

Modelling economic high-frequency time series

Stefan Lundbergh

AKADEMISK AVHANDLING
som för avläggande av ekonomie doktorsexamen
vid Handelshögskolan i Stockholm
framläggs för offentlig granskning
måndagen den 6 september 1999, kl 13.15
i sal Ragnar, Handelshögskolan, Sveavägen 65



Modelling economic high-frequency time series



STOCKHOLM SCHOOL OF ECONOMICS

EFI, THE ECONOMIC RESEARCH INSTITUTE

EFI Mission

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.

EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization

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

ORGANIZATION AND MANAGEMENT

Management and Organisation; (A)

Center for Ethics and Economics; (CEE)

Public Management; (F)

Information Management; (I)

Man and Organisation; (PMO)

Industrial Production; (T)

Prof Sven-Erik Sjöstrand

Adj Prof Hans de Geer

Prof Nils Brunsson

Prof Mats Lundberg

Acting Prof Jan Löwstedt

Prof Christer Karlsson

MARKETING

Center for Information and Communication

Research; (CIC)

Center for Consumer Marketing; (CCM)

Marketing, Distribution and Industrial

Dynamics; (D)

Adj Prof Bertil Thorngren

Associate Prof Magnus Söderlund

Prof Lars-Gunnar Mattsson

ECONOMIC PSYCHOLOGY

Center for Risk Research; (CFR)

Economic Psychology; (P)

Prof Lennart Sjöberg

Prof Lennart Sjöberg

ACCOUNTING, CONTROL AND CORPORATE FINANCE

Accounting and Managerial Finance; (B)

Managerial Economics; (C)

Prof Lars Östman

Prof Peter Jennergren

FINANCE

Finance; (FI)

Prof Clas Bergström

ECONOMICS

Center for Health Economics; (CHE)

International Economics and Geography; (IEG)

Economics; (S)

Prof Bengt Jönsson

Prof Mats Lundahl

Prof Lars Bergman

ECONOMICS STATISTICS

Economic Statistics; (ES)

Prof Anders Westlund

LAW

Law; (RV)

Prof Erik Nerep

Chairman of the Board: Prof Sven-Erik Sjöstrand. *Director:* Associate Prof Bo Sellstedt,

Adress

EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/

Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Modelling economic high-frequency time series

Stefan Lundbergh



STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE



Dissertation for the degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics, 1999

© EFI and the author 1999

ISBN 91-7258-517-X

Keywords:

GARCH, misspecification tests, model specification, nonlinear time series, parameter constancy, smooth transition, STAR, time series.

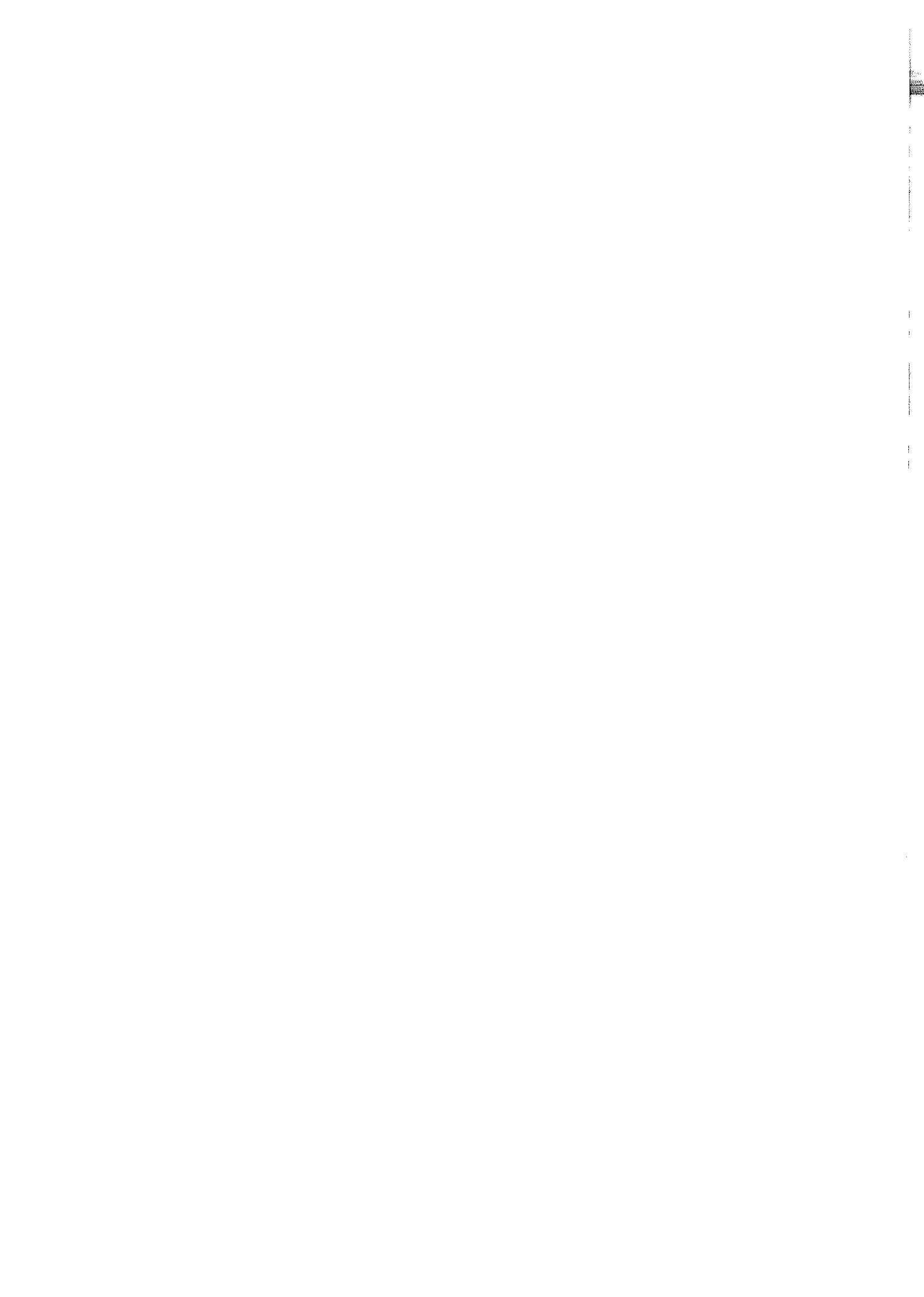
Distributed by:

EFI, Stockholm School of Economics
Box 6501, SE-113 83 Stockholm, Sweden

Stockholm School of Economics, Stockholm, 1999

Contents

Acknowledgements	vii
Introduction and summary	1
Modelling economic high-frequency time series with STAR-STGARCH models	15
Evaluating GARCH models	63
A GARCH model with time-varying parameters	87
A smooth transition autoregressive target zone model	111

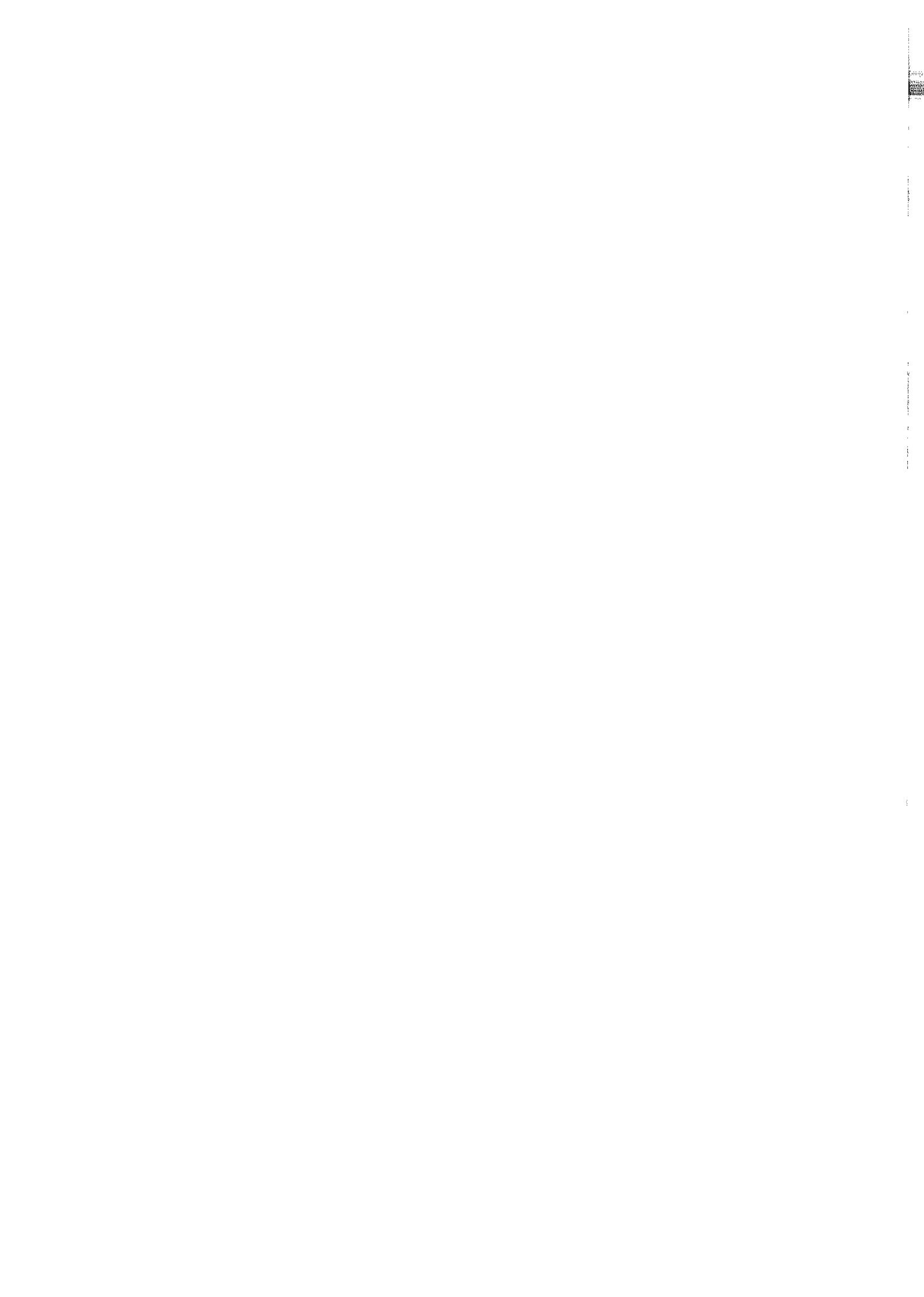


Acknowledgements

First of all I want to thank my supervisor Professor Timo Teräsvirta for always finding the time to answer my questions and for discussing ideas. During the work with this thesis he has been given me valuable support and inspiration that significantly improved this work. I also want to thank my girlfriend Sigrid Schmidt and our two cats for the inspiration to this work.

The Department of Economic Statistics provides a stimulating atmosphere and research environment. There has always been time for discussions at the coffee table. During these discussions, and at other occasions as well, valuable comments and suggestions on my work has been given by fellow students and members of the department. At conferences and seminars I have received comments and suggestions whose value can not be underestimated. Any remaining errors are my own.

The author gratefully acknowledges generous financial support from the Tore Browaldh's Foundation and the Swedish Research Council for Humanities and Social Sciences.



Chapter 1

Introduction and summary

1. Background

Nonlinear (asymmetric) behaviour has been suggested by economic theory and observed in economic data. As a simple example of nonlinearity one may consider the dynamic response to a shock in an economic system. The nonlinearity arises when the dynamic response to a large shock differs from the dynamic response to a shock of moderate size. It is not always realistic to assume that the dynamic response to a positive shock will be a mirror image of the dynamic response to a negative shock. The idea that the dynamic response to a shock differs according to the sign of the shock is a simple example of asymmetry.

When nonlinearity is suggested by macro-economic theory, the switching regression has been used. Some special cases of the switching regression model may be viewed as a threshold model. The idea goes back at least to Quandt (1960). In the literature on time series, the idea of threshold autoregressive models was introduced much later. However, the model has become very popular and has also been applied to describe nonlinear economic behaviour; see Tong (1990) for a thorough account of these models. In the most common case the threshold models may be viewed as a two-regime system where a linear model may describe each of the regimes. Any change between these regimes is assumed to be abrupt. On the other hand, the smooth transition autoregressive (STAR) model and its multivari-

ate single-equation counterpart, the smooth transition regression (STR) model, see for example Teräsvirta (1994, 1998), have also proved to be useful tools for modellers of nonlinear economic behaviour. The ST(A)R model is a generalization of the two-regime system threshold model. In this model the transition between the two extreme regimes is smooth. The idea of a smooth transition was introduced by Bacon and Watts (1971). One advantage with the STAR or STR model is that despite its relatively simple structure it provides the modeller with a very flexible model. Another advantage is that the STR model can often be interpreted in economic terms due to the regime switching structure. This type of behaviour is often suggested by economic theory, in which case it is natural to consider the ST(A)R model. In practice the model has performed well when modelling macro-economic series. To mention a few examples: house prices in Teräsvirta (1998), asymmetries in business cycles in Anderson and Teräsvirta (1992), Skalin and Teräsvirta (1999) and Osborn and Öcal (1999), real exchange rates in Michael, Nobay and Peel (1997) and Taylor and Sarno (1999) and interest rates in Franses and van Dijk (1999). An advantage compared to threshold models is that the log-likelihood of a smooth transition model is continuously everywhere in the parameter space and differentiable. This makes it possible, at least in principle, to construct evaluation tests that rely on standard asymptotic theory.

Hitherto, the ST(A)R model has been applied assuming that the conditional variance is constant. This assumption is relaxed here and therefore a flexible parameterization of the conditional variance is needed. The idea of conditional autoregressive heteroskedasticity (ARCH; Engle, 1982) has led to a large number of extensions of the original model and applications. Bollerslev, Engle and Nelson (1994) and Palm (1996) are examples of recent surveys of the area. Another variant of ARCH, the so-called Stochastic Volatility (SV) model (Taylor, 1986, p. 73ff.) has gained popularity recently, and the latest developments have been surveyed in Ghysels, Harvey and Renault (1996). The original ARCH model and its most important extension, the generalized ARCH (GARCH), are symmetric: while the size of the shock matters, the sign does not. The same is true for the (SV) model. Many authors have argued that shocks may have asymmetric effects on volatility. The dynamic response to a positive shock is not necessarily the mirror image of the response to a negative shock of the same size. Pagan (1996), in his survey of developments in financial econometrics, provided a useful review of models that can handle this type of asymmetry. Glosten, Jagannathan and Runkle (1993) introduced an asymmetric GARCH model which is a model in the threshold spirit and

may be viewed as a predecessor to the Smooth Transition GARCH (STGARCH). The STGARCH is a very flexible nonlinear GARCH model where the ideas behind STAR models in the conditional mean are adopted to the conditional variance; see Hagerud (1997) and González-Rivera (1998).

In this thesis the focus is on modelling nonlinearities in both the conditional mean and the conditional variance. The models are designed for and applied to univariate time series but may be extended to multivariate situations. In the second chapter a nonlinear model is suggested which combines the STAR model for the conditional mean with the STGARCH model for the conditional variance. To evaluate the conditional variance model, a battery of tests are employed. A unified framework for testing the adequacy of an estimated GARCH model is suggested in Chapter 3. All tests are tests against a parametric alternative. The advantage with this approach is that if the null hypothesis is rejected the alternative may be estimated. This possibility is pursued in Chapter 4 where a GARCH model with time-varying parameters is suggested. Finally in Chapter 5, a new nonlinear empirical target zone model, the smooth transition autoregressive target zone (STARTZ) model, is introduced.

A summary of the contents and the main results from each of the following four chapters of the thesis are given in the following sections.

