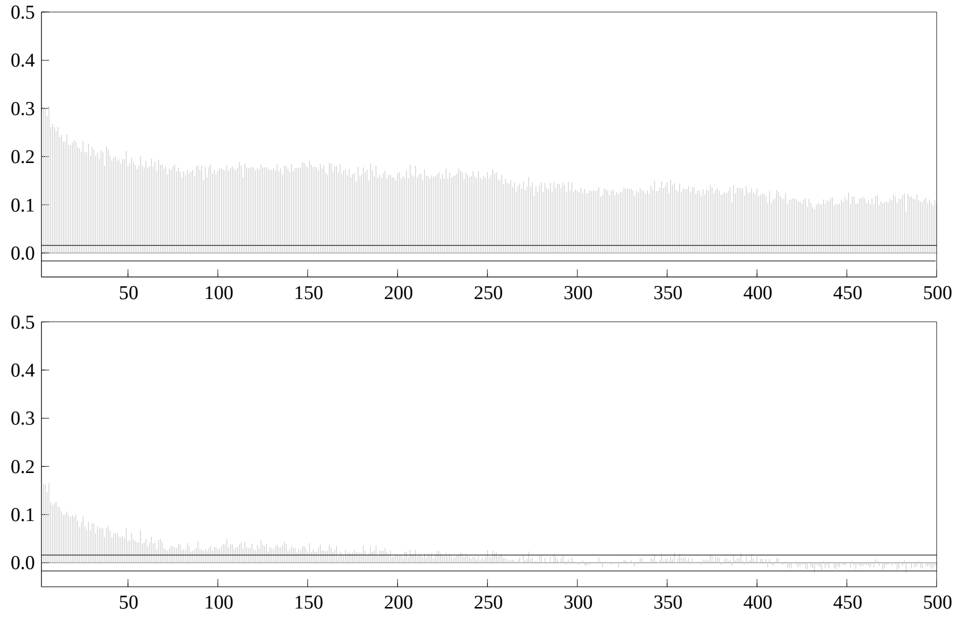


Figure 3.3 First panel, shows the sample autocorrelation functions of the absolute values of the DJIA daily returns. Second panel, shows the sample autocorrelation functions of the standardized variable $|\varepsilon_t| / \widehat{g}_t^{1/2}$. The horizontal lines are the corresponding 95% confidence interval under the iid normality assumption.

lowed by positive autocorrelations around a stable level at long lags. On the contrary, the autocorrelations of $|\varepsilon_t| / \widehat{g}_t^{1/2}$, plotted in the lower panel of Figure 3.3, decay very quickly with the lag length and only the first 70 autocorrelation estimates or so are significantly different from zero judging from the 95% confidence bounds drawn under the assumption that the errors are normal and independent. The decay rate looks more or less exponential, and the persistence indicator now equals 0.97. The results show that modelling the changes in the unconditional variance strongly reduces the amount of evidence for long-memory. This can also be seen from the Geweke and Porter-Hudak (1983) (GPH) estimates of the long-memory parameter in Table 3.4. Of course, the GPH parameter estimates are different for different bandwidths but, overall, the table indicates that the daily DJIA return series is either nonstationary (for the bandwidth choices $m = T^{0.4}$ and $m = T^{0.5}$) or is very close to the nonstationary region (for $m = T^{0.6}$). However, when the movements in the unconditional variance component are taken into account the GPH estimates have the remarkable low values of $0.1198, 0.2340$ and 0.3050 for these three bandwidths.

A similar conclusion can be reached by looking at the estimated conditional at the estimated conditional standard deviations from the GJR-GARCH(1,1) model of

Table 3.4 GPH estimates of the long-memory parameter: full sample

	$d_{GPH}(m = T^{0.4})$	$d_{GPH}(m = T^{0.5})$	$d_{GPH}(m = T^{0.6})$
ε_t	0.7364 (0.0614)	0.5470 (0.0511)	0.4588 (0.0322)
$\varepsilon_t / \widehat{g}_t^{1/2}$	0.1198 (0.1035)	0.2340 (0.0576)	0.3050 (0.0333)

Notes: The numbers in parentheses are the standard errors. The bandwidth m equals $T^{\alpha}, \alpha \in \{0.4, 0.5, 0.6\}$ where T is the number of observations.

Table 3.5 Diagnostic tests: p-values of the test of no ARCH in GARCH

	Lag order				
Model	1	2	3	4	5
GJR-GARCH(1,1)	0.197	0.266	0.448	0.612	0.718
TV-GJR-GARCH(1,1)	0.366	0.360	0.561	0.717	0.835

ε_t and $\varepsilon_t / \widehat{g}_t^{1/2}$. The lower panel of Figure 3.2 displays both series. The (almost) stationary behaviour of the conditional standard deviation of $\varepsilon_t / \widehat{g}_t^{1/2}$ (black curve) contrasts with the nonstationary behaviour of the conditional standard deviation of ε_t (grey line). It shows that the conditional variance of $\varepsilon_t / \widehat{g}_t^{1/2}$ is considerably smaller than that of ε_t from the GJR-GARCH(1,1) model. For illustration, we also show in Figure 3.9 (Appendix A) the estimated conditional standard deviations generated from both models separately for each subperiod.

The adequacy of the estimated TV-GJR-GARCH(1,1) model is checked using the diagnostic tests proposed by Amado and Teräsvirta (2008). We perform tests against remaining ARCH in the standardized residuals, additional transitions in g_t, TV-GJR-GARCH(1,2) and TV-GJR-GARCH(2,1) models, and ST-GJR-GARCH(1,1) model of order 1. The p-values of the tests are given in Tables 3.5-3.7. For comparison we also show the test results for the estimated GJR-GARCH(1,1) model. The results indicate no evidence of remaining ARCH in the standardized residuals, nor can argue in favour of additional transitions in g_t; see Tables 3.5 and 3.6. However, the tests against TV-GJR-GARCH(1,2) and TV-GJR-GARCH(2,1) reject the null hypothesis at the 5% significance level; see Table 3.7. Moreover, the TV-GJR-GARCH(1,1) model is strongly rejected against ST-GJR-GARCH(1,1) model. The results suggests that the TV-GJR-GARCH(1,1) model is an inadequate parameterization, and a higher lag in the GJR-GARCH component or a nonlinear GARCH model should be employed. Modelling the short-run dynamics of volatility over a long time series does need more work. But then, the focus of this paper is on the modelling of changes in the long-run volatility component and refinements in the modelling of h_t are left for further work.

Table 3.6 Diagnostic tests: p-values of tests of no additional transition in g_t

Model	LM_1	LM_2	LM_3
GJR-GARCH(1,1)	0.905	4×10^{-5}	1×10^{-4}
TV-GJR-GARCH(1,1)	0.092	0.186	0.321

Table 3.7 Diagnostic tests: p-values of tests against models of higher orders and against a nonlinear structure

	Alternative model		
Model	GJR(1,2)	GJR(2,1)	ST-GJR ($K = 1$)
GJR-GARCH(1,1)	0.0094	0.006	1×10^{-6}
TV-GJR-GARCH(1,1)	0.0086	0.036	3×10^{-4}

3.5 Monte Carlo experiment

In this section, we further investigate the effects of ignoring deterministic changes in the unconditional variance on the estimation of two GARCH-type models. This is done by conducting a small Monte-Carlo experiment. The purpose of the experiment is to illustrate how such shifts may bias the GARCH parameters and the persistence indicator. We generate data from two models. The first model is a TV-GARCH(1,1) model, whereas the second one is a TV-GJR-GARCH(1,1) model. The data-generating process is defined as follows:

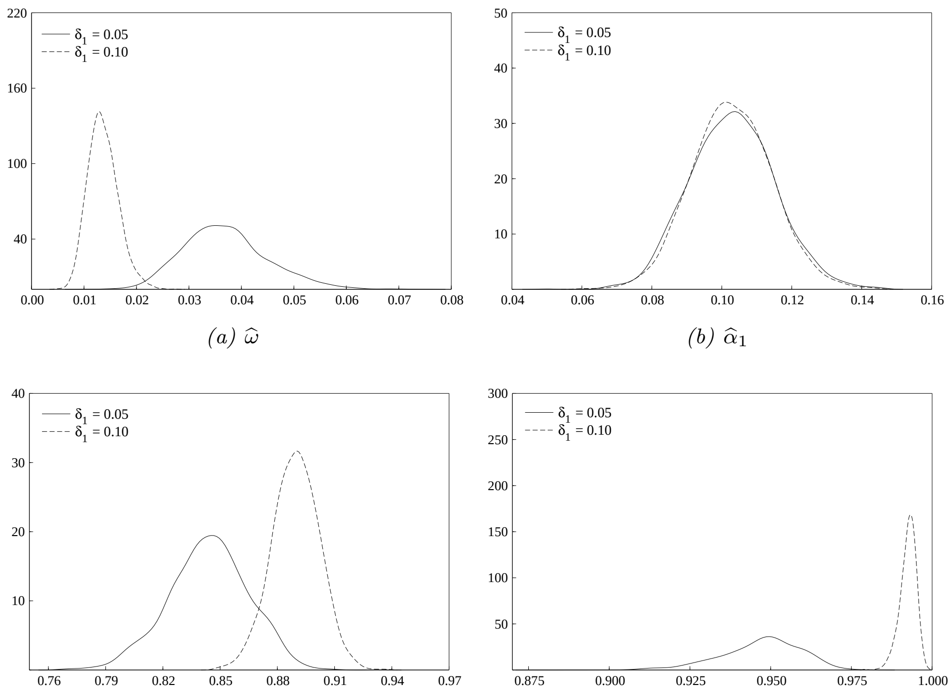
$$\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}, \qquad \varepsilon_t | \mathfrak{F}_{t-1} \sim N(0, h_t g_t) \tag{3.15}$$

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \kappa_1 \varepsilon_{t-1}^2 I_{t-1}(\varepsilon_{t-1} < 0) + \beta_1 h_{t-1} \tag{3.16}$$

$$g_t = 1 + \delta_1 (1 + \exp\{-\gamma_1 (t/T - c_1)\})^{-1} \tag{3.17}$$

with $\omega = 0.05, \alpha_1 = \{0.1, 0.05\}, \kappa_1 = \{0, 0.05\}, \beta_1 = 0.8, \delta_1 = \{0.05, 0.10\}$, and $c_1 = 0.5$ in each experiment. The simulations differ according to the pair $\{\alpha_1, \kappa_1\}$ and to the values of the slope parameter γ_1 which varies in the interval $\gamma_1 = \{10, 50\}$. The first 1000 observations of each generated series have been discarded to avoid initialization effects. For each experiment, the number of replications equals 2000 with a sample size of 5000 observations.

Figure 3.4 contains the estimated density of the estimated GARCH parameters and the persistence measured by the sum $\widehat{\alpha}_1 + \widehat{\beta}_1$ when the true model is (3.15)-(3.17) with $\{\alpha_1, \kappa_1\} = \{0.1, 0\}$ and $\gamma_1 = 10$. The figure shows that the probability mass of the empirical distribution of the parameter ω is shifted to the left and is very close to zero and that the empirical distribution of the parameter β_1 is shifted to the right when $\delta_1 = 0.05$. These results are even more striking for a large change in the unconditional variance, i.e. for $\delta_1 = 0.1$. The empirical distribution of the persistence of volatility shocks measured by $\widehat{\alpha}_1 + \widehat{\beta}_1$ is very close to 0.95 when small changes occur in the unconditional variance. Of course, the probability of $\widehat{\alpha}_1 + \widehat{\beta}_1$ being very close to one increases with the size of the deterministic change. For changes in the unconditional variance well approximated by a step function, the bias in measuring the persistence of volatility shocks is particularly severe. The plots of the estimated densities for $\gamma_1 = 50$ can be found in Figure 3.5. These findings are in agreement with the well documented results in the GARCH literature that shifts in the unconditional variance lead to an upward bias in the persistence of volatility shocks (see e.g. Lamoureux and Lastrapes (1990)).



(a) $\widehat{\omega}$ (b) $\widehat{\alpha}_1$ (c) $\widehat{\beta}_1$ (d) $\widehat{\alpha}_1 + \widehat{\beta}_1$

Figure 3.4 Estimated densities of the estimated GARCH(1,1) parameters when the DGP is a TV-GARCH(1,1) model with a single transition. The observations are generated by the process $\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}$ where $h_t = 0.05 + 0.1\varepsilon_{t-1}^2 + 0.8h_{t-1}$ and $g_t = 1 + \delta_1(1 + \exp\{-10(t/T - 0.5)\})^{-1}$. The results are based on 2000 replications with 5000 observations.

In the second design, data generated from a TV-GJR-GARCH(1,1) model for $\{\alpha_1, \kappa_1\} = \{0.05, 0.05\}$ is fitted to a GJR-GARCH(1,1) model. The results are presented in Figures 3.6-3.7. The findings in the asymmetric model are very similar to the symmetric case. The empirical distributions of the parameters ω and κ_1 are shifted to the left and that the empirical distributions of the parameters α_1 and β_1 are shifted to the right due to changes in the unconditional variance. Again, the probability of finding the persistence indicator $\widehat{\alpha}_1 + \widehat{\kappa}_1/2 + \widehat{\beta}_1$ very close to one is very large. This

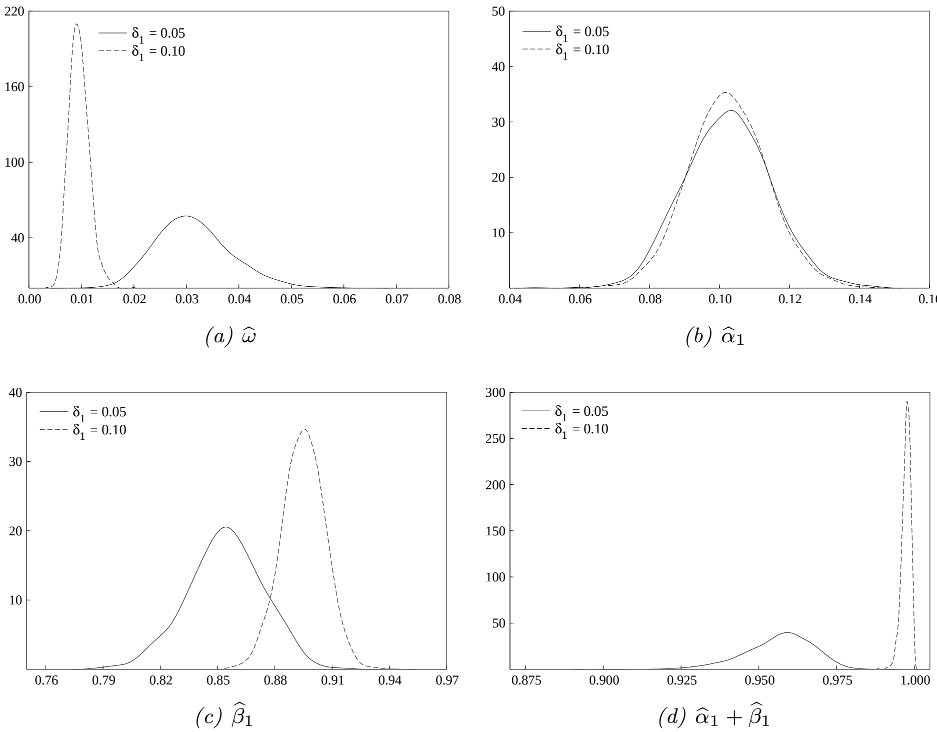


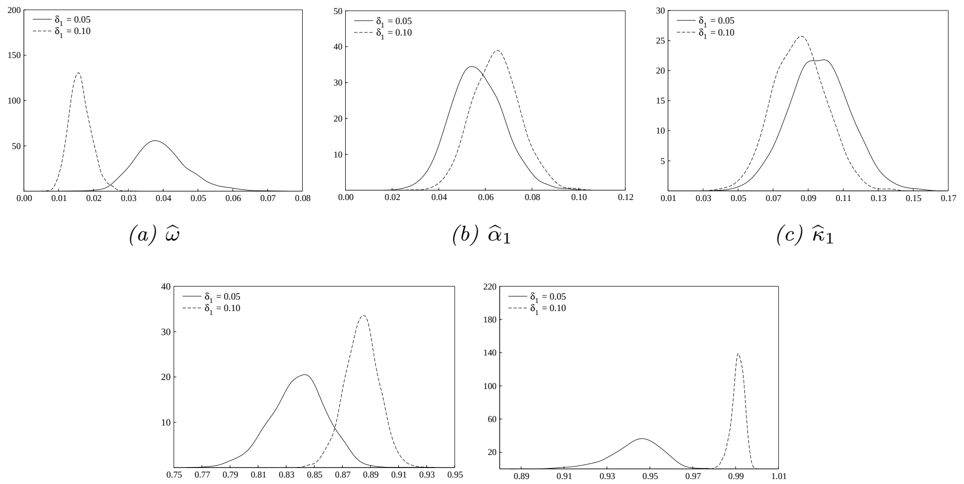
Figure 3.5 Estimated densities of the estimated GARCH(1,1) parameters when the DGP is a TV-GARCH(1,1) model with a single transition. The observations are generated by the process $\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}$ where $h_t = 0.05 + 0.1\varepsilon_{t-1}^2 + 0.8h_{t-1}$ and $g_t = 1 + \delta_1(1 + \exp\{-50(t/T - 0.5)\})^{-1}$. The results are based on 2000 replications with 5000 observations.

effect is more remarkable when the changes in the unconditional variance are abrupt. Moreover, the simulations agree with the results obtained in the application.

It may be of interest to investigate the behaviour of the LM-type tests involved in the modelling strategy when the data has been generated by a long-memory process. For this purpose, we study the empirical power properties of the test statistic using as data-generating process the FIGARCH model of Baillie, Bollerslev, and Mikkelsen (1996). In particular, the artificial series is generated according to a FIGARCH(1,d,1) model as follows:

$$\begin{aligned} \varepsilon_t &= \zeta_t h_t^{1/2}, \qquad\qquad \varepsilon_t|\mathfrak{F}_{t-1} \sim N(0, h_t) \\ [1 - \beta L]h_t &= \omega + [1 - \beta L - \phi L(1 - L)^d]\varepsilon_t^2 \end{aligned}$$

where $\omega = 0.05, \phi = 0.5$ and $\beta = 0.7$. We use sample sizes of 1000, 2500 and 5000 observations with 2000 replications. The actual rejection frequencies of the test at



(a) $\widehat{\omega}$ *(b)* $\widehat{\alpha}_1$ *(c)* $\widehat{\kappa}_1$

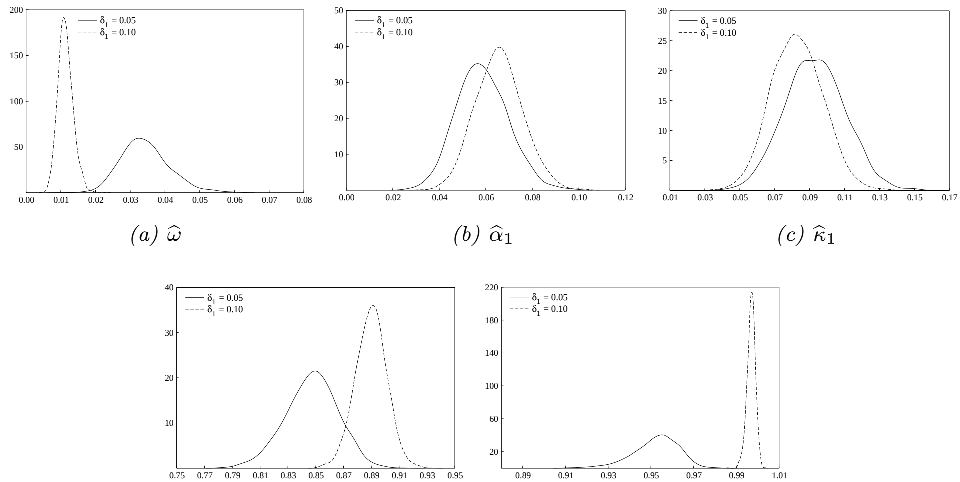
(d) $\widehat{\beta}_1$ *(e)* $\widehat{\alpha}_1 + \widehat{\kappa}_1/2 + \widehat{\beta}_1$

Figure 3.6 Estimated densities of the estimated GJR-GARCH(1,1) parameters when the DGP is a TV-GJR-GARCH(1,1) model with a single transition. The observations are generated by the process $\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}$ where $h_t = 0.05 + 0.05\varepsilon_{t-1}^2 + 0.1\varepsilon_{t-1}^2 I_{t-1}(\varepsilon_{t-1} < 0) + 0.8h_{t-1}$ and $g_t = 1 + \delta_1(1 + \exp\{-10(t/T - 0.5)\})^{-1}$. The results are based on 2000 replications with 5000 observations.

1% and 5% critical values are reported in Table 3.8. Since we focus on the power results when the long-memory parameter d is located in the nonstationary region we use $d = 0.5, 0.6, 0.7$ and 0.8. The power of the tests is moderate for the LM_1-type test for the sample size of 2500 observations. The LM_3-type test is, however, more powerful than the LM_1-test and it has better power properties for the larger sample size. Interestingly, the power is higher at low values of the long memory parameter d it is at higher values of this parameter.

3.6 Conclusions

In this paper we develop a testing and modelling procedure for describing the long-term movements in stock market returns over very long time periods. This is done by multiplicatively decomposing the variance of a GARCH model into a conditional and an unconditional component, in which the unconditional variance is allowed to change smoothly over time. The proposed model is the Time-Varying GARCH model as in Amado and Teräsvirta (2008). The model building strategy relies on statistical inference, making use of a sequence of Lagrange-multiplier type specification tests. Because of the length of the observation period, the time series is divided into non-overlapping subperiods with the aim of alleviating the model building procedure. One



(a) $\widehat{\omega}$ (b) $\widehat{\alpha}_1$ (c) $\widehat{\kappa}_1$

(d) $\widehat{\beta}_1$ (e) $\widehat{\alpha}_1 + \widehat{\kappa}_1/2 + \widehat{\beta}_1$

Figure 3.7 Estimated densities of the estimated GJR-GARCH(1,1) parameters when the DGP is a TV-GJR-GARCH(1,1) model with a single transition. The observations are generated by the process $\varepsilon_t = \zeta_t h_t^{1/2} g_t^{1/2}$ where $h_t = 0.05 + 0.05\varepsilon_{t-1}^2 + 0.1\varepsilon_{t-1}^2 I_{t-1}(\varepsilon_{t-1} < 0) + 0.8h_{t-1}$ and $g_t = 1 + \delta_1(1 + \exp\{-50(t/T - 0.5)\})^{-1}$. The results are based on 2000 replications with 5000 observations.

Table 3.8 Actual rejection frequencies of the standard LM test of constant unconditional variance

		$d = 0.5$		$d = 0.6$		$d = 0.7$		$d = 0.8$	
	α	LM_1	LM_3	LM_1	LM_3	LM_1	LM_3	LM_1	LM_3
$T = 1000$	1%	18.20	27.80	19.55	29.75	18.70	28.90	14.80	23.80
	5%	33.15	48.35	33.85	50.45	31.90	48.85	27.30	41.95
$T = 2500$	1%	29.85	49.20	29.60	49.80	25.95	45.40	19.20	35.15
	5%	45.25	67.55	44.75	68.40	38.30	62.10	30.40	53.50
$T = 5000$	1%	37.65	66.30	35.50	64.85	28.95	54.05	20.50	38.00
	5%	50.95	78.80	48.15	76.25	42.45	69.65	32.65	56.00

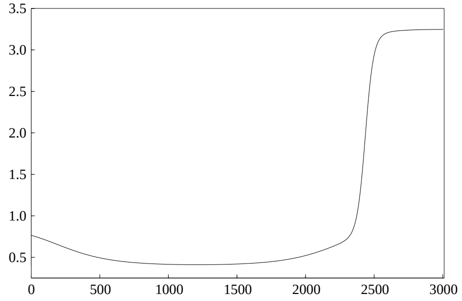
Notes: Monte Carlo results in percentage of the non-robust LM parameter constancy test based on 2000 replications. The artificial series is generated according to a FIGARCH(1,d,1) model $y_t = \varepsilon_t, \varepsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t)$ and $[1 - \beta L]h_t = \omega + [1 - \beta L - \phi L(1 - L)^d]\varepsilon_t^2$. Results are shown for the LM test at the 1%, 5% and 10% nominal significance levels. The columns 'LM_1' and 'LM_3' correspond to the test procedure based on the first-order and third-order Taylor expansions, respectively.

advantage of this device is that it provides a modelling framework particularly useful in applications involving very long time series.

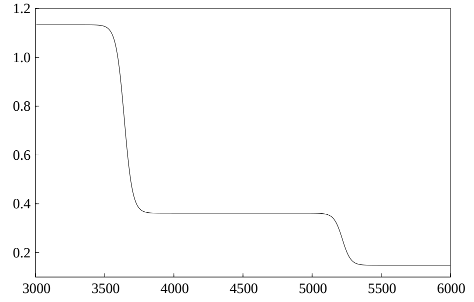
An empirical example applied to the long daily DJIA return series shows how the technique works in practice. Our results show that the dependence structure of the series is best explained by deterministic changes in the unconditional variance, and consequently the hypothesis of constant unconditional variance turns out to be inappropriate. We also show empirically and with a small Monte Carlo experiment how unmodelled deterministic changes in the unconditional variance reproduce the long-memory property in the variance. Based on the diagnostic tests, we claim that the nonstationary TV-GARCH model should be preferred to the stationary model in applications using long financial data.

Moreover, the results indicate that the first-order GJR-GARCH model is inadequate to describe the short-run dynamics of volatility over long return series, and another type of nonlinear model should be considered. Further improvements in the modelling of the conditional variance over long time series are needed, but this problem is left for further research.