2. Modelling economic high-frequency time series with STAR-STGARCH models

Many of the financial time series are "high frequency", that is, they consist of weekly, daily or even intradaily observations. As many high frequency series show little or no linear dependence, the focus has been on modelling the conditional variance. It has been argued, however, that despite the absence of linear dependence there may be nonlinear dependence in the conditional mean. This should then be appropriately modelled in order to avoid misspecification of the conditional variance. A natural idea would be to combine an asymmetric/nonlinear parameterization of the conditional variance with a nonlinear model for the conditional mean. Li and Li (1996) did exactly that by defining a double threshold autoregressive heteroskedastic (DTARCH) time series model. A DTARCH model has a SETAR-type conditional mean. The conditional variance is parameterized similarly, and the authors called their specification the threshold ARCH (TARCH) model. Li and Li (1996) also provided a comprehensive modelling strategy for DTARCH models. Recently, Lee

and Li (1998) generalized the DTARCH model by allowing the transition between the two regimes to be smooth. They called this model the Smooth Transition Double Threshold model.

In Chapter 2 we introduce another smooth transition model, the STAR-STGARCH model. It is a model that can capture nonlinear behaviour in both the conditional mean and the conditional variance. A modelling cycle for this family of models consists of three stages: specification, estimation and evaluation. This STAR-STGARCH model may be seen as an extension of Lee and Li's work in the sense that a rather flexible specification is simultaneously allowed not only for the conditional mean but for the conditional variance as well. Besides, misspecification testing will receive plenty of attention in Chapter 2.

The conditional mean is thus specified as a Smooth Transition Autoregressive (STAR) model; see, for example, Chan and Tong (1986), Granger and Teräsvirta (1993) and Teräsvirta (1994). The conditional variance is specified as a Smooth Transition GARCH (STGARCH) model; see Hagerud (1997) and González-Rivera (1998). The construction of a complete modelling cycle for the STAR-STGARCH family of models requires a coherent specification strategy of the kind introduced, for example, in Box and Jenkins (1970), Li and Li (1996), Tsay (1989, 1998) and Teräsvirta (1994, 1998). The conditional mean is specified first, followed by the conditional variance. The modelling scheme proceeds from restricted models to more general ones to avoid estimating unidentified models. A sequence of tests is computed to investigate the validity of the assumptions imposed in the estimation of parameters once the parameters of the STAR-STGARCH model (or a submodel) have been estimated. The tests are Lagrange Multiplier (LM)-type tests. The tests for evaluation of a STAR model in Teräsvirta (1994) are modified to this environment and the tests for evaluating the conditional variance are discussed in detail in Chapter 3.

The STAR-STGARCH model is illustrated by the Swedish OMX index and the exchange rate between the Japanese Yen and the US Dollar. These empirical examples indicate that there are nonlinear structures in the conditional mean that has to be modelled. In the case of the OMX index this leads to improved one-step-ahead forecasts of the changes in the index. For the JPY/USD exchange rate return series the nonlinear parameters only characterize some extreme events in the series. Because such events by definition are rare, the forecast accuracy, when measured from a large number of forecasts, is not improved by these extra parameters in the conditional mean specification.

3. Evaluating GARCH models

A unified framework for testing the adequacy of an estimated GARCH model is suggested in Chapter 3. This Chapter contains a number of new tests while some existing ones fit into this framework as well. A simulation study is conducted to find out the performance of the tests in finite samples and to compare the tests with existing ones.

When modelling the conditional mean, at least when it is a linear function of parameters, the estimated model is regularly subjected to a battery of misspecification tests to check its adequacy. The hypotheses of no (conditional) heteroskedasticity, no error autocorrelation, linearity, and parameter constancy, to name a few, are tested using various methods. In models of conditional variance, such as the popular GARCH model, testing the adequacy of the estimated model has been much less common in practice. But then, misspecification tests do exist in the literature for GARCH models also. For example, Bollerslev (1986) already suggested a score or Lagrange Multiplier (LM) test for testing a GARCH model of a given order against a higher-order GARCH model. Li and Mak (1994) derived a portmanteau type test for testing the adequacy of a GARCH model. Engle and Ng (1993) considered testing the GARCH specification against asymmetry using the so-called sign-bias test. Chu (1995) derived a test of parameter constancy against a single structural break. This test has a nonstandard asymptotic null distribution, but Chu provided tables for critical values.

The suggested tests in Chapter 3 are LM-type tests and require only standard asymptotic distribution theory. They may be obtained from the same "root" by merely changing the definitions of the elements of the score vector corresponding to the alternative hypothesis. This makes testing easy as the sample counterparts of the analytical first and second order derivatives of the logarithmic likelihood function may be computed without difficulty using the results in Fiorentini, Calzolari and Panattoni (1996). As a result, misspecification of a GARCH model may be detected quite easily at low computational cost.

First a test of the null hypothesis of no error autocorrelation in the squared residuals is suggested. This test may be viewed as a general but possibly parsimonious misspecification test along the lines in Bollerslev (1986). Secondly, the test of symmetry/linearity against the smooth transition GARCH, see Hagerud (1997) and González-Rivera (1998), fits into the unified framework. Finally, a parameter constancy test is suggested. The alternative to constant parameters in the condi-

tional variance is that the parameters, or a subset of them, change smoothly over time. Lin and Teräsvirta (1994) applied this idea to testing parameter constancy in the conditional mean. Because the tests of symmetry and parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This helps the model builder to find out the weaknesses of the estimated specification and may give useful ideas of how the current specification could be improved.

It is also shown that a test of no ARCH in the standardized error process is asymptotically equivalent to a portmanteau test of Li and Mak (1994). This links the work of these authors to the suggested framework and indicates that the null hypothesis of no remaining ARCH can be tested in different ways while the asymptotic distribution theory remains the same.

A Monte Carlo simulation shows that the proposed tests have reasonable power, that is, they compare favourably with the tests currently available in the literature. Applications of these tests can be found in Chapters 2, 4 and 5 of this thesis.

4. A GARCH model with time-varying parameters

The case of time-varying parameters in a GARCH model is studied in Chapter 4. The time-varying parameter GARCH (TVGARCH) model is suggested. In this model the parameters are allowed to vary smoothly and deterministically over time.

In practice, when modelling the conditional mean, it is difficult to distinguish between an autoregressive process with a step and a random-walk (unit root) process, see for example Hendry and Neale (1991). This is of importance when modelling empirical series, since a regime shift may generate random-walk-like behaviour in piecewise stationary autoregressive time series. This may put the modeller in a situation where a choice has to be made between unit roots or stationarity with structural breaks. A vast literature has emerged for this topic, starting with Perron (1989). An equivalent problem is apparent when modelling the conditional variance. Parameter nonconstancy may manifest itself as an apparent lack of weak stationarity (IGARCH); see, for example, Diebold (1986) and Lamoureux and Lastrapes (1990).

The parameter constancy test suggested in Chapter 3 serves as a starting point for modelling time-varying parameters. The alternative to parameter constancy is that the parameters change smoothly over time. The idea is to parameterize this alternative by allowing a smooth change from one GARCH model into another over time. Technically, the change is parameterized using a logistic function with time

as the argument located at the centre of the change. A major advantage with this approach is that the modeller does not have to know when the change occurs and how rapid it is. These characteristics can be estimated directly from the data. A modelling cycle is constructed for the time-varying GARCH model including specification, estimation and evaluation. The misspecification tests for evaluation of the estimated TVGARCH model are obtained by modifying the corresponding tests suggested in Chapter 3.

To illustrate what happens when the parameters of a GARCH model are non-constant but are assumed fixed, a small simulation study is performed. It is found that a (smooth) shift in the parameters can be mistaken for an integrated GARCH model when a standard fixed-parameter GARCH model is estimated from data generated by a TVGARCH model. As the size of the shift increases the indication for lack of weak stationarity (IGARCH) becomes more apparent.

To illustrate the TVGARCH model it is applied to both Affärsvärldens monthly share index (1919-1995) and the daily exchange rate between the US Dollar and the British Pound (1987-1997). Both series appear to lack weak stationarity in the conditional variance when a standard fixed-parameter GARCH model is fitted to them. The misspecification test detects time variation in the parameters of the standard fixed-parameter GARCH model for both series, and after this has been properly accounted for there is no evidence suggesting that the series are integrated. The Affärsvärldens share index is found to be less volatile between the mid 1920s and early 1970s than elsewhere during the observation period. For the exchange rate between the US dollars and the British Pound a structural shift is detected in September 1993. At that date, the unconditional variance for the corresponding GARCH process decreases rather abruptly.

5. A smooth transition autoregressive target zone model

A new nonlinear empirical target zone model, the smooth transition autoregressive target zone (STARTZ) model, is introduced in Chapter 5. The idea is to modify an AR-GARCH model to fit the target zone environment. This is done by introducing nonlinearities of smooth transition type near the boundaries of the target zone in both the conditional mean and the conditional variance. The model is illustrated by estimating the exchange rate of two Nordic currencies in the second half of the 1980s.

The basic target zone model introduced by Krugman (1991) has been followed

by a vast literature. The empirical implications of the theoretical model have been studied and rejected, see for example Bertola and Caballero (1992), Flood, Rose and Mathieson (1991) and Lindberg and Söderlind (1994). The empirical failure of the basic Krugman model has prompted some extensions. The model is constructed under two crucial assumptions: perfect credibility of the target zone and interventions only at the margin of the exchange rate band. In the second generation of target zone models these assumptions are relaxed. One extension of the target zone model has been to introduce imperfect credibility, see Bertola and Svensson (1993) and Bertola and Caballero (1992). The other extension is to allow for intramarginal interventions, see Froot and Obstfeld (1991) and Delgado and Dumas (1991). It is worth noting that all these target zone models imply mean reverting behaviour for the exchange rate.

The model proposed here is a rather flexible, empirical, time series model that is capable of capturing the dynamic behaviour of an exchange rate implied by various theoretical target zone models. The present framework also enables the modeller to estimate the bounds of an implicit band, should such a band exist. To model empirical data in a systematic way a coherent modelling strategy is the key, and such a strategy, including specification, estimation, and evaluation, is designed and applied to data in Chapter 5. An advantage of the proposed strategy is that the misspecification tests only require standard asymptotic theory and are easy to perform. These tests are obtained by modifying the corresponding tests suggested in Chapter 3.

The empirical examples indicate that there is structure in data that corresponds to theoretical target zone models. For the Swedish krona, 1985-91, the behaviour of the currency within the target zone is in line with what theory suggests for a currency for which the Central Bank intervenes intramarginally. As regards the Norwegian krone, the Central Bank of Norway intervened only at the edges of the band in 1986-88. The behaviour of the Norwegian krone according to the estimated model for this period agrees with the implications of the basic target zone model. For the Norwegian krone, 1988-90 it has been argued, Mundaca (1998), that there existed an unofficial implicit band. The STARTZ model estimated under this assumption leads one to conclude that the boundaries of the implicit band were soft, i.e. the band was not strictly enforced by the Central Bank of Norway.

6. Extensions

In the four chapters of the thesis the modelling cycles are constructed under the assumption that the errors are independently normally distributed. The normality assumption may not always be credible in high-frequency financial time series. If this assumption does not hold the parameter estimates may be obtained using a quasi-maximum likelihood framework, see Bollerslev and Wooldridge (1992). They also show that Lagrange Multiplier (LM)-type tests may easily be robustified against non-normality. The robust version makes use of a sandwich estimator based on the expectation of the outer product and the expectation of the Hessian, whereas the version that assumes normal errors makes use of the expectation of the Hessian. Extending the chapters in this thesis to accommodate scope with this possibility is a topic for further work.

One possible extension of the STAR-STGARCH model in Chapter 2 is to abandon the restriction that the parametrization of conditional variance does not depend on parametrization of the conditional mean. The time-varying GARCH in Chapter 4 may be extended in the same way as the TVSTAR model for the conditional mean in Lundbergh, Teräsvirta and van Dijk (1999) by allowing for both time-varying GARCH and smooth transition GARCH in the same parameterization. This will result in a very flexible parameterization of the conditional variance. The ideas presented in Chapter 5 regarding the target zone model may be applied to other areas as well where an explicit or implicit zone or boundary is assumed to exist. For example, in both unemployment series when the economy approaches full employment and interest rate series when the interest rate becomes sufficiently low, one might expect to find a lower bound of this kind.

References

- Anderson, H. M. and Teräsvirta, T. (1992). Characterizing nonlinearities in business cycles using smooth transition autoregressive models, *Journal of Applied Econometrics* **7**: S119–S136.
- Bacon, D. W. and Watts, D. G. (1971). Estimating the transition between two intersecting straight lines, *Biometrika* **58**: 525–534.
- Bertola, G. and Caballero, R. J. (1992). Target zones and realignments, *Review of Economic Studies* **82**: 520–526.
- Bertola, G. and Svensson, L. E. O. (1993). Stochastic devaluation risk and the empirical fit of target-zone models, *Review of Economic Studies* **60**: 689–712.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**: 307–327.
- Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994). ARCH models, in R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, Elsevier, Amsterdam, pp. 2959–3038.
- Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews* **11**: 143–172.
- Box, G. and Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco.
- Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models, *Journal of Time Series Analysis* **7**: 178–190.
- Chu, C.-S. J. (1995). Detecting parameter shift in GARCH models, *Econometric Reviews* **14**: 241–266.
- Delgado, F. and Dumas, B. (1991). Target zones, broad and narrow, in P. Krugman and M. Miller (eds), *Exchange Rate Targets and Currency Bands*, Cambridge University Press, Cambridge, pp. 35–56.
- Diebold, F. (1986). Modeling the persistence of conditional variances: A comment, *Econometric Reviews* **5**: 51–56.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**: 987–1007.

- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *Journal of Finance* **48**: 1749–1777.
- Fiorentini, G., Calzolari, G. and Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics* **11**: 399–417.
- Flood, R. P., Rose, A. K. and Mathieson, D. J. (1991). An empirical exploration of exchange-rate target-zones, *Carnegie-Rochester Series on Public Policy* **35**: 7–65.
- Franses, P. H. and van Dijk, D. (1999). Nonlinear error-correction models for interest rates in the Netherlands, in W. A. Barnett, D. F. Hendry, S. Hylleberg, T. Teräsvirta, D. Tjøstheim and A. Würtz (eds), *Nonlinear Econometric Modelling of Time Series*, Cambridge University Press, Cambridge. (forthcoming).
- Froot, K. A. and Obstfeld, M. (1991). Exchange-rate dynamics under stochastic regime shifts: A unified approach, *Journal of International Economics* **31**: 203–229.
- Ghysels, E., Harvey, A. C. and Renault, E. (1996). Stochastic volatility, in G. S. Maddala and C. R. Rao (eds), *Handbook of Statistics*, Vol. 14, Elsevier, Amsterdam, pp. 119–191.
- Glosten, L., Jagannathan, R. and Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**: 1779–1801.
- González-Rivera, G. (1998). Smooth transition GARCH models, *Studies in Nonlinear Dynamics and Econometrics* **3**: 61–78.
- Granger, C. W. J. and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oxford.
- Hagerud, G. E. (1997). *A New Non-Linear GARCH Model*, EFI, Stockholm School of Economics.
- Hendry, D. F. and Neale, A. J. (1991). A monte carlo study of the effects of structural breaks on tests for unit roots, in P. Hackl and A. H. Westlund (eds), *Economic Structural Change, Analysis and Forecasting*, Springer-Verlag, Berlin, pp. 95–119.

- Krugman, P. R. (1991). Target zones and exchange rate dynamics, *Quarterly journal of economics* **106**: 669–682.
- Lamoureux, C. G. and Lastrapes, W. G. (1990). Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**: 225–234.
- Lee, Y. N. and Li, W. K. (1998). On smooth transition double threshold models. Department of Statistics, University of Hong Kong, Research Report No. 198.
- Li, C. W. and Li, W. K. (1996). On a double threshold autoregressive heteroscedastic time series model, *Journal of Applied Econometrics* **11**: 253–274.
- Li, W. K. and Mak, T. K. (1994). On the squared residual autocorrelations in nonlinear time series with conditional heteroskedasticity, *Journal of Time Series Analysis* **15**: 627–636.
- Lin, C.-F. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**: 211–228.
- Lindberg, H. and Söderlind, P. (1994). Testing the basic target zone model on Swedish data 1982–1990, *European Economic Review* **38**: 1441–1469.
- Lundbergh, S., Teräsvirta, T. and van Dijk, D. (1999). Time-varying smooth transition regression models. Manuscript in preparation.
- Michael, P., Nobay, A. R. and Peel, D. A. (1997). Transactions costs and nonlinear adjustment in real exchange rates: An empirical investigation, *Journal of Political Economy* **105**: 862–879.
- Mundaca, G. (1998). The effect of interventions on the probabilities of realignments. Working Paper, Research Department, Central Bank of Norway, Oslo, Norway.
- Osborn, D. and Öcal, N. (1999). Business cycle nonlinearities in UK consumption and production, *Journal of Applied Econometrics* (forthcoming) .
- Pagan, A. (1996). The econometrics of financial markets, *Journal of Empirical Finance* **3**: 15–102.
- Palm, F. (1996). GARCH models of volatility, in G. Maddala and C. Rao (eds), *Handbook of Statistics*, Vol. 14, Elsevier, Amsterdam, pp. 209–240.
- Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis, *Econometrica* **57**: 1361–1401.

- Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes, *Journal of the American Statistical Association* **55**: 324–330.
- Skalin, J. and Teräsvirta, T. (1999). Another look at Swedish business cycles, 1861–1988, *Journal of Applied Econometrics* (forthcoming) .
- Taylor, M. P. and Sarno, L. (1999). Nonlinearities in European real exchange rate adjustment. Unpublished manuscript.
- Taylor, S. (1986). *Modelling Financial Time Series*, Wiley, Chichester.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**: 208–218.
- Teräsvirta, T. (1998). Modelling economic relationships with smooth transition regressions, in A. Ullah and D. E. A. Giles (eds), *Handbook of Applied Economic Statistics*, Dekker, New York, pp. 507–552.
- Tong, H. (1990). *Non-Linear Time Series. A Dynamical System Approach*, Oxford University Press, Oxford.
- Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes, *Journal of the American Statistical Association* **84**: 231–240.
- Tsay, R. S. (1998). Testing and modeling multivariate threshold models, *Journal of the American Statistical Association* **93**: 1188–1202.

Chapter 2

Modelling economic high-frequency time series with STAR-STGARCH models

1. Introduction

Modelling financial time series has recently received considerable attention. Many of these series are "high frequency", that is, they consist of weekly, daily or even intradaily observations. In this chapter, "high frequency" simply means that the conditional variance of the process is not constant over time. As many high frequency series show little or no linear dependence, the focus has been on modelling the conditional variance. The idea of conditional autoregressive heteroskedasticity (ARCH; Engle, 1982) has led to a large number of extensions of the original model and applications. Bollerslev, Engle and Nelson (1994) and Palm (1996) are examples of recent surveys of the area. Another variant of ARCH, the so-called Stochastic Volatility model (Taylor, 1986, p. 73ff.) has gained popularity recently, and the latest developments were surveyed in Ghysels, Harvey and Renault (1996). It has been argued, however, that despite the absence of linear dependence there may be nonlinear dependence in the conditional mean. This should then be appropriately modelled in order to avoid misspecification of the conditional variance. Tong (1990, p. 116) suggested combining the Self Exciting Threshold Autoregressive (SETAR) model for the conditional mean with an ARCH model for the conditional variance.

Li and Lam (1995) followed this suggestion and also devised a specification strategy for building SETAR-ARCH models. They applied their model to the daily Hong Kong Hang Seng stock index and reported nonlinearities in the conditional mean.

The original ARCH model and its most important extension, the generalized ARCH (GARCH), are symmetric: while the size of the shock matters, the sign does not. Many authors have argued that shocks may have asymmetric effects on volatility: the dynamic response to a positive shock is not necessarily the mirror image of the response to a negative shock of the same size. Pagan (1996), in his survey of developments in financial econometrics, provided a useful review of models that can handle this type of asymmetry. A natural idea would be to combine such a parameterization of the conditional variance with a nonlinear model for the conditional mean. Li and Li (1996) did exactly that by defining a double threshold autoregressive heteroskedastic (DTARCH) time series model. A DTARCH model has a SETAR-type conditional mean. The conditional variance is parameterized similarly, and the authors called their specification the threshold ARCH (TARCH) model. Note, however, that it differs from the TARCH model of Zakoïan (1994) in that the latter is a parameterization of the conditional standard deviation. Li and Li (1996) also provided a comprehensive modelling strategy for DTARCH models. It was based on the idea of ordered autoregressions which Tsay (1989) successfully applied to the specification of SETAR models. The authors fitted their DTARCH specification to the daily Hong Kong Hang Seng stock index.

Recently, Lee and Li (1998) generalized the DTARCH model by allowing the transition of the first and second regime to be smooth. They called this model the Smooth Transition Double Threshold model. In this chapter we follow Lee and Li (1998) by adopting the idea of smooth transition in the conditional mean which first appeared in Bacon and Watts (1971). This chapter may be seen as an extension of Lee and Li's work in the sense that we simultaneously allow a rather flexible specification for the conditional variance as well. In addition, misspecification testing will receive plenty of attention in this chapter. The conditional mean is thus specified as a Smooth Transition Autoregressive (STAR) model; see, for example, Chan and Tong (1986), Granger and Teräsvirta (1993) and Teräsvirta (1994). The conditional variance is specified as a Smooth Transition GARCH (STGARCH) model; see Hagerud (1997) and González-Rivera (1998). Our STGARCH model is a generalization of the GJR-GARCH model (Glosten, Jagannathan and Runkle, 1993) and the Generalized Quadratic ARCH model of Sentana (1995). It allows plenty of scope for explaining asymmetries in volatility. Our aim is to construct a

complete modelling cycle for our STAR-STGARCH family of models, consisting of three stages: specification, estimation and evaluation.

The plan of the chapter is as follows. We define the model in Section 2 and discuss its specification, estimation and evaluation in Sections 3 and 4. In Section 5 we apply the STAR-STGARCH model to daily returns of the Swedish OMX index and the daily JPY/USD exchange rate and study properties of one-step-ahead out-of-sample forecasts. Section 6 concludes.

2. The model

The logistic smooth transition version of the $AR(m)$ -GARCH(p,q) parameterization is a special case of the following additive nonlinear model in which the conditional mean has the following structure:

$$y_t = \boldsymbol{\varphi}' \mathbf{w}_t + f(\mathbf{w}_t; \boldsymbol{\theta}) + u_t \quad (2.1)$$

where $\boldsymbol{\varphi} = (\varphi_0, \varphi_1, \dots, \varphi_m)'$ is the parameter vector for the autoregressive part of the model and $\mathbf{w}_t = (1, y_{t-1}, \dots, y_{t-m})'$ is the corresponding lag vector. Function $f(\mathbf{w}_t; \boldsymbol{\theta})$ is nonlinear and assumed to be at least twice continuously differentiable for $\boldsymbol{\theta} \in \Theta$ everywhere in $\mathbf{w}_t \in \mathfrak{R}^m$. The error process of this model is parametrized as

$$u_t = \varepsilon_t \sqrt{h(\mathbf{w}_t, \boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\zeta})} \quad (2.2)$$

where $\{\varepsilon_t\} \sim \text{nid}(0, 1)$ and $h_t = h(\mathbf{w}_t, \boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\zeta}) = \boldsymbol{\eta}' \mathbf{z}_t + g(\mathbf{z}_t; \boldsymbol{\zeta})$ is the conditional variance not dependent on ε_t and positive for every t with probability one. Definition (2.2) precludes linear dependence in $\{u_t\}$. Setting $\boldsymbol{\eta} = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ and $\mathbf{z}_t = (1, u_{t-1}^2, \dots, u_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$, where $h_t > 0$ for all t with probability one, makes the linear part of (2.2) a GARCH(p,q) model. Furthermore, $u_t = y_t - \boldsymbol{\varphi}' \mathbf{w}_t - f(\mathbf{w}_t; \boldsymbol{\theta})$ such that neither $\boldsymbol{\varphi}$ nor $\boldsymbol{\theta}$ is assumed to depend on either $\boldsymbol{\eta}$ or $\boldsymbol{\zeta}$. Function $g(\mathbf{z}_t; \boldsymbol{\zeta})$ is nonlinear and at least twice continuously differentiable for $\boldsymbol{\zeta} \in \Xi$ everywhere in $\mathbf{z}_t \in \mathfrak{R}^{p+q+1}$. The normality assumption of errors $\{\varepsilon_t\}$ is not necessary but is retained for inference. It is also assumed that the moments necessary for the inference exist and that the parameters are subject to restrictions such that the model is stationary and ergodic. The usual restrictions imposed on $\boldsymbol{\eta}$ to ensure nonnegative conditional variance have been $\alpha_0 > 0, \alpha_j \geq 0, j = 1, \dots, q-1, \alpha_q > 0, \beta_j \geq 0, j = 1, \dots, p$. They can be relaxed as in Nelson and Cao (1992); see also He and Teräsvirta (1999b).

In order to define $f(\mathbf{w}_t; \boldsymbol{\theta})$ and $g(\mathbf{z}_t; \boldsymbol{\zeta})$, let

$$H_n(s_t; \gamma, \mathbf{c}) = \left(1 + \exp(-\gamma \prod_{l=1}^n (s_t - c_l)) \right)^{-1}, \gamma > 0, c_1 \leq \dots \leq c_n \quad (2.3)$$

where s_t is the transition variable, γ is a slope parameter and $\mathbf{c} = (c_1, \dots, c_n)'$ a location vector. The parameter restrictions $\gamma > 0$ and $c_1 \leq \dots \leq c_n$ are identifying restrictions. The value of the logistic function (2.3), which is bounded between a and 1, $0 \leq a \leq 1/2$ depends on the transition variable s_t . Note that for $\gamma = 0$, $H_n(s_t; \gamma, \mathbf{c}) \equiv 1/2$ and when $\gamma \rightarrow \infty$, and $n = 1$, $H_n(s_t; \gamma, \mathbf{c})$ becomes a step function. It becomes a "multistep" function as $\gamma \rightarrow \infty$, if $n > 1$.

In this chapter, $f(\mathbf{w}_t; \boldsymbol{\theta})$ is defined as a product of the logistic function (2.3) of order n and another linear combination including lags of y_t . Setting $\boldsymbol{\theta} = (\boldsymbol{\phi}', \gamma, \mathbf{c}')'$, the function can be written as

$$f(\mathbf{w}_t; \boldsymbol{\theta}) = \boldsymbol{\phi}' \mathbf{w}_t H_n(s_t; \gamma, \mathbf{c}) \quad (2.4)$$

Function (2.4) is bounded only in probability. In this chapter the transition variable is $s_t = y_{t-d}$ most of the time but $s_t = t$ is another important case. Other definitions such as s_t being a linear combination of variables are possible as well. By inserting (2.4) into (2.1) we obtain the LSTAR(n) model which by definition becomes linear if $\gamma = 0$. Setting $n = 2$ and $c_1 = c_2$ yields a model that closely approximates the exponential STAR or ESTAR model; see Teräsvirta (1994). STAR models are capable of characterizing asymmetric series and series with sudden upswings and downturns. Chappell and Peel (1998) recently showed that they can also generate realizations that appear chaotic. Bacon and Watts (1971) were the first to apply the idea of smooth transition to statistical modelling.

Our conditional variance specification is a generalization of the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). We make the transition between the extreme regimes smooth by assuming that $g(\mathbf{z}_t; \boldsymbol{\zeta})$ has the same logistic structure as $f(\mathbf{w}_t; \boldsymbol{\theta})$. This is a natural extension of the idea of smooth transition to modelling conditional variance. Thus by setting $\boldsymbol{\zeta} = (\boldsymbol{\alpha}^*, \delta, \mathbf{k}')$ where $\boldsymbol{\alpha}^* = (\alpha_{01}, \dots, \alpha_{0q}, \alpha_{21}, \dots, \alpha_{2q})'$, the nonlinear function $g(\mathbf{z}_t; \boldsymbol{\zeta})$ may be written as

$$g(\mathbf{z}_t; \boldsymbol{\zeta}) = \sum_{j=1}^q \alpha_{0j} H_n^*(u_{t-j}; \delta, \mathbf{k}) + \sum_{j=1}^q \alpha_{2j} H_n^*(u_{t-j}; \delta, \mathbf{k}) u_{t-j}^2 \quad (2.5)$$

In practice we restrict ourselves to cases $n^* = 1, 2$. No nonlinear structure is imposed

on h_{t-j} , $j = 1, \dots, p$, since the model is very flexible even without such an extension. The model can characterize processes with an asymmetric response to shocks with the same magnitude but opposite signs. With $n = 2$ and $k_1 = k_2$, a symmetric but nonlinear response may be characterized as well. The first sum on the right-hand side is a nonlinear variant of the corresponding structure in the GQARCH model of Sentana (1995). Conditions for positivity of the conditional variance are simpler than in the GQARCH model: $\alpha_0 > 0$, $\alpha_0 + \sum_{j=1}^q \alpha_{0j} > 0$, $\alpha_j \geq 0$, $\alpha_j + \alpha_{2j} \geq 0$, $j = 1, \dots, q$; $\beta_j \geq 0$, $j = 1, \dots, p$, form a set of sufficient conditions. We denote (2.5) inserted into (2.2) as the Logistic Smooth Transition GARCH (LSTGARCH(n^*)) model which by definition collapses into the standard GARCH model if $\delta = 0$. The nonlinear GARCH model in Hagerud (1997) may be viewed as a special case of this parametrization with $\alpha_{01} = \dots = \alpha_{0q} = 0$.

Assuming normally distributed errors, Engle (1982) showed that the information matrix of the conditional mean-ARCH model is block-diagonal if some regularity conditions hold and if the parameterization of the conditional variance is symmetric in the sense that the model responds similarly to positive and negative inputs of the same size. This in turn implies that if the conditional mean is estimated with a consistent estimator, the conditional variance can be estimated from the residuals of the conditional mean model without loss of asymptotic efficiency. The classical GARCH parameterization is symmetric and satisfies the regularity conditions, see Bollerslev (1986), so that the LSTAR-linear GARCH model has this property. But then, the general smooth transition GARCH model may not be symmetric, in which case the usual two-stage estimation strategy leads to consistent but not asymptotically efficient estimates. However, if $n^* = 2$ and $c_1 = -c_2$ in (2.3), the STGARCH model is symmetric.

3. Specification and estimation of a STAR-STGARCH model

The nonlinear STAR-STGARCH model defined in (2.1-2.5) is the most general parameterization considered in this chapter. It is nevertheless possible that the time series under consideration may be adequately characterized by a submodel of the general STAR-STGARCH one. For instance, the conditional mean may be linear or the conditional variance constant. Furthermore, even if we eventually select a general model there are still choices to be made that have to be based on the data. The delay d in the conditional mean usually has to be specified from the data as well as the maximum lag length and the type of transition function ($n = 1$ or 2).

We also have to select the lag length and the type of transition function ($n^* = 1$ or 2) in the STGARCH specification of the conditional variance. All this requires a coherent specification strategy such as, for example, in Box and Jenkins (1970), Li and Li (1996), Tsay (1989, 1998) and Teräsvirta (1994).

Our general rule is to specify the conditional mean first, followed by the conditional variance. The reason is that we may estimate the parameters of the conditional mean consistently even if the conditional variance is not specified, that is, even if it is assumed constant. On the other hand, it is not possible to estimate the parameters of the conditional variance consistently if the conditional mean is misspecified. The specification of the STAR-STGARCH model consists of the following stages:

1. Test linearity of conditional mean and, if rejected, choose d and n .
2. Estimate the parameters of the conditional mean assuming that the conditional variance remains constant and test the null hypothesis of no linear ARCH against ARCH of a given order. If the hypothesis of no ARCH is rejected, tentatively assume that the conditional variance follows a low-order standard GARCH process.
3. Estimate the parameters of the STAR-GARCH model and test the adequacy of the STAR (conditional mean) and the GARCH (conditional variance) specifications by various misspecification tests. If rejected, specify a STAR-STGARCH model.
4. Estimate the parameters of the STAR-STGARCH model and test the adequacy of both the conditional mean and the conditional variance of that specification by appropriate misspecification tests.
5. If the model passes the tests tentatively accept it. In the opposite case try another specification search or choose another family of models.

It should be noted that by following the above modelling scheme we proceed from restricted models to more general ones. This may be simpler than starting from the most general model and gradually reducing its size, but there is also a statistical rationale behind this choice of direction. If the conditional mean is linear then no STAR specification is identified. As for the conditional variance, the same is true for any STGARCH specification if the linear GARCH is already a valid parameterization. The lack of identification leads to lack of consistency in

the parameter estimation, which, in turn, is likely to create numerical difficulties in estimation. See Hansen (1996) for a recent discussion of this problem. To avoid estimating unidentified models we have to proceed from specific to general. In the following we consider the specification stages in detail.

3.1. Testing linearity of the conditional mean

We begin the modelling cycle with the specification of the conditional mean. In order to carry out the linearity tests we have to determine the maximum lag, m , of the linear AR model.

Following Teräsvirta (1994), linearity against a logistic STAR model of order n is tested with an LM-type test, where the null hypothesis is a linear AR model and the alternative an LSTAR(n) model. The null hypothesis is $\gamma = 0$. As mentioned above, function $f(\mathbf{w}_t; \boldsymbol{\theta})$ is not identified under the null hypothesis. To circumvent this problem, (2.4) is Taylor-expanded around $\gamma = 0$. Setting $s_t = y_{t-d}$ in (2.4), assuming $d \leq m$ without loss of generality, leads to

$$y_t = \boldsymbol{\pi}'_0 \mathbf{w}_t + \boldsymbol{\pi}'_1 \tilde{\mathbf{w}}_t y_{t-d} + \boldsymbol{\pi}'_2 \tilde{\mathbf{w}}_t y_{t-d}^2 + \dots + \boldsymbol{\pi}'_n \tilde{\mathbf{w}}_t y_{t-d}^n + R_1(\mathbf{w}_t; \boldsymbol{\theta}) + u_t \quad (3.1)$$

where $\tilde{\mathbf{w}}_t = (y_{t-1}, \dots, y_{t-m})'$, $\boldsymbol{\pi}_i$ is a function of γ such that $\boldsymbol{\pi}_i = \mathbf{0}$, $i = 1, \dots, n$, when $\gamma = 0$ and $R_1(\mathbf{w}_t; \boldsymbol{\theta})$ is the remainder. The new null hypothesis is thus $H_0 : \boldsymbol{\pi}_i = \mathbf{0}$, $i = 1, \dots, n$. Note that under H_0 , $R_1(\mathbf{w}_t; \boldsymbol{\theta}) \equiv 0$, so that the remainder does not affect the distribution theory when the test is based on the LM-principle. When the conditional variance is constant the LM-type statistic with an asymptotic χ^2 distribution (we assume that the necessary moments exist) under the null hypothesis can be computed by two auxiliary regressions. The F -version of the test is often recommended as it has better small sample properties than the χ^2 -version; see, for example, Granger and Teräsvirta (1993, p. 66). The sample sizes in the analysis of high-frequency series are usually so large, however, that in the present context this recommendation has no practical value. Both tests may be carried out in stages as follows.

1. Regress y_t on \mathbf{w}_t and compute the sum of squared residuals, SSR_0 .
2. Regress y_t on $\mathbf{w}_t, \tilde{\mathbf{w}}_t y_{t-d}, \tilde{\mathbf{w}}_t y_{t-d}^2, \dots, \tilde{\mathbf{w}}_t y_{t-d}^n$ and compute the sum of squared residuals, SSR_1 .
3. Compute the F -version of the test statistic $F = \frac{(SSR_0 - SSR_1)/mn}{SSR_1/(T - m(n+1) - 1)}$, or the χ^2 -version, $\chi^2 = T \frac{(SSR_0 - SSR_1)}{SSR_0}$

This linearity test assumes constant conditional variance, and is therefore not robust against conditional heteroskedasticity. If H_0 cannot be rejected then the conclusion is that the conditional mean is linear. The problem arises when H_0 is rejected because then we do not in principle know if this is because of nonlinearity in the conditional mean or because of conditional heteroskedasticity. However, when the heteroskedasticity is of GARCH type, the size of the test may in some cases be affected. This would suggest using a robust version of the linearity test such as that in Granger and Teräsvirta (1993, p. 69). We consider this possibility by a small simulation study whose results can be found in Appendix A. The results do show some size distortion in situations where the GARCH-type error process has fat tails, i.e. the kurtosis is high. On the other hand, the power simulations indicate that in those cases the robustification may remove most of the power, so that existing nonlinearity may remain undetected by a robustified linearity test. Consequently since our objective is to find and model any existing nonlinearity, including any in the conditional mean, robustification cannot be recommended. We expect to discover false rejections of the null hypothesis of linearity, due to heteroskedasticity of GARCH type, at the evaluation stage of model building.

In order to carry out the linearity test(s) we have to determine the order of the linear stationary AR model representing the conditional mean under the null hypothesis. Teräsvirta (1994) suggested that the order could be determined by an order selection criterion such as the AIC, see Akaike (1974). The problem is that for high-frequency economic series the usual order selection criterion would typically select a model with no lags because there is normally little or no linear dependence in the series. To avoid the problem, the maximum lag ($m \geq 6$ is used for daily observations) is fixed in advance. If the null hypothesis is not rejected we assume that the conditional mean is linear.

3.2. Specification and estimation of the conditional mean model and testing for ARCH

The specification of the STAR model for the conditional mean is carried out as follows. First, the linearity test above is used to select the delay parameter, d . This is followed by choosing n , that is, selecting the type of the STAR(n) model. After that, the estimated model is tested for ARCH in the error process.

The linearity test (3.1) is computed for different values of d , and the one for which the null hypothesis, $\gamma = 0$, has the smallest p -value is selected, see Teräsvirta (1994). This requires the smallest p -value to be lower than a pre-determined value

chosen by the researcher. After fixing the delay parameter the order, $n \leq 2$, of the LSTAR model is selected. We can choose between the LSTAR(1) and the LSTAR(2) model by testing a sequence of nested hypotheses. The sequence is defined within (3.1) assuming $n = 3$:

$$\begin{aligned} H_{04} : \pi'_3 &= 0 \\ H_{03} : \pi'_2 = 0 | \pi'_3 &= 0 \\ H_{02} : \pi'_1 = 0 | \pi'_2 = \pi'_3 &= 0. \end{aligned}$$

If H_{03} is most strongly rejected the LSTAR(2) model is selected. In the other two cases the choice is the LSTAR(1) model; for the rationale of this rule, see Teräsvirta (1994). This selection rule is not balanced as it sometimes has a tendency to favour the LSTAR(1) model. In practice this happens when the true model is LSTAR(2) and there are no (or very few) observations in one of the tails of the transition function. In such cases, the LSTAR(1) model is a good approximation to the LSTAR(2) model in the relevant subset of the sample space so that an erroneous choice has little practical significance. Escribano and Jordá (1996) recently presented a rule that purports to remedy this problem, and it could be applied here.

The idea of these selection rules has been to avoid estimating a possibly large number of nonlinear models, but with the steady increase in computational power which an average modeller has at his/her disposal this step is no longer crucial. If nonlinear estimation is not considered a time-consuming task one may simply estimate an LSTAR(n) model for $n = 1, 2$, and make the choice between the LSTAR(1) and LSTAR(2) families at the evaluation stage. This can be done by considering the estimation results themselves and the results of the misspecification tests, to be discussed in Section 4.1. The logistic part (2.3) of $f(w_t; \theta)$ should also be examined when determining n . If an element in the estimated location vector $c = (c_1, c_2)'$ of the LSTAR(2) model does not in practice affect the values of $f(w_t; \theta)$ in the sample then n may be reduced from two to one.

The estimation of the parameters of the STAR model is carried out by maximum likelihood. Our algorithm also checks if the size of the model can be reduced. The autoregressive parameters whose estimates are insignificant according to a predetermined level are removed using a backward elimination algorithm. In practice this is done by repetitively removing the parameter corresponding to the least significant (if nonsignificant) parameter estimate and reestimating the reduced model. This algorithm also considers restrictions of the form $\varphi_i = -\phi_j$, which are exclusion restrictions comparable to $\varphi_i = 0$. They allow an autoregressive parameter to be cancelled out smoothly in the transition between the two regimes. The backward

elimination terminates when all the remaining autoregressive parameter estimates are significant.

The assumption that the error sequence $\{u_t\}$ in (2.1) has a constant conditional variance is normally not realistic when modelling high-frequency financial series, and this has to be tested. We first test it against the alternative that $\{u_t\}$ follows an ARCH(s) process. As a GARCH(p, q) model can be adequately approximated by a long ARCH specification we choose a large s (≥ 8) for the alternative. Engle (1982) suggested an LM-test which is used here. The asymptotically equivalent test of McLeod and Li (1983) is another possibility. If the null hypothesis is not rejected and the estimated conditional mean model passes the appropriate evaluation tests the modelling sequence terminates. On the other hand, if the null is rejected, as we generally expect when dealing with high-frequency economic series, we continue by assuming that the conditional variance follows a low-order GARCH process.

3.3. Estimation of the STAR-GARCH model

If the conditional variance is not constant the next step is to fit a STAR-GARCH model to the data. The usual way of obtaining the estimates for the conditional mean and the conditional variance when the latter has a standard GARCH representation is to make use of the block-diagonality of the information matrix. The conditional mean model is estimated first. This is followed by the estimation of the conditional variance model using the residuals from estimating the conditional mean. This procedure yields consistent estimates, but in this chapter all parameters are ultimately estimated simultaneously. One advantage with simultaneous estimation is that it may lead to more parsimonious models, at least if the series are not very long. The two-step estimation has a tendency to yield over-parameterized models because some effects due to the nonconstant conditional variance may at first be captured by the estimated conditional mean. The autoregressive parameters turning out to be redundant are eliminated during joint estimation by applying the previous backward elimination algorithm. The estimation is carried out using analytical second derivatives, which gives numerically reliable estimates for the information matrix. This is needed at the evaluation stage when the estimated model is tested for misspecification.

However, the two-step estimation is useful for obtaining initial values for the joint estimation. We proceed as follows. First estimate the STAR model for the conditional mean and then estimate a GARCH(1, 1) model for the residuals. As a first-order GARCH model has very often been found to be adequate in practice, it

is only expanded if necessary. The decision to do this is based on a misspecification test of the functional form which together with other evaluation procedures is discussed in Section 4. To force the conditional variance generated by any higher-order GARCH model to be nonnegative, the constraints in Nelson and Cao (1992) for parameters of such models are imposed. The validity of restrictions constraining linear combinations of estimated parameters is verified after the estimation.

3.4. Specification and estimation of the STAR-STGARCH model

The STAR-GARCH model has to be subjected to misspecification tests. We postpone the discussion of such tests to Section 4. At this stage we assume that the linear GARCH specification is rejected in favour of STGARCH and proceed to discuss STAR-STGARCH models. We have to consider specification, estimation and evaluation of these models and begin with specification.

For the smooth transition type alternative (2.5) the problem of selecting the transition variable is not present. We only have to select the order, $n^* \leq 2$, of the logistic part in (2.5). One way of doing this is to apply a decision rule similar to that suggested for the STAR model. But then, one may instead simply estimate the STGARCH(n^*) model for $n^* = 1, 2$ and make the choice on the basis of the results, including the results of the misspecification tests in Section 4. The logistic function in $g(\mathbf{z}_t; \zeta)$ should also be examined when determining n^* . If an element in the estimated location vector $\mathbf{k} = (k_1, k_2)'$ of the STGARCH(2) model does not in practice affect the values of $g(\mathbf{z}_t; \zeta)$ in the sample, then n^* may be reduced from two to one. Furthermore, the same types of exclusion restrictions that were considered for the STAR model are relevant for the STGARCH model.

When estimating STAR-STGARCH models it is not certain that the conditional variance is symmetric with respect to the error terms (most often it is not) and therefore the assumption of block diagonality of the information matrix may not hold. This implies that the two-stage estimation algorithm does not yield asymptotically efficient estimates. We maintain our previous strategy and ultimately estimate the conditional mean and the conditional variance jointly.

4. Model evaluation

As discussed above, the validity of the assumptions used in the estimation of parameters must be investigated once the parameters of the STAR-STGARCH model (or a submodel) have been estimated. These assumptions include:

1. The errors and the squared (and standardized) errors of the model are not serially correlated.
2. The parameters of the model are constant.
3. The squared (and standardized) errors of the model are independent and identically distributed.

These assumptions are testable. Furthermore, it is useful to find out whether or not there are any nonlinearities left in the process after fitting a STAR-STGARCH model to the series under consideration. In the present chapter, this possibility is investigated by testing the hypothesis of no additive nonlinearity against this type of nonlinearity.

As for the three testable assumptions, the first two may be tested following Eitrheim and Teräsvirta (1996). These authors have also suggested a test of no additive nonlinearity for the conditional mean. We only have to generalize these tests to the case where the conditional variance follows a STGARCH process. As regards the independence hypothesis, the BDS test, see Brock, Dechert, Scheinkman and LeBaron (1996), is applicable if the number of observations is sufficiently large.

4.1. Misspecification tests

4.1.1. General

This section follows Eitrheim and Teräsvirta (1996) and Chapter 3. Consider the estimated additive STAR-STGARCH model as defined in (2.1) and (2.2). An additive extension of the model may be written as

$$\begin{aligned} y_t &= A(\mathbf{w}_t; \boldsymbol{\pi}_a) + \boldsymbol{\varphi}' \mathbf{w}_t + f(\mathbf{w}_t; \boldsymbol{\theta}) + u_t \\ u_t &= \varepsilon_t \sqrt{\boldsymbol{\eta}' \mathbf{z}_t + g(\mathbf{z}_t; \boldsymbol{\zeta}) + B(\mathbf{z}_t; \boldsymbol{\pi}_b)} \end{aligned} \quad (4.1)$$

where $f(\mathbf{w}_t; \boldsymbol{\theta})$ and $g(\mathbf{z}_t; \boldsymbol{\zeta})$ are defined in (2.4) and (2.5) respectively and $\{\varepsilon_t\}$ is a sequence of independent standard normal variables. Model (4.1) forms a unifying framework for our tests. Set $\boldsymbol{\omega} = (\boldsymbol{\varphi}', \boldsymbol{\theta}', \boldsymbol{\eta}', \boldsymbol{\zeta}')$ which comprises all the parameters of the model. Functions $A(\mathbf{w}_t; \boldsymbol{\pi}_a)$ and $B(\mathbf{z}_t; \boldsymbol{\pi}_b)$ are assumed twice continuously differentiable for all $\boldsymbol{\pi}_a$ and $\boldsymbol{\pi}_b$ everywhere in the corresponding sample spaces. For notational simplicity and without loss of generality we assume $A(\mathbf{w}_t; \boldsymbol{\pi}_a) \equiv 0$ for $\boldsymbol{\pi}_a = \mathbf{0}$ and $B(\mathbf{z}_t; \boldsymbol{\pi}_b) \equiv 0$ for $\boldsymbol{\pi}_b = \mathbf{0}$. Tests for various types of misspecification are obtained by different parameterizations of A and B . It is assumed that the

maximum likelihood estimator of ω is consistent and asymptotic normal under any null hypothesis to be considered, which implies that $\{y_t\}$ satisfies the regularity conditions for stationarity and ergodicity. Also, the necessary moments for $\{u_t\}$ that are required for the asymptotic distribution theory to work are assumed to exist. The null hypothesis of no additional structure is $H_0 : \pi_a = \mathbf{0}$ and $\pi_b = \mathbf{0}$. The Lagrange multiplier (or score) test statistic is defined as

$$\text{LM} = T \begin{pmatrix} \frac{1}{T} \sum \frac{\partial l_t}{\partial \pi_a} |_{H_0} \\ 0 \\ \frac{1}{T} \sum \frac{\partial l_t}{\partial \pi_b} |_{H_0} \end{pmatrix}' \widehat{\mathbf{I}}(\pi_a, \widehat{\omega}, \pi_b)^{-1} |_{H_0} \begin{pmatrix} \frac{1}{T} \sum \frac{\partial l_t}{\partial \pi_a} |_{H_0} \\ 0 \\ \frac{1}{T} \sum \frac{\partial l_t}{\partial \pi_b} |_{H_0} \end{pmatrix} \quad (4.2)$$

where $\widehat{\mathbf{I}}$ is a consistent estimator of the information matrix under the null hypothesis. We use the estimated negative expectation of the Hessian as our estimator of the information matrix. The partial derivatives forming the Hessian may be found in Appendix B. The test statistic (4.2) is asymptotically χ^2 -distributed with $\dim(\pi_a) + \dim(\pi_b)$ degrees of freedom under the null hypothesis. If the information matrix is block-diagonal the test statistic may be computed simply by two artificial regressions; see Chapter 3. This approach does not always apply to the STAR-STGARCH model, because the GARCH component may be asymmetric. Note that by letting either $A(\mathbf{w}_t; \pi_a)$ or $B(\mathbf{z}_t; \pi_b)$ be identically equal to zero also under the alternative hypothesis amounts to testing the conditional mean and the conditional variance specifications separately. From the modeller's point of view, this is often the most practical alternative. The above structure may also be used for evaluating submodels within the STAR-STGARCH parameterization. For example, by setting $g(\mathbf{z}_t; \zeta) \equiv 0$ in (4.1) we can, among other things, test a STAR-GARCH specification against a STAR-STGARCH model nesting the former.