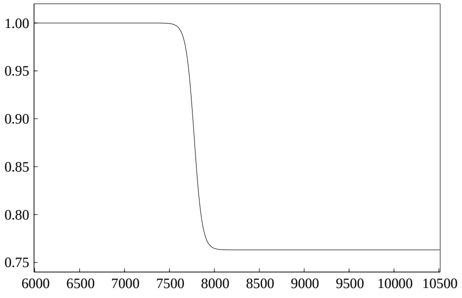
Appendix A: Figures



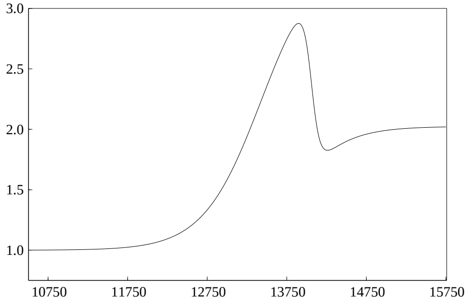
(a) Subperiod 1



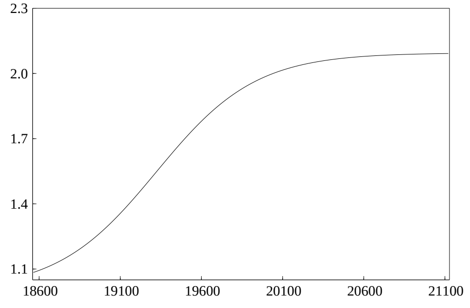
(b) Subperiod 2



(c) Subperiod 3

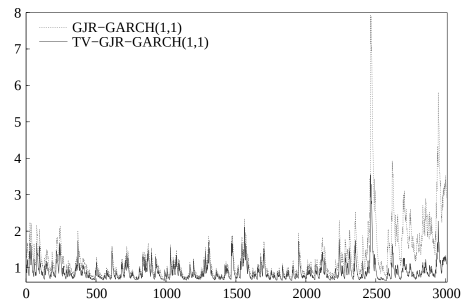


(d) Subperiod 4

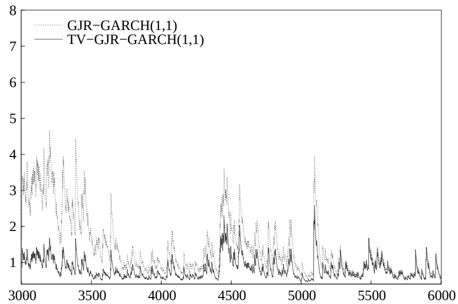


(e) Subperiod 6

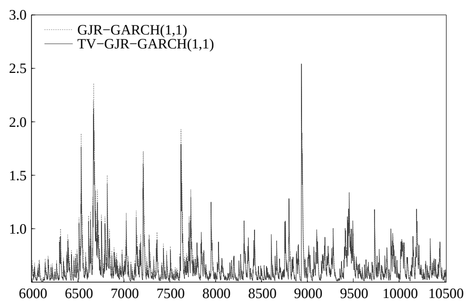
Figure 3.8 Estimated g_t functions for the five subperiods.



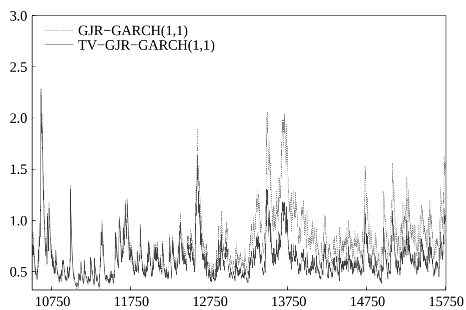
(a) Subperiod 1



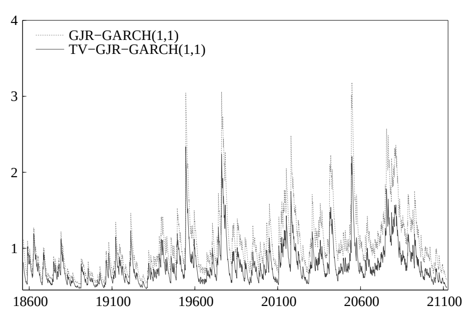
(b) Subperiod 2



(c) Subperiod 3



(d) Subperiod 4



(e) Subperiod 6

Figure 3.9 Conditional standard deviations of the GJR-GARCH(1,1) and the TV-GJR-GARCH(1,1) model for the five subperiods.

Appendix B: Tables

Table 3.9 Summary statistics of the subperiod return series

Periods	Min	Max	Mean	S.D.	Skew	Ex.Kr.	Rob.Sk.	Rob.Kr.
Subsample 1 (02/01/1920 – 31/12/1931): T=3004								
ε_t	-14.47	13.86	-0.011	1.486	-0.402	13.97	-0.073	0.173
$\varepsilon_t/\widehat{g}_t^{1/2}$	-5.669	4.704	0.017	0.876	-0.435	2.899	-0.030	0.118
Subsample 2 (04/01/1932 – 31/12/1943): T=2995								
ε_t	-8.778	14.27	0.019	1.652	0.198	7.178	-0.015	0.389
$\varepsilon_t/\widehat{g}_t^{1/2}$	-4.785	5.824	0.011	0.860	-0.173	3.684	0.008	0.266
Subsample 3 (04/01/1944 – 29/12/1961): T=4511								
ε_t	-6.766	4.048	0.037	0.702	-0.869	6.378	0.008	0.145
$\varepsilon_t/\widehat{g}_t^{1/2}$	-6.766	4.048	0.036	0.660	-0.818	6.054	0.013	0.143
Subsample 4 (01/01/1962 – 16/11/1982): T=5241								
ε_t	-5.882	4.952	0.006	0.844	0.251	2.908	-0.015	0.123
$\varepsilon_t/\widehat{g}_t^{1/2}$	-5.878	4.576	0.005	0.656	0.185	3.822	-0.002	0.068
Subsample 5 (17/11/1982 – 31/12/1993): T=2813								
ε_t	-25.63	9.666	0.047	1.087	-4.768	115.57	-0.020	0.187
Subsample 6 (03/01/1994 – 31/12/2003): T=2557								
ε_t	-7.455	6.155	0.040	1.111	-0.261	4.131	0.030	0.232
$\varepsilon_t/\widehat{g}_t^{1/2}$	-5.655	4.262	0.033	0.814	-0.287	3.653	0.040	0.148

Notes: The table contains summary statistics for each of the subperiod series. The sample periods are indicated in parentheses. The statistic ‘S.D.’ is the standard deviation, ‘Skew’ is the coefficient of skewness and the statistic ‘Ex.Kr’ is the value of the excess kurtosis. ‘Rob.Sk.’ denotes the robust measure for skewness and ‘Rob.Kr.’ denotes the robust centred coefficient for kurtosis. ‘Rob.Sk.’ is computed as $SK = (Q_3 + Q_1 - 2Q_2)/(Q_3 - Q_1)$ where Q_i is the i th quartile of the returns and ‘Rob.Kr.’ is computed as $KR = (E_7 - E_5 + E_3 - E_1)/(E_6 - E_2) - 1.23$ where E_i is the i th octile (see Kim and White (2004) for details).

Table 3.10 Estimation results of the TV-GJR-GARCH(1,1) model: subperiods

Subsample 1

$$h_t = \underset{(0.0223)}{0.0668} + \underset{(0.0132)}{0.0032}\varepsilon^{*2}_{t-1} + \underset{(0.0405)}{0.8244}h_{t-1} + \underset{(0.0383)}{0.1575}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -3677.8 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9063$$

$$g_t = \underset{(0.0128)}{0.3738} + \underset{(0.0627)}{0.4779}(1 + \exp\{-\underset{(8.3423)}{24.379}(t^* - \underset{(0.0036)}{0.0856})(t^* - \underset{(0.0118)}{0.7228})\})^{-1}$$
$$+\underset{(0.2051)}{2.3966}(1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0023)}{0.8109})\})^{-1}$$

Subsample 2

$$h_t = \underset{(0.0065)}{0.0205} + \underset{(0.0098)}{0.0338}\varepsilon^{*2}_{t-1} + \underset{(0.0220)}{0.8876}h_{t-1} + \underset{(0.0277)}{0.1041}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -3524.3 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9735$$

$$g_t = \underset{(0.0585)}{1.1331} - \underset{(0.0599)}{0.7717}(1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0031)}{0.2128})\})^{-1}$$
$$-\underset{(0.0152)}{0.2130}(1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0020)}{0.7395})\})^{-1}$$

Subsample 3

$$h_t = \underset{(0.0255)}{0.0423} + \underset{(0.0109)}{0.0007}\varepsilon^{*2}_{t-1} + \underset{(0.0793)}{0.8331}h_{t-1} + \underset{(0.05278)}{0.1348}\, I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -4298.5 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9012$$

$$g_t = 1 - \underset{(0.0207)}{0.2369}(1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0039)}{0.3937})\})^{-1}$$

Subsample 4

$$h_t = \underset{(0.0016)}{0.0076} + \underset{(0.0056)}{0.0203}\varepsilon^{*2}_{t-1} + \underset{(0.0094)}{0.9191}h_{t-1} + \underset{(0.0116)}{0.0884}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -4769.2 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9836$$

$$g_t = 1 + \underset{(0.2498)}{2.4727}(1 + \exp\{-\underset{(4.9947)}{14.373}(t^* - \underset{(0.0059)}{0.5568})\})^{-1}$$
$$-\underset{(0.2590)}{1.4488}(1 + \exp\{-\underset{(-)}{100}(t^* - \underset{(0.0040)}{0.6775})\})^{-1}$$

Subsample 6

$$h_t = \underset{(0.0065)}{0.0197} + \underset{(0.0225)}{0.9026}h_{t-1} + \underset{(0.0341)}{0.1400}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

$$\text{Log-Lik} = -2883.1 \qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9726$$

$$g_t = 1 + \underset{(0.0741)}{1.0949}(1 + \exp\{-\underset{(9.5905)}{8.3780}(t^* - \underset{(0.0128)}{0.2967})\})^{-1}$$

Notes: The table contains the parameter estimates from the TV-GJR(1,1) model for each of the subperiods of the DJIA daily returns from January 2, 1920 until December 31, 2003. The estimated model has the form of the equations (3.5)-(3.8). The numbers in parentheses are the standard errors.

Table 3.11 Estimation results of the GJR-GARCH(1,1) model: subperiods

Subsample 1

$$h_t = \underset{(0.0110)}{0.0467} + \underset{(0.0093)}{0.0230}\varepsilon^{*2}_{t-1} + \underset{(0.0176)}{0.8752}h_{t-1} + \underset{(0.0255)}{0.1399}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

Log-Lik $= -4675.49$ $\qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9680$

Subsample 2

$$h_t = \underset{(0.0063)}{0.0168} + \underset{(0.0111)}{0.0349}\varepsilon^{*2}_{t-1} + \underset{(0.0192)}{0.9126}h_{t-1} + \underset{(0.0260)}{0.0930}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

Log-Lik $= -4911.1$ $\qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9940$

Subsample 3

$$h_t = \underset{(0.0274)}{0.0332} + \underset{(0.0105)}{0.0015}\varepsilon^{*2}_{t-1} + \underset{(0.0881)}{0.8732}h_{t-1} + \underset{(0.0642)}{0.1111}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

Log-Lik $= -4542.8$ $\qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9302$

Subsample 4

$$h_t = \underset{(0.0012)}{0.0045} + \underset{(0.0055)}{0.0252}\varepsilon^{*2}_{t-1} + \underset{(0.0077)}{0.9310}h_{t-1} + \underset{(0.0110)}{0.0814}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

Log-Lik $= -5908.2$ $\qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9969$

Subsample 5

$$h_t = \underset{(0.0313)}{0.0359} + \underset{(0.0136)}{0.0279}\varepsilon^{*2}_{t-1} + \underset{(0.0633)}{0.8932}h_{t-1} + \underset{(0.0799)}{0.0941}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

Log-Lik $= -3736.4$ $\qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9681$

Subsample 6

$$h_t = \underset{(0.0062)}{0.0189} + \underset{(0.0095)}{0.0059}\varepsilon^{*2}_{t-1} + \underset{(0.0177)}{0.9169}h_{t-1} + \underset{(0.0313)}{0.1299}I_{t-1}(\varepsilon^{*}_{t-1} < 0)\varepsilon^{*2}_{t-1}$$

Log-Lik $= -3587.5$ $\qquad \widehat{\alpha}_1 + \widehat{\beta}_1 + \widehat{\kappa}_1/2 = 0.9878$

Notes: The table contains the parameter estimates from the GJR(1,1) model for each of the subperiods of the DJIA daily returns from January 2, 1920 until December 31, 2003. The estimated model has the form $h_{it} = \omega_i + \alpha_{i1}\varepsilon^2_{it-1} + \beta_{i1}h_{it-1} + \kappa_{i1}I_{it-1}(\varepsilon_{it-1})\varepsilon^2_{it-1}$, where $I_{it}(\varepsilon_{it}) = 1$ if $\varepsilon_{it} < 0$ (and 0 otherwise) for all i. The numbers in parentheses are the Bollerslev-Wooldridge robust standard errors.

Table 3.12 GPH estimates of the long-memory parameter

Periods	$d_{GPH}(m=T^{0.4})$		$d_{GPH}(m=T^{0.5})$		$d_{GPH}(m=T^{0.6})$	
	ε_t	$\varepsilon_t/\widehat{g}_t^{1/2}$	ε_t	$\varepsilon_t/\widehat{g}_t^{1/2}$	ε_t	$\varepsilon_t/\widehat{g}_t^{1/2}$
Subsample 1	0.3688 (0.1434)	−0.171 (0.1214)	0.4237 (0.0866)	0.1059 (0.0837)	0.4050 (0.0539)	0.2001 (0.04865)
Subsample 2	0.7285 (0.1199)	0.2173 (0.1918)	0.6442 (0.0889)	0.2924 (0.1035)	0.5240 (0.0651)	0.3725 (0.0683)
Subsample 3	0.2958 (0.1618)	0.2454 (0.1493)	0.3134 (0.0885)	0.2892 (0.0836)	0.2898 (0.0466)	0.2802 (0.0454)
Subsample 4	0.4457 (0.1330)	0.2228 (0.1179)	0.4728 (0.0893)	0.3198 (0.0773)	0.4873 (0.0605)	0.4333 (0.0520)
Subsample 5	0.2996 (0.0974)	0.2996 (0.0974)	0.4420 (0.0819)	0.4420 (0.0819)	0.2902 (0.0446)	0.2902 (0.0446)
Subsample 6	0.4776 (0.1543)	0.3804 (0.1820)	0.4250 (0.0924)	0.3552 (0.0968)	0.4015 (0.0551)	0.3515 (0.0572)

Notes: The numbers in parentheses are the standard errors. The bandwidth m equals $T^{\alpha}, \alpha \in \{0.4, 0.5, 0.6\}$ where T is the number of observations.

Bibliography

Amado, C., and T. Teräsvirta (2008): "Modelling Unconditional and Conditional Heteroskedasticity with Smoothly Time-Varying Structure," SSE/EFI Working Paper Series in Economics and Finance No. 691.

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

Baillie, R. T., and C. Morana (2007): "Modeling Long Memory and Structural Breaks in Conditional Variances: An Adaptive FIGARCH Approach," Working Paper No. 593, Queen Mary University of London, Department of Economics.

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

Diebold, F. X. (1986): "Comment on "Modelling the Persistence of Conditional Variance", by R. Engle and T. Bollerslev," *Econometric Reviews*, 5, 51–56.

Doornik, J. A. (2002): *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press, see also www.nuff.ox.ac.uk/Users/Doornik, 3rd edn.

Engle, R. F., and J. Gonzalo Rangel (2008): "The Spline-GARCH Model for Low-Frequency Volatility and its Global Macroeconomic Causes," *Review of Financial Studies*, 21, 1187–1222.

Engle, R. F., and G. G. J. Lee (1999): "A Long-Run and Short-Run Component Model of Stock Return Volatility," in *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, ed. by R. F. Engle, and H. White, pp. 475–497. Oxford University Press.

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

Gallant, R. (1984): "The Fourier Flexible Form," *American Journal of Agricultural Economics*, 66, 204–208.

Geweke, J., and S. Porter-Hudak (1983): "The Estimation and Application of Long-Memory Time-Series Models," *Journal of Time Series Analysis*, 4, 221–238.

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

Granger, C. W. J., and N. Hyung (2004): "Occasional Structural Breaks and Long Memory with an Application to the SP 500 Absolute Stock Returns," *Journal of Empirical Finance*, 11, 399–421.

Kim, T.-H., and H. White (2004): "On More Robust Estimation of Skewness and Kurtosis," *Finance Research Letters*, 1, 56–73.

Lamoureux, C. G., and W. D. Lastrapes (1990): "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business & Economic Statistics*, 8, 225–234.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

Mikosch, T., and C. Stărică (2004): "Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects," *The Review of Economics and Statistics*, 86, 378–390.

Song, P. X., Y. Fan, and J. D. Kalbfleisch (2005): "Maximization by Parts in Likelihood Inference," *Journal of the American Statistical Association*, 100, 1145–1158.

Stărică, C., and C. W. J. Granger (2005): "Nonstationarities in Stock Returns," *Review of Economics and Statistics*, 87, 503–522.

Chapter 4

Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations

Conditional Correlation Models of Autoregressive Conditional Heteroskedasticity with Nonstationary GARCH Equations [1]

Abstract

In this paper we investigate the effects of careful modelling the long-run dynamics of the volatilities of stock market returns on the conditional correlation structure. To this end we allow the individual unconditional variances in Conditional Correlation GARCH models to change smoothly over time by incorporating a nonstationary component in the variance equations. The modelling technique to determine the parametric structure of this time-varying component is based on a sequence of specification Lagrange multiplier-type tests derived in Amado and Teräsvirta (2008). The variance equations combine the long-run and the short-run dynamic behaviour of the volatilities. The structure of the conditional correlation matrix is assumed to be either time independent or to vary over time. We apply our model to pairs of seven daily stock returns belonging to the S&P 500 composite index and traded at the New York Stock Exchange. The results suggest that accounting for deterministic changes in the unconditional variances considerably improves the fit of the multivariate Conditional Correlation GARCH models to the data. The effect of careful specification of the variance equations on the estimated correlations is variable: in some cases rather small, in others more discernible.

[1] This paper is a joint work with Timo Teräsvirta.

Acknowledgements: The first author would like to acknowledge financial support from the Louis Fraenckels Stipendiefond. Part of this research was done while the first author was visiting CREATES, University of Aarhus, whose kind hospitality is gratefully acknowledged. The Center for Research in Econometric Analysis of Time Series, CREATES, is funded by the Danish National Research Foundation. The responsibility for any errors and shortcomings in this paper remains ours.

4.1 Introduction

Many financial issues, such as hedging and risk management, portfolio selection and asset allocation rely on information about the covariances or correlations between the underlying returns. This has motivated the modelling of volatility using multivariate financial time series rather than modelling individual returns separately. A number of multivariate generalized autoregressive conditional heteroskedasticity (GARCH) models have been proposed, and some of them have become standard tools for financial analysts. For recent surveys of Multivariate GARCH models see Bauwens, Laurent, and Rombouts (2006) and Silvennoinen and Teräsvirta (2008).

In the univariate setting, volatility models have been extensively investigated. Many modelling proposals of univariate financial returns have suggested that nonstationarities in return series may cause the extreme persistence of shocks observed through estimated GARCH models. In particular, Mikosch and Stărică (2004) found that the long-range dependence and the 'integrated GARCH effect' can also be explained by unaccounted structural breaks in the unconditional variance. Previously, Diebold (1986) and Lamoureux and Lastrapes (1990) have also argued that spurious long memory may be detected from a time series with structural breaks.

The problem of structural breaks in the conditional variance can be tackled by introducing nonstationarity in the volatility equations. In the univariate context, Dahlhaus and Subba Rao (2006) proposed a locally time-varying ARCH process for modelling nonstationary. Engle and Gonzalo Rangel (2008) and, independently, Amado and Teräsvirta (2008) proposed an approach in which the volatility is modelled by a multiplicative decomposition of both conditional and unconditional variance. More specifically, the unconditional variance component is modelled by a slowly varying function: Engle and Gonzalo Rangel (2008) used an exponential spline for this purpose. As an alternative, Amado and Teräsvirta (2008) described the nonstationary component of volatility by a linear combination of logistic transition functions. The authors proposed a modelling technique for determining the parametric structure of the time-varying component from the data.

Despite the growing literature on multivariate GARCH models, little attention has been devoted to modelling multivariate financial data by explicitly allowing for nonstationarity. Recently, Hafner and Linton (2008) proposed a semiparametric generalization of the scalar multiplicative model of Engle and Gonzalo Rangel (2008). Their multivariate GARCH model is a first-order BEKK model with a deterministic nonstationary component. The authors suggested an estimation procedure for the parametric and nonparametric components and established semiparametric efficiency of their estimators.

In this paper, we propose a parametric extension of the univariate multiplicative GARCH model of Amado and Teräsvirta (2008) to the multivariate case. We investigate the effects of careful modelling of the time-varying unconditional variance on the correlation structure in Conditional Correlation GARCH (CC-GARCH) models. To this end, we allow the individual unconditional variances in the multivariate GARCH models to change smoothly over time by incorporating a nonstationary component in the variance equations. The empirical analysis consists of fitting bivariate conditional correlation GARCH models to pairs of daily return series and comparing the results from models with the time-varying unconditional variance component to models without such a component. As a by-product, we extend the concept of news impact surfaces of Kroner and Ng (1998) to the case where both the variances and conditional correlations are fluctuating over time. These surfaces illustrate how the impact of news to covariances between asset returns depends both on the state of the market and the time-varying dependence between the returns.

This paper is organized as follows. In Section 4.2 we describe the Conditional Correlation GARCH model in which the individual unconditional variances change smoothly over time. Estimation of parameters of these models is discussed in Section 4.3. Details regarding the modelling strategy are considered in Section 4.4. Section 4.5 contains the empirical results of fitting bivariate CC-GARCH models to 21 pairs of seven daily stock return series belonging to the S&P 500 composite index. Conclusions can be found in Section 4.6.

4.2 The model

4.2.1 The general framework

Consider a $N \times 1$ vector of return time series $\{\mathbf{y}_t\}, t = 1, \ldots, T$, described by the following vector process:

$$\mathbf{y}_t = \mathsf{E}(\mathbf{y}_t|\mathfrak{F}_{t-1}) + \boldsymbol{\varepsilon}_t \tag{4.1}$$

where $\mathfrak{F}_{t-1}$ is the sigma-algebra generated by the available information up until $t-1$. For simplicity, we assume that the conditional expectation of the returns $\mathsf{E}(\mathbf{y}_t|\mathfrak{F}_{t-1}) = \mathbf{0}$. The $N-$dimensional vector of innovations (or returns) $\{\boldsymbol{\varepsilon}_t\}$ is defined as

$$\boldsymbol{\varepsilon}_t = \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{\zeta}_t \tag{4.2}$$

where the conditional covariance matrix $\boldsymbol{\Sigma}_t = [\sigma_{ijt}]$ of $\boldsymbol{\varepsilon}_t$, given the information set $\mathfrak{F}_{t-1}$, is a positive-definite $N \times N$ matrix. The error vectors $\boldsymbol{\zeta}_t$ form a sequence of independent and identically distributed variables with mean zero and a positive definite correlation matrix $\mathbf{P}_t$. Furthermore, the vector of standardized errors $\mathbf{P}_t^{-1/2}\boldsymbol{\zeta}_t \sim \mathrm{iid}(\mathbf{0}, \mathbf{I}_N)$. Under these assumptions, the error vector $\boldsymbol{\varepsilon}_t$ satisfies the following moments conditions:

$$\begin{aligned} &\mathsf{E}(\boldsymbol{\varepsilon}_t|\mathfrak{F}_{t-1}) = \mathbf{0} \\ &\mathsf{E}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'|\mathfrak{F}_{t-1}) = \boldsymbol{\Sigma}_t = \mathbf{D}_t\mathbf{P}_t\mathbf{D}_t' \end{aligned} \tag{4.3}$$

where $\mathbf{D}_t$ is a diagonal matrix of standard deviations. It is now assumed that $\mathbf{D}_t$ consists of a conditionally heteroskedastic component and a deterministic time-dependent one such that

$$\mathbf{D}_t = \mathbf{S}_t\mathbf{G}_t \tag{4.4}$$

where $\mathbf{S}_t = \mathrm{diag}(h_{1t}^{1/2}, \ldots, h_{Nt}^{1/2})$ contains the conditional standard deviations $h_{it}^{1/2}, i = 1, \ldots, N$, and $\mathbf{G}_t = \mathrm{diag}(g_{1t}^{1/2}, \ldots, g_{Nt}^{1/2})$. The elements $g_{it}, i = 1, \ldots, N$, are positive-valued deterministic functions of time, whose structure will be defined in a moment.

Equations (4.3) and (4.4) jointly define the time-varying covariance matrix

$$\mathbf{\Sigma}_t = \mathbf{S}_t\mathbf{G}_t\mathbf{P}_t\mathbf{G}_t\mathbf{S}_t. \tag{4.5}$$

It follows that

$$\sigma_{ijt} = \rho_{ijt}(h_{it}g_{it})^{1/2}(h_{jt}g_{jt})^{1/2},\ i \neq j \tag{4.6}$$

and that

$$\sigma_{iit} = h_{it}g_{it},\ i = 1, \ldots, N. \tag{4.7}$$

From (4.7) it follows that $h_{it} = \sigma_{iit}/g_{it} = \mathsf{E}(\varepsilon_{it}^{*}\varepsilon_{it}^{*\prime}|\mathfrak{F}_{t-1})$, where $\varepsilon_{it}^{*} = \varepsilon_{it}/g_{it}^{1/2}$. When $\mathbf{G}_t \equiv \mathbf{I}_N$ and the conditional correlation matrix $\mathbf{P}_t \equiv \mathbf{P}$, these assumptions define the Constant Conditional Correlation (CCC-) GARCH model of Bollerslev (1990). More generally, when $\mathbf{G}_t \equiv \mathbf{I}_N$ and $\mathbf{P}_t$ is a time-varying correlation matrix, the model belongs to the family of Conditional Correlation GARCH models.

The diagonal elements of the matrix $\mathbf{G}_t$ are defined as follows:

$$g_{it} = 1 + \sum_{l=1}^{r} \delta_{il}G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il}) \tag{4.8}$$

with $\gamma_{il} > 0, i = 1, \ldots, N, l = 1, \ldots, r$. Each g_{it} varies smoothly over time satisfying the conditions $\inf_{t=1,\ldots,T} g_{it} > 0$, and $\delta_{il} \leq M_\delta < \infty, l = 1, \ldots, r$, for $i = 1, \ldots, N$. The parametric form of (4.8), introduced in Amado and Teräsvirta (2008), allows the unconditional variance to change smoothly over time according to the transition function $G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$. The function $G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$ is a generalized logistic function, that is,

$$G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il}) = \left(1 + \exp\left\{-\gamma_{il}\prod_{j=1}^{k}(t/T - c_{ilj})\right\}\right)^{-1},\ \gamma_{il} > 0,\ c_{il1} \leq \ldots \leq c_{ilk}. \tag{4.9}$$

Function (4.9) is by construction continuous and bounded between zero and one. The parameters, c_{ilj} and γ_{il}, determine the location and the speed of the transition between regimes.

The parametric form of (4.8) with (4.9) is very flexible and capable of describing smooth deterministic changes in volatility. Under $\delta_{i1} = \ldots = \delta_{ir} = 0, i = 1, \ldots, N$, in (4.8), the unconditional volatility becomes constant. Assuming either $r > 1$ or $k > 1$ or both with $\delta_{il} \neq 0$ adds flexibility to the unconditional variance component g_{it}. In the simplest case, $r = 1$ and $k = 1$, g_{it} increases monotonically over time when $\delta_{il} > 0$ and decreases monotonically when $\delta_{il} < 0$. The slope parameter γ_{i1} in (4.9) controls the degree of smoothness of the transition: the larger γ_{i1}, the faster the transition is between the extreme regimes. As $\gamma_{i1} \to \infty, g_{it}$ approaches a step function. For small values of γ_{i1}, the transition between regimes is very smooth.

In this work we shall account for potentially asymmetric responses of volatility to positive and negative shocks or returns by modelling the conditional variances by the

GJR-GARCH process of Glosten, Jagannathan, and Runkle (1993). In the present context,

$$h_{it} = \omega_i + \sum_{j=1}^{q} \alpha_{ij} \varepsilon_{i,t-j}^{*2} + \sum_{j=1}^{q} \kappa_{ij} I(\varepsilon_{i,t-j}^{*} < 0) \varepsilon_{i,t-j}^{*2} + \sum_{j=1}^{p} \beta_{ij} h_{i,t-j}, \tag{4.10}$$

where the indicator function $I(A) = 1$ when A is valid, otherwise $I(A) = 0$.

4.2.2 The structure of the conditional correlations

In this work we shall investigate the effects of modelling changes in the unconditional variances on conditional correlations. The idea is to compare models in which the nonstationary component is left unmodelled with ones relying on the decomposition (4.5) with $\mathbf{G}_t \neq \mathbf{I}_N$. As to modelling the time-variation in the correlation matrix $\mathbf{P}_t$, several choices exist. As already mentioned, the simplest multivariate correlation model is the CCC-GARCH model in which $\mathbf{P}_t \equiv \mathbf{P}$. With h_{it} specified as in (4.10), this model will be called the CCC-TVGJR-GARCH model. When $g_{it} \equiv 1$, (4.10) defines the ith conditional variance of the CCC-GJR-GARCH model.

The CCC-GARCH model has considerable appeal due to its computational simplicity, but in many studies the assumption of constant correlations has been found to be too restrictive. There are several ways of relaxing this assumption using parametric representations for the correlations. Engle (2002) introduced the so-called Dynamic CC-GARCH (DCC-GARCH) model in which the conditional correlations are defined through GARCH(1,1) type equations. Tse and Tsui (2002) presented a rather similar model. In the DCC-GARCH model, the coefficient of correlation ρ_{ijt} is a typical element of the matrix $\mathbf{P}_t$ with the dynamic structure

$$\mathbf{P}_t = \{\mathrm{diag}\mathbf{Q}_t\}^{-1/2} \mathbf{Q}_t \{\mathrm{diag}\mathbf{Q}_t\}^{-1/2} \tag{4.11}$$

where

$$\mathbf{Q}_t = (1 - \theta_1 - \theta_2)\overline{\mathbf{Q}} + \theta_1 \boldsymbol{\zeta}_{t-1} \boldsymbol{\zeta}_{t-1}' + \theta_2 \mathbf{Q}_{t-1} \tag{4.12}$$

with the scalars θ_1 and θ_2 satisfying $\theta_1 > 0$ and $\theta_2 \geq 0$ such that $\theta_1 + \theta_2 < 1$, $\overline{\mathbf{Q}}$ is the unconditional correlation matrix of the standardized errors $\zeta_{it}, i = 1, \ldots, N$, and $\boldsymbol{\zeta}_t = (\zeta_{1t}, \ldots, \zeta_{Nt})'$. In our case, each $\zeta_{it} = \varepsilon_{it}/(h_{it} g_{it})^{1/2}$, and this version of the model will be called the DCC-TVGJR-GARCH model. Accordingly, when $g_{it} \equiv 1$, the model becomes the DCC-GJR-GARCH model.

Another way of introducing time-varying correlations is to assume that the conditional correlation matrix $\mathbf{P}_t$ varies smoothly over time between two extreme states of correlations $\mathbf{P}_{(1)}$ and $\mathbf{P}_{(2)}$; see Berben and Jansen (2005) and Silvennoinen and Teräsvirta (2005, 2007). The correlation matrix is a convex combination of these two matrices such that:

$$\mathbf{P}_t = \{1 - G(s_t; \gamma, c)\}\mathbf{P}_{(1)} + G(s_t; \gamma, c)\mathbf{P}_{(2)} \tag{4.13}$$

where $\mathbf{P}_{(1)}$ and $\mathbf{P}_{(2)}$ are positive definite matrices and $\mathbf{P}_{(1)} \neq \mathbf{P}_{(2)}$.$G(s_t;\gamma,c)$ is a monotonic function bounded between zero and one, where the stochastic or deterministic transition variable s_t controls the correlations. More specifically,

$$G(s_t;\gamma,c) = (1+\exp\{-\gamma(s_t-c)\})^{-1}, \ \gamma>0 \tag{4.14}$$

where, as in (4.9), the parameter γ determines the smoothness and c the location of the transition between the two correlation regimes. In this work, $s_t = t^* = t/T$, and we call the resulting model with (4.10) the Time-Varying CC-TVGJR-GARCH (TVCC-TVGJR-GARCH) model when the equations for h_{it} are parameterized using a TVGJR specification. When $g_{it} \equiv 1$, (4.13) reduces to the conditional covariance of the TVCC-GJR-GARCH model.

4.3 Estimation of parameters

4.3.1 Estimation of DCC-TVGJR-GARCH models

In this section, we assume that $\omega_i = 1, i = 1,\ldots,N$, in (4.10) and that (4.8) has the form

$$g_{it} = \delta_{i0} + \sum_{l=1}^{r} \delta_{il} G_{il}(t/T;\gamma_{il},\mathbf{c}_{il})$$

where $\delta_{i0} > 0$. This facilitates the notation but does not change the argument. Under the assumption of normality, $\boldsymbol{\varepsilon}_t|\mathfrak{F}_{t-1} \sim N(\mathbf{0},\boldsymbol{\Sigma}_t)$, the conditional log-likelihood function for observation t is defined as

$$\begin{aligned}
\ell_t(\boldsymbol{\theta}) &= -(N/2)\ln 2\pi - (1/2)\ln|\boldsymbol{\Sigma}_t| - (1/2)\boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\varepsilon}_t \\
&= -(N/2)\ln(2\pi) - (1/2)\ln|\mathbf{S}_t\mathbf{G}_t\mathbf{P}_t\mathbf{G}_t\mathbf{S}_t| - (1/2)\boldsymbol{\varepsilon}_t'\mathbf{S}_t^{-1}\mathbf{G}_t^{-1}\mathbf{P}_t^{-1}\mathbf{G}_t^{-1}\mathbf{S}_t^{-1}\boldsymbol{\varepsilon}_t \\
&= -(N/2)\ln(2\pi) - \ln|\mathbf{S}_t\mathbf{G}_t| - (1/2)\ln|\mathbf{P}_t| - (1/2)\boldsymbol{\zeta}_t'\mathbf{P}_t^{-1}\boldsymbol{\zeta}_t \\
&= -(N/2)\ln(2\pi) - \ln|\mathbf{G}_t| - (1/2)\widetilde{\boldsymbol{\varepsilon}}_t'\mathbf{G}_t^{-2}\widetilde{\boldsymbol{\varepsilon}}_t - \ln|\mathbf{S}_t| - (1/2)\boldsymbol{\varepsilon}_t^{*\prime}\mathbf{S}_t^{-2}\boldsymbol{\varepsilon}_t^{*} \\
&\quad +\boldsymbol{\zeta}_t'\boldsymbol{\zeta}_t - (1/2)\ln|\mathbf{P}_t| - (1/2)\boldsymbol{\zeta}_t'\mathbf{P}_t^{-1}\boldsymbol{\zeta}_t
\end{aligned} \tag{4.15}$$

where $\boldsymbol{\theta}= (\boldsymbol{\psi}',\boldsymbol{\varphi}',\boldsymbol{\phi}')'$ is the vector of all parameters of the model, and

$$\begin{aligned}
\widetilde{\boldsymbol{\varepsilon}}_t &= \mathbf{S}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}/\{h_{1t}(\boldsymbol{\psi}_1,\boldsymbol{\varphi}_1)\}^{1/2},\ldots,\varepsilon_{Nt}/\{h_{Nt}(\boldsymbol{\psi}_N,\boldsymbol{\varphi}_N)\}^{1/2})' \\
\boldsymbol{\varepsilon}_t^{*} &= \mathbf{G}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}/\{g_{1t}(\boldsymbol{\psi}_1)\}^{1/2},\ldots,\varepsilon_{Nt}/\{g_{Nt}(\boldsymbol{\psi}_N)\}^{1/2})' \\
\boldsymbol{\zeta}_t &= \mathbf{G}_t^{-1}\mathbf{S}_t^{-1}\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}/\{g_{1t}(\boldsymbol{\psi}_1)h_{1t}(\boldsymbol{\psi}_1,\boldsymbol{\varphi}_1)\}^{1/2},\ldots,\varepsilon_{Nt}/\{g_{Nt}(\boldsymbol{\psi}_N)h_{Nt}(\boldsymbol{\psi}_N,\boldsymbol{\varphi}_N)\}^{1/2})'.
\end{aligned}$$

Equation (4.15) implies the following decomposition of the log-likelihood function for observation t :

$$\ell_t(\boldsymbol{\psi},\boldsymbol{\varphi},\boldsymbol{\phi}) = \ell_t^U(\boldsymbol{\psi}) + \ell_t^V(\boldsymbol{\psi},\boldsymbol{\varphi}) + \ell_t^C(\boldsymbol{\psi},\boldsymbol{\varphi},\boldsymbol{\phi})$$

where first, $\boldsymbol{\psi}= (\boldsymbol{\psi}_1',\ldots,\boldsymbol{\psi}_N')'$, and

$$\ell_t^U(\boldsymbol{\psi}) = \sum_{i=1}^{N} \ell_{it}^U(\boldsymbol{\psi}_i) \tag{4.16}$$

with $\boldsymbol{\psi}_i = (\delta_{i0}, \boldsymbol{\delta}_i', \boldsymbol{\gamma}_i', \mathbf{c}_i')'$, $\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{ir})'$, $\boldsymbol{\gamma}_i = (\gamma_{i1}, \ldots, \gamma_{ir})'$, $\mathbf{c}_i = (\mathbf{c}_{i1}', \ldots, \mathbf{c}_{ir}')'$, $i = 1, \ldots, N$, and

$$\ell_{it}^U(\boldsymbol{\psi}_i) = -(1/2)\{\ln g_{it}(\boldsymbol{\psi}_i) + \tilde{\varepsilon}_{it}^2 / g_{it}(\boldsymbol{\psi}_i)\}.$$

Second,

$$\ell_t^V(\boldsymbol{\psi},\boldsymbol{\varphi}) = \sum_{i=1}^{N} \ell_{it}^V(\boldsymbol{\psi}_i,\boldsymbol{\varphi}_i) \tag{4.17}$$

where $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1', \ldots, \boldsymbol{\varphi}_N')'$, and

$$\ell_{it}^V(\boldsymbol{\psi}_i,\boldsymbol{\varphi}_i) = -(1/2)\{\ln h_{it}(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i) + \varepsilon_{it}^{*2} / h_{it}(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i)\}.$$

with $\boldsymbol{\varphi}_i = (\alpha_{i1}, \ldots, \alpha_{iq}, \kappa_{i1}, \ldots, \kappa_{iq}, \beta_{i1}, \ldots, \beta_{ip})'$, $i = 1, \ldots, N$. Finally,

$$\ell_t^C(\boldsymbol{\psi},\boldsymbol{\varphi},\boldsymbol{\phi}) = -(1/2)\{\ln |\mathbf{P}_t(\boldsymbol{\psi},\boldsymbol{\varphi},\boldsymbol{\phi})| + \boldsymbol{\zeta}_t' \mathbf{P}_t^{-1}(\boldsymbol{\psi},\boldsymbol{\varphi},\boldsymbol{\phi})\boldsymbol{\zeta}_t - 2\boldsymbol{\zeta}_t'\boldsymbol{\zeta}_t\}. \tag{4.18}$$

The GARCH equations are estimated separately using maximization by parts. The first iteration consists of the following:

1. Maximize

$$L_{iT}^U(\boldsymbol{\psi}) = \sum_{t=1}^{T} \ell_{it}^U(\boldsymbol{\psi}) = -(1/2) \sum_{t=1}^{T} \{\ln g_{it}(\boldsymbol{\psi}_i) + \tilde{\varepsilon}_{it}^2 / g_{it}(\boldsymbol{\psi}_i)\}$$

   for each $i, i = 1, \ldots, N$, separately, assuming $\tilde{\varepsilon}_{it} = \varepsilon_{it}$, that is, setting $h_{it}(\boldsymbol{\psi}_i, \boldsymbol{\varphi}_i) \equiv 1$. The resulting estimators are $\widehat{\boldsymbol{\psi}}_i^{(1)}, i = 1, \ldots, N$.

2. Making use of $\widehat{\boldsymbol{\psi}}_i^{(1)}, i = 1, \ldots, N$, maximize

$$L_{iT}^V(\widehat{\boldsymbol{\psi}}_i^{(1)},\boldsymbol{\varphi}_i) = \sum_{t=1}^{T} \ell_{it}^V(\widehat{\boldsymbol{\psi}}_i^{(1)},\boldsymbol{\varphi}_i) = -(1/2) \sum_{t=1}^{T} \{\ln h_{it}(\widehat{\boldsymbol{\psi}}_i^{(1)}, \boldsymbol{\varphi}_i) + \varepsilon_{it}^{*2} / h_{it}(\widehat{\boldsymbol{\psi}}_i^{(1)}, \boldsymbol{\varphi}_i)\}$$

   with respect to $\boldsymbol{\varphi}_i$ assuming $\varepsilon_{it}^* = \varepsilon_{it} / g_{it}^{1/2}(\widehat{\boldsymbol{\psi}}_i^{(1)})$, for each $i, i = 1, \ldots, N$, separately. Call the resulting estimators $\widehat{\boldsymbol{\varphi}}_i^{(1)}, i = 1, \ldots, N$.

The second iteration is as follows:

1. Maximize

$$L_{iT}^U(\boldsymbol{\psi}) = \sum_{t=1}^{T} \ell_{it}^U(\boldsymbol{\psi}_i) = -(1/2) \sum_{t=1}^{T} \{\ln g_{it}(\boldsymbol{\psi}_i) + \tilde{\varepsilon}_{it}^2 / g_{it}(\boldsymbol{\psi}_i)\}$$

   assuming $\tilde{\varepsilon}_{it} = \varepsilon_{it} / h_{it}^{1/2}(\widehat{\boldsymbol{\psi}}_i^{(1)}, \widehat{\boldsymbol{\varphi}}_i^{(1)})$, for each $i, i = 1, \ldots, N$. Call the resulting estimators $\widehat{\boldsymbol{\psi}}_i^{(2)}$.

2. Maximize

$$L_{iT}^{V}(\widehat{\boldsymbol{\psi}}_i^{(2)},\boldsymbol{\varphi}_i) = \sum_{t=1}^{T} \ell_{it}^{V}(\widehat{\boldsymbol{\psi}}_i^{(2)},\boldsymbol{\varphi}_i) = -(1/2)\sum_{t=1}^{T}\{\ln h_{it}(\widehat{\boldsymbol{\psi}}_i^{(2)},\boldsymbol{\varphi}_i)+\varepsilon_{it}^{*2}/h_{it}(\widehat{\boldsymbol{\psi}}_i^{(2)},\boldsymbol{\varphi}_i)\}$$

with respect to $\boldsymbol{\varphi}_i$ for each $i, i = 1,\ldots,N$, separately, assuming $\varepsilon_{it}^{*} = \varepsilon_{it}/g_{it}(\widehat{\boldsymbol{\psi}}_i^{(2)})$. This yields $\widehat{\boldsymbol{\varphi}}_i^{(2)}, i = 1,\ldots,N$.

Iterate until convergence. Call the resulting estimators $\widehat{\boldsymbol{\psi}}_i$ and $\widehat{\boldsymbol{\varphi}}_i, i = 1,\ldots,N$, and set $\widehat{\boldsymbol{\psi}} = (\widehat{\boldsymbol{\psi}}_1',\ldots,\widehat{\boldsymbol{\psi}}_N')'$ and $\widehat{\boldsymbol{\varphi}} = (\widehat{\boldsymbol{\varphi}}_1',\ldots,\widehat{\boldsymbol{\varphi}}_N')'$.

Maximization is carried out in the usual fashion by solving the equations

$$\frac{\partial}{\partial\boldsymbol{\psi}_i}L_{iT}^{U}(\boldsymbol{\psi}_i) = (1/2)\sum_{t=1}^{T}(\frac{\tilde{\varepsilon}_{it}^2}{g_{it}(\boldsymbol{\psi}_i)}-1)\frac{1}{g_{it}(\boldsymbol{\psi}_i)}\frac{\partial g_{it}(\boldsymbol{\psi}_i)}{\partial\boldsymbol{\psi}_i} = \mathbf{0}$$

for $\boldsymbol{\psi}_i$ and

$$\frac{\partial}{\partial\boldsymbol{\varphi}_i}L_{iT}^{V}(\boldsymbol{\varphi}_i) = (1/2)\sum_{t=1}^{T}(\frac{\varepsilon_{it}^{*2}}{h_{it}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i)}-1)\frac{1}{h_{it}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i)}\frac{\partial h_{it}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i)}{\partial\boldsymbol{\varphi}_i} = \mathbf{0}$$

for $\boldsymbol{\varphi}_i$ in the nth iteration. Writing $G_{ilt} = G(t^*,\gamma_{il},\mathbf{c}_{il})$, we have

$$\frac{\partial g_{it}(\boldsymbol{\psi}_i)}{\partial\boldsymbol{\psi}_i} = (1, G_{i1t}, G_{i1t}^{(\gamma)}, G_{i1t}^{(c)},\ldots,G_{irt},G_{irt}^{(\gamma)},G_{irt}^{(c)})'$$

where, for $k = 1$ in (4.9),

$$\begin{aligned} G_{ilt}^{(\gamma)} &= \frac{\partial G_{ilt}}{\partial\gamma_{il}} = \delta_{il}G_{ilt}(1-G_{ilt})(t^*-c_{il}) \\ G_{ilt}^{(c)} &= \frac{\partial G_{ilt}}{\partial c_{ilj}} = -\gamma_{il}\delta_{il}G_{ilt}(1-G_{ilt}) \end{aligned}$$

where c_{ilj} denotes the jth element in the parameter vector $\mathbf{c}_{il}, l = 1,\ldots,r$, and

$$\begin{aligned} \frac{\partial h_{it}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i)}{\partial\boldsymbol{\varphi}_i} &= (1,\varepsilon_{i,t-1}^{*2},\ldots,\varepsilon_{i,t-q}^{*2},\varepsilon_{i,t-1}^{*2}I(\varepsilon_{i,t-1}^{*}<0),\ldots,\varepsilon_{i,t-q}^{*2}I(\varepsilon_{i,t-q}^{*}<0), \\ &\quad h_{i,t-1}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i),\ldots,h_{i,t-p}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i))' + \sum_{j=1}^{p}\beta_{ij}\frac{\partial h_{i,t-j}(\widehat{\boldsymbol{\psi}}_i^{(n)},\boldsymbol{\varphi}_i)}{\partial\boldsymbol{\varphi}_i} \end{aligned}$$

when the conditional variance h_{it} is defined in (4.10).

After estimating the TVGARCH equations, estimate $\boldsymbol{\phi}$ given $\widehat{\boldsymbol{\psi}}_i$ and $\widehat{\boldsymbol{\varphi}}_i$ by maximizing

$$L_T^{C}(\boldsymbol{\phi}) = \sum_{t=1}^{T}\ell_t^{C}(\boldsymbol{\phi}) = -(1/2)\sum_{t=1}^{T}\{\ln|\mathbf{P}_t(\boldsymbol{\phi})| + \boldsymbol{\zeta}_t'\mathbf{P}_t^{-1}(\boldsymbol{\phi})\boldsymbol{\zeta}_t - 2\boldsymbol{\zeta}_t'\boldsymbol{\zeta}_t\}$$

where $\boldsymbol{\zeta}_t = (\zeta_{1t}, \ldots, \zeta_{Nt})'$ with $\zeta_{it} = \varepsilon_{it}/\{h_{it}(\widehat{\boldsymbol{\psi}}_i, \widehat{\boldsymbol{\varphi}}_i)g_{it}(\widehat{\boldsymbol{\psi}}_i)\}^{1/2}, i = 1, \ldots, N$, and

$$\frac{\partial}{\partial \boldsymbol{\phi}} L_T^C(\boldsymbol{\phi}) = -(1/2)\sum_{t=1}^{T} \frac{\partial vec(\mathbf{P}_t)'}{\partial \boldsymbol{\phi}} vec(\mathbf{P}_t^{-1} - \mathbf{P}_t^{-1}\boldsymbol{\zeta}_t\boldsymbol{\zeta}_t'\mathbf{P}_t^{-1}).$$

All computations in this paper have been performed using Ox, version 3.40 (see Doornik (2002)) and a modified version of Matteo Pelagatti's source code[2].

This approach is computationally attractive. Engle and Sheppard (2001) only estimate the GARCH equations once and show that for $\mathbf{G}_t = \mathbf{I}_N$, the maximum likelihood estimators $\widehat{\boldsymbol{\varphi}}_i, i = 1, \ldots, N$, (in their framework $g_{it}(\boldsymbol{\psi}_i) \equiv 1$) are consistent. The two-step estimator is, however, asymptotically less efficient than the full maximum likelihood estimator. Further iteration in order to obtain efficient estimators is possible, see Fan, Pastorello, and Renault (2007) for discussion, but it has not been undertaken here.

4.3.2 Estimation of TVCC-TVGJR-GARCH models

The maximum likelihood estimation of the parameters of the model TVCC-GJR-GARCH model can be carried out in three steps as in Silvennoinen and Teräsvirta (2005, 2007). The log-likelihood function can be decomposed as before. The components (4.16) and (4.17) remain the same, whereas (4.18) becomes

$$\ell_t^C(\boldsymbol{\psi},\boldsymbol{\varphi},\boldsymbol{\varpi}) = -(1/2)\{\ln|\mathbf{P}_t(\boldsymbol{\varpi})| + \boldsymbol{\zeta}_t'\mathbf{P}_t^{-1}(\boldsymbol{\varpi})\boldsymbol{\zeta}_t - 2\boldsymbol{\zeta}_t'\boldsymbol{\zeta}_t\}$$

where the $\{N(N-1)+2\} \times 1$ vector $\boldsymbol{\varpi} = (\text{vecl}(\mathbf{P}_{(1)})', \text{vecl}(\mathbf{P}_{(2)})', \gamma, c)'$. (The vecl operator stacks the columns below the main diagonal into a vector.) In their scheme, the parameter vectors $\boldsymbol{\psi}$ and $\boldsymbol{\varphi}$ of the GARCH equations are estimated first, followed by the conditional correlations in $\mathbf{P}_{(1)}$ and $\mathbf{P}_{(2)}$, given the transition function parameters γ and c in (4.14). Finally, γ and c are estimated given $\boldsymbol{\psi}$,$\boldsymbol{\varphi}$,$\mathbf{P}_{(1)}$ and $\mathbf{P}_{(2)}$. The next iteration begins by re-estimating $\boldsymbol{\varphi}$ given the previous estimates of $\mathbf{P}_{(1)}, \mathbf{P}_{(2)}, \gamma$ and c. The only modification required for the estimation of TVCC-TVGJR-GARCH models compared to Silvennoinen and Teräsvirta (2005) is that for each main iteration there is an inside loop for iterative estimation (maximization by parts) of $\boldsymbol{\psi}$ and $\boldsymbol{\varphi}$. In practice, compared to the two-step estimates, the extra iterations do not change the estimates very much, but the estimators become fully efficient.

Asymptotic properties of the maximum likelihood estimators of the TVCC-TVGJR-GARCH model are not yet known. The existing results only cover the CCC-GARCH model; see Ling and McAleer (2003). Deriving corresponding asymptotic results for the TVCC-TVGJR-GARCH model is a nontrivial problem and beyond the scope of the present paper.

[2]The Ox estimation package is freely available at
http://www.statistica.unimib.it/utenti/p_matteo/Ricerca/research.html

4.4 Modelling with TVGJR-GARCH models

4.4.1 Specifying the unconditional variance component

In applying a model belonging to the family of CC-TVGJR-GARCH models, there are two specification problems. First, one has to determine p and q in (4.10) and r in (4.8). Furthermore, if $r \geq 1$, one also has to determine k for each transition function (4.9). Second, one has to test the null of conditional correlations against either the DCC- or TVCC-GARCH model. We shall concentrate on the first set of issues. It appears that in applications involving DCC-GARCH models, the null hypothesis of constant correlations is never tested, and we shall adhere to that practice. In applications of the STCC-GARCH model, constancy of correlations is typically tested before applying the larger model, see Silvennoinen and Teräsvirta (2005, 2007). The test can be extended to the current situation in which the GARCH equations are TVGJR-GARCH ones instead of plain GJR-GARCH ones. Nevertheless, in this work we assume that the correlations do vary over time as is done in the context of DCC-GARCH models and apply the TVCC-GARCH model without a constancy test.

We shall thus concentrate on the first set of specification issues. We choose $p = q = 1$ and test for higher orders at the evaluation stage. As to selecting r and k, we shall follow Amado and Teräsvirta (2008) and briefly review their procedure. The functions g_{it} are specified equation by equation under the assumption $h_{it} \equiv \alpha_{i0} > 0$, which means that the specification is carried out by assuming that the conditional variances remain constant. For the ith equation, the first hypothesis to be tested is H_{01}: $\gamma_{i1} = 0$ against H_{11}: $\gamma_{i1} > 0$ in

$$g_{it} = \alpha_{i0}^{-1}\{1 + \delta_{i1} G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1})\} = \alpha_{i0}^{-1} + \delta_{i1}^{*} G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1})$$

where $\delta_{i1}^{*} = \alpha_{i0}^{-1}\delta_{i1}$. The standard test statistic has a non-standard asymptotic distribution because δ_{i1}^{*} and $\mathbf{c}_{i1}$ are unidentified nuisance parameters when H_{01} is true. This lack of identification may be circumvented by following Luukkonen, Saikkonen, and Teräsvirta (1988). This means that $G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1})$ is replaced by its mth-order Taylor expansion around $\gamma_{i1} = 0$. Choosing $m = 3$, this yields

$$g_{it} = \alpha_0^{*} + \sum_{j=1}^{3} \delta_{ij}^{**}(t/T)^j + R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1}) \tag{4.19}$$

where $\delta_{ij}^{**} = \gamma_{i1}\widetilde{\delta}_{ij}^{*}$ with $\widetilde{\delta}_{ij}^{*} \neq 0$, and $R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1})$ is the remainder. The new null hypothesis based on this approximation is H_{01}': $\delta_{i1}^{**} = \delta_{i2}^{**} = \delta_{i3}^{**} = 0$ in (4.19). In order to test this null hypothesis, we use the Lagrange multiplier (LM) test. Furthermore, $R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1}) \equiv 0$ under H_{01}, so the asymptotic distribution theory is not affected by the remainder. As discussed in Amado and Teräsvirta (2008), the LM-type test statistic has an asymptotic χ^2-distribution with three degrees of freedom when H_{01} holds.