4.1.2. Test against serial dependence

To test the joint null hypothesis of no serial dependence in either the conditional mean or the conditional variance or both, the alternative is stated as remaining serial dependence, of order p in the ordinary error process and of order p^* in the squared (and standardized) errors. In the general case, this gives the extended model (4.1) with $A(\mathbf{w}_t; \pi_a) = \pi_a' \mathbf{v}_t$ and $B(\mathbf{w}_t; \pi_b) = \pi_b' \mathbf{v}_t^*$ where $\pi_a = (\pi_{a,1}, \dots, \pi_{a,p})'$, $\mathbf{v}_t = (u_{t-1}, \dots, u_{t-p})'$, $\pi_b = (\pi_{b,1}, \dots, \pi_{b,p^*})'$, and $\mathbf{v}_t^* = (h_{t-1}, \dots, h_{t-p^*})'$. The null hypothesis of no remaining serial dependence in either the conditional mean or the conditional variance is equivalent to $\pi_a = \mathbf{0}$ and $\pi_b = \mathbf{0}$. Under the null hypothesis

and assuming that the necessary moments exist, the LM-statistic (4.2) is asymptotically χ^2 -distributed with $\dim(\boldsymbol{\pi}_a) + \dim(\boldsymbol{\pi}_b)$ degrees of freedom. The details of an LM-test of this hypothesis for the squared standardized errors are given in Chapter 3.

4.1.3. Test against nonconstant parameters

We assume that the alternative to constant parameters in either the conditional mean or the conditional variance or both is that the parameters change smoothly over time, see Lin and Teräsvirta (1994) and Chapter 3. This gives rise to the following model:

$$\begin{aligned} y_t &= \boldsymbol{\varphi}(t)' \mathbf{w}_t + f(\mathbf{w}_t; \boldsymbol{\theta}(t)) + u_t \\ u_t &= \varepsilon_t \sqrt{\boldsymbol{\eta}(t)' \mathbf{z}_t + g(\mathbf{z}_t; \boldsymbol{\zeta}(t))} \end{aligned} \quad (4.3)$$

where $\boldsymbol{\theta}(t) = (\boldsymbol{\phi}(t)', \gamma, \mathbf{c}')'$ and $\boldsymbol{\zeta}(t) = (\boldsymbol{\psi}(t)', \delta, \mathbf{k}')'$. All time-varying parameter vectors are assumed to have the same structure. For instance, $\boldsymbol{\varphi}(t) = \boldsymbol{\varphi}^* + \lambda_\varphi \bar{H}_{n_\varphi}(t; \gamma_\varphi, \mathbf{c}_\varphi)$ where $\bar{H}_{n_\varphi} = H_{n_\varphi} - \frac{1}{2}$, subtracting 1/2 is a notational convenience in deriving the test and do not affect the generality of the argument. The transition function $H_{n_\varphi}(t; \gamma_\varphi, \mathbf{c}_\varphi)$ is a logistic function of order n_φ defined in (2.3) with $s_t \equiv t$. If $\gamma_\varphi \rightarrow \infty$ while $n_\varphi = 1$, function $H_1(t; \gamma_\varphi, \mathbf{c}_\varphi)$ becomes a step function and the alternative to parameter constancy is a single structural break. The null hypothesis of parameter constancy can be stated as $H_0 : \gamma_\varphi = \gamma_\eta = \gamma_\psi = 0$ under which $\boldsymbol{\varphi}^* = \boldsymbol{\varphi}$, $\boldsymbol{\phi}^* = \boldsymbol{\phi}$, $\boldsymbol{\eta}^* = \boldsymbol{\eta}$ and $\boldsymbol{\psi}^* = \boldsymbol{\psi}$. We can then write this alternative as a special case of (4.1) by setting $A(\mathbf{w}_t; \boldsymbol{\pi}_a) = (\lambda'_\varphi \mathbf{w}_t) \bar{H}_{n_\varphi}(t; \gamma_\varphi, \mathbf{c}_\varphi) + (\lambda'_\phi \mathbf{w}_t) H_{n_\phi}(s_t; \hat{\gamma}, \hat{\mathbf{c}}) \bar{H}_{n_\phi}(t; \gamma_\phi, \mathbf{c}_\phi)$ and $B(\mathbf{w}_t; \boldsymbol{\pi}_b) = (\lambda'_\eta \mathbf{w}_t) \bar{H}_{n_\eta}(t; \gamma_\eta, \mathbf{c}_\eta) + (\lambda'_\psi \mathbf{w}_t) H_{n^*}(s_t; \hat{\delta}, \hat{\mathbf{k}}) \bar{H}_{n^*}(t; \gamma_\psi, \mathbf{c}_\psi)$. For notational simplicity we denote the logistic functions for the estimated STAR-STGARCH model with $H_n = H_n(s_t; \hat{\gamma}, \hat{\mathbf{c}})$ and $H_{n^*} = H_{n^*}(s_t; \hat{\delta}, \hat{\mathbf{k}})$. We circumvent the identification problem under the null hypothesis as before by expanding $\bar{H}_{n_i}(t; \gamma_i, \mathbf{c}_i)$ into a Taylor series around $\gamma_i = 0$. A realistic assumption in a test situation is to assume that $n_\varphi = n_\phi = l_1$ and $n_\eta = n_\psi = l_2$. Using the first-order expansion we obtain, after reparameterization, the extended model (4.1) with $A(\mathbf{w}_t; \boldsymbol{\pi}_a) = \tilde{\boldsymbol{\pi}}_a' \mathbf{v}_t^a + R_1(\mathbf{w}_t; \boldsymbol{\pi}_a, \boldsymbol{\omega})$ and $B(\mathbf{w}_t; \boldsymbol{\pi}_b) = \tilde{\boldsymbol{\pi}}_b' \mathbf{v}_t^b + R_2(\mathbf{z}_t; \boldsymbol{\pi}_b, \boldsymbol{\omega})$ where $\tilde{\boldsymbol{\pi}}_a = (\boldsymbol{\varkappa}'_{a,1}, \dots, \boldsymbol{\varkappa}'_{a,l_1}, \boldsymbol{\varkappa}'_{a,l_1+2}, \dots, \boldsymbol{\varkappa}'_{a,2l_1+1})'$, $\tilde{\boldsymbol{\pi}}_b = (\boldsymbol{\varkappa}'_{b,1}, \dots, \boldsymbol{\varkappa}'_{b,l_2}, \boldsymbol{\varkappa}'_{b,l_2+2}, \dots, \boldsymbol{\varkappa}'_{b,2l_2+1})'$, $\mathbf{v}_t^a = (\mathbf{w}_t t, \dots, \mathbf{w}_t t^{l_1}, \mathbf{w}_t t H_n, \dots, \mathbf{w}_t t^{l_1} H_n)'$ and $\mathbf{v}_t^b = (\mathbf{z}_t t, \dots, \mathbf{z}_t t^{l_2}, \mathbf{z}_t t H_{n^*}, \dots, \mathbf{z}_t t^{l_2} H_{n^*})'$.

The joint null hypothesis of parameter constancy in both the conditional mean

and the conditional variance consists of the restrictions $\tilde{\pi}_a = \mathbf{0}$ and $\tilde{\pi}_b = \mathbf{0}$. Furthermore, then the two remainder terms $R_1(\mathbf{w}_t; \boldsymbol{\pi}_a, \boldsymbol{\omega}) \equiv R_2(\mathbf{z}_t; \boldsymbol{\pi}_b, \boldsymbol{\omega}) \equiv 0$ so that they do not affect the asymptotic distribution theory. Under the null hypothesis, the LM-type teststatistic (4.2) is asymptotically χ^2 -distributed with $\dim(\tilde{\pi}_a) + \dim(\tilde{\pi}_b)$ degrees of freedom. Again, it is useful to test the constancy of the parameters of the conditional mean and the conditional variance separately. More details about the test of the latter hypothesis and its finite-sample properties can be found in Chapter 3. These tests may also be applied to a given subset of parameters by assuming that the remaining ones are constant even under the alternative hypothesis. This often helps locate the nonconstancy if it exists.

4.1.4. Test against remaining nonlinearity

As we are searching for an adequate nonlinear specification for our time series it is of considerable interest to try and check whether or not our estimated parametrization adequately characterizes all nonlinearity in the series. To keep things simple, we focus on the null hypothesis of no remaining additive nonlinearity. The alternative to this null hypothesis is assumed to be an additive smooth transition component of the same type as before. This alternative may be written as

$$\begin{aligned} y_t &= A(\mathbf{w}_t; \boldsymbol{\pi}_a) + \boldsymbol{\varphi}' \mathbf{w}_t + f(\mathbf{w}_t; \boldsymbol{\theta}) + u_t \\ u_t &= \varepsilon_t \sqrt{\boldsymbol{\eta}' \mathbf{z}_t + g(\mathbf{z}_t; \boldsymbol{\zeta}) + B(\mathbf{z}_t; \boldsymbol{\pi}_b)} \end{aligned} \quad (4.4)$$

where the functions $A(\mathbf{w}_t; \boldsymbol{\pi}_a) = f_a(\mathbf{w}_t; \boldsymbol{\pi}_a)$ and $B(\mathbf{z}_t; \boldsymbol{\pi}_b) = g_b(\mathbf{z}_t; \boldsymbol{\pi}_b)$ represent any remaining nonlinearity and have almost the same parameter structure as (2.4) and (2.5). The only difference is that $\bar{H}_n = H_n - \frac{1}{2}$ is used instead of H_n . Set $\boldsymbol{\pi}_a = (\boldsymbol{\phi}'_a, \gamma_a, \mathbf{c}'_a)'$ and $\boldsymbol{\pi}_b = (\boldsymbol{\psi}'_b, \delta_b, \mathbf{k}'_b)'$. The null hypothesis of no remaining nonlinearity can be written as $H_0 : \gamma_a = \delta_b = 0$. Again, functions $f_a(\mathbf{w}_t; \boldsymbol{\pi}_a)$ and $g_b(\mathbf{z}_t; \boldsymbol{\pi}_b)$ are Taylor-expanded to circumvent the identification problem, and we assume that the orders of the smooth transitions are l_1 and l_2 , respectively. This yields, after reparameterization, the transformed model (4.1) with $A(\mathbf{w}_t; \boldsymbol{\pi}_a) = \tilde{\boldsymbol{\pi}}'_a \mathbf{v}_t^a + R_3(\mathbf{w}_t; \boldsymbol{\pi}_a, \boldsymbol{\omega})$ where $\tilde{\boldsymbol{\pi}}_a = (\boldsymbol{\varkappa}'_{a,0}, \boldsymbol{\varkappa}'_{a,1}, \dots, \boldsymbol{\varkappa}'_{a,l_1})'$ and $\mathbf{v}_t^a = (\mathbf{w}_t^{lin}, \mathbf{w}_t^{nc} y_{t-c}, \dots, \mathbf{w}_t^{nc} y_{t-c}^{l_1})'$. Any linear terms (lags) that are not included in the estimated smooth transition AR part of the model are denoted by \mathbf{w}_t^{lin} whereas $\mathbf{w}_t = (1, (\mathbf{w}_t^{nc})')'$. We assume that the nonlinearity in the conditional variance part is of smooth transition type and therefore given by $B(\mathbf{z}_t; \boldsymbol{\pi}_b) = \tilde{\boldsymbol{\pi}}'_b \mathbf{v}_t^b + R_4(\mathbf{z}_t; \boldsymbol{\pi}_b, \boldsymbol{\omega})$ where $\tilde{\boldsymbol{\pi}}_b = (\boldsymbol{\varkappa}_{b,1}, \dots, \boldsymbol{\varkappa}_{b,(1+l_2)q})'$ and $\mathbf{v}_t^b = (u_{t-1}, u_{t-1}^3, \dots, u_{t-1}^{l_2+2}, \dots, u_{t-q}, u_{t-q}^3, \dots, u_{t-q}^{l_2+2})'$.

The two remainders $R_3(\mathbf{w}_t; \boldsymbol{\pi}_a, \boldsymbol{\omega})$ and $R_4(\mathbf{z}_t; \boldsymbol{\pi}_b, \boldsymbol{\omega})$ do not affect the distribution theory because both are identically equal to zero under H_0 . The null hypothesis is $\tilde{\boldsymbol{\pi}}_a = \mathbf{0}$ and $\tilde{\boldsymbol{\pi}}_b = \mathbf{0}$, under which the LM-type teststatistic (4.2) is asymptotically χ^2 -distributed with $\dim(\tilde{\boldsymbol{\pi}}_a) + \dim(\tilde{\boldsymbol{\pi}}_b)$ degrees of freedom. As for the conditional mean, the test statistic is the same as in Eitrheim and Teräsvirta (1996). The corresponding test for the adequacy of the conditional variance is discussed in detail in Chapter 3.

If we assume $g_b(\mathbf{z}_t; \boldsymbol{\pi}_b) = f(\mathbf{w}_t; \boldsymbol{\theta}) \equiv 0$ in (4.4) and only test $H_0 : \gamma_a = 0$ then the test is a linearity test for the conditional mean under nonconstant GARCH-type conditional variance. Wong and Li (1997) recently presented a linearity test for a related situation where the conditional mean follows a threshold autoregressive process under the alternative hypothesis.

4.2. Test of independent, identically distributed disturbances

The misspecification tests of the previous section are tests against a parametric alternative. This is useful in model building because a rejection or a set of rejections normally contain information about the nature of the misspecification. Nevertheless, we may complete our set of tests by a general test of independence of the (standardized) errors. The BDS statistic (Brock et al., 1996) appears to be a suitable large-sample test for this purpose. It is based on the m^{th} correlation integral which represents the fraction of all possible pairs of m consecutive points of the series that are closer than $\varepsilon > 0$ to each other. Closeness ε is defined by the Euclidean norm. Let T denote the number of observations and $C_{m,T}(\varepsilon)$ the m^{th} correlation integral. The test statistic is

$$W_{m,T}(\varepsilon) = \sqrt{T} \frac{C_{m,T}(\varepsilon) - C_{1,T}(\varepsilon)^m}{\hat{\sigma}_{m,T}(\varepsilon)} \quad (4.5)$$

where $\hat{\sigma}_{m,T}(\varepsilon)$ is an estimate of the standard deviation under the null hypothesis. Statistic (4.5) has an asymptotic $N(0, 1)$ distribution under this hypothesis. For details see Brock et al. (1996). In this chapter, we have used the C-code of LeBaron (1997) to compute the values of the test statistic.

The null distribution of the statistic depends on the two nuisance parameters, m and ε . In this chapter the BDS statistic is computed with $m = 2$ and ε is chosen as the standard deviation of the residual series. These nuisance parameters cause problems in small samples because then the size of the test is a function of these two parameters. In order to have the size of the test under control we perform a

simple bootstrap to generate an empirical null distribution of the test statistic for the selected combination of m and ε . However, at the sample sizes (≥ 1000) we are primarily interested in, this precaution seems no longer necessary.

5. Application to two high-frequency series

As an illustration we apply the STAR-STGARCH modelling strategy to two series consisting of daily observations. The first series is the Swedish OMX index which consists of the values of the 30 most traded stocks at the Stockholm Stock Exchange. The observation period is December 30, 1983 to September 30, 1998, with a total of 3693 observations. The period until October 4, 1994 is used for estimating the model, whereas the remaining period of 1000 observations is reserved for studying the predictive properties of the model. The second series is the exchange rate for the Japanese yen (JPY) against the US dollar (USD). The observations extend from December 28, 1978 to September 30, 1997, a total of 4756 observations. The period until December 31, 1991 is the estimation period, and the remaining period of 1460 observations is used for a study of the predictive properties of the estimated STAR-STGARCH model.

Both series are transformed to percentage changes in continuously compounded rates. This is done by differencing the logarithms of the original series. Plots of the closing-bid and the percentage changes for both series can be found in Figures C.1-C.4.

5.1. Swedish OMX index

If the errors of a linear model are IID, then the skewness and leptokurtosis may be directly characterized by a nonsymmetric fat-tailed distribution; see for instance Mittnik and Rachev (1993). In such a case there is no need for further modelling. To consider linear dependence in the OMX series, we specified a linear AR model. This AR model was augmented by weekday and holiday dummies to account for possible weekday and holiday effects. An LM-test of no autocorrelation of order at most m , Breusch and Pagan (1980), was then applied to the residuals and there was no evidence of remaining serial dependence. On the other hand, the errors were found to be leptokurtic. Characteristics of the errors of the linear model can be found in Table C.1. The BDS test applied to the residuals from the linear model heavily rejected the null hypothesis of IID errors; see Table C.1. This rejection indicates that there may be some nonlinear structure to be modelled in the OMX

index.

5.1.1. Estimation and evaluation

Following the specification strategy outlined above a linearity test, assuming constant conditional variance, was performed where the maximum lag of the linear autoregressive part was 7 and the alternative was an LSTAR(3) model. The results of this linearity test for delay parameter values $0 < d \leq 5$ appear in Table C.1. The test strongly rejects linearity. The strongest rejection occurs at $d = 1$, which is the delay we select. The STAR-STGARCH model we consider has the following form:

$$\begin{aligned}
 y_t &= D_{Mo}Mo + D_{Tu}Tu + D_{We}We + D_{Th}Th + D_{Hol}Hol & (5.1) \\
 &+ \varphi_0 + \varphi_1 y_{t-1} + \dots + \varphi_4 y_{t-4} \\
 &+ (\phi_0 + \phi_1 y_{t-1} + \dots + \phi_4 y_{t-4})(1 + \exp(-\gamma(y_{t-d} - c)/\hat{\sigma}(y)))^{-1} + \varepsilon_t \sqrt{h_t} \\
 h_t &= \alpha_0 + \alpha_{11} u_{t-1}^2 + \beta_1 h_{t-1} & (5.2) \\
 &+ (\alpha_{01} + \alpha_{21} u_{t-1}^2)(1 + \exp(-\delta(u_{t-1} - k)/\hat{\sigma}(u)))^{-1}
 \end{aligned}$$

where $\hat{\sigma}(y)$ and $\hat{\sigma}(u)$ are the sample standard deviations of $\{y_t\}$ and $\{u_t\}$ making γ and δ scale-free.

We first estimated the STAR-STGARCH model ($\alpha_{01} = \alpha_{21} = 0$ in (5.2)) for the returns and tested it against the STAR-STGARCH specification. The results in Table C.5 show that the standard GARCH(1,1) model for the conditional variance is inadequate so that an extension to smooth transition GARCH was necessary. Model (5.2) was thus re-estimated without the symmetry restrictions, and the maximum likelihood estimates (standard deviations in parentheses) of the parameters based on analytical first and second derivatives are reported in Table C.2. A missing value in Table C.2 means that the corresponding parameter in (5.1) or (5.2) has been set to zero.

A few characteristic features of the standardized residuals of the estimated model can be found in Table C.3. The standardized residuals are less leptokurtic than the original observations. The results of the misspecification tests for the conditional mean of the STAR-STGARCH model can be found in Table C.4. The null hypothesis of the LM-test of no remaining autocorrelation cannot be rejected and the results of the parameter constancy tests indicate that parameters are constant, except for the daily dummies. This rejection has not been followed up, however. As for the LM-test of no additional nonlinearity against additional nonlinearity of

LSTAR(3) type, the hypothesis of no additional nonlinearity cannot be rejected. This is remarkable given the very low p -values of the linearity tests.

The results of the misspecification tests for the conditional variance of the model appear in Table C.6. The null hypothesis of no remaining multiplicative ARCH structure, which is asymptotically equivalent to the test of Li and Mak (1994), in the squared and standardized errors cannot be rejected. The test of the functional form indicates no remaining serial dependence of GARCH type. There is still some evidence of nonlinearity in the conditional variance, see Table C.6, but it is now very weak compared to the previous results in Table C.5. The parameter constancy test does not indicate nonconstancy at the 1% level of significance. Finally, it can be seen from Table C.3 that the hypothesis of IID errors cannot be rejected for the standardized errors of the STAR-STGARCH model.

5.1.2. Interpretation

Having obtained a satisfactory model for the OMX index we proceed to interpret the estimation results. A conspicuous feature in the model is that both a Monday and a Tuesday effect seem to exist at a 5 percent significance level. In general there appears to be a weak tendency for the index to display growth towards the end of each week. No holiday effect is found. The conditional mean model is asymmetric. A sufficiently large negative shock causes a secondary negative effect after four days. For positive shocks there is no similar positive secondary effect because $\hat{\varphi}_4 + \hat{\theta}_4$ changes sign and becomes slightly negative as the value of the transition function increases towards unity. Figure C.5 shows that both extreme regimes are invoked quite often. As to the conditional variance, the STGARCH model responds asymmetrically to positive and negative errors from the conditional mean part of the model. For positive residuals the behaviour is locally represented by a standard GARCH model, but for negative residuals the local dynamics are close to the dynamic behaviour of an IGARCH process.

Under the assumption of constant conditional variance, a model spectrum for each value of the observed logistic transition function in $f(w_t; \hat{\theta})$ is plotted in Figure C.6. This 'sliced' spectrum was introduced by Skalin and Teräsvirta (1999). It is a model spectrum conditional on the value of the transition function and describes the change in the local dynamics of the conditional mean with the transition from one of the two extreme regimes to the other in the estimated STAR model. The slight peak at the frequency corresponding to four days reflects the situation at $\hat{H}_1 = 0$. This peak fades away as $\hat{H}_1 \rightarrow 1$, and thus we obtain another description of the

asymmetry in the conditional mean discussed above. Note, however, that the sliced spectrum characterizes local behaviour and cannot be interpreted as representing the "global" dynamic behaviour of the series. If one wants to estimate the global spectrum of the conditional mean process it has to be done numerically; see Skalin and Teräsvirta (1999) for discussion.

It should also be noted that the conditional variance of the model is clearly asymmetric. Positive and negative shocks, i.e. news, of the same size do not have the same impact on the conditional variance.

5.1.3. Forecasting one business day ahead

We computed 1000 one-day-ahead predictions of the conditional mean. The root mean square error (RMSE) equals 0.010 for the conditional mean part of the estimated STAR-STGARCH model while the corresponding linear model has an RMSE of 0.012. We evaluate the predictive properties of the model in two ways. First we compare the forecasts from our model with those from a linear AR(7) model by applying the test in Granger and Newbold (1986, pp. 278-280) which is based on the correlation coefficient, r , of the sums and differences of the forecast errors. If this correlation coefficient equals zero then both forecast processes have the same RMSE. The null hypothesis of no correlation is $H_0 : r = 0$ and the alternative $H_1 : r > 0$. The alternative corresponds to the situation where the forecasts from the nonlinear model have a smaller RMSE than those from the linear model. For inference we use the well-known transformation

$$w = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) \quad (5.3)$$

where $w \sim N(0, \frac{1}{N-3})$ under the null hypothesis, N being equal to the number of one-step-ahead forecasts. The test provides very strong evidence of a positive correlation, see Table C.7. It seems that the nonlinear structure captured by the model is useful in forecasting one step ahead.

Another way of evaluating the one-step-ahead forecasts is to compare them with the true outcomes using ordinal data. Positive returns and forecasts that are greater than 0.002 are given the value 1. If a forecast or realization is less than -0.002 it is given the value -1. The remaining observations have the value zero. This classification may be useful if we think that an agent facing transaction costs only acts upon a forecast if it deviates sufficiently from zero. It may then be of interest to consider prediction accuracy in such a framework. The ordinal data obtained this

way are cross-tabulated in Table C.8. Following Agresti (1984, pp. 156-165), the association between the one-step-ahead forecasts and the corresponding outcome is measured using concordant and discordant pairs of observations. For example, a forecast-outcome combination (1,1) forms a concordant pair with any (0,0), (0,-1), (-1,0) or (-1,-1) combination. The discordant pairs are those where one of the two elements is higher and the other lower than the corresponding element in the other observation forming the pair. The remaining paired observations are ties. Two measures of association that are based on the differences between concordant and discordant pairs, Somers' d and Kendall's tau- b , equal 0.65 and 0.66, respectively. This indicates a positive association between the sign of the one-step-ahead forecasts and the true outcome. An ordinal test of independence based on concordant and discordant pairs can be found in Agresti (1984, pp. 180-181). If we denote the number of concordant and discordant pairs by C and D the test statistic is

$$z = \frac{C - D}{\sigma(C - D)} \quad (5.4)$$

where $\sigma(C - D)$ is the standard error of $C - D$. Under the null hypothesis of no association the statistic (5.4) is approximately normally distributed in large samples with zero mean and unit variance. For the OMX index the value of this statistic equals $\hat{z} = 22.9$, which is strong evidence in favour of positive ordinal association between the forecasts and outcomes.

5.2. JPY/USD exchange rate

Here again, the analysis was started by estimating an AR model including weekday dummies and a constant term for the series. No remaining serial dependence was found. Various characteristics of the (leptokurtic) errors for the linear models can be found in Table C.1 as well as the result that the BDS test heavily rejects the null hypothesis of IID errors.

5.2.1. Estimation and evaluation

As before, testing linearity was the first step of the specification procedure. In these tests the maximum lag of the linear autoregressive part was 7 and the alternative was an LSTAR(3) model. The results from these linearity tests for delay parameter $0 < d \leq 5$ appear in Table C.1, and show that linearity is rejected.

As before, we first estimated a STAR-GARCH model without asymmetry in the conditional variance. As in the linear model, the weekday effects in the conditional

mean are represented by daily dummies and a dummy for holiday effects. The parameter estimates (standard deviations in parentheses) of the model are reported in Table C.2 and show that linearity is rejected. A few characteristic features of the standardized residuals of the estimated model can be found in Table C.3. The leptokurtosis has increased compared to that in the original series. This is because with our univariate model we cannot predict the actual shock, but we can model the average response to it. The actual shock may thus become more conspicuous in the standardized residuals than in the original data, which leads to increased leptokurtosis in the residuals. The condition for the error process having a finite fourth moment, given in Bollerslev (1986) or, more generally, He and Teräsvirta (1999a) is valid for the estimated parameter combination.

The results of the misspecification tests in Section 4 for the conditional mean can be found in Table C.4. There is no remaining autocorrelation, the parameters seem constant, and there is no evidence of remaining additional nonlinearity either. The results of the misspecification tests for the conditional variance appear in Table C.6. The test against remaining multiplicative ARCH structure indicates that there is some structure left that the STAR-GARCH model does not capture. On the other hand, the test of the functional form does not indicate remaining serial dependence of GARCH type and the results of the parameter constancy test are satisfactory. The BDS test in Table C.3 does not reject the IID hypothesis of the errors. Finally, the linearity test against smooth transition GARCH in Table C.6 does not reject the null of symmetry. Thus we tentatively accept the STAR-GARCH model and do not consider the STAR-STGARCH one.

5.2.2. Interpretation

A feature of the results is that the coefficient estimates of all weekday dummies are negative and significant at a 5 percent level which suggests a positive "Friday effect". The conditional mean of the model is asymmetric. Most of the time there exists some linear structure in the process. However, a sufficiently large positive shock causes a nonlinear response in the series. This is clearly seen from the sliced spectrum in Figure C.8 which demonstrates the emergence of a local cycle with a period of about 8 days. On the other hand, such a large shock is a relatively rare event. This is best seen from the graph of the transition function in Figure C.7 where every circle represents a single observation. As discussed above, there is no asymmetry to be modelled in the conditional variance.

5.2.3. Forecasting one business day ahead

We computed 1460 one-step-ahead predictions of the returns of the JPY/USD exchange rate both with a linear and a nonlinear model. We did not find any linear dependence in the conditional mean and assume therefore that in the linear case the process is a random walk. The RMSE for the conditional mean part of the STAR-GARCH model and the deviation of the actual observations from zero are both 0.0070. We computed the Granger and Newbold (1986) RMSE test using statistic (5.3). It is seen from Table C.7 that we cannot reject the null hypothesis $H_0 : r = 0$ against $H_1 : r > 0$. As nonlinearity is only required to characterize the response of the process to large shocks, there is no general improvement in the predictive performance compared to the random walk model.

We use the same ordinal observations as when comparing one-step-ahead forecasts with the true outcomes. The boundaries for "zero" are now -0.0005 and 0.0005. These ordinal data are cross-tabulated in Table C.9. The Somers' d and the Kendall's tau- b equal 0.02 and 0.01 respectively, and do not suggest any association between the direction of the one-step-ahead forecasts and that of the true outcome. For the JPY/USD exchange rate the value of the ordinal test statistic (5.4) is $\hat{z} = 0.57$, which strengthens this conclusion.

6. Conclusions

Our STAR-STGARCH model is intended to help us characterize the behaviour of high-frequency economic time series. Many modellers of such series tend to ignore the first moment, but in this chapter the first and the second moment are modelled jointly. A coherent modelling strategy is a key to doing this in a systematic way, and such a strategy is designed and applied to data here. An advantage of the proposed strategy is that the specification and misspecification tests we use only require standard asymptotic theory and are easy to perform. The tests for the conditional variance are discussed in detail in the Chapter 3.

The empirical examples indicate that there is nonlinear structure in the conditional mean to be modelled. In the case of the OMX index this leads to improved forecasts. For the JPY/USD exchange rate return series the nonlinear parameters only characterize some extreme events in the series. Because such events by definition are rare, the forecast accuracy, when measured from a large number of forecasts, is not improved by extra parameters. The aftermath of a large positive shock is the only occasion on which the estimated STAR-STGARCH model may

generate better forecasts than a linear autoregressive model for this series.

References

- Agresti, A. (1984). *Analysis of Ordinal Categorical Data*, Wiley, New York.
- Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions of Automatic Control* **19**: 716–723.
- Bacon, D. W. and Watts, D. G. (1971). Estimating the transition between two intersecting straight lines, *Biometrika* **58**: 525–534.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**: 307–327.
- Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994). ARCH models, in R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, Elsevier, Amsterdam, pp. 2959–3038.
- Box, G. and Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco.
- Breusch, T. S. and Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies* **47**: 239–253.
- Brock, W. A., Dechert, W., Scheinkman, J. and LeBaron, B. (1996). A test for independence based on the correlation dimension, *Econometric Reviews* **15**: 197–235.
- Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models, *Journal of Time Series Analysis* **7**: 178–190.
- Chappell, D. and Peel, D. A. (1998). A note on some properties of the ESTAR model, *Economics Letters* **60**: 311–315.
- Eitrheim, Ø. and Teräsvirta, T. (1996). Testing the adequacy of smooth transition autoregressive models, *Journal of Econometrics* **74**: 59–75.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**: 987–1007.
- Escribano, A. and Jordá, O. (1996). Improved testing and specification of smooth transition regression models. Universidad Carlos III de Madrid, Working Paper.