If the null hypothesis is rejected, the model builder also faces the problem of selecting the order $k \leq 3$ in the exponent of $G_{il}(\gamma_{il}; \mathbf{c}_{il}, t/T)$. It is solved by carrying out a short test sequence within (4.19); for details see Amado and Teräsvirta (2008).

The next step is then to estimate the alternative with the chosen k, add another transition, and test the hypothesis $\gamma_{i2} = 0$ in

$$g_{it} = \alpha_0^{-1} + \delta_{i1}^* G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1}) + \delta_{i2}^* G_{i1}(t/T; \gamma_{i2}, \mathbf{c}_{i2})$$

using the same technique as before. Testing continues until the first non-rejection of the null hypothesis.

4.4.2 The modelling cycle

After specifying the model, its parameters are estimated and the estimated model evaluated. In short, building TVGJR-GARCH models for the elements of $\mathbf{D}_t = \mathbf{S}_t\mathbf{G}_t$ of the CC-GARCH model defined by equations (4.3) and (4.4) proceeds as follows:

1. First assume $h_t \equiv \alpha_{i0}$ and test H_{01}: $g_{it} \equiv \alpha_{i0}^{-1}$ (constant) against H_{11}: $g_{it} = \alpha_{i0}^{-1} + \delta_{i1} G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1})$, for $i = 1, \ldots, N$, at the significance level $\alpha^{(1)}$. In case of a rejection, select k and test H_{02}: $g_{it} = \alpha_{i0}^{-1} + \delta_{i1}^* G_{i1}(t/T; \gamma_{i1}, \mathbf{c}_{i1})$ against H_{12}: $g_{it} = \alpha_{i0}^{-1} + \sum_{l=1}^{2} \delta_{il}^* G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$ at the significance level $\alpha^{(2)} = \tau\alpha^{(1)}$, where $\tau \in (0, 1)$. More generally, $\alpha^{(j)} = \tau\alpha^{(j-1)}, j = 2, 3, \ldots$. The significance level is lowered at each stage for reasons of parsimony. (We choose $\tau = 1/2$ but note that in our application, the results are quite robust to the choice $\alpha^{(1)}$ and τ in the sense that a wide range of these parameters yield the same r.) Testing is continued until the first non-rejection of the null hypothesis.

2. After specifying and estimating g_{it}, test for conditional heteroskedasticity in $\{\varepsilon_{it}^*\}$. If the null hypothesis of no ARCH is rejected, then model the conditional variance h_t as in (4.10) with $p = q = 1$. In applications to financial return series of sufficiently high frequency, the test may be omitted and the TVGJR-GARCH model for σ_{ijt} estimated directly using maximization by parts.

3. Evaluate the estimated individual TVGJR-GARCH equations by means of LM and LM-type diagnostic tests. For relevant misspecification tests for TV-GARCH models (they are directly applicable to testing TVGJR-GARCH models), see Amado and Teräsvirta (2008). This includes testing for higher orders of p and q in (4.10). If the models pass the tests, they will be incorporated in multivariate CC-GARCH models. If the multivariate model is the TVCC-GJR-GARCH model, the GARCH equations will be re-estimated as described in Section 4.3.

We shall now apply the modelling cycle to individual daily return series. As already indicated, the interest lies in how careful modelling of nonstationarity in return series affects correlation estimates. This will be investigated by a set of bivariate CC-GARCH models.

4.5 Empirical analysis

4.5.1 Data

The effects of modelling the nonstationarity in return series on the conditional correlations are studied with price series of seven stocks of the S&P 500 composite index traded at the New York Stock Exchange. The time series are available at the website *Yahoo! Finance.* They consist of daily closing prices of American International Group Inc. (AIG), American Express (AXP), Boeing Company (BA), Ford Motor Company (F), Intel Corporation (INTC), JPMorgan Chase & Co. (JPM) and AT&T Inc. (T). The seven companies belong to different industries that are insurance services (AIG), consumer finance (AXP), aerospace and defence (BA), automotive (F), semiconductors (INTC), banking (JPM) and telecommunications services (T). A bivariate analysis of returns of these companies may give some idea of how different the correlations between firms representing different industries can be. The observation period starts in October 1, 1993 and ends in September 30, 2003, yielding a total of 2518 observations. All stock prices are converted into continuously compounded rates of returns, whose values are plotted in Figure 4.1. A common pattern is evident in the seven return series. There is a less volatile period from the beginning until the middle of the observation period and a more volatile period starting around 1998 that continues until the end of the sample. Moreover, as expected, all seven return series exhibit volatility clustering.

Descriptive statistics for the individual return series can be found in Table 4.1. Conventional measures for skewness and kurtosis and also their robust counterparts are provided for all series. The conventional estimates indicate both non-zero skewness and excess kurtosis: both are typically found in financial asset returns. However, conventional measures of skewness and kurtosis are sensitive to outliers and should therefore be viewed with caution. Kim and White (2004) suggested to look at robust estimates of these quantities. The robust measures for skewness are all positive but very close to zero indicating that the return distributions show very little skewness. All robust kurtosis measures are positive, which suggests some excess kurtosis (the kurtosis equals zero for normally distributed returns) but less than what the conventional measures indicate. The estimates are strictly univariate and any correlations between the series are ignored.

4.5.2 Modelling the unconditional variances

We shall now construct an adequate parametric model for the unconditional variance of each of the seven return series as discussed in Section 4.4.1. First we shall test the hypothesis of constant unconditional variance against a smoothly time-varying unconditional variance with one transition in g_{it}. We choose $\alpha^{(1)} = 0.05$. The test results are reported in the second column of Table 4.3. As already mentioned, they are quite insensitive to this choice of significance level. The null hypothesis is rejected very strongly for each of the seven return series as the largest p-value equals 3×10^{-10}. The tests of the last three columns correspond to a sequence of nested tests based on (4.19) for choosing k in (4.9). If H_{0i} is rejected most strongly, measured by the p-value, of the three hypotheses, one selects $k = i$. For details, see Amado and Teräsvirta (2008). Table 4.3 shows that $k = 1$ in (4.9) for all seven series. After fitting the TV-GARCH model with one transition function and setting $k = 1$, the estimated model is tested against a double transition model. The results indicate that the null hypothesis is rejected at $\alpha^{(2)} = 0.025$ in three out of the seven return series. For two of these three price series, $k = 1$ appears to be the right alternative, whereas for INTC $k = 2$ is the appropriate choice. The p-values for testing for another transition for the three series appear at the bottom of Table 4.3. Now $\alpha^{(3)} = 0.0125$, but the decision not to reject the null hypothesis could be made at all conventional significance levels. Consequently, the TV-GARCH model with two transitions is tentatively selected as the final model for the AXP, F and INTC.

Table 4.4 contains the final estimates for the functions g_{it} from the TVGJR-GARCH(1,1) models. Plots of the time-varying unconditional variances appear in Figure 4.2. With the exception of INTC, the volatility has increased around mid-1997 in most of the series, which coincides with the beginning of the East Asian financial crisis. The level of volatility has then remained high, except for the AXP and INTC returns. For these two stocks, the general level of volatility has decreased towards the end of the observation period. It is interesting to note that the transition between the extreme volatility regimes is quite rapid for all series. The maximum value of the slope parameter γ has been set to 100 to save computing time. It should be noted that the standard error estimates reported in Table 4.4 have been computed conditionally on $\gamma = 100$. Descriptive statistics of the standardized returns ε_{it}^{*}(or $\varepsilon_{it}/\widehat{g}_{it}^{1/2}$) can be found in 4.2.

For comparison, we have also fitted the stationary GJR-GARCH(1,1) model to our return series, and the results can be found in Table 4.6. The stationarity condition for this model is $\alpha_1+\kappa_1/2+\beta_1 < 1$. In all cases the estimated models show high persistence as $\widehat{\alpha}_1 + \widehat{\kappa}_1/2 + \widehat{\beta}_1$ is very close to one. A look at the sample autocorrelation functions of $|\varepsilon_{it}|$ plotted in Figure 4.4 leads to the same conclusion. The autocorrelations decay at a rate that appears clearly slower than the exponential rate.

Table 4.5 contains the results (other than the ones involving $\widehat{g}_{it}$) from fitting a TV-GJR-GARCH(1,1) model to the series. The persistence measured by $\widehat{\alpha}_1 + \widehat{\kappa}_1/2 + \widehat{\beta}_1$ is in all seven cases lower than indicated by the GJR-GARCH(1,1) model. In two occasions, remarkably low values, 0.740 for BA and 0.907 for INTC, are obtained. For the remaining series the reduction in persistence is smaller but the values are still distinctly different from the corresponding ones in Table 4.6. The autocorrelations functions of $|\varepsilon_{it}^{*}|$ shown in Figure 4.6 are in line with these findings. The autocorrelations decay very quickly with the lag length and only a few first of them exceed the 95% confidence bounds drawn under the assumption that the errors are normal and independent. Moreover, the decay rate of the autocorrelogram appears close to exponential for all seven series. This is what we would expect after the unconditional variance component has absorbed the long-run movements in the series. These findings justify at an empirical level that the low level of persistence is exclusively due to the modelling of the changes in the unconditional variance.

Finally, Figure 4.3 shows the estimated conditional standard deviations obtained from the GJR-GARCH model. The behaviour of these series looks clearly nonstationary. The 'baseline volatility' clearly increases around 1998. The observed evidence of long-memory in Figure 4.4 accords with the behaviour of the series in Figure 4.3. The conditional standard deviations from the GJR model for ε_{it}^* can be found in Figure 4.5. These plots, in contrast to the ones in Figure 4.3, do not show signs of nonstationarity. The deterministic component g_{it} is able to handle the changing baseline volatility, and only volatility clustering is left to be parameterized by h_t.

In order to assess the validity of the GJR-GARCH model several LM misspecification tests were carried out (see Amado and Teräsvirta (2008) for details). The p-values are reported in Table 4.7. The hypothesis of no transition ($g_{it} \equiv 1$) is strongly rejected for the seven return series. Thus, even if the modelling had been begun by first fitting a GJR-GARCH(1,1) model to the data, the need for transitions would have been discovered at the evaluation stage. With three exceptions, the remaining diagnostic tests do not show signs of misspecification. The same misspecification tests applied to the TVGJR-GARCH model can be found in Table 4.8. The model passes the misspecification tests for the seven return series with a single exception. For F, it seems that the asymmetry in the effect of shocks is not satisfactorily described by the GJR-GARCH model as the test against the Smooth Transition GARCH has the p-value 0.015. This is not, however, a very small p-value, and no action is taken here. This means that the model is indeed able to capture the time-variation in the unconditional variance. The main conclusion is that the TVGJR-GARCH model adequately captures the most conspicuous features in our set of daily return series. We shall therefore retain these models for modelling the conditional correlations.

4.5.3 Effects of modelling the long-run dynamics of volatility on the conditional correlations

In this section, we shall investigate the effects of modelling nonstationary volatility equations on the conditional correlations. From equation (4.3) we can expect that ignoring the nonstationary component $\mathbf{G}_t$ may affect the correlation estimates, but the magnitude of the effect is not known. We consider three bivariate Conditional Correlation GARCH models, the CCC-, the DCC-, and the TVCC-GJR-GARCH(1,1) model. They were defined in Section 4.2.2. Two specifications will be estimated for each model. One is the first-order GJR-GARCH model that corresponds to $\mathbf{G}_t \equiv \mathbf{I}_2$, whereas the other one is the TVGJR-GARCH model for which $\mathbf{G}_t \neq \mathbf{I}_2$ in (4.3). As already discussed, the TVGJR-GARCH model can also account for slow movements in volatility, whereas the GJR-GARCH model is designed for only modelling volatility clustering.

The log-likelihood values of the estimated models for the 21 pairs of return series are reported in Table 4.9. The maximum values of the log-likelihood function across models are shown in boldface. Three remarks are in order. First, as may be expected, the CC-GARCH model with time-varying correlations outperforms the CCC-GARCH model for all pairs of assets. No formal test has been carried out, but this result suggests that the conditional correlations are time-varying. Second, the in-sample fit of the multivariate models when the univariate GARCH component is specified as a TVGJR-GARCH(1,1) model is vastly superior to the fit obtained by the standard GJR-GARCH(1,1) model. Third, the DCC-TVGJR-GARCH and TVCC-TVGJR-GARCH models provide the best in-sample fit to the bivariate data. Overall, the TVCC-TVGJR model fits the data best in 16 pairs out of 21.

In order to save space, the results for the estimated correlations for the CCC-GJR and CCC-TVGJR models are not shown. A general finding is that the values of the correlations from the CCC-TVGJR-GARCH model remain equal to the ones obtained from the CCC-GJR-GARCH model. This tells us that while careful modelling of the GARCH equations alone considerably improves the in-sample fit, the correlations remain unaffected.

The differences between the estimated conditional correlations obtained from the DCC-GJR-GARCH model and the DCC-TVGJR-GARCH model for pairs of asset returns are plotted in Figure 4.7. The effect of careful modelling of the individual GARCH equations on the estimated correlations is generally rather small. In some occasions the effect seems to be systematic such that the correlations decrease over time, but the magnitude of the change remains small. The AIG (insurance) and AXP (consumer finance) pair constitutes the only exception: the difference reaches 0.13 in the beginning of the observation period and lies around -0.08 at the end. A rather general conclusion is that in the DCC-GJR-GARCH model, the nonmodelled nonstationarity in the variances only has a small effect on time-varying correlations. Thus, if the focus on the analysis is on estimating time-varying correlations and the model is a DCC-GJR-GARCH model, the simpler GJR-GARCH model for the conditional variance may be preferred to the more sophisticated TVGJR-GARCH model. However, the results show that the magnitude of such effect varies across different stock series and according to the level of the correlations. Pairs in which at least one asset belongs to the consumer finance, banking and semiconductors with higher than average estimated correlations display a tendency for greater than average differences between the estimated correlations from the two models. On the other hand, pairs where at least one asset belongs to the telecommunication services are responsible for smallest differences. These pairs of assets also coincide with the ones that have rather low (but positive) correlations. To summarize, the empirical evidence suggests that the magnitude of the effect on the estimated correlations implied by the DCC-GJR-GARCH model when accounting for time-variation in the unconditional variances is driven by certain stock returns, in particular, those having higher values for the correlations.

The results are somewhat different when the DCC-GJR-CARCH model is replaced by the TVCC-GJR-GARCH model. The results from fitting a TVCC-GJR-GARCH model to all pairs of asset returns appear in Table 4.10. They include the estimated correlations and the estimates of the parameters for the smooth transition function (4.14). For the majority of the estimated models the estimate of the slope transition parameter γ attains its upper bound of 500. For these cases, the transition function is close to a step function. The estimates of the location parameter c lie in the range $0.7-0.9$ with the exception of the pair AXP-JPM. This range roughly corresponds to the years $2000-2002$. For all pairs of assets, the estimated conditional correlations increase over time. This is in agreement with the rather frequent observation that correlations between stock returns increase with the degree of market turbulence. Note, however, that the increase in volatility seen in Figure 4.1 is not immediately reflected in the correlations that increase later than the volatilities.

The estimated results for the TVCC-TVGJR-GARCH model are shown in Table 4.11. A comparison of Tables 4.10 and 4.11 show some differences between the results from TVCC-GJR-GARCH and TVCC-TVGJR-GARCH models. In a number of cases, the changes from the low correlation regime to the high correlation one often becomes smoother than it was estimated in the TVCC-GJR-GARCH model. This happens in 13 cases out of 21, whereas the opposite occurs only twice. The two sets of correlations over time are graphed in Figure 4.8. The largest changes in conditional correlations involve the returns of Ford (F). For the pairs AIG-F and AXP-F the location $\widehat{c}$ of the change in correlations changes to close to the one-third of the observations, and on the average correlations become smaller. A more modest but still distinct change in location is also observed for the INTC-F pair. Furthermore, the change becomes clearly smoother in the JPM-F pair than it is when estimated from a TVCC-GJR-GARCH model.

In Figures 4.4 and 4.6, we showed the autocorrelations of $|\varepsilon_t|$ and $|\widehat{\varepsilon}_t^*|$ for all seven return series. Somewhat analogously, we shall compare moving correlations between ε_{it} and ε_{jt} on one hand and $\widehat{\varepsilon}_{it}^*$ and $\widehat{\varepsilon}_{jt}^*$ on the other. Figure 4.9 contains the pairwise correlations of the between the former (grey dotted curve) and the latter (black solid curve) computed over 100 trading days. Modelling the nonstationary component in ε_t has a strong effect on the correlations. Typically, the conditional correlations between $\widehat{\varepsilon}_{it}^*$ and $\widehat{\varepsilon}_{jt}^*$ are generally lower towards the end of the sample than the corresponding correlations between ε_{it} and ε_{jt} and look 'stationary' overall. The increase observed in the latter cannot be seen in the former. It appears that the observed increase over time in correlations between the raw returns is due to the systematic increased in volatility towards the end of the period. This increase is left unmodelled in TVCC-GJR-GARCH models.

4.5.4 Time-varying news impact surfaces

Next, we shall consider the impact of unexpected shocks to the asset returns on the estimated covariances. This is done by employing a generalization of the univariate *news impact curve* of Engle and Ng (1993) to the multivariate case introduced by Kroner and Ng (1998). The so-called *news impact surface* is the plot of the conditional covariance against a pair of lagged shocks, holding the past conditional covariances constant at their unconditional sample mean levels. The news impact surfaces of the multivariate correlation models with the volatility equations modelled as TVGJR-GARCH models are time-varying because they depend on the component g_{it-1}. Therefore, these will be called as *time-varying news impact surfaces.* The time-varying news impact surface for h_{ijt} is the three dimensional graph of the function

$$h_{ijt} = f(\varepsilon_{it-1}, \varepsilon_{jt-1}, g_{it-1}; \mathbf{h}_{t-1})$$

where $\mathbf{h}_{t-1}$ is a vector of conditional covariances at time $t-1$ defined at their unconditional sample means. As an example, Figure 4.10 contains the time-varying news impact surface for the covariance generated by the CCC-TVGJR-GARCH model for the pair AIG-AXP. The choice of this single pair of assets is merely illustrative, but the same shapes of the surfaces can be found for other pairs as well. From the figure we see how the surface can vary over time due to the nonstationary component g_{it-1}. We are able to distinguish different reaction levels of covariance estimates to past shocks during tranquil and turbulent times. It shows that the response to the news of a given size on the estimated covariances is clearly stronger during periods of calm in the market than it is during periods of high turbulence. According to the results, when calm prevails a minor piece of 'bad news' (unexpected negative shock) is rather big news compared to a big piece of 'good news' (unexpected positive shock) during turbulent periods. This is seen from the asymmetric bowl-shaped impact surface.

Figure 4.11 contains the time-varying news impact surfaces under low and high volatility from the CCC-TVGJR-GARCH model for the conditional variance of AIG when there is no shock to AXP. Figure 4.12 contains a similar graph for AXP when there is no shock to AIG. The asymmetric shape shows that a negative return shock has a greater impact than a positive return shock of the same size. Furthermore, as is already obvious from Figure 4.10, a piece of news of a given size has a stronger effect on the conditional variance when volatility is low than when it is high.

Estimated news impact surfaces from the TVCC-TVGJR-GARCH model, allowing for time-variation in correlations, are plotted in Figure 4.13. These news impact surfaces are able to distinguish between responses low and high variance as well as low and high correlation levels. It is seen from Figure 4.13 that not only the degree of turbulence in the market but also the level of the correlations affect the impact of past shocks on the covariances. This means that both factors play an important role in assessing the effect of shocks on the covariances according to the TVCC-TVGJR-GARCH model. It is evident from the figure that high covariance estimates are related to strong correlations and a high degree of turbulence in the market.

4.6 Conclusions

In this paper, we extend the univariate multiplicative TV-GARCH model of Amado and Teräsvirta (2008) to the CC-GARCH framework. The model allows the individual unconditional volatilities to vary smoothly over time according to the logistic transition function. We also develop a modelling technique for specifying the parametric structure of the deterministic time-varying component that involves a sequence of Lagrange multiplier-type tests. In this respect, our model differs from the semiparametric model of Hafner and Linton (2008).

We consider a set of CC-GJR-GARCH models to investigate the effects of nonstationary variance equations on the conditional correlation matrix. The models are applied to pairs of seven daily stock returns belonging to the S&P 500 composite index. We find that in our examples, modelling the time-variation of the unconditional variances considerably improves the fit of the CC-GJR-GARCH models. The results show that multivariate correlation models combining both time-varying correlations and time-varying unconditional variances provide the best in-sample fit. They also indicate that modelling the nonstationary component in the variance has relatively little effect on correlation estimates when the conditional correlation model is the DCC-GJR-GARCH model, whereas the results are different for the STCC-GJR-GARCH model of Silvennoinen and Teräsvirta (2005, 2007). In a number of occasions, the correlations estimated from this model with time as the sole transition variable (TVCC-GJR-GARCH) are quite different from what they are when the GJR-GARCH equations are implicitly assumed stationary. The most conspicuous difference is that the time-varying correlations estimated from the TVCC-TVGJR-GARCH model are often smoother than the ones obtained from the TVCC-GJR-GARCH model.

With the TVGJR-GARCH equations we are also able to consider the effect of the nonstationary variance component on the moving correlations. For many pairs of returns, the fact that correlations between raw returns increase over time can be attributed to increasing volatility. This conclusion is based on the observation that the same moving correlations computed from returns with constant unconditional variance do not increase over time.

The TVGJR-GARCH approach also gives us the opportunity to generalize the news impact surfaces introduced by Kroner and Ng (1998) such that they can vary over time. In the TVCC-TVGJR-GARCH model, the impact of news (shocks) on the covariances between returns is a function of both time-varying variances and time-varying correlations. As in the univariate case already considered in Amado and Teräsvirta (2008), it is seen that the impact of a piece of news of a given size is larger when the market is calm than when it is when during periods of high volatility. In the present multivariate case we can also conclude that high conditional correlation between to returns adds to the impact as compared to the situation in which the correlation is low. We also reproduce the old result that negative shocks or news have a stronger effect on volatility than positive news of the same size.

An extension of this methodology to the case in which the conditional correlations are also controlled by a stochastic variable is available through the Double STCC-GJR-GARCH model. This makes it possible to model for example asymmetric responses of conditional correlations to functions of past returns. This, however, is a topic left for future research.

Appendix A: Tables

Table 4.1 Descriptive statistics of the asset returns

Asset	Min	Max	Mean	Std.dev.	Skew	Ex.Kurt	Rob.Sk.	Rob.Kr.
AIG	-9.419	10.44	0.054	1.871	0.192	2.418	0.036	0.107
AXP	-14.63	12.02	0.064	2.259	-0.056	2.439	0.046	0.111
BA	-19.38	11.01	0.029	2.165	-0.608	7.332	0.005	0.106
F	-15.88	14.62	0.016	2.276	0.120	3.879	4×10^{-4}	0.095
INTC	-24.87	18.32	0.072	3.041	-0.395	4.890	0.045	0.010
JPM	-19.97	14.86	0.046	2.404	0.117	4.549	-0.006	0.215
T	-13.54	10.64	0.013	2.053	-0.066	2.598	0.031	0.161

Notes: The table contains summary statistics for the seven stock returns of the S&P 500 composite index. The sample period is from October 1, 1993 until September 30, 2003 (2518 observations). Rob.Sk. denotes the robust measure for skewness based on quantiles proposed by Bowley and the Rob.Kr. denotes the robust centred coefficient for kurtosis proposed by Moors; see Kim and White (2004) for details.

Table 4.2 Descriptive statistics of the standardized returns

Asset	Min	Max	Mean	Std.dev.	Skew	Ex.Kurt	Rob.Sk.	Rob.Kr.
AIG	-6.281	6.964	0.045	1.365	0.184	1.494	0.055	0.063
AXP	-10.07	8.664	0.062	1.778	0.052	1.689	0.060	0.063
BA	-12.33	7.964	0.034	1.568	-0.318	4.874	0.005	0.122
F	-9.761	8.605	0.024	1.688	0.141	1.695	0.015	0.097
INTC	-12.84	9.385	0.062	1.937	-0.296	2.544	0.073	-0.063
JPM	-12.89	10.07	0.047	1.781	0.125	2.748	-0.005	0.166
T	-8.132	6.894	0.019	1.446	0.015	1.392	0.024	0.195

Notes: The table contains summary statistics for the standardized returns of the seven stocks of the S&P 500 composite index. The standardized returns are obtained dividing the raw returns by the estimate of the function g_t. The sample period is from October 1, 1993 until September 30, 2003 (2518 observations). Rob.Sk. denotes the robust measure for skewness based on quantiles proposed by Bowley and the Rob.Kr. denotes the robust centred coefficient for kurtosis proposed by Moors; see Kim and White (2004) for details.

Table 4.3 Sequence of tests of constant unconditional variance against a time-varying GARCH model with multiplicative structure

Transitions in the alternative model	H_0	H_{03}	H_{02}	H_{01}
Single transition				
AIG	1×10^{-21}	0.0120	0.0044	1×10^{-20}
AXP	6×10^{-25}	2×10^{-10}	1×10^{-5}	3×10^{-14}
BA	3×10^{-10}	0.0019	0.0015	1×10^{-7}
F	1×10^{-20}	0.0359	0.0491	8×10^{-21}
INTC	2×10^{-20}	2×10^{-8}	0.0627	6×10^{-15}
JPM	1×10^{-19}	5×10^{-5}	0.0513	2×10^{-17}
T	2×10^{-25}	8×10^{-7}	0.3812	3×10^{-22}
Double transition				
AIG	0.1127	0.0418	0.5284	0.2297
AXP	4×10^{-4}	0.2735	0.0073	0.0016
BA	0.2419	0.2362	0.3391	0.1712
F	0.0084	0.3381	0.3607	0.0016
INTC	3×10^{-5}	0.8513	1×10^{-5}	0.0277
JPM	0.1421	0.9199	0.0352	0.3175
T	0.0939	0.0174	0.7878	0.4116
Triple transition				
AIG	–	–	–	–
AXP	0.2405	0.1536	0.3026	0.2932
BA	–	–	–	–
F	0.3205	0.1264	0.8555	0.2869
INTC	0.0433	0.2563	0.0210	0.2170
JPM	–	–	–	–
T	–	–	–	–

Notes: The entries are the p-values of the LM-type tests of constant unconditional variance applied to the seven stock returns of the S&P 500 composite index. The appropriate order k in (4.9) is chosen from the short sequence of hypothesis as follows: If the smallest p-value of the test corresponds to H_{02}, then choose $k = 2$. If either H_{01} or H_{03} are rejected more strongly than H_{02}, then select either $k = 1$ or $k = 3$. See Amado and Teräsvirta (2008) for further details.

Table 4.4 Estimation results for the univariate TV-GJR-GARCH models

Asset	$\widehat{\delta}_1$	$\widehat{\gamma}_1$	$\widehat{c}_{11}$	$\widehat{\delta}_2$	$\widehat{\gamma}_2$	$\widehat{c}_{21}$	$\widehat{c}_{22}$	r
				g_t component				
AIG	1.2550 (0.0803)	100 (–)	0.3709 (0.0028)	—	—	—	—	1
AXP	1.4280 (0.1218)	100 (–)	0.4815 (0.0025)	-0.8940 (0.1568)	100 (–)	0.8022 (0.0032)	—	2
BA	1.4708 (0.0902)	100 (–)	0.4041 (0.0027)	—	—	—	—	1
F	1.0413 (0.0979)	100 (–)	0.4426 (0.0030)	0.8489 (0.2003)	100 (–)	0.7872 (0.0038)	—	2
INTC	0.2720 (0.0310)	100 (–)	0.1650 (0.0020)	-0.6862 (0.0209)	100 (–)	0.6180 (0.0106)	0.9240 (0.0118)	2
JPM	1.3984 (0.0938)	100 (–)	0.4814 (0.0026)	—	—	—	—	1
T	1.7708 (0.1048)	100 (–)	0.4442 (0.0023)	—	—	—	—	1

Notes: The table contains the parameter estimates of the g_{it} component from the TV-GJR-GARCH(1,1) model for the seven stocks of the S&P 500 composite index, over the period October 1, 1993 - September 30, 2003. The estimated model has the form $g_{it} = 1 + \sum_{l=1}^{r} \delta_{il} G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$, where $G_{il}(t/T; \gamma_{il}, \mathbf{c}_{il})$ is defined in (4.9) for all i. The numbers in parentheses are the standard errors.

Table 4.5 Estimation results for the univariate TV-GJR-GARCH models

Asset	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\kappa}_1$	$\widehat{\beta}_1$	$\widehat{\alpha}_1 + \frac{\widehat{\kappa}_1}{2} + \widehat{\beta}_1$
			h_t component		
AIG	0.0472	0.0273	0.0514	0.9235	0.9765
	(0.0164)	(0.0105)	(0.0162)	(0.0163)	
AXP	0.1777	0.0222	0.1183	0.8675	0.9488
	(0.0488)	(0.0118)	(0.0261)	(0.0242)	
BA	0.6447	0.0829	0.0743	0.6195	0.7396
	(0.1912)	(0.0381)	(0.0518)	(0.0927)	
F	0.0849	–	0.0547	0.9444	0.9717
	(0.0423)		(0.0139)	(0.0198)	
INTC	0.3594	–	0.1139	0.8502	0.9071
	(0.2094)		(0.0500)	(0.0760)	
JPM	0.1055	–	0.0905	0.9241	0.9693
	(0.0328)		(0.0220)	(0.0177)	
T	0.0865	0.0337	0.0617	0.8962	0.9608
	(0.0308)	(0.0119)	(0.0244)	(0.0264)	

Notes: The table contains the parameter estimates of the h_{it} component from the TV-GJR-GARCH(1,1) model for the seven stocks of the S&P 500 composite index, over the period October 1, 1993 - September 30, 2003. The estimated model has the form $h_{it} = \omega_i + \alpha_{i1}\varepsilon^{*2}_{it-1} + \kappa_{i1} I_{it-1}(\varepsilon^{*}_{it-1})\varepsilon^{*2}_{it-1} + \beta_{i1} h_{it-1}$, where $\varepsilon^{*}_{it} = \varepsilon_{it}/g_{it}^{1/2}$ and $I_{it}(\varepsilon^{*}_{it}) = 1$ if $\varepsilon^{*}_{it} < 0$ (and 0 otherwise) for all i. The numbers in parentheses are the Bollerslev-Wooldridge robust standard errors.

Table 4.6 Estimation results for the univariate GJR-GARCH models

Asset	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\kappa}_1$	$\widehat{\beta}_1$	$\widehat{\alpha}_1 + \frac{\widehat{\kappa}_1}{2} + \widehat{\beta}_1$
AIG	0.0396 (0.0145)	0.0312 (0.0105)	0.0535 (0.0163)	0.9325 (0.0133)	0.9905
AXP	0.1083 (0.0328)	0.0282 (0.0114)	0.1124 (0.0276)	0.8997 (0.0186)	0.9842
BA	0.0685 (0.0615)	0.0424 (0.0193)	0.0337 (0.0318)	0.9292 (0.0371)	0.9885
F	0.0402 (0.0191)	0.0160 (0.0095)	0.0415 (0.0152)	0.9564 (0.0118)	0.9932
INTC	0.2339 (0.1233)	0.0286 (0.0186)	0.0879 (0.0439)	0.9049 (0.0344)	0.9774
JPM	0.0562 (0.0208)	0.0099 (0.0071)	0.0790 (0.0222)	0.9431 (0.0127)	0.9925
T	0.0337 (0.0173)	0.0355 (0.0113)	0.0458 (0.0222)	0.9355 (0.0188)	0.9938

Notes: The table contains the parameter estimates from the GJR-GARCH(1,1) model for the seven stocks of the S&P 500 composite index, over the period October 1, 1993 - September 30, 2003. The estimated model has the form $h_{it} = \omega_i + \alpha_{i1}\varepsilon_{it-1}^2 + \kappa_{i1} I_{it-1}(\varepsilon_{it-1})\varepsilon_{it-1}^2 + \beta_{i1} h_{it-1}$, where $I_{it}(\varepsilon_{it}) = 1$ if $\varepsilon_{it} < 0$ (and 0 otherwise) for all i. The numbers in parentheses are the Bollerslev-Wooldridge robust standard errors.

Table 4.7 Misspecification tests for the GJR-GARCH models

(a) LM test of no ARCH in the standardized residuals							
Return	AIG	AXP	BA	F	INTC	JPM	T
$r = 1$	0.133	0.180	0.243	0.018	0.786	0.435	0.215
$r = 5$	0.568	0.670	0.314	0.170	0.822	0.776	0.700
$r = 10$	0.253	0.727	0.362	0.302	0.812	0.902	0.766
(b) LM test of GJR-GARCH(1,1) vs. GJR-GARCH(1,2) model							
Return	AIG	AXP	BA	F	INTC	JPM	T
	0.404	0.773	0.001	0.374	0.320	0.308	0.965
(c) LM test of GJR-GARCH(1,1) vs. GJR-GARCH(2,1) model							
Return	AIG	AXP	BA	F	INTC	JPM	T
	0.490	0.607	0.007	0.026	0.049	0.272	0.482
(d) LM type test of no additional transition in the function g_t							
Return	AIG	AXP	BA	F	INTC	JPM	T
	0.003	2×10^{-4}	0.008	0.001	10×10^{-5}	0.016	0.003
(e) LM type test of no ST-GJR-GARCH model of order 1							
Return	AIG	AXP	BA	F	INTC	JPM	T
	0.795	0.685	0.403	0.007	0.286	0.165	0.626

Notes: The entries are the p-values of the LM-type misspecification tests in Amado and Teräsvirta (2008). The diagnostic tests are the following: (a) test of no ARCH-in-GARCH against remaining ARCH of order r in the standardized residuals; (b) test of a GJR-GARCH(1,1) model against a GJR-GARCH(1,2) model; (c) test of a GJR-GARCH(1,1) model against a GJR-GARCH(2,1) model; (d) test of no additional transition against another transition function in g_t; (e) test of no remaining nonlinearity against a Smooth Transition GJR-GARCH (ST-GJR-GARCH) of order $k = 1$.

Table 4.8 Misspecification tests for the TV-GJR-GARCH models

(a) LM test of no ARCH in the standardized residuals

Return	AIG	AXP	BA	F	INTC	JPM	T
$r = 1$	0.112	0.077	0.907	0.104	0.993	0.429	0.390
$r = 5$	0.641	0.429	0.956	0.360	0.881	0.871	0.759
$r = 10$	0.392	0.570	0.725	0.340	0.557	0.960	0.749

(b) LM test of GJR-GARCH(1,1) vs. GJR-GARCH(1,2) model

Return	AIG	AXP	BA	F	INTC	JPM	T
	0.579	0.700	0.208	0.411	0.191	0.487	0.819

(c) LM test of GJR-GARCH(1,1) vs. GJR-GARCH(2,1) model

Return	AIG	AXP	BA	F	INTC	JPM	T
	0.878	0.943	0.741	0.419	0.071	0.702	0.557

(d) LM type test of no additional transition in the function g_t

Return	AIG	AXP	BA	F	INTC	JPM	T
	0.660	0.085	0.128	0.718	0.855	0.587	0.453

(e) LM type test of no ST-GJR-GARCH model of order 1

Return	AIG	AXP	BA	F	INTC	JPM	T
	0.630	0.265	0.123	0.015	0.442	0.170	0.846

Notes: The entries are the p-values of the LM-type misspecification tests in Amado and Teräsvirta (2008). The diagnostic tests are the following: (a) test of no ARCH-in-GARCH against remaining ARCH of order r in the standardized residuals; (b) test of a GJR-GARCH(1,1) model against a GJR-GARCH(1,2) model; (c) test of a GJR-GARCH(1,1) model against a GJR-GARCH(2,1) model; (d) test of no additional transition against another transition function in g_t; (e) test of no remaining nonlinearity against a Smooth Transition GJR-GARCH (ST-GJR-GARCH) of order $k = 1$.

Table 4.9 Log-likelihood values from the bivariate normal density for the CC-GJR-GARCH estimated models

Pairs of Assets	CCC		DCC		TVCC	
	GJR	TV-GJR	GJR	TV-GJR	GJR	TV-GJR
AIG−AXP	−9965.1	−8834.7	−9937.2	**−8798.7**	−9934.4	−8801.2
AIG−BA	−10176.6	−8805.8	−10166.8	−8795.9	−10162.2	**−8789.6**
AIG−F	−10256.1	−8999.0	−10250.2	**−8989.9**	−10249.6	−8991.8
AIG−INTC	−11007.0	−9366.1	−10991.0	−9348.7	−10991.7	**−9345.1**
AIG−JPM	−10084.6	−8851.0	−10050.6	**−8813.2**	−10058.6	−8821.3
AIG−T	−9947.4	−8556.8	−9941.6	−8552.4	−9938.6	**−8547.9**
AXP−BA	−10658.2	−9461.8	−10623.3	−9430.5	−10623.5	**−9429.8**
AXP−F	−10712.7	−9624.7	−10697.6	**−9606.7**	−10704.2	−9613.4
AXP−INTC	−11459.6	−9988.8	−11432.9	−9961.0	−11428.2	**−9953.9**
AXP−JPM	−10445.9	−9381.6	−10388.2	**−9315.7**	−10409.4	−9334.6
AXP−T	−10466.0	−9248.4	−10454.1	−9233.2	−10447.2	**−9228.3**
BA−F	−10741.8	−9422.1	−10734.3	−9415.0	−10726.1	**−9405.5**
BA−INTC	−11481.5	−9780.2	−11463.6	−9761.9	−11464.2	**−9756.7**
BA−JPM	−10769.5	−9470.4	−10752.4	−9455.9	−10750.1	**−9446.8**
BA−T	−10485.0	−9037.9	−10475.3	−9029.6	−10471.0	**−9022.6**
F−INTC	−11537.5	−9937.9	−11522.4	−9924.7	−11526.9	**−9924.5**
F−JPM	−10786.9	−9593.6	−10768.7	−9576.5	−10771.5	**−9572.8**
F−T	−10578.2	−9237.6	−10575.4	−9233.6	−10569.8	**−9226.9**
INTC−JPM	−11540.4	−9966.7	−11527.4	−9951.2	−11522.7	**−9941.8**
INTC−T	−11314.7	−9594.2	−11306.7	−9585.9	−11298.6	**−9575.5**
JPM−T	−10569.5	−9247.9	−10555.7	−9233.4	−10546.5	**−9224.3**

Notes: The table contains the log-likelihood values for each of the bivariate CC-GJR-GARCH model. The conditional variances are modelled as GJR-GARCH(1,1). The GJR column indicates that the unconditional variances are time-invariant functions. The TV-GJR column indicates that the unconditional variances vary over time according to function (4.8). The maximized values for the log-likelihood are shown in boldface.

Table 4.10 Estimation results for the bivariate TVCC-GJR-GARCH models

Pairs of assets	$\rho_{(1)}$	$\rho_{(2)}$	γ	c
AIG−AXP	0.4564 (0.0338)	0.7394 (0.0177)	500 (−)	0.88 (0.001)
AIG−BA	0.2193 (0.0379)	0.4907 (0.0454)	15.06 (66.91)	0.85 (0.09)
AIG−F	0.2647 (0.0211)	0.4490 (0.0604)	500 (−)	0.88 (0.001)
AIG−INTC	0.1992 (0.0274)	0.5000 (0.0667)	9.51 (6.05)	0.85 (0.06)
AIG−JPM	0.4422 (0.0216)	0.6948 (0.0244)	500 (−)	0.86 (0.000)
AIG−T	0.2906 (0.0219)	0.4785 (0.0397)	500 (−)	0.85 (0.001)
AXP−BA	0.1839 (0.0272)	0.4954 (0.0305)	139.98 (114.18)	0.71 (0.01)
AXP−F	0.2880 (0.0223)	0.4436 (0.0383)	500 (−)	0.75 (0.000)
AXP−INTC	0.2159 (0.0257)	0.5137 (0.0297)	44.61 (52.01)	0.74 (0.02)
AXP−JPM	0.3842 (0.0330)	0.6241 (0.0163)	500 (−)	0.34 (0.001)
AXP−T	0.2445 (0.0230)	0.5377 (0.0443)	32.43 (21.82)	0.86 (0.02)
BA−F	0.1789 (0.0241)	0.4136 (0.0419)	500 (−)	0.78 (0.002)
BA−INTC	0.1390 (0.0273)	0.3820 (0.0435)	500 (−)	0.75 (0.001)
BA−JPM	0.1984 (0.0250)	0.4445 (0.0426)	500 (−)	0.74 (0.001)
BA−T	0.1488 (0.0214)	0.4217 (0.0420)	33.64 (15.056)	0.85 (0.02)
F−INTC	0.2412 (0.0226)	0.4743 (0.0531)	500 (−)	0.88 (0.001)
F−JPM	0.3200 (0.0331)	0.5888 (0.0554)	15.56 (64.01)	0.87 (0.07)
F−T	0.1657 (0.0237)	0.3572 (0.0429)	500 (−)	0.82 (0.001)
INTC−JPM	0.2701 (0.0265)	0.4962 (0.0428)	500 (−)	0.78 (0.000)
INTC−T	0.1197 (0.0235)	0.3834 (0.0567)	159.43 (105.88)	0.82 (0.02)
JPM−T	0.2396 (0.0222)	0.5473 (0.0401)	500 (−)	0.86 (0.001)

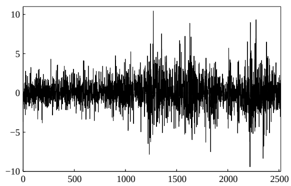
Notes: The table contains the estimation results for each of the bivariate TVCC-GJR-GARCH model. The conditional variances are modelled as GJR-GARCH(1,1) and the unconditional variances are time-invariant functions. The numbers in parentheses are the standard errors.

Table 4.11 Estimation results for the bivariate TVCC-TVGJR-GARCH models

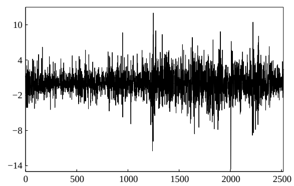
Pairs of assets	$\rho_{(1)}$	$\rho_{(2)}$	γ	c
AIG−AXP	0.4504	0.7459	500	0.88
	(0.0175)	(0.0224)	(−)	(0.001)
AIG−BA	0.2202	0.5390	6.54	0.84
	(0.0235)	(0.1174)	(4.09)	(0.070)
AIG−F	0.1824	0.3303	53.85	0.33
	(0.0341)	(0.0249)	(51.68)	(0.036)
AIG−INTC	0.1872	0.4554	144.64	0.75
	(0.0240)	(0.0443)	(37.18)	(0.004)
AIG−JPM	0.4355	0.7052	109.72	0.86
	(0.0218)	(0.0236)	(52.88)	(0.007)
AIG−T	0.2954	0.4786	500	0.84
	(0.0211)	(0.0387)	(−)	(0.001)
AXP−BA	0.1910	0.5069	29.57	0.73
	(0.0273)	(0.0373)	(21.12)	(0.025)
AXP−F	0.2033	0.3852	15.37	0.37
	(0.0392)	(0.0256)	(9.46)	(0.053)
AXP−INTC	0.2137	0.5327	40.15	0.74
	(0.0240)	(0.0306)	(19.92)	(0.011)
AXP−JPM	0.3705	0.6397	23.58	0.36
	(0.0338)	(0.0173)	(16.22)	(0.033)
AXP−T	0.2414	0.5413	25.43	0.85
	(0.0230)	(0.0626)	(29.86)	(0.032)
BA−F	0.1746	0.4172	500	0.75
	(0.0230)	(0.0413)	(−)	(0.001)
BA−INTC	0.1366	0.4342	144.19	0.76
	(0.0251)	(0.0436)	(38.35)	(0.005)
BA−JPM	0.1975	0.4753	500	0.74
	(0.0241)	(0.0409)	(−)	(0.001)
BA−T	0.1554	0.7707	6.62	0.94
	(0.0211)	(0.2668)	(2.06)	(0.043)
F−INTC	0.2258	0.4369	500	0.76
	(0.0234)	(0.0373)	(−)	(0.001)
F−JPM	0.3040	0.7842	4.56	0.91
	(0.0222)	(0.5354)	(4.05)	(0.174)
F−T	0.1623	0.3826	500	0.82
	(0.0232)	(0.0452)	(−)	(0.001)
INTC−JPM	0.2537	0.5291	24.54	0.74
	(0.0267)	(0.0468)	(11.48)	(0.019)
INTC−T	0.1183	0.4100	500	0.82
	(0.0228)	(0.0559)	(−)	(0.004)
JPM−T	0.2410	0.5462	500	0.84
	(0.0223)	(0.0377)	(−)	(0.001)

Notes: The table contains the estimation results for each of the bivariate TVCC-TVGJR-GARCH model. The conditional variances are modelled as GJR-GARCH(1,1) and the unconditional variances vary over time according to function (4.8). The numbers in parentheses are the standard errors.

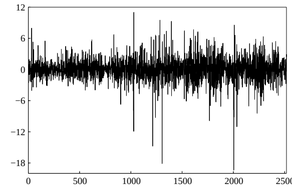
Appendix B: Figures



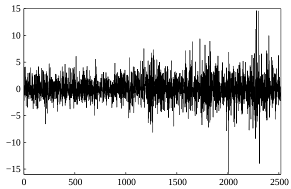
(a) AIG returns



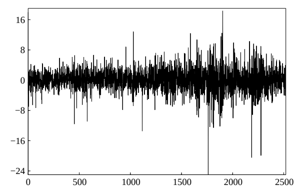
(b) AXP returns



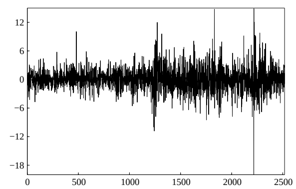
(c) BA returns



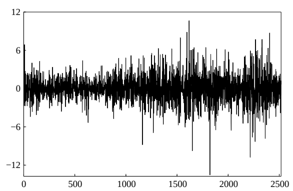
(d) F returns



(e) INTC returns

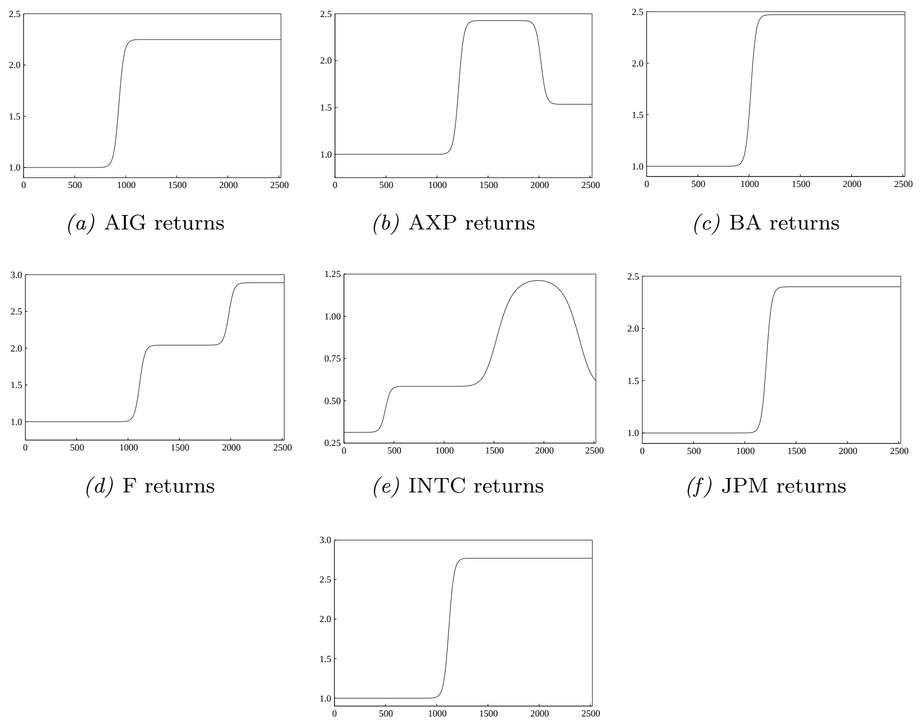


(f) JPM returns



(g) T returns

Figure 4.1 The seven stock returns of the S&P 500 composite index from October 1, 1993 until September 30, 2003 (2518 observations).

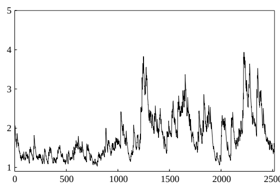


(a) AIG returns *(b)* AXP returns *(c)* BA returns

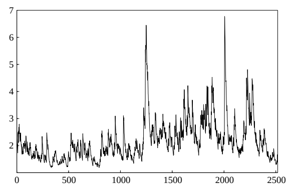
(d) F returns *(e)* INTC returns *(f)* JPM returns

(g) T returns

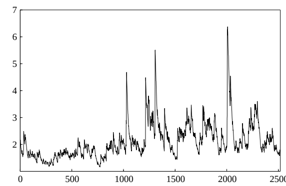
Figure 4.2 Estimated g_t functions for the seven stock returns of the S&P 500 composite index.



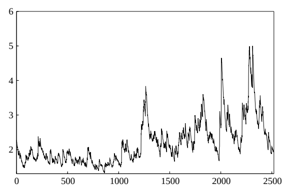
(a) AIG returns



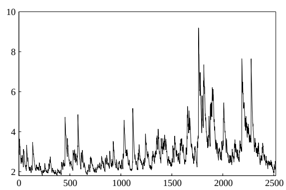
(b) AXP returns



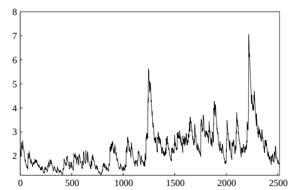
(c) BA returns



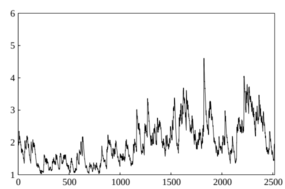
(d) F returns



(e) INTC returns

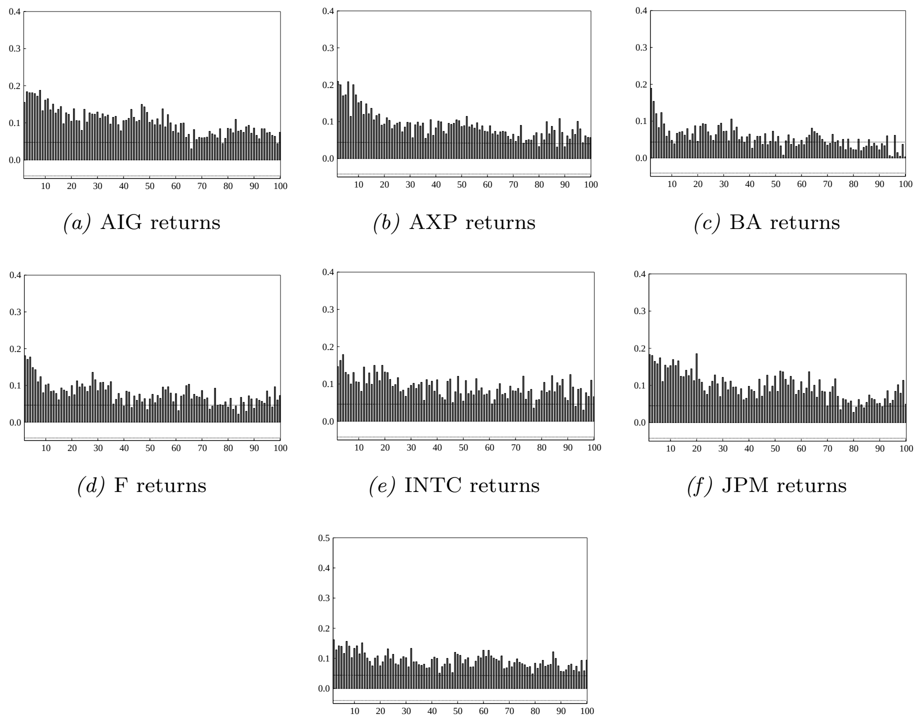


(f) JPM returns



(g) T returns

Figure 4.3 Estimated conditional standard deviations from the GJR(1,1) model for the seven stock returns of the S&P 500 composite index.