- Ghysels, E., Harvey, A. C. and Renault, E. (1996). Stochastic volatility, in G. S. Maddala and C. R. Rao (eds), *Handbook of Statistics*, Vol. 14, Elsevier, Amsterdam, pp. 119–191.
- Glosten, L., Jagannathan, R. and Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**: 1779–1801.
- González-Rivera, G. (1998). Smooth transition GARCH models, *Studies in Non-linear Dynamics and Econometrics* **3**: 61–78.
- Granger, C. W. J. and Newbold, P. (1986). *Forecasting Economic Time Series*, Academic Press, London and San Diego.
- Granger, C. W. J. and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oxford.
- Hagerud, G. E. (1997). *A New Non-Linear GARCH Model*, EFI, Stockholm School of Economics.
- Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* **64**: 413–430.
- He, C. and Teräsvirta, T. (1999a). Fourth moment structure of the GARCH(p, q) process, *Econometric Theory* (forthcoming) .
- He, C. and Teräsvirta, T. (1999b). Properties of the autocorrelation function of squared observations for second order GARCH processes under two sets of parameter constraints, *Journal of Time Series Analysis* **20**: 23–30.
- LeBaron, B. (1997). A fast algorithm for the BDS statistic, *Studies in Nonlinear Dynamics and Econometrics* **2**: 53–59.
- Lee, Y. N. and Li, W. K. (1998). On smooth transition double threshold models. Department of Statistics, University of Hong Kong, Research Report No. 198.
- Li, C. W. and Li, W. K. (1996). On a double threshold autoregressive heteroscedastic time series model, *Journal of Applied Econometrics* **11**: 253–274.
- Li, W. K. and Lam, K. (1995). Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model, *The Statistician* **44**: 333–341.

- Li, W. K. and Mak, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity, *Journal of Time Series Analysis* **15**: 627–636.
- Lin, C.-F. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**: 211–228.
- McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations, *Journal of Time Series Analysis* **4**: 269–273.
- Mittnik, S. and Rachev, S. T. (1993). Modeling asset returns with alternative stable distributions, *Econometric Reviews* **12**: 261–330.
- Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics* **10**: 229–235.
- Pagan, A. (1996). The econometrics of financial markets, *Journal of Empirical Finance* **3**: 15–102.
- Palm, F. (1996). GARCH models of volatility, in G. Maddala and C. Rao (eds), *Handbook of Statistics*, Vol. 14, Elsevier, Amsterdam, pp. 209–240.
- Sentana, E. (1995). Quadratic ARCH models, *Review of Economic Studies* **62**: 639–661.
- Skalin, J. and Teräsvirta, T. (1999). Another look at Swedish business cycles, 1861–1988, *Journal of Applied Econometrics* (forthcoming) .
- Taylor, S. (1986). *Modelling Financial Time Series*, Wiley, Chichester.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**: 208–218.
- Tong, H. (1990). *Non-Linear Time Series. A Dynamical System Approach*, Oxford University Press, Oxford.
- Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes, *Journal of the American Statistical Association* **84**: 231–240.
- Tsay, R. S. (1998). Testing and modeling multivariate threshold models, *Journal of the American Statistical Association* **93**: 1188–1202.

Wong, C. S. and Li, W. K. (1997). Testing for threshold autoregression with conditional heteroscedasticity, *Biometrika* **84**: 407–418.

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

A. The linearity test in the presence of GARCH

The linearity test in Section 3 is derived under the assumption of constant variance. The aim with this simple simulation study is to examine how the test performs when the conditional variance follows a GARCH-type process.

We let the conditional mean follow an AR(4) model, with parameter values that correspond to one of the estimated regimes found in the Swedish OMX index. This yields the conditional mean model $y_t = 0.0055 - 0.038y_{t-4} + u_t$ where $u_t = \sqrt{h_t}\varepsilon_t$. The IID error term, ε_t , is assumed $N(0, 1)$. In the simulations we let the conditional variance follow the asymmetric GJR-GARCH specification suggested by Glosten et al. (1993). This model may be written as

$$h_t = \alpha_0 + \alpha_1 [|u_{t-1}| + \omega u_{t-1}]^2 + \beta_1 h_{t-1} \tag{A.1}$$

where ω is the asymmetry parameter. If $\omega = 0$ the model reduces to the standard GARCH(1,1) model, and with $\alpha_0 > 0, \alpha_1 = \beta_1 = \omega = 0$ the model has a constant conditional variance. The values of the parameters match the estimated ones for the Swedish OMX-index. The parameter values for the different conditional variance DGPs can be found in Table A.1, where a missing value denotes that the corresponding parameter value in (A.1) equals zero.

DGP	Parameter				
	α_0	α_1	β_1	$\alpha_1 + \beta_1$	ω
DGP0	2×10^{-4}
DGP1	7×10^{-6}	0.15	0.84	0.99	.
DGP2	7×10^{-6}	0.06	0.84	0.90	.
DGP3	7×10^{-6}	0.15	0.84	0.90 or 0.99	-0.10

Table A.1: Simulation design for the GARCH model (A.1).

The number of replications in the simulation study is set to 5000. The length of the generated time series is 1000 observations after removing the first 500 observations from the beginning of the series to eliminate the effects of the initial values. For each replicate we compute two versions of the linearity test (3.1), one described in Section 3 and another one mentioned in Granger and Teräsvirta (1993, p. 69) which is robust against unspecified heteroskedasticity.

The empirical size of the standard linearity test can be found in Table A.2. When the conditional variance is generated by DGP2 it is found that the size is only marginally affected. As regards DGP1 and DGP3, the standard linearity test quite often erroneously detects nonlinear structure in the conditional mean, whereas

the corresponding robust version is slightly undersized.

$T=1000$	Standard version			Robust version		
	1 %	5 %	10 %	1 %	5 %	10 %
DGP0	0.008	0.043	0.085	0.004	0.032	0.069
DGP1	0.166	0.276	0.356	0.003	0.023	0.062
DGP2	0.011	0.054	0.100	0.004	0.035	0.079
DGP3	0.189	0.312	0.391	0.003	0.024	0.063

Table A.2: Empirical size of the linearity test. Each cell represents the proportion of rejections at the given nominal significance level. The alternative to linearity is a nonlinearity of LSTAR(3)-type. The transition variable used in the linearity test is y_{t-1} .

To see how the two versions of the linearity test behave when a nonlinearity is present in the conditional mean we generated data from an LSTAR model whose parameter values match the estimated ones for the Swedish OMX index, see Table C.2. The model is

$$y_t = 0.072y_{t-4} + (0.0055 - 0.11y_{t-4})(1 + \exp(-6.1(y_{t-1} - 0.0039)))/\hat{\sigma}(y))^{-1} + \sqrt{h_t}\varepsilon_t$$

where $\hat{\sigma}(y) = 0.013$. The conditional variance is generated by the DGPs in Table A.1. In this situation we find that the robustification against heteroskedasticity partly absorbs the nonlinearity in the conditional mean, see Table A.3. In the case when the conditional variance follows DGP1 or DGP3 the robust version of the linearity test has very little power against the nonlinearity in the conditional mean. Because of this disadvantage we shall not apply the robust linearity test in this chapter.

$T=1000$	Standard version			Robust version		
	1 %	5 %	10 %	1 %	5 %	10 %
DGP0	0.10	0.26	0.38	0.065	0.22	0.36
DGP1	0.81	0.90	0.93	0.014	0.073	0.15
DGP2	0.40	0.63	0.73	0.13	0.33	0.49
DGP3	0.82	0.91	0.94	0.012	0.069	0.15

Table A.3: Empirical power of the linearity test. Two versions are computed, the standard linearity test for the conditional mean and a version robust against unspecified heteroskedasticity. The alternative to linearity is a nonlinearity of LSTAR(3)-type. The transition variable used in the linearity test is y_{t-1} .

B. Gradient and Hessian of the log-likelihood function

The analytical gradient and the analytical Hessian of the STAR-GARCH model are reported here. The specification tests in Section 4.1 consist of, at most, two additional linear terms and are thus straightforward to implement.

Consider the STAR-STGARCH model defined in (2.1) and (2.2),

$$\begin{aligned} y_t &= \boldsymbol{\varphi}' \mathbf{w}_t + f(\mathbf{w}_t; \boldsymbol{\theta}) + u_t \\ u_t &= \varepsilon_t \sqrt{h(\mathbf{w}_t; \boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\zeta})} \end{aligned}$$

where $h(\mathbf{w}_t, \boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\zeta}) = h_t = \boldsymbol{\eta}' \mathbf{z}_t + g(\mathbf{z}_t; \boldsymbol{\zeta})$. The nonlinear functions $f(\mathbf{w}_t; \boldsymbol{\theta})$ and $g(\mathbf{z}_t; \boldsymbol{\zeta})$ are defined in (2.4) and (2.5) respectively. Assuming that the sequence $\{\varepsilon_t\}$ is identically normal distributed, the log-likelihood function at time t is

$$l_t = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{u_t^2}{h_t}$$

where $u_t = y_t - \boldsymbol{\varphi}' \mathbf{w}_t - f(\mathbf{w}_t; \boldsymbol{\theta})$, since neither $\boldsymbol{\varphi}$ nor $\boldsymbol{\theta}$ is assumed to depend on either $\boldsymbol{\eta}$ or $\boldsymbol{\zeta}$. The derivatives of the log-likelihood function are reported in B.1 and the derivatives of the nonlinear functions $f(\mathbf{w}_t; \boldsymbol{\theta})$ and $g(\mathbf{z}_t; \boldsymbol{\zeta})$ are reported in B.2 and B.3.

B.1. Partial derivatives of the log-likelihood function

The first and second order partial derivatives are to be found in this section. The second order partial derivatives are also given in expectation, which is useful when computing the specification tests. Also recall that $E[u_t] = 0$ and $E[u_t^2] = E[g_t] =$ unconditional variance.

B.1.1. First order partial derivative of l_t

The gradient of the log-likelihood function at time t is given by the derivative with respect to the parameters of the nonlinear function, the linear part and the variance model.

$$G_t = \left(\frac{\partial l_t}{\partial \boldsymbol{\theta}'} \quad \frac{\partial l_t}{\partial \boldsymbol{\varphi}'} \quad \frac{\partial l_t}{\partial \boldsymbol{\eta}'} \quad \frac{\partial l_t}{\partial \boldsymbol{\zeta}'} \right)$$

where

$$\frac{\partial l_t}{\partial \boldsymbol{\theta}'} = \frac{u_t}{h_t} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\theta}'}$$

$$\begin{aligned}\frac{\partial l_t}{\partial \boldsymbol{\varphi}'} &= \frac{u_t}{h_t} \mathbf{w}'_t + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \\ \frac{\partial l_t}{\partial \boldsymbol{\eta}'} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\eta}'} \\ \frac{\partial l_t}{\partial \boldsymbol{\zeta}'} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\zeta}'}.\end{aligned}$$

B.1.2. Second order partial derivative of l_t

The Hessian of the log-likelihood function at time t is given by:

$$H_t = \begin{pmatrix} \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\theta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\zeta}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\zeta}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\theta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\zeta}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\theta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}'} \end{pmatrix}.$$

We first report the derivatives with respect only to the parameters in the conditional mean block, followed by the derivatives with respect only to the parameters of the conditional variance block. Then we report the cross-term derivatives.

The derivative with respect to the parameters of the linear part in the conditional mean

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} &= -\frac{1}{h_t} \mathbf{w}_t \mathbf{w}'_t - \frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \boldsymbol{\varphi}} \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} - \frac{u_t}{h_t^2} \left(\frac{\partial h_t}{\partial \boldsymbol{\varphi}} \mathbf{w}'_t + \mathbf{w}_t \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \right) \\ &\quad + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} - \frac{1}{g_t} \frac{\partial h_t}{\partial \boldsymbol{\varphi}} \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \right)\end{aligned}$$

and the expectation of it,

$$E \left[\frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} \right] = -E \left[\frac{1}{h_t} \mathbf{w}_t \mathbf{w}'_t + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\varphi}} \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \right].$$

The derivative with respect to the parameters of the nonlinear function in the conditional mean

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} &= -\frac{1}{h_t} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{u_t}{h_t} \left(\frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} - \frac{u_t}{2h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\theta}} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right) \\ &\quad - \frac{u_t}{h_t^2} \left(\frac{\partial h_t}{\partial \boldsymbol{\theta}} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right) \\ &\quad + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} - \frac{1}{g_t} \frac{\partial h_t}{\partial \boldsymbol{\theta}} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right)\end{aligned}$$

and the expectation of it,

$$E \left[\frac{\partial^2 l_t}{\partial \theta \partial \theta'} \right] = -E \left[\frac{1}{h_t} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\theta}} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right].$$

The cross-term derivative within the conditional mean

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \varphi \partial \boldsymbol{\theta}'} &= -\frac{1}{h_t} \mathbf{w}_t \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} - \frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} - \frac{u_t}{h_t^2} \left(\frac{\partial h_t}{\partial \varphi} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \mathbf{w}_t \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right) \\ &\quad + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \varphi \partial \boldsymbol{\theta}'} - \frac{1}{h_t} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right) \end{aligned}$$

and the expectation of it,

$$E \left[\frac{\partial^2 l_t}{\partial \varphi \partial \boldsymbol{\theta}'} \right] = -E \left[\frac{1}{h_t} \mathbf{w}_t \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \boldsymbol{\theta}'} \right].$$

The derivative with respect to the parameters of the linear part of the conditional variance

$$\frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} = -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \boldsymbol{\eta}} \frac{\partial h_t}{\partial \boldsymbol{\eta}'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} - \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\eta}} \frac{\partial h_t}{\partial \boldsymbol{\eta}'} \right)$$

and the expectation of it,

$$E \left[\frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} \right] = -E \left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\eta}} \frac{\partial h_t}{\partial \boldsymbol{\eta}'} \right].$$

The derivative with respect to the parameters of the nonlinear part of the conditional variance

$$\frac{\partial^2 l_t}{\partial \zeta \partial \zeta'} = -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \zeta} \frac{\partial h_t}{\partial \zeta'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \zeta \partial \zeta'} - \frac{1}{h_t} \frac{\partial h_t}{\partial \zeta} \frac{\partial h_t}{\partial \zeta'} \right)$$

and the expectation of it,

$$E \left[\frac{\partial^2 l_t}{\partial \zeta \partial \zeta'} \right] = -E \left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \zeta} \frac{\partial h_t}{\partial \zeta'} \right].$$

The cross-term derivative within the conditional variance

$$\frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \zeta'} = -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \boldsymbol{\eta}} \frac{\partial h_t}{\partial \zeta'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \boldsymbol{\eta} \partial \zeta'} - \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\eta}} \frac{\partial h_t}{\partial \zeta'} \right)$$

and the expectation of it,

$$E \left[\frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \zeta'} \right] = -E \left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\eta}} \frac{\partial h_t}{\partial \zeta'} \right].$$

Cross-term derivatives between the conditional mean and the conditional variance:

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \varphi \partial \eta'} &= -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \eta'} - \frac{u_t}{h_t^2} \mathbf{w}_t' \frac{\partial h_t}{\partial \eta'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \varphi \partial \eta'} - \frac{1}{h_t} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \eta'} \right) \\ \frac{\partial^2 l_t}{\partial \varphi \partial \zeta'} &= -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \zeta'} - \frac{u_t}{h_t^2} \mathbf{w}_t' \frac{\partial h_t}{\partial \zeta'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \varphi \partial \zeta'} - \frac{1}{h_t} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \zeta'} \right) \\ \frac{\partial^2 l_t}{\partial \theta \partial \eta'} &= -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \eta'} - \frac{u_t}{h_t^2} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \theta} \frac{\partial h_t}{\partial \eta'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \theta \partial \eta'} - \frac{1}{g_t} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \eta'} \right) \\ \frac{\partial^2 l_t}{\partial \theta \partial \zeta'} &= -\frac{u_t^2}{2h_t^3} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \zeta'} - \frac{u_t}{h_t^2} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \theta} \frac{\partial h_t}{\partial \zeta'} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \left(\frac{\partial^2 h_t}{\partial \theta \partial \zeta'} - \frac{1}{g_t} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \zeta'} \right)\end{aligned}$$

and the corresponding expectations of them,

$$\begin{aligned}E \left[\frac{\partial^2 l_t}{\partial \varphi \partial \eta'} \right] &= -E \left[\frac{1}{2g_t^2} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \eta'} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \varphi \partial \zeta'} \right] &= -E \left[\frac{1}{2g_t^2} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \zeta'} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \theta \partial \eta'} \right] &= -E \left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \eta'} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \theta \partial \zeta'} \right] &= -E \left[\frac{1}{2h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \zeta'} \right].\end{aligned}$$

If the conditional variance model is symmetric and satisfies certain regularity conditions then the expectation of these matrices will be zero; see Theorem 4 in Engle (1982).