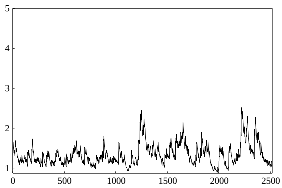


(a) AIG returns *(b)* AXP returns *(c)* BA returns

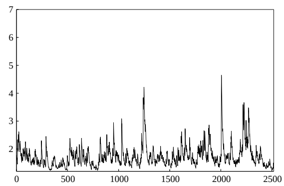
(d) F returns *(e)* INTC returns *(f)* JPM returns

(g) T returns

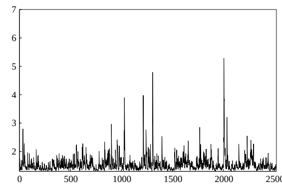
Figure 4.4 Sample autocorrelation functions of the absolute value for the seven stock returns of the S&P 500 composite index. The horizontal lines are the corresponding 95% confidence interval under the iid normality assumption.



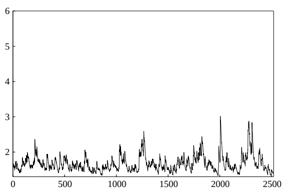
(a) AIG returns



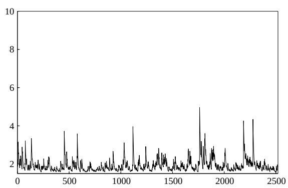
(b) AXP returns



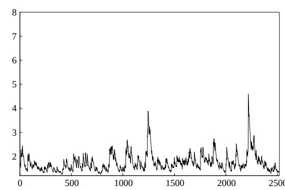
(c) BA returns



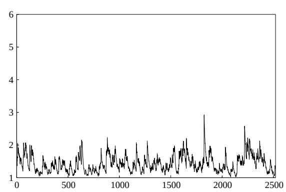
(d) F returns



(e) INTC returns

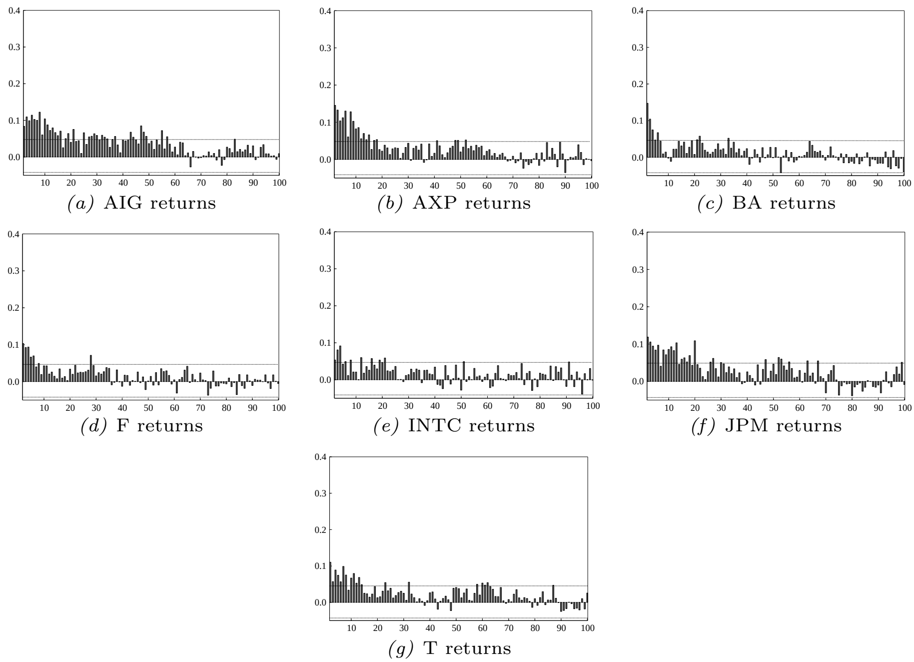


(f) JPM returns



(g) T returns

Figure 4.5 Estimated conditional standard deviations from the GJR(1,1) model for the standardized variable $\varepsilon_t/\hat{g}_t^{1/2}$ for the seven stock returns of the S&P 500 composite index.



(a) AIG returns *(b)* AXP returns *(c)* BA returns *(d)* F returns *(e)* INTC returns *(f)* JPM returns *(g)* T returns

Figure 4.6 Sample autocorrelation functions of the absolute value of the standardized variable $\varepsilon_t/\hat{g}_t^{1/2}$ for the seven stock returns of the S&P 500 composite index. The horizontal lines are the corresponding 95% confidence interval under the iid normality assumption.

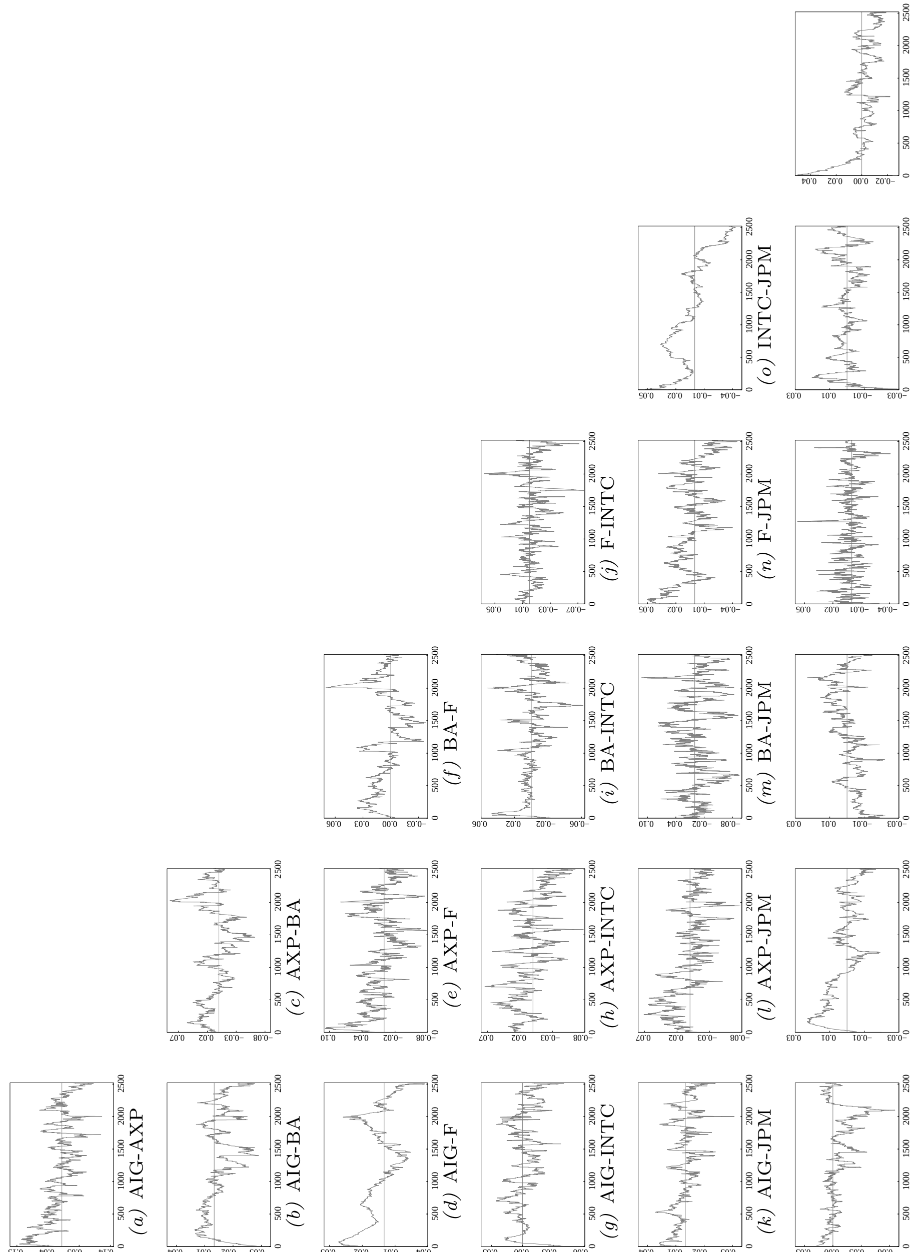


Figure 4.7 Difference between the estimated conditional correlations obtained from the bivariate DCC-GJR-GARCH and the bivariate DCC-TVGJR-GARCH models for the asset returns.

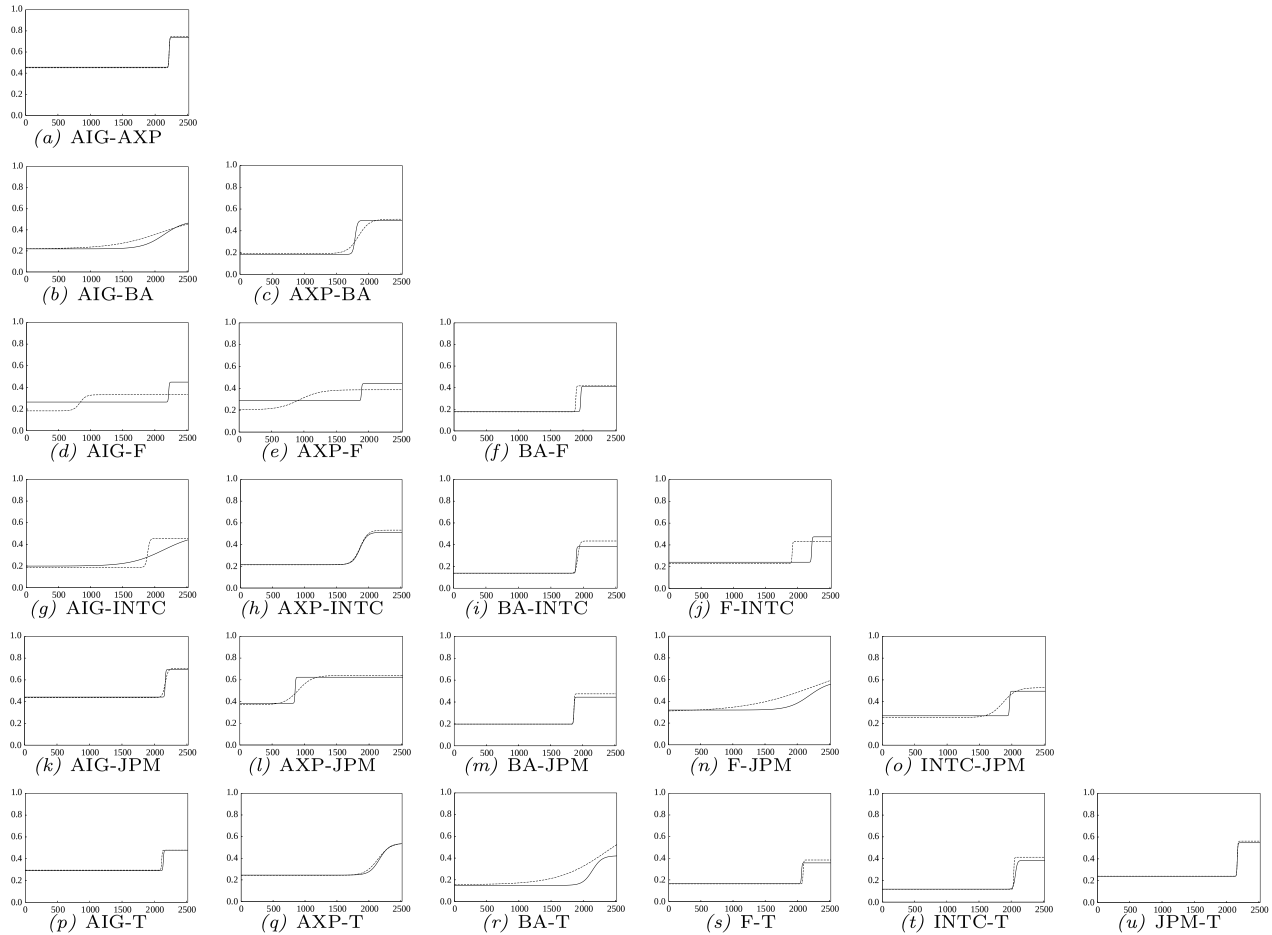


Figure 4.8 The estimated conditional correlations obtained from the bivariate TVCC-GJR-GARCH (solid curve) and the bivariate TVCC-TVGJR-GARCH (dotted curve) models for the asset returns.

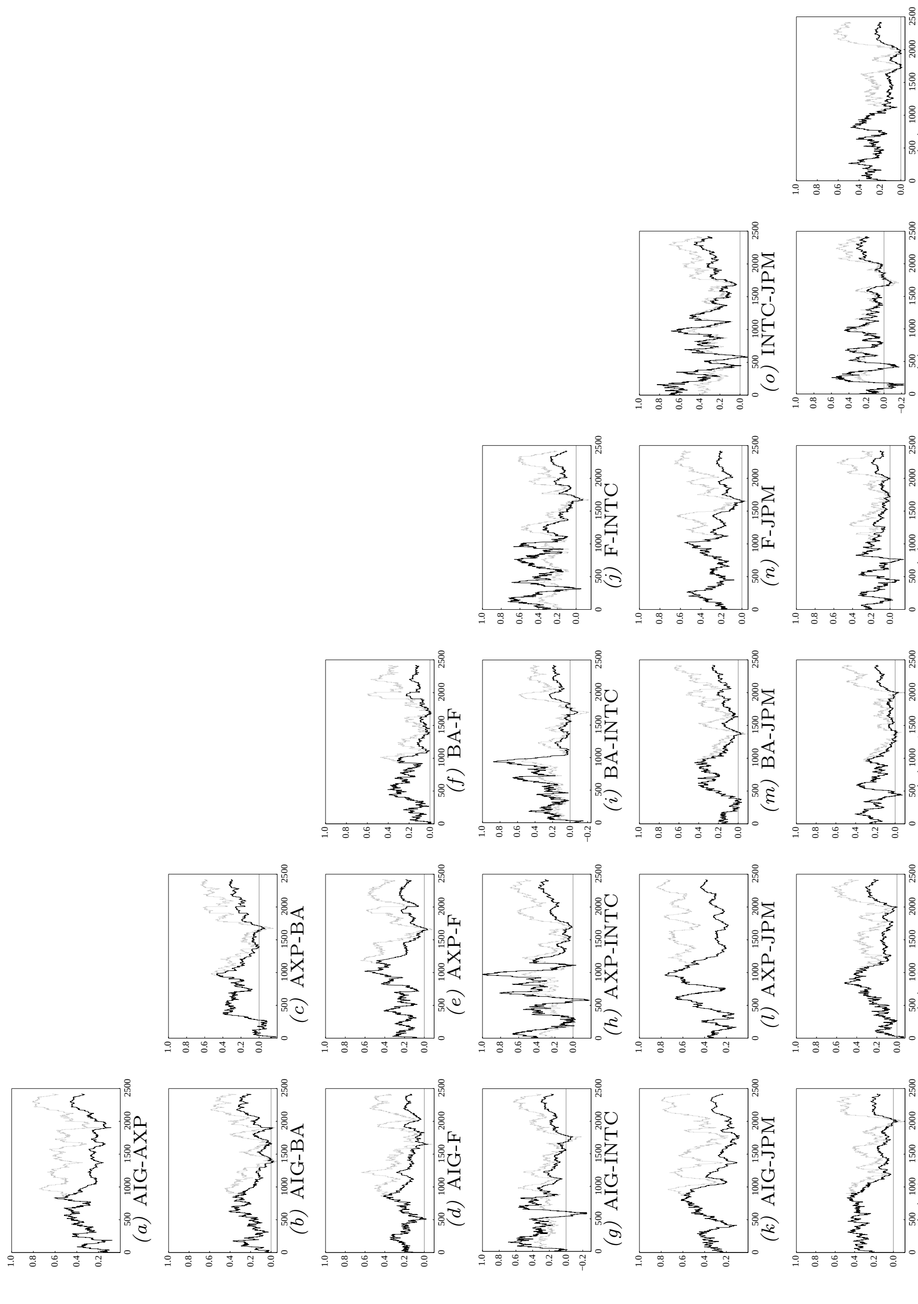
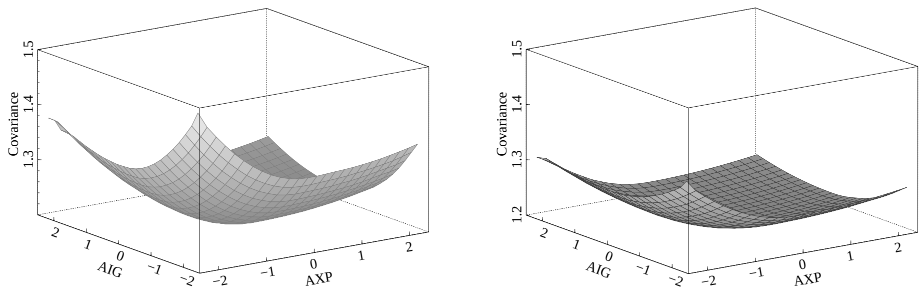
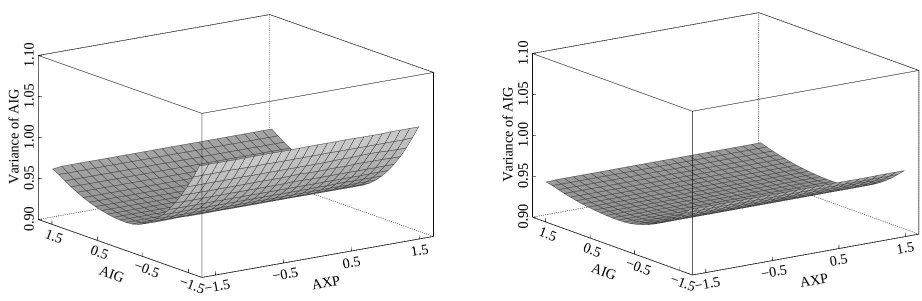


Figure 4.9 Difference between the estimated rolling correlation coefficients for pairs of the raw returns (grey dotted curve) and pairs of the standardized returns (black solid curve).



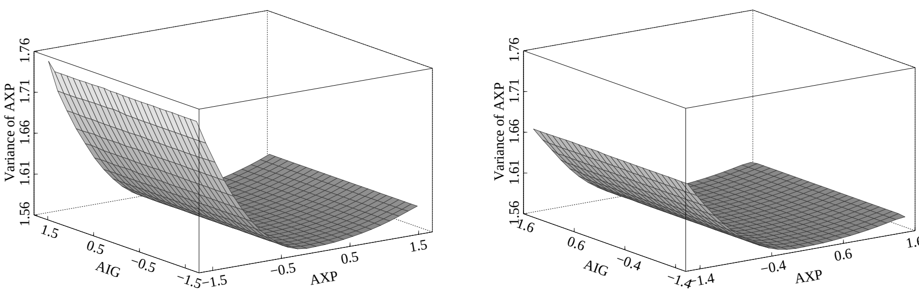
(a) lower regime *(b)* upper regime

Figure 4.10 Estimated estimated time-varying news impact surfaces for the covariance between the AXP and AIG returns under the CCC-TVGJR-GARCH model in the (a) lower regime and in the (b) upper regime of volatility.



(a) lower regime *(b)* upper regime

Figure 4.11 Estimated time-varying news impact surfaces for the conditional variance of the AIG returns under the CCC-TVGJR-GARCH model in the (a) lower regime and (b) upper regime of volatility.



(a) lower regime   *(b)* upper regime

Figure 4.12 Estimated time-varying news impact surfaces for the conditional variance of the AXP returns under the CCC-TVGJR-GARCH model in the (a) lower regime and (b) upper regime of volatility.

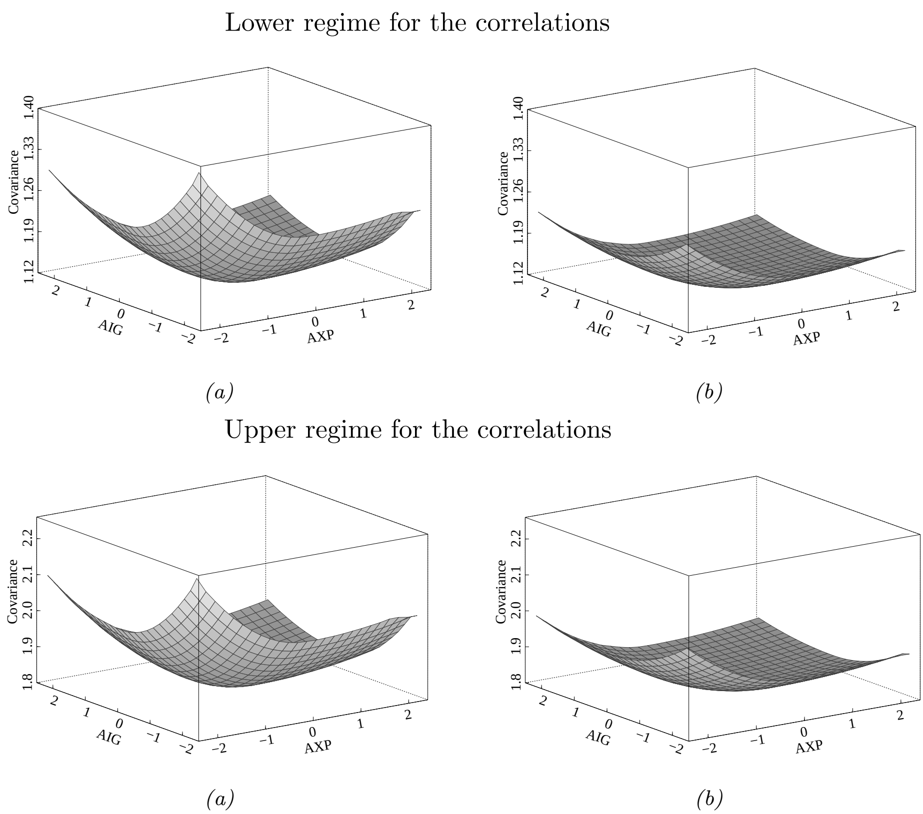


Figure 4.13 Estimated time-varying news impact surfaces for the covariance between the AXP and AIG returns under the TVCC-TVGJR model in the (a) lower regime and in the (b) upper regime of volatility.

Bibliography

Amado, C., and T. Teräsvirta (2008): "Modelling Unconditional and Conditional Heteroskedasticity with Smoothly Time-Varying Structure," SSE/EFI Working Paper Series in Economics and Finance No. 691.

Bauwens, L., S. Laurent, and J. Rombouts (2006): "Multivariate GARCH Models: A Survey," *Journal of Applied Econometrics*, 21, 79–109.

Berben, R.-P., and W. J. Jansen (2005): "Comovements in International Equity Markets: A Sectorial View," *Journal of International Money and Finance*, 24, 832–857.

Bollerslev, T. (1990): "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *The Review of Economic and Statistics*, 72, 498–505.

Dahlhaus, R., and S. Subba Rao (2006): "Statistical Inference for Time-Varying ARCH Processes," *Annals of Statistics*, 34, 1075–1114.

Diebold, F. X. (1986): "Comment on "Modelling the Persistence of Conditional Variance", by R. Engle and T. Bollerslev," *Econometric Reviews*, 5, 51–56.

Doornik, J. A. (2002): *Object-Oriented Matrix Programming Using Ox.* Timberlake Consultants Press, see also www.nuff.ox.ac.uk/Users/Doornik, 3rd edn.

Engle, R. F. (2002): "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business and Economic Statistics*, 20, 339–350.

Engle, R. F., and J. Gonzalo Rangel (2008): "The Spline-GARCH Model for Low-Frequency Volatility and its Global Macroeconomic Causes," *Review of Financial Studies*, 21, 1187–1222.

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

Fan, Y., S. Pastorello, and E. Renault (2007): "Maximization by Parts in Extremum Estimation," Unpublished paper.

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

Hafner, C., and O. Linton (2008): "Efficient Estimation of a Multivariate Multiplicative Model," Unpublished paper.

Kim, T.-H., and H. White (2004): "On More Robust Estimation of Skewness and Kurtosis," *Finance Research Letters*, 1, 56–73.

Kroner, K. F., and V. K. Ng (1998): "Modeling Asymmetric Comovements of Asset Returns," *Review of Financial Studies*, 11, 817–844.

Lamoureux, C. G., and W. D. Lastrapes (1990): "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business* & *Economic Statistics*, 8, 225–234.

Ling, S., and M. McAleer (2003): "Asymptotic Theory for a New Vector ARMA-GARCH Model," *Econometric Theory*, 19, 280–310.

Lütkepohl, H. (1996): *Handbook of Matrices.* John Wiley & Sons, Chichester.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

Mikosch, T., and C. Stărică (2004): "Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects," *Review of Economics and Statistics*, 86, 378–390.

Silvennoinen, A., and T. Teräsvirta (2005): "Multivariate Autoregressive Conditional Heteroskedasticity with Smooth Transitions in Conditional Correlations," SSE/EFI Working Paper Series in Economics and Finance No. 577.

——— (2007): "Modelling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model," SSE/EFI Working Paper Series in Economics and Finance No. 652.

——— (2008): "Multivariate GARCH Models," in *Handbook of Financial Time Series*, ed. by T. G. Andersen, R. A. Davis, J.-P. Kreiss, and T. Mikosch. New York: Springer.

Tse, Y., and A. K. Tsui (2002): "A Multivariate GARCH Model with Time-Varying Correlations," *Journal of Business* & *Economic Statistics*, 20, 351–362.

Chapter 5

A Smooth Transition Approach to Modelling Diurnal Variation in Models of Autoregressive Conditional Duration

A Smooth Transition Approach to Modelling Diurnal Variation in Models of Autoregressive Conditional Duration[1]

Abstract

This paper introduces a new approach for adjusting the diurnal variation in the trade durations. The model considers that durations are multiplicatively decomposed into a deterministic time-of-day and a stochastic component. The parametric structure of the diurnal component allows the duration process to change smoothly over the time-of-day. In addition, a testing framework consisting of Lagrange multiplier tests is proposed for specifying the diurnal component. Our methodology is applied to the IBM transaction durations traded at the New York Stock Exchange.

[1]This paper is a joint work with Timo Teräsvirta.

Acknowledgements: This research has been supported by the Danish National Research Foundation. The responsibility for any errors and shortcomings in this article remains ours.

5.1 Introduction

The automated trading in financial markets and the development in computing power have made available intraday datasets containing recorded information of the transactions at the exchanges. Because transaction data arrive in irregularly spaced time intervals, standard econometric methods are no longer applicable. The so-called high-frequency financial duration models were first introduced by Engle and Russell (1998) for tackling this inherent feature of the transaction data. These authors developed the class of Autoregressive Conditional Duration (ACD) models, in which the time elapsed between two market events (or duration) is the object of modelling. Their model considers an autoregressive structure on the conditional expected durations whose dynamics resembles the GARCH process for modelling the conditional variances.

An important feature of financial durations is the evidence of a strong diurnal variation over the trading day. Several studies have documented that trading activity is usually more intensive (shorter durations) near the opening and the closing of the market, and less intensive (longer durations) around lunchtime. Therefore, prior to using the ACD model any daily periodicity should be removed from the financial durations. This observation has been first reported in Engle and Russell (1998) whose procedure is often used for taking into account the diurnal variation in the durations. Their approach consists of decomposing the durations into a deterministic component, that accounts for the diurnal variation, and a stochastic component for modelling the durations dynamics. A common practice is to parameterize the diurnal component according to a spline function; see Engle and Russell (1998) and Bauwens and Giot (2000). Some others methods have been considered in previous studies. McCulloch and Tsay (2001) suggested a smooth quadratic function, whereas Zhang, Russell, and Tsay (2001) have estimated the diurnal variation using the super smoother method of Friedman (1984). As an alternative, Rodriguez-Poo, Veredas, and Espasa (2008) suggested a joint estimation of the deterministic and stochastic components in which the diurnal variation is estimated nonparametrically.

In this paper we follow Engle and Russell (1998) and let the durations be multiplicatively decomposed into a deterministic and a stochastic component. We propose a new parameterization for the diurnal component in which the duration process is allowed to change smoothly over the time-of-day. In addition, we provide a testing framework for specifying the structure of the diurnal component by a sequence of Lagrange multiplier tests. The empirical results suggest that the diurnal variation may not always have the inverted U-shaped pattern for the trade durations as documented in earlier studies. For this reason, one should proceed with care in the modelling of the time-of-day component.

The outline of the paper is as follows. In Section 5.2 we briefly review the ACD model. In Section 5.3 we present our method for removing the diurnal variation. Section 5.4 introduces the testing strategy for specifying the diurnal variation component. Section 5.5 contains the application for the IBM trade durations. Section 5.6 concludes.

5.2 The ACD framework

Let t_i be the time (measured in seconds) at which the $i-$th trade occurs and the duration $x_i = t_i - t_{i-1}$ be the time interval between two consecutive transactions occurring at times t_i and t_{i-1}. Let $\mathcal{F}_{i-1}$ be the information set consisting of past durations available at time t_{i-1}. Following Engle and Russell (1998), the basic assumption of the ACD model is that the time dependence in the durations be captured by their conditional expectation such that

$$\frac{x_i}{\mathsf{E}(x_i|\mathcal{F}_{i-1})} \equiv \varepsilon_i \sim \text{i.i.d. } \mathfrak{D}(\boldsymbol{\varpi}) \tag{5.1}$$

where $\mathsf{E}(x_i|\mathcal{F}_{i-1}) = \psi_i(x_{i-1}, \ldots, x_1; \boldsymbol{\theta}_\psi)$ is the conditional mean duration on $\mathcal{F}_{i-1}$, and $\mathfrak{D}$ is a general distribution with positive support and parameter vector $\boldsymbol{\varpi}$. In the simplest ACD model, the durations are defined in terms of a multiplicative error term as

$$x_i = \psi_i \varepsilon_i, \tag{5.2}$$
$$\psi_i = \psi_i(x_{i-1}, \ldots, x_1; \boldsymbol{\theta}_\psi), \tag{5.3}$$

and

$$\varepsilon_i \sim \text{i.i.d. } \exp(1) \tag{5.4}$$

for $i = 1, \ldots, n$. Equations (5.2)-(5.4) define the Exponential ACD model. The expectation of the duration conditional on $\mathcal{F}_{i-1}$ is specified as

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}. \tag{5.5}$$

The condition $1 - \sum_{j=1}^{m} \alpha_j - \sum_{j=1}^{q} \beta_j < 1$ is necessary and sufficient for the existence of $\mathsf{E}(x_i)$. The parameter restrictions $\omega > 0, \alpha_j \geq 0, j = 1, \ldots, m$, and $\beta_j \geq 0, j = 1, \ldots, q$, are sufficient for the positivity of the conditional durations. For $m = q = 1$ they are also necessary. Bauwens and Giot (2000) proposed a class of logaritmic ACD models to ensure positiveness of the conditional durations without parametric constraints. The autoregressive structure in equation (5.5) allows the model to account for clustering of durations. Its dynamic structure resembles that of the GARCH model of Bollerslev (1986): it is a linear autoregressive process of past durations and conditional expectations. Consequently, many results and properties of the GARCH literature can be adapted to the ACD context.

The distribution for the errors of the ACD model (5.1)-(5.3) is not limited to the exponential density as defined in (5.4). Other choices of distribution are possible such as the Weibull, Burr, or generalized gamma distribution. Gourieroux, Monfort, and Trognon (1984) showed that the Quasi-Maximum Likelihood (QML) method produces consistent estimators of a correctly specified conditional mean model if and only if the QML is based on a distribution belonging to the linear exponential family, and this holds even when the density is misspecified. For this reason, the QML estimators based on the exponential distribution will be consistent regardless the true error distribution, whereas the QML estimation based on densities not belonging to the linear exponential family such as the Weibull, Burr, or generalized gamma distributions will not produce consistent estimators. As an alternative, Drost and Werker (2004) showed that the QML estimators are consistent based on the standard gamma distribution. The score function of this log-likelihood, however, is proportional to the one from the exponential density, and using either one of these two distributions thus yields identical estimators.

5.3 Adjusting diurnal variation with smooth transitions

It is well documented in the literature that trading market activity is subjected to systematic variations over the time of the day. This is mainly because of the institutional features of the exchanges. Trading activity is usually more intense at the beginning and at the end of the day than it is around lunchtime. The high trading activity after the opening of the market occurs because traders want to adjust their positions to the information accumulated before the opening of the exchange. The frequency of transactions is also high near the closing of the market as traders want to close their positions before the trading session ends. This leads to an inverted U-shaped pattern for the average intertrade durations over the trading day.

In order to account for this diurnal variation, the durations may be multiplicatively decomposed into a deterministic and a stochastic component. The deterministic component describes the intradaily pattern of the durations and the stochastic one represents the dynamics of the ACD-type model. The deterministic effect may be removed from the data prior to estimating the ACD model. Engle and Russell (1998) suggested to 'diurnally adjust' the duration series by

$$\tilde{x}_i = x_i / \phi(t_{i-1}; \boldsymbol{\theta}_\phi) \tag{5.6}$$

in which $\tilde{x}_i$ is the adjusted duration and $\phi(t_{i-1}; \boldsymbol{\theta}_\phi)$ is the deterministic time-of-day component. The expected duration equals

$$\mathsf{E}(x_i | \mathcal{F}_{i-1}) = \phi(t_{i-1}; \boldsymbol{\theta}_\phi) \psi_i(\tilde{x}_{i-1}, \ldots, \tilde{x}_1; \boldsymbol{\theta}_\psi) \tag{5.7}$$

The component ψ_i is interpreted as the expected proportion above or below the normal duration level at that time of day.

The usual practice is to 'diurnally adjust' the duration series by estimating the average durations using a linear or a cubic spline function conditioned on the time-of-day and then remove this diurnal component from the original durations. The diurnal pattern is estimated by averaging the durations over thirty minute intervals, and using a cubic spline over the course of the day to smooth the durations; see Bauwens and Giot (2000). The duration series is 'diurnally adjusted' by dividing the original durations by the estimated diurnal component as in (5.6). This is done separately for each trading day as the time-of-day component may vary according to the day of the week. Alternatively, the parameter vectors $\boldsymbol{\theta}_\psi$ and $\boldsymbol{\theta}_\phi$ may be estimated jointly by maximum likelihood. However, Engle and Russell (1998) pointed out that the two-step procedure and the joint estimation by ML yield almost the same results for large samples.

Alternative ways of parameterizing the time-of-day effect exist. McCulloch and Tsay (2001) have developed a linear regression method to remove the diurnal pattern using quadratic time functions and indicator variables. Furthermore, Rodriguez-Poo, Veredas, and Espasa (2008) suggested a semiparametric approach to the problem using a simple transformation of the Nadaraya–Watson estimator to remove diurnal variation from the process.

In this paper we propose to 'diurnally adjust' the durations by parameterizing the diurnal component as follows:

$$\phi(t_{i-1};\boldsymbol{\theta}_\phi) = \delta_0 + \sum_{l=1}^{r} \delta_l G_l(t_{i-1};\gamma_l,\mathbf{c}_l) \tag{5.8}$$

where $G_l(t_{i-1};\gamma_l,\mathbf{c}_l), l = 1,\ldots,r$, is a transition function bounded between zero and one. This implies that the duration process is assumed to change smoothly over the time-of-day. In this context, the transition variable t_{i-1} is the intraday time (measured in seconds from the beginning of the trading day) which is for convenience rescaled to run from zero to one. The transition function is the logistic function

$$G_l(t_{i-1};\gamma_l,\mathbf{c}_l) = \left(1+\exp\left\{-\gamma_l \prod_{j=1}^{k}(t_{i-1}-c_{lj})\right\}\right)^{-1}, \ l = 1,\ldots,r, \tag{5.9}$$

satisfying the identification restrictions $\gamma_l > 0, c_{l1} \le c_{l2} \le ... \le c_{lk}$. The slope parameter γ_l controls the degree of smoothness of the transition function: the larger γ_l, the faster the transition is between the extreme regimes. When $\gamma_l \longrightarrow \infty$, (5.9) becomes a step function, and the process switches instantaneously from one regime to the other at $t_{i-1} = c_{lj}$. Typical choices for the transition function in practice are $k = 1$ and $k = 2$. When $k = 2$ the model can describe the aforementioned situation in which trading activity is higher in the beginning and then at the end of the session than it is in the middle of the day. Note that when $\delta_1 = ... = \delta_r = 0$ in (5.8) there are no systematic changes in the durations during the day. In this case, the durations tend to be uniformly exponentially distributed around their "normal" level. This special case is included in our model.

5.4 Specification tests for the diurnal component

In our framework, the structure of the diurnal component is determined by a sequence of specification tests. These are based on statistical inference and they consist on two specification tests. The first one tests the hypothesis of no diurnal variation against durations that change smoothly over the time-of-day. The other tests are for testing whether yet another transition function is required in the definition of the diurnal component.

5.4.1 Testing for no diurnal variation

The test of no diurnal variation is an important tool for checking the presence of systematic changes over the time-of-day. The starting-point is that there is no systematic change over the time-of-day while the alternative is that durations are varying smoothly during the day. In order to consider this testing problem let $r = 1$ in (5.8) and assume that the conditional duration process $\psi_i = \omega$, i.e.

$$x_i = \psi_i \phi_i \varepsilon_i, \qquad \varepsilon_i \sim \text{i.i.d. } \exp(1) \tag{5.10}$$

$$\psi_i \phi_i = \omega\{1 + \delta_1 G_1(t_{i-1}; \gamma_1, \mathbf{c}_1)\} = \omega + \delta_1^* G_1(t_{i-1}; \gamma_1, \mathbf{c}_1). \tag{5.11}$$

where $\psi_i \equiv \psi_i(\tilde{x}_{i-1}, \ldots, \tilde{x}_1; \boldsymbol{\theta}_\psi), \phi_i \equiv \phi(t_{i-1}; \boldsymbol{\theta}_\phi)$ and $\delta_1^* = \omega\delta_1$. The null hypothesis of no diurnal variation is $\gamma_1 = 0$ and the alternative $\gamma_1 > 0$. The testing problem is nonstandard as δ_1^* and $\mathbf{c}_1$ are unidentified nuisance parameters when $\gamma_1 = 0$. Following Luukkonen, Saikkonen, and Teräsvirta (1988) we solve the identification problem by approximating $G_1(t_{i-1}; \gamma_1, \mathbf{c}_1)$ with its third-order Taylor expansion around $\gamma_1 = 0$. After reparameterizing, we obtain

$$\psi_i \phi_i = \omega + \sum_{k=1}^{3} \lambda_k t_{i-1}^k + R_3(t_{i-1}; \gamma_1, \mathbf{c}_1) \tag{5.12}$$

where $\lambda_j = \gamma_1 \widetilde{\delta_j^*}$ and $R_3(t/T; \gamma_{i1}, \mathbf{c}_{i1})$ is the remainder. Under $\mathrm{H}_0, R_3(t_{i-1}; \gamma_1, \mathbf{c}_1) \equiv 0$, so the asymptotic theory of the LM test statistic is not affected by this approximation. The null hypothesis of no diurnal variation becomes $\mathrm{H}_0^{'} : \lambda_1 = \lambda_2 = \lambda_3 = 0$. This hypothesis can be tested by an LM test as follows:

1. Estimate the conditional duration process under the assumption that $\psi_i = \omega$, and compute $\hat{u}_i = x_i/\hat{\omega} - 1, i = 1, \ldots, n$, and $SSR_0 = \sum_{i=1}^{n} \hat{u}_i^2$.

2. Regress $\hat{u}_i$ on $\hat{\mathbf{x}}_{1i}' = \hat{\omega}^{-1}$ and $\hat{\mathbf{x}}_{2i}' = (t_{i-1}, t_{i-1}^2, t_{i-1}^3)', i = 1, \ldots, n$, and compute SSR_1.

3. Then, under the null hypothesis and the assumption $\psi_i = \omega$, the test statistic

$$LM = n(SSR_0 - SSR_1)/SSR_0$$

has an asymptotic χ^2 distribution with three degrees of freedom.

We shall call our LM test statistic LM_k, where k indicates the order of the polynomial in the exponent of the transition function. The rejection of the null hypothesis raises the problem of choosing k. In order to select k, we carry out a short sequence of nested tests following Teräsvirta (1994) and Lin and Teräsvirta (1994). This is done as follows. If parameter constancy is rejected at the significance level $\alpha^{(1)}$, test the following sequence of hypotheses:

$$\begin{aligned}
\mathrm{H}_{03} &: \lambda_3 = 0, \\
\mathrm{H}_{02} &: \lambda_2 = 0 \mid \lambda_3 = 0, \\
\mathrm{H}_{01} &: \lambda_1 = \lambda_2 = \lambda_3 = 0 \mid \lambda_2 = \lambda_3 = 0,
\end{aligned}$$

in (5.12), by means of LM-type tests. If H_{0i} is rejected most strongly, measured by the p-value, of the three hypotheses, one selects $k = i$. Visual inspection of the series may sometimes also be helpful choosing k and can be used in parallel with the tests.

5.4.2 Testing the hypothesis of no additional transitions

After estimating the diurnal component with a single transition the next step is to investigate the possibility of remaining diurnal variation in the durations. In order to do that, the previous test must be generalized to the case where we test $r = 1$ against $r \geq 2$ in (5.8). To derive the test, consider the model

$$x_i = \psi_i \phi_i \varepsilon_i, \qquad \varepsilon_i \sim \text{i.i.d. } \exp(1) \tag{5.13}$$

$$\psi_i \phi_i = \omega + \delta_1^* G_1(t_{i-1}; \gamma_1, \mathbf{c}_1) + \delta_2^* G_2(t_{i-1}; \gamma_2, \mathbf{c}_2). \tag{5.14}$$

The hypothesis of no additional transition is $\gamma_2 = 0$. Again, the parameters δ_2^* and $\mathbf{c}_2$ are only identified under the alternative. The identification problem is solved as before, using a Taylor series approximation of $G_2(t_{i-1}; \gamma_2, \mathbf{c}_2)$ around $\gamma_2 = 0$. After rearranging terms we have

$$\psi_i \phi_i = \omega + \delta_1^* G_1(t_{i-1}; \gamma_1, \mathbf{c}_1) + \sum_{k=1}^{3} \theta_k t_{i-1}^k + R_3^*(t_{i-1}; \gamma_2, \mathbf{c}_2) \tag{5.15}$$

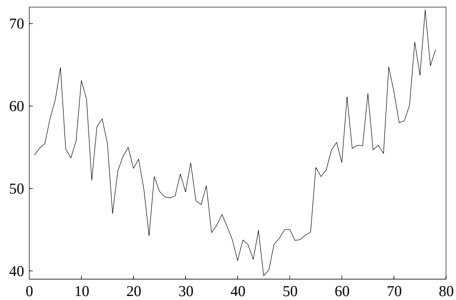
Under the null, the remainder $R_3^*(t_{i-1}; \gamma_2, \mathbf{c}_2) \equiv 0$, so it does not affect the asymptotic theory. The new null hypothesis becomes $H_{02}^{'} : \theta_1 = \theta_2 = \theta_3 = 0$. The significance level is now reduced by a factor $\tau \in (0, 1)$ in order to favour parsimony: $\alpha^{(2)} = \tau \alpha^{(1)}$. In the application, we set $\tau = 0.5$. Assuming $\psi_i = \omega$, this hypothesis can be tested using an LM test as before with $u_i = x_i / \psi_i \phi_i - 1$, $\mathbf{x}_{1i}' = \psi_i^{-1}(\phi_i + \psi_i \frac{\partial \phi_i}{\partial \delta})$, $\frac{\partial \phi_i}{\partial \delta} = (\frac{\partial \phi_i}{\partial \delta_1^*}, \frac{\partial \phi_i}{\partial \gamma_1}, \frac{\partial \phi_i}{\partial \mathbf{c}_1'})'$, and $\mathbf{x}_{2i}' = \phi_i^{-1}(t_{i-1}, t_{i-1}^2, t_{i-1}^3)'$, $i = 1, \ldots, n$.

5.5 An application to the IBM trade durations

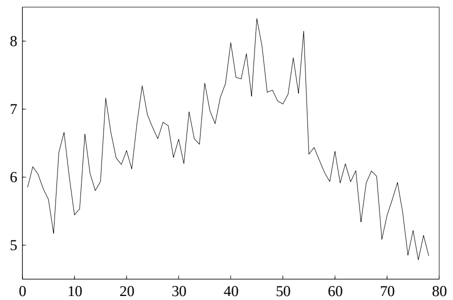
5.5.1 Data

The time series in this application consist of intertrade durations between transactions of IBM shares traded at the New York Stock Exchange (NYSE). The original series were extracted from the Trade and Quote (TAQ) database available from the NYSE, and the sample period covers the entire month of December 2002. Besides detailed information about volume, transaction prices, and bid and ask quotes at the time of the trade, the database contains a time stamp, measured in seconds after midnight and indicating the time when the transaction occurred.

Prior to using the data, we remove the irregular transactions using the correction indicator attached to each trade. Trades that occur before 9:30 AM and after 4:00 PM are also excluded. Because the market was partly closed on December 24th, this day has been removed from the data set. Furthermore, the overnight durations are ignored as the trades are treated consecutively from day to day. This leaves us with



(a) Transactions: 5-min time intervals



(b) Durations: 5-min time intervals

Figure 5.1 Plots of the IBM transactions data in December 2002: (a) average number of transactions in 5-minute time intervals; (b) average durations in 5-minute time intervals.

20 trading days with a total of 82011 transactions at 76823 unique times. Multiple transactions within a second are considered as a single trade. Thus the minimum time between events is one second, whereas the maximum observed duration turns out to be 142 seconds or 2 minutes and 22 seconds. The average duration between sucessive events is 6.09 seconds with a standard deviation of 6.22 seconds. Moreover, there is strong autocorrelation in the durations as the Ljung-Box statistic of serial correlation up to the 15th order equals $Q(15) = 1244$ with p-value $= 7 \times 10^{-256}$. For modelling such dependence, we follow Engle and Russell (1998) and use the linear ACD model. At this stage, where the focus is on modelling the diurnal variation, we consider the simplest version of the model. A better alternative such as a nonlinear ACD-type model would probably be needed to adequately capture the dynamic behaviour of the durations. Evidence of nonlinearity in the durations has been reported in Zhang, Russell, and Tsay (2001) and Meitz and Teräsvirta (2006), among others.

Figure 5.1(a) displays the average number of transactions within 5-minute time intervals over the 20 trading days. Each trading day contains 78 intervals. The plot exhibits a U-shaped pattern indicating active trading at the opening and the closing of the market and slower pace around the lunch hour. The average intraday durations in 5-minute time intervals plotted in Figure 5.1(b) reinforces this feature. The durations are usually shorter at the beginning and the end of the trading day, and longer around midday, which yields an inverted U-shaped in the durations. As the intraday variation is repeated systematically for every trading day, the ACF of the number of transactions in 5-minute intervals is characterized by the periodicity shown in Figure 5.2.

Our analysis differs from the usual practice in the sense that we consider every week in the sample separately. This way we can accommodate possible calendar effects over the time-of-week. However, a formal test is needed for testing the hypothesis of no systematic pattern over the days of the week. In case of rejection, the day-of-the week effect should be explicitly incorporated in the model. This problem is,

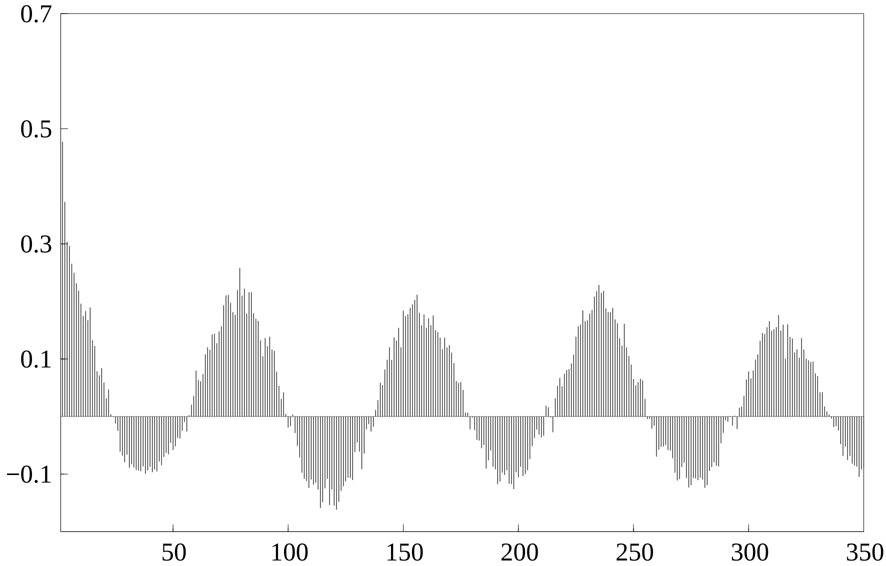


Figure 5.2 Sample autocorrelation function of the number of transactions in 5-minute time intervals for the IBM stock during the month of December 2002.

however, left for further research. Instead, we shall model the intraday pattern on complete five-day weeks to account for day-of-the-week effects. This leaves us with three complete weeks. Summary statistics of the unadjusted durations for December 2002 are presented in Table 5.1. The observed durations of the first week are plotted in Figure 5.3 (Appendix A). It shows the clustering effect in the durations, i.e. short (large) durations tend to be followed by durations of the same kind. The strong diurnal variation for each trading day is clearly visible as well.

5.5.2 Modelling smooth daily periodicities

Prior to modelling the durations the diurnal component should be removed from the series separately for each day of the sample. This raises the question of how to proceed. Does the time-of-day always affect the structure of the durations? If this happens, which adjustment method should be used to remove the intraday pattern? To find out whether the durations are affected by the time-of-day we shall make use of the tests suggested in Section 5.4.

The test results of no intraday pattern against smooth diurnal variation are presented in the second column of Table 5.2. Here, we choose $\alpha^{(1)} = 0.01$. Intraday variation clearly seems to be an inherent feature in the durations: the null hypothesis is rejected very strongly for almost every trading day. The exceptions are December 23 and December 31 in which the test does not reject at 1% significance level, although

Table 5.1 Summary statistics of the IBM durations

Durations	Min	Max	Mean	S.D.	Skew.	Ex.Kurt.
	Week 1(02/12/02 - 06/12/02): T=20811					
Raw	1	142	5.618	5.554	3.631	33.26
Adjusted Spline	0.125	21.43	1.000	0.941	3.083	22.10
Adjusted Smooth	0.130	19.01	0.933	0.895	3.069	19.90
	Week 2(09/12/02 - 13/12/02): T=19248					
Raw	1	81	6.074	6.101	2.893	13.71
Adjusted Spline	0.113	13.19	1.000	0.966	2.768	13.01
Adjusted Smooth	0.116	12.10	0.967	0.949	2.778	12.56
	Week 3(16/12/02 - 20/12/02): T=18818					
Raw	1	82	6.215	6.220	2.701	11.64
Adjusted Spline	0.0951	12.18	0.999	0.961	2.533	10.07
Adjusted Smooth	0.1226	13.66	0.956	0.949	2.698	11.97

Notes: Summary statistics for transaction duration data in seconds for the IBM stock traded at NYSE in December 2002.

the p-values still remain below the 5% level. This is explained by the irregular trading activity of these days due to their proximity to the Christmas Eve and the New Year's Day. Note, however, that the test based on the first-order Taylor approximation does reject the null hypothesis (it is H_{01}) at the 1% level.

The tests of the last three columns correspond to a sequence of nested tests based on (5.12) for choosing k in (5.9). In this context, k is related with the shape of diurnal pattern. As mentioned aboved, if H_{0i} is rejected most strongly, measured by the p-value, of the three hypotheses, one selects $k = i$. The table shows $k = 2$ for sixteen out of nineteen days, whereas for the remaining days $k = 1$ is the appropriate choice. As already mentioned, the test of H_{01} is the test of no diurnal variation based on the first-order Taylor approximation. It has power against the single logistic transition $k = 1$. Because the diurnal variation typically has an inverted U-shape, we fit the diurnal component ϕ_i with a single transition and then test for another transition. The resulting p-values are presented in Table 5.3. The hypothesis of one transition against another transition in the diurnal component is still rejected at $\alpha^{(2)} = 0.005$ for one trading day, but the rejection is much weaker than before. Consequently, the function ϕ_i with two transitions is chosen as the diurnal component for December 16th.

For comparison, tests for October 2002 are also reported in Table 5.4 (Appendix B). Two outcomes deserve particular attention. First, the null hypothesis of no diurnal

Table 5.2 p-values of the LM tests of no diurnal variation (December 2002)

Day	H_0	H_{03}	H_{02}	H_{01}
02/12/2002	3×10^{-15}	0.0169	$\mathbf{3 \times 10^{-13}}$	5×10^{-4}
03/12/2002	9×10^{-13}	0.2428	$\mathbf{8 \times 10^{-13}}$	0.0097
04/12/2002	6×10^{-15}	0.1887	2×10^{-7}	$\mathbf{2 \times 10^{-10}}$
05/12/2002	6×10^{-22}	0.3468	$\mathbf{1 \times 10^{-17}}$	1×10^{-7}
06/12/2002	4×10^{-7}	0.1053	$\mathbf{9 \times 10^{-5}}$	1×10^{-4}
09/12/2002	1×10^{-11}	0.0927	$\mathbf{1 \times 10^{-12}}$	0.4913
10/12/2002	3×10^{-20}	0.9704	$\mathbf{1 \times 10^{-21}}$	0.1005
11/12/2002	1×10^{-11}	7×10^{-4}	$\mathbf{2 \times 10^{-9}}$	0.0095
12/12/2002	3×10^{-11}	0.0071	$\mathbf{2 \times 10^{-11}}$	0.8108
13/12/2002	6×10^{-5}	0.1327	$\mathbf{9 \times 10^{-6}}$	0.5924
16/12/2002	3×10^{-10}	1×10^{-4}	0.0021	$\mathbf{1 \times 10^{-6}}$
17/12/2002	1×10^{-20}	2×10^{-4}	$\mathbf{4 \times 10^{-19}}$	0.1782
18/12/2002	5×10^{-15}	0.7886	$\mathbf{1 \times 10^{-12}}$	1×10^{-5}
19/12/2002	1×10^{-15}	0.0038	$\mathbf{4 \times 10^{-13}}$	1×10^{-4}
20/12/2002	1×10^{-17}	0.1157	$\mathbf{4 \times 10^{-18}}$	0.0340
23/12/2002	0.0164	0.0666	**0.0086**	0.9627
26/12/2002	4×10^{-8}	0.6262	$\mathbf{5 \times 10^{-9}}$	0.1069
27/12/2002	2×10^{-13}	0.2733	$\mathbf{1 \times 10^{-9}}$	1×10^{-6}
31/12/2002	0.0360	0.5900	0.4609	**0.0055**

Notes: The table contains p-values of the LM tests of no diurnal variation against smoothly time-varying diurnal pattern in the durations for December 2002. The p-values shown in boldface indicate the lowest rejection rate.

pattern is not rejected for seven out of 23 trading days. The intraday pattern is thus considerably less conspicuous in October than in December, which suggests that the durations may also vary systematically over the year besides the time of day. Second, when the null hypothesis is rejected, the tests support the usual inverted U-shape of diurnal variation only in four out of fifteen days. This suggests the diurnal pattern may not always have the shape proposed earlier: short durations early and late in the day and lower activity in the middle. For this reason, one should proceed with care when specifying the diurnal component. The estimation of the ACD model should be preceded by a specification search to determine the diurnal variation.

For the empirical analysis, durations are 'diurnally adjusted' using two different estimators for the diurnal component. The first estimator is cubic spline of Engle and Russell (1998), and the second one our smooth transition component. The adjusted durations are obtained by dividing the raw durations by one of these deterministic

Table 5.3 *p*-values of the LM tests of a single transition against a double transition in the diurnal component

Day	H_0	H_{03}	H_{02}	H_{01}
04/12/2002	0.0161	0.2310	0.0272	0.0455
16/12/2002	9×10^{-5}	0.0476	0.0282	4×10^{-4}
31/12/2002	0.0967	0.4336	0.5983	0.0197

Notes: The table contains p-values of the LM tests of no diurnal variation against smoothly time-varying diurnal pattern in the durations for December 2002. The p-values shown in boldface indicate the lowest rejection rate.

diurnal factors. The diurnal components estimated by cubic splines for each day of the second week of December are shown in Figure 5.4 whereas the ones estimated using the smooth transition approach can be found in Figure 5.5. The durations change systematically through the day, and the pattern of trading activity accords to what has been reported in earlier studies. Durations tend to be shorter at the beginning and at the end of the day and longer around lunch time. Some summary statistics of the adjusted durations can be found in Table 5.1.

Table 5.5 presents the estimation results for the two parameterizations of ϕ_i. The parameter estimates are significant in the three weeks. The smooth diurnal component has a larger value of $\hat{\alpha}_1 + \hat{\beta}_1$ than the cubic spline method. Similar results were obtained by Rodriguez-Poo, Veredas, and Espasa (2008) who employed a semiparametric model for the diurnal factor. Furthermore, the results indicate that the smooth transition parameterization attains higher likelihood values, and therefore seems to fit durations better than the cubic spline approach.

5.6 Conclusions

In this paper we have introduced the smooth transition method for parameterizing the diurnal variation in the intertrade durations. This is done by multiplicatively decomposing durations into a deterministic and an stochastic component, in which the durations are allowed to change smoothly over the time-of-day. A testing framework is also provided for determining the structure of the diurnal component using a sequence of specification tests.

Our modelling technique is illustrated with an application to IBM stock transaction data. The test results indicate that the diurnal variation may not always have the documented inverted U-shaped pattern for the trade durations. In addition, the results suggest that our method fits durations better than the cubic spline approach.

A possible extension of this framework is to jointly estimate the parametric conditional duration and diurnal variation components. This may be done by maximisation by parts as in Song, Fan, and Kalbfleisch (2005) in which the resulting estimator coincides with the ML estimator. This consideration is, however, left for future research.

Appendix A: Figures

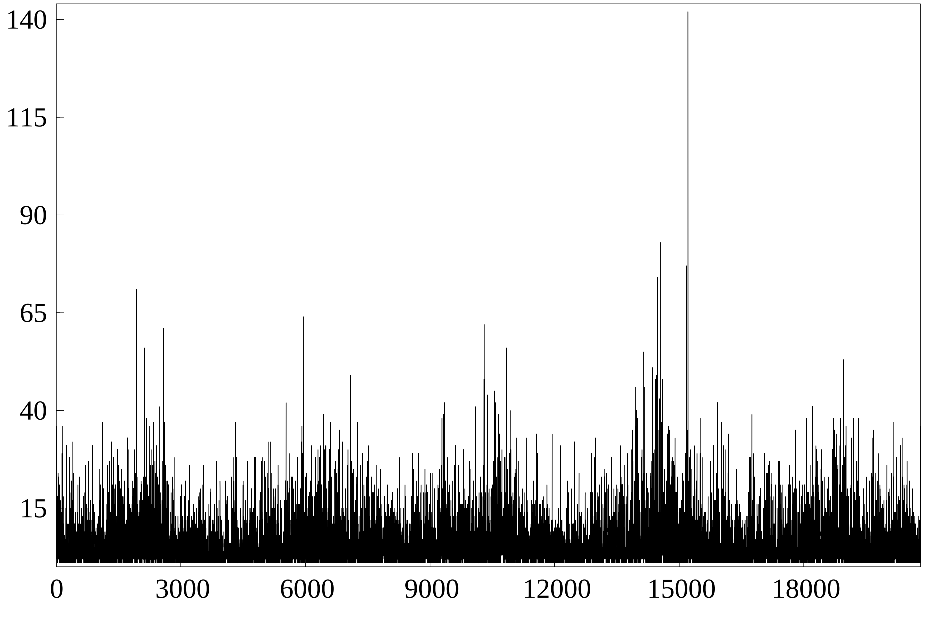
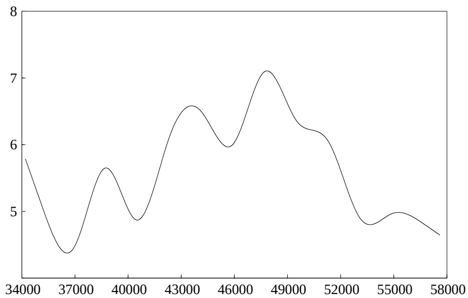
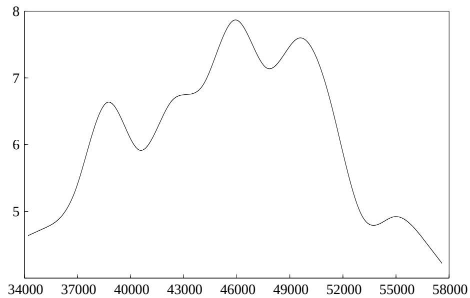


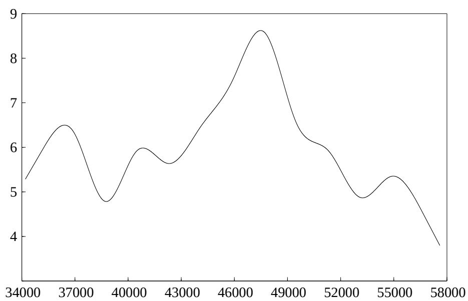
Figure 5.3 Durations for IBM traded in the first five trading days of December 2002.



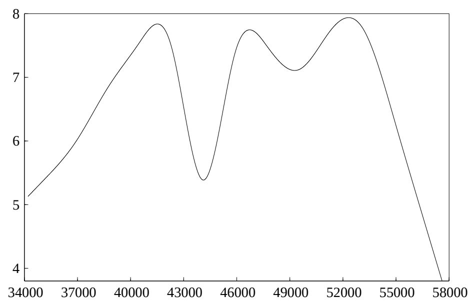
(a) Monday



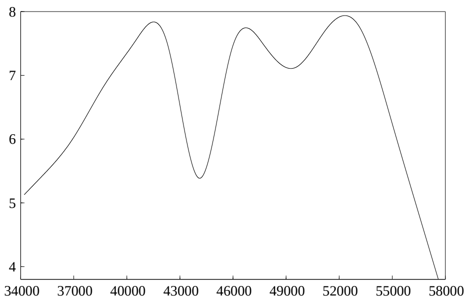
(b) Tuesday



(c) Wednesday

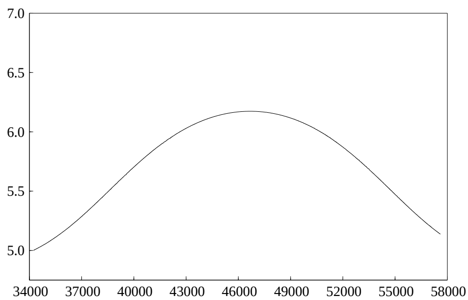


(d) Thursday

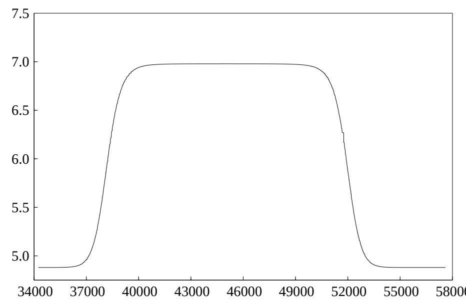


(e) Friday

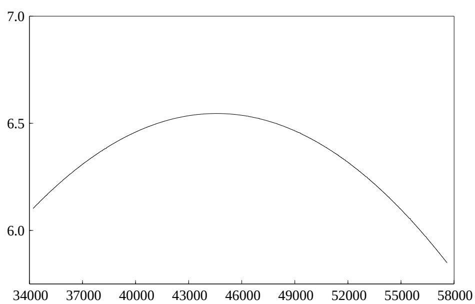
Figure 5.4 Estimated diurnal variation using the cubic spline for the trade durations of IBM in the second week of December 2002. The scale of the x-axis is time measured in seconds after midnight.



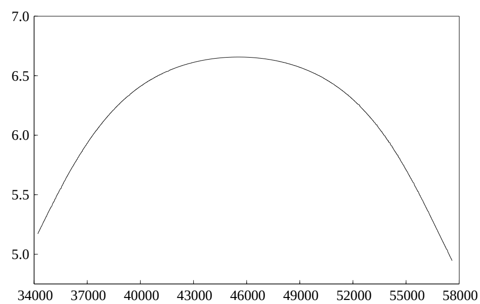
(a) Monday



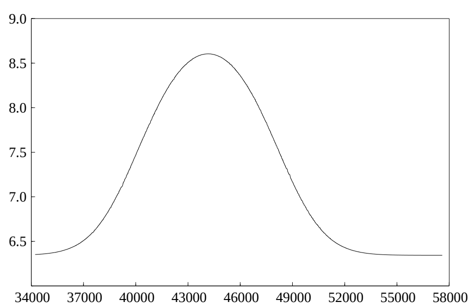
(b) Tuesday



(c) Wednesday



(d) Thursday



(e) Friday

Figure 5.5 Estimated diurnal variation using the smooth transition for the trade durations of IBM in the second week of December 2002. The scale of the x-axis is time measured in seconds after midnight.

Appendix B: Tables

Table 5.4 p-values of the LM tests of no diurnal variation (October 2002)

Day	H_0	H_{03}	H_{02}	H_{01}
01/10/2002	9×10^{-4}	0.6022	$\mathbf{2 \times 10^{-4}}$	0.1166
02/10/2002	9×10^{-4}	**0.0041**	0.1707	0.0109
03/10/2002	6×10^{-17}	0.0025	$\mathbf{7 \times 10^{-17}}$	0.8382
04/10/2002	0.3204	0.1032	0.4342	0.6270
07/10/2002	0.4354	0.5866	0.2659	0.2741
08/10/2002	0.0985	0.2842	0.5659	0.0283
09/10/2002	3×10^{-4}	0.0327	0.5361	$\mathbf{2 \times 10^{-4}}$
10/10/2002	3×10^{-17}	3×10^{-5}	1×10^{-5}	$\mathbf{3 \times 10^{-11}}$
11/10/2002	4×10^{-8}	0.1502	0.0655	$\mathbf{2 \times 10^{-8}}$
14/10/2002	7×10^{-10}	$\mathbf{6 \times 10^{-7}}$	6×10^{-6}	0.6683
15/10/2002	0.2632	0.7177	0.3549	0.0833
16/10/2002	0.0797	0.1412	0.9622	0.0319
17/10/2002	9×10^{-12}	0.3052	0.2504	$\mathbf{5 \times 10^{-13}}$
18/10/2002	0.5046	0.7125	0.1597	0.6318
21/10/2002	0.0023	0.0508	0.2786	**0.0021**
22/10/2002	0.0205	0.6768	0.0517	**0.0157**
23/10/2002	9×10^{-13}	0.0932	$\mathbf{7 \times 10^{-8}}$	2×10^{-7}
24/10/2002	9×10^{-9}	$\mathbf{7 \times 10^{-6}}$	2×10^{-5}	0.2146
25/10/2002	1×10^{-7}	0.5485	$\mathbf{2 \times 10^{-5}}$	4×10^{-5}
28/10/2002	0.0018	$\mathbf{3 \times 10^{-4}}$	0.2108	0.7034
29/10/2002	0.7720	0.7551	0.7723	0.3323
30/10/2002	2×10^{-9}	0.0193	0.9384	$\mathbf{6 \times 10^{-10}}$
31/10/2002	2×10^{-8}	0.6035	0.3705	$\mathbf{1 \times 10^{-9}}$

Notes: The table contains p-values of the LM tests of no diurnal variation against a smooth diurnal pattern in the durations for October 2002. The p-values shown in boldface indicate the lowest rejection rate of the null hypothesis.

Table 5.5 ACD Estimation results for the Exponential ACD model (robust standard errors in parentheses)

	$\widehat{\omega}$	$\widehat{\alpha}_1$	$\widehat{\beta}_1$	$\widehat{\alpha}_1 + \widehat{\beta}_1$	Log-Lik
		Week 1: 02/12/02 - 06/12/02			
Spline	0.2095 (0.0227)	0.0865 (0.0072)	0.7042 (0.0268)	0.7907	-20673.7
Smooth	0.1111 (0.0237)	0.0843 (0.0100)	0.7969 (0.0341)	0.8812	-19150.2
		Week 2: 9/12/02 - 13/12/02			
Spline	0.1289 (0.0251)	0.0640 (0.0075)	0.8071 (0.0311)	0.8711	-19128.6
Smooth	0.0553 (0.0141)	0.0560 (0.0078)	0.8868 (0.0216)	0.9429	-18382.0
		Week 3: 16/12/02 - 20/12/02			
Spline	0.1612 (0.0277)	0.0781 (0.0080)	0.7608 (0.0339)	0.8389	-18669.7
Smooth	0.0308 (0.0100)	0.0551 (0.0094)	0.9129 (0.0195)	0.9679	-17629.3

Bibliography

Bauwens, L., and P. Giot (2000): "The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks," *Annales d'economie et de statistique*, 60, 117–149.

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

Drost, F. C., and B. J. M. Werker (2004): "Semiparametric Duration Models," *Journal of Business and Economic Statistics*, 22, 40–50.

Engle, R. F., and J. R. Russell (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transition Data," *Econometrica*, 66, 1127–1162.

Friedman, J. H. (1984): "A Variable Span Smoother," Technical Report No. 5, Laboratory for Computational Statistics, Department of Statistics, Stanford University.

Gourieroux, C., A. Monfort, and A. Trognon (1984): "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 52, 681–700.

Lin, C.-F. J., and T. Teräsvirta (1994): "Testing the Constancy of Regression Parameters Against Continuous Structural Change," *Journal of Econometrics*, 62, 211–228.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

McCulloch, R. E., and R. S. Tsay (2001): "Nonlinearity in High-Frequency Financial Data and Hierarchical Models," *Studies in Nonlinear Dynamics and Econometrics*, 5, 1–17.

Meitz, M., and T. Teräsvirta (2006): "Evaluating Models of Autoregressive Conditional Duration," *Journal of Business and Economic Statistics*, 24, 104–124.

Rodriguez-Poo, J. M., D. Veredas, and A. Espasa (2008): "Semiparametric Estimation for Financial Durations," in *High Frequency Financial Econometrics: Recent Developments (Studies in Empirical Economics)*, ed. by L. Bauwens, W. Pohlmeier, and D. Veredas, pp. 225–251. Springer Verlag.

Song, P. X., Y. Fan, and J. D. Kalbfleisch (2005): "Maximization by Parts in Likelihood Inference," *Journal of the American Statistical Association*, 100, 1145–1158.

Teräsvirta, T. (1994): "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89, 208–218.

Zhang, M. Y., J. R. Russell, and R. S. Tsay (2001): "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data," *Journal of Econometrics*, 104, 179–207.

EFI, The Economic Research Institute

Published in the language indicated by the title.
A complete publication list can be found at *www.hhs.se/efi*
Books and dissertations can be ordered from EFI via e-mail: EFI.Publications@hhs.se

Reports since 2006

2009

Dissertations

Hernant, Mikael. *Profitability performance of supermarkets: the effects of scale of operation, local market conditions, and conduct on the economic performance of supermarkets.*
Lakomaa, Erik. *The economic psychology of the welfare state.*

2008

Books

Breman, Anna. *Forskning om filantropi. Varför skänker vi bort pengar?* Forskning i Fickformat.
Einarsson, Torbjörn. *Medlemskapet i den svenska idrottsrörelsen: En studie av medlemmar i fyra idrottsföreningar.* EFI Civil Society Reports.
Helgesson, Claes-Fredrik and Hans Winberg (eds). *Detta borde vårddebatten handla om.*
Jennergren, Peter, Johnny Lind, Walter Schuster and Kenth Skogsvik (eds). *Redovisning i fokus.* EFI:s Årsbok 2008. EFI/Studentlitteratur.
Kraus, Kalle. *Sven eller pengarna? Styrningsdilemman i äldrevården.* Forskning i Fickformat.
Petrelius Karlberg, Pernilla. *Vd under press – om medialiseringen av näringslivets ledare.* Forskning i Fickformat.
Portnoff, Linda. *Musikbranschens styrningsproblematik.* Forskning i Fickformat.
Sjöstrand, Sven-Erik. *Management: från kontorsteknik till lednings- och organisationsteori: utveck-lingen på Handelshögskolan under 100 år: 1909–2009.*
Östman, Lars. *Den finansiella styrningens realiteter och fiktioner: de finansiella styrformernas svenska historia, berättelser om Petersson och "Ericsson", finansiell styrning – en ansats till generell teori.*
Östman, Lars. *Mycket hände på vägen från Buchhaltung till Accounting: delar av Handelshögskolan under 100 år.*

Dissertations

Axelson, Mattias. *Enabling Knowledge Communication between Companies – the role of integration mechanisms in product development collaborations.*

Benson, Ilinca. *Organisering av övergångar på arbetsmarknaden: en studie av omställningsprogram.*
Elhouar, Mikael. *Essays on interest rate theory.*
Farooqi Lind, Raana. *On capital structure and debt placement in Swedish companies.*
Granström, Ola. *Aid, drugs, and informality: essays in empirical economics.*
Hvenmark, Johan. *Reconsidering membership: a study of individual members' formal affiliation with democratically governed federations.*
Höglin, Erik. *Inequality in the labor market: insurance, unions, and discrimination.*
Johansson, Marjana. *Engaging resources for cultural events: a performative view.*
Kallenberg, Kristian. *Business at risk. Four studies on operational risk anagement.*
Kviselius, Niklas Z. *Trust-building and communication in SME internationalization: a study of Swedish-Japanese business relations.*
Landberg, Anders. *New venture creation: resistance, coping and energy.*
Pemer, Frida. *Framgång eller fiasko? En studie av hur konsultprojekt värderas i klientorganisationer.*
Rosengren, Sara. *Facing Clutter: On Message Competition in Marketing Communication.*
Schilling, Annika. *Kan konsulter fusionera?: en studie av betydelsen av identitet vid en fusion mellan konsultföretag.*
Schriber, Svante. *Ledning av synergirealisering i fusioner och förvärv.*
Sjödin, Henrik. *Tensions of extensions: adverse effects of brand extension within consumer relationship.*
Strandqvist, Kristoffer. *Kritiska år: formativa moment för den svenska flygplansindustrin 1944–1951.*
Strömqvist, Maria. *Hedge funds and international capital flow.*
Söderström, Johan. *Empirical studies in market efficiency.*
Sölvell, Ingela. *Formalization in high-technology ventures.*
Thorsell, Håkan. *The pricing of corporate bonds and determinants of financial structure.*
Ulbrich, Frank. *The adoption of IT-enabled management ideas: insights from shared services in government agencies.*
Östling, Robert. *Bounded rationality and endogenous preferences.*

2007

Books

Andersson, Per, Ulf Essler and Bertil Thorngren (eds). *Beyond mobility.* EFI Yearbook 2007. EFI/Studentlitteratur.
Einarsson, Torbjörn and Filip Wijkström. *Analysmodell för sektorsöverskridande statistik: fallet vård och omsorg.* EFI Civil Society Reports.
Ericsson, Daniel. *Musikmysteriet: organiserade stämningar och motstämningar.*
Samuelson, Lennart (ed). *Bönder och bolsjeviker: den ryska landsbygdens historia 1902–1939.*

Dissertations

Ahlersten, Krister. *Empirical asset pricing and investment strategies.*
Alexius, Susanna. *Regelmotståndarna: om konsten att undkomma regler.*
Andersson, Magnus. *Essays in empirical finance.*
Berg, Bengt Åke. *Volatility, integration and grain bank: studies in harvests, rye prices and institutional development of the parish magasins in Sweden in the 18th and 19th centuries.*
Bianchi, Milo. *Of speculators, migrants and entrepreneurs: essays on the economics of trying your fortune.*
Brodin, Karolina. *Consuming the commercial break: an ethnographic study of the potential audiences for television advertising.*
Elger, Max. *Three essays on investment-specific technical change.*
Hagberg, Axel. *Bankkrishantering: aktörer, marknad och stat.*
Hinnerich, Mia. *Derivatives pricing and term Structure modeling.*
Hjalmarson, Hanna. *En växande marknad: studie av nöjdheten med konsumtionsrelaterade livsområden bland unga konsumenter.*
Hjelström, Tomas. *The closed-end investment company premium puzzle: model development and empirical tests on Swedish and British data.*
Kraus, Kalle. *Sven, inter-organisational relationships and control: a case study of domestic care of the elderly.*
Lindqvist, Erik. *Essays on privatization, identity, and political polarization.*
Macquet, Monica. *Partnerskap för hållbar utveckling: systrar av Oikos och guvernanten som blev diplomat.*
Melian, Catharina. *Progressive open source.*
Nilsson, Daniel. *Transactions in cyberspace: the continued use of Internet banking.*
Petrelius Karlberg, Pernilla. *Den medialiserade direktören.*
Portnoff, Linda. *Control, cultural production and consumption: theoretical perspectives, empirical dilemmas, and Swedish music industry practices.*
Sköld, Martin. *Synergirealisering: realisering av produktsynergier efter företags- sammanslagningar.*
Sonnerby, Per. *Contract-theoretic analyses of consultants and trade unions.*
Tyrefors, Björn. *Institutions, policy and quasi-experimental evidence.*
Valiente, Pablo. *Re-innovating the existing: a study of wireless IS capabilities to support mobile workforces.*

2006

Books

Lundeberg, Mats, Pär Mårtensson and Magnus Mähring (eds). *IT & business performance: a dynamic relationship.* EFI Yearbook 2006. EFI / Studentlitteratur.
Thodenius, Björn. *Organisering av kunskap: en studie av Wallenberg Consortium North.* EFI Civil Society Reports.
Wijkström, Filip and Torbjörn Einarsson. *Från nationalstat till näringsliv?: det civila samhällets organisationsliv i förändring.*

Wijkström, Filip, Stefan Einarsson and Ola Larsson. *Staten och det civila samhället: idétraditioner och tankemodeller i den statliga bidragsgivningen till ideella organisationer.*
Östman, Lars. *Lysande ögonblick och finansiella kriser: Dramaten under ett sekel.*

Dissertations

Argenton, Cedric. *Quality provision in duopoly.*
Beckerman, Carina. *The clinical eye: constructiong and computerizing an anesthesia patient record.*
Borglund, Tommy. *Aktievärden i fokus: internationell påverkan på intressentrelationer vid förvärv och fusion.*
Breman, Anna. *The Economics of altruism, paternalism and self-control.*
Edquist, Harald. *Technological breakthroughs and productivity growth.*
Eklund, Jana. *Essays on forecasting and bayesian model averaging.*
Frostenson, Magnus. *Legitimitetskontrollen: en studie av etiska värderingars roll i gränsöverskridande förvärv och fusioner.*
Gaspar, Raquel M. *Credit risk and forward price models.*
Gustafsson, Peter. *Essays on trade and technological change.*
Hopkins, Elisabeth. *Is a higher degree of local currency pricing associated with lower exchange rate pass-through?: a study of import pricing in 51 Swedish industries.*
Kling, Ragnar. *Developing product development in times of brutal change.*
Langenskiöld, Sophie. *Peer influence on smoking: causation or correlation?*
Lychnell, Lars-Olof. *"Och fungerar det inte, gör vi på något annat sätt" en klinisk fallstudie av IT-relaterat förändringsarbete i småföretag.*
Meitz, Mika. *Five contributions to econometric theory and the econometrics of ultra-high-frequency data.*
Mendicino, Caterina. *Financial market imperfections, business cycle fluctuations and economic growth.*
Ovanfors, Anna. *Essays on nonlinear time series analysis and health economics.*
Paltseva, Elena. *Essays on commitment and inefficiency in political economy.*
Rogberg, Martin. *Den modeföljande organisationen: om acceptansen av TQM och andra populära managementmodeller.*
Silvennoinen, Annastiina. *Essays on autoregressive conditional heteroskedasticity.*
Sjögren, Ebba. *Reasonable drugs: making decisions with ambiguous knowledge.*
Slinko, Irina. *Essays in option pricing and interest rate models.*
Wilander, Fredrik. *Essays on exchange rates and prices.*