B.2. Partial derivatives of the nonlinear function $f(\mathbf{w}_t; \boldsymbol{\theta})$

As neither φ nor $\boldsymbol{\theta}$ is assumed to depend on either $\boldsymbol{\eta}$ or $\boldsymbol{\zeta}$, the partial derivatives of the conditional mean do not depend on the parameterization of the conditional variance. Assume that the nonlinear function $f(\mathbf{w}_t; \boldsymbol{\theta})$ in the conditional mean is parameterized as in (2.4):

$$f(\mathbf{w}_t; \boldsymbol{\theta}) = \boldsymbol{\phi}' \mathbf{w}_t \left(1 + \exp(-\gamma \prod_{l=1}^n (y_{t-d} - c_l)) \right)^{-1}.$$

To simplify the calculation of the derivatives rewrite function $f(\mathbf{w}_t; \boldsymbol{\theta})$ as:

$$f(\mathbf{w}_t; \boldsymbol{\theta}) = \boldsymbol{\phi}' \mathbf{w}_t \frac{1}{2 \exp(-\frac{\gamma}{2} \prod_{l=1}^n (y_{t-d} - c_l)) \cosh(\frac{\gamma}{2} \prod_{l=1}^n (y_{t-d} - c_l))}$$

Let $\xi_t(y_{t-d}; \gamma, c_1, \dots, c_n) = \frac{\gamma}{2} \prod_{l=1}^n (y_{t-d} - c_l)$, drop the arguments and write the

function as

$$f(\mathbf{w}_t; \boldsymbol{\theta}) = \boldsymbol{\phi}' \mathbf{w}_t \frac{e^{\xi_t}}{2 \cosh \xi_t}.$$

B.2.1. First order partial derivative of $f(\mathbf{w}_t; \boldsymbol{\theta})$

The derivative with respect to the parameters of the autoregressive part, the slope and the location vector:

$$\frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left(\frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\phi}'} \quad \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma} \quad \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c}'} \right)$$

where

$$\begin{aligned} \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\phi}'} &= \mathbf{w}_t' \frac{e^{\xi_t}}{2 \cosh \xi_t} \\ \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma} &= \frac{\boldsymbol{\phi}' \mathbf{w}_t}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \gamma} \\ \frac{\partial f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c}'} &= \frac{\boldsymbol{\phi}' \mathbf{w}_t}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \mathbf{c}'} \end{aligned}$$

By letting c_i denote the i^{th} element in the location vector \mathbf{c} , we can write the first order derivative of ξ_t as

$$\frac{\partial \xi_t}{\partial \gamma} = \frac{1}{2} \prod_{l=1}^n (y_{t-d} - c_l), \quad \frac{\partial \xi_t}{\partial c_i} = -\frac{\gamma}{2} \prod_{l=1, l \neq i}^n (y_{t-d} - c_l).$$

B.2.2. Second order partial derivative of $f(\mathbf{w}_t; \boldsymbol{\theta})$

These derivatives are only interesting when computing the full Hessian; they are not needed for the expectation.

$$\frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{pmatrix} \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'} & \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \gamma} & \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \mathbf{c}'} \\ \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma \partial \boldsymbol{\phi}'} & \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma \partial \gamma} & \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma \partial \mathbf{c}'} \\ \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c} \partial \boldsymbol{\phi}'} & \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c} \partial \gamma} & \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c} \partial \mathbf{c}'} \end{pmatrix} \quad (\text{B.1})$$

where the elements in (B.1) are given by:

$$\begin{aligned} \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'} &= 0 \\ \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma \partial \boldsymbol{\phi}'} &= \frac{\mathbf{w}_t'}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \gamma} \\ \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c} \partial \boldsymbol{\phi}'} &= \frac{\mathbf{w}_t'}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \mathbf{c}} \\ \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \gamma \partial \gamma} &= \frac{\boldsymbol{\phi}' \mathbf{w}_t}{\cosh^2 \xi_t} \left(\frac{1}{2} \frac{\partial^2 \xi_t}{\partial \gamma \partial \gamma} - \frac{\partial \xi_t}{\partial \gamma} \frac{\partial \xi_t}{\partial \gamma} \tanh \xi_t \right) \end{aligned}$$

$$\begin{aligned}\frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c} \partial \gamma} &= \frac{\boldsymbol{\phi}' \mathbf{w}_t}{\cosh^2 \xi_t} \left(\frac{1}{2} \frac{\partial^2 \xi_t}{\partial \mathbf{c} \partial \gamma} - \frac{\partial \xi_t}{\partial \mathbf{c}} \frac{\partial \xi_t}{\partial \gamma} \tanh \xi_t \right) \\ \frac{\partial^2 f(\mathbf{w}_t; \boldsymbol{\theta})}{\partial \mathbf{c} \partial \mathbf{c}'} &= \frac{\boldsymbol{\phi}' \mathbf{w}_t}{\cosh^2 \xi_t} \left(\frac{1}{2} \frac{\partial^2 \xi_t}{\partial \mathbf{c} \partial \mathbf{c}'} - \frac{\partial \xi_t}{\partial \mathbf{c}} \frac{\partial \xi_t}{\partial \mathbf{c}'} \tanh \xi_t \right).\end{aligned}$$

Let c_i, c_j denote the i^{th} and j^{th} element respectively of the location vector \mathbf{c} . The second order derivative of ξ_t is then

$$\frac{\partial^2 \xi_t}{\partial \gamma \partial \gamma} = 0, \quad \frac{\partial^2 \xi_t}{\partial \gamma \partial c_i} = -\frac{1}{2} \prod_{l=1, l \neq i}^n (y_{t-d} - c_l), \quad \frac{\partial^2 \xi_t}{\partial c_j \partial c_i} = \frac{\gamma}{2} \prod_{l=1, l \neq i, l \neq j}^n (y_{t-d} - c_l).$$

B.3. Partial derivative of the conditional variance h_t .

Assume that the conditional variance, $h(w_t; \boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\zeta}) = h_t$, is parameterized as in (2.2) and (2.5):

$$\begin{aligned}h_t &= \boldsymbol{\eta}' \mathbf{z}_t + g(\mathbf{z}_t; \boldsymbol{\zeta}) \\ &= \alpha_{01} + \sum_{j=1}^q \alpha_{0j} H_{n^*}(u_{t-j}) + \sum_{j=1}^q (\alpha_{1j} + \alpha_{2j} H_{n^*}(u_{t-j})) u_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j}.\end{aligned}$$

To initialize, the iterative computation pre-sample values of h_t are estimated by the sample (unconditional) variance. This is done for all $t \leq 0$ by setting $h_t = u_t^2 = \frac{1}{T} \sum_{s=1}^T u_s^2$ where $u_s = y_s - \boldsymbol{\varphi}' \mathbf{w}_s - f(\mathbf{w}_s; \boldsymbol{\theta})$.

B.3.1. First order partial derivative of h_t

The first order derivatives may be computed iteratively by using the following expressions

$$\begin{aligned}\frac{\partial h_t}{\partial \boldsymbol{\theta}'} &= \sum_{j=1}^q (\alpha_{0j} + \alpha_{2j} u_{t-j}^2) \frac{\partial H_{n^*}(u_{t-j})}{\partial \boldsymbol{\theta}'} \\ &\quad - 2 \sum_{j=1}^q (\alpha_{1j} + \alpha_{2j} H_{n^*}(u_{t-j})) u_{t-j} \frac{\partial f(\mathbf{w}_{t-j}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\theta}'} \\ \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} &= \sum_{j=1}^q (\alpha_{0j} + \alpha_{2j} u_{t-j}^2) \frac{\partial H_{n^*}(u_{t-j})}{\partial \boldsymbol{\varphi}'} \\ &\quad - 2 \sum_{j=1}^q (\alpha_{1j} + \alpha_{2j} H_{n^*}(u_{t-j})) u_{t-j} \mathbf{w}'_{t-j} + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\varphi}'} \\ \frac{\partial h_t}{\partial \boldsymbol{\eta}'} &= \mathbf{z}'_t + \sum_{j=1}^q (\alpha_{0j} + \alpha_{2j} u_{t-j}^2) \frac{\partial H_{n^*}(u_{t-j})}{\partial \boldsymbol{\eta}'}\end{aligned}$$

$$\begin{aligned} & + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \eta'} \\ \frac{\partial h_t}{\partial \zeta'} & = \sum_{j=1}^q \frac{\partial(\alpha_{0j} + \alpha_{2j} u_{t-j}^2)}{\partial \zeta'} + \sum_{j=1}^q (\alpha_{0j} + \alpha_{2j} u_{t-j}^2) \frac{\partial H_{\eta^*}(u_{t-j})}{\partial \zeta'} \\ & + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \zeta'} \end{aligned}$$

where the pre-sample values are given by $-2u_t \frac{\partial f(\mathbf{w}_t; \theta)}{\partial \theta'} = \frac{\partial h_t}{\partial \theta'} = -\frac{2}{T} \sum_{s=1}^T u_s \frac{\partial f(\mathbf{w}_s; \theta)}{\partial \theta'}$,
 $-2u_t \mathbf{w}'_t = \frac{\partial h_t}{\partial \varphi'} = -\frac{2}{T} \sum_{s=1}^T u_s \mathbf{w}'_s$, $\frac{\partial h_t}{\partial \eta'} = 0$.

B.3.2. Second order partial derivative of $h_t = h(w_t; \varphi, \theta, \eta, \zeta)$

$$\frac{\partial^2 h_t}{\partial \theta \partial \theta'} = 2 \sum_{i=1}^q \alpha_i \left(\frac{\partial f(\mathbf{w}_{t-i}; \theta)}{\partial \theta} \frac{\partial f(\mathbf{w}_{t-i}; \theta)}{\partial \theta'} - u_{t-i} \frac{\partial^2 f(\mathbf{w}_{t-i}; \theta)}{\partial \theta \partial \theta'} \right) + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}}{\partial \theta \partial \theta'}$$

where the pre-sample values are given by $2 \left(\frac{\partial f(\mathbf{w}_t; \theta)}{\partial \theta} \frac{\partial f(\mathbf{w}_t; \theta)}{\partial \theta'} - u_{t-i} \frac{\partial^2 f(\mathbf{w}_t; \theta)}{\partial \theta \partial \theta'} \right) =$
 $\frac{\partial^2 h_t}{\partial \theta \partial \theta'} = \frac{2}{T} \sum_{s=1}^T \left(\frac{\partial f(\mathbf{w}_s; \theta)}{\partial \theta} \frac{\partial f(\mathbf{w}_s; \theta)}{\partial \theta'} - u_s \frac{\partial^2 f(\mathbf{w}_s; \theta)}{\partial \theta \partial \theta'} \right)$.

$$\frac{\partial^2 h_t}{\partial \varphi \partial \varphi'} = 2 \sum_{i=1}^q \alpha_i \mathbf{w}_{t-i} \mathbf{w}'_{t-i} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}}{\partial \varphi \partial \varphi'}$$

where the pre-sample values are given by $2 \mathbf{w}_t \mathbf{w}'_t = \frac{\partial^2 h_t}{\partial \varphi \partial \varphi'} = \frac{2}{T} \sum_{s=1}^T \mathbf{w}_s \mathbf{w}'_s$.

$$\frac{\partial^2 h_t}{\partial \eta \partial \eta'} = \frac{\partial \mathbf{z}'_t}{\partial \eta} + \sum_{j=1}^p \left(\frac{\partial \beta_j}{\partial \eta} \frac{\partial h_{t-j}}{\partial \eta'} + \beta_j \frac{\partial^2 h_{t-j}}{\partial \eta \partial \eta'} \right)$$

where the pre-sample values are given by $\frac{\partial^2 h_{t-j}}{\partial \eta \partial \eta'} = 0$. The partial derivative of \mathbf{z}_t is: $\frac{\partial \mathbf{z}'_t}{\partial \eta} = \left(0, 0, \dots, 0, \frac{\partial h_{t-1}}{\partial \eta}, \dots, \frac{\partial h_{t-p}}{\partial \eta} \right)$ and the derivative of β_i is: $\frac{\partial \beta_i}{\partial \eta} = (0, 0, \dots, 1^*, \dots, 0)'$ where 1^* correspond to element $1 + q + i$.

$$\frac{\partial^2 h_t}{\partial \varphi \partial \theta'} = 2 \sum_{i=1}^q \alpha_i \mathbf{w}_{t-i} \frac{\partial f(\mathbf{w}_{t-i}; \theta)}{\partial \theta'} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}}{\partial \varphi \partial \theta'}$$

where the pre-sample values are given by $2 \mathbf{w}_t \frac{\partial f(\mathbf{w}_t; \theta)}{\partial \theta'} = \frac{\partial^2 h_{t-j}}{\partial \varphi \partial \theta'} = \frac{2}{T} \sum_{s=1}^T \mathbf{w}_s \frac{\partial f(\mathbf{w}_s; \theta)}{\partial \theta'}$.

$$\frac{\partial^2 h_t}{\partial \varphi \partial \eta'} = \frac{\partial \mathbf{z}'_t}{\partial \varphi} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}}{\partial \varphi \partial \eta'}$$

where the pre-sample values are given by $\frac{\partial^2 h_{t-j}}{\partial \varphi \partial \eta'} = 0$. The partial derivative of \mathbf{z}_t is: $\frac{\partial \mathbf{z}'_t}{\partial \varphi} = \left(0, -2u_{t-1} \mathbf{w}_{t-1}, \dots, -2u_{t-q} \mathbf{w}_{t-q}, \frac{\partial h_{t-1}}{\partial \varphi}, \dots, \frac{\partial h_{t-p}}{\partial \varphi}\right)$.

$$\frac{\partial^2 h_t}{\partial \theta \partial \eta'} = \frac{\partial \mathbf{z}'_t}{\partial \theta} + \sum_{j=1}^p \beta_j \frac{\partial^2 h_{t-j}}{\partial \theta \partial \eta'}$$

where the pre-sample values are given by $\frac{\partial^2 h_{t-j}}{\partial \theta \partial \eta'} = 0$. The partial derivative of \mathbf{z}_t is: $\frac{\partial \mathbf{z}'_t}{\partial \theta} = \left(0, -2u_{t-1} \frac{\partial f(\mathbf{w}_{t-1}; \theta)}{\partial \theta'}, \dots, -2u_{t-q} \frac{\partial f(\mathbf{w}_{t-q}; \theta)}{\partial \theta'}, \frac{\partial h_{t-1}}{\partial \varphi}, \dots, \frac{\partial h_{t-p}}{\partial \varphi}\right)$.

B.4. Partial derivatives of the logistic function $H_n(s_t; \gamma, \mathbf{c})$

The logistic function (2.3) is the key term in the parameterization of the nonlinearities in both the conditional mean and the conditional variance. On the one hand, the nonlinear function in the conditional mean is constructed as $f(\mathbf{w}_t; \theta) = \phi' \mathbf{w}_t H_n(s_t; \gamma, \mathbf{c})$, with $s_t = y_{t-d}$. On the other hand, the logistic function $H_{n^*}(s_t; \delta, \mathbf{k})$ that imposes the nonlinearity in the conditional variance is parameterized with $s_t = u_{t-d}$, which depends on φ and θ . The logistic function is defined as:

$$H_n(s_t; \gamma, \mathbf{c}) = \left(1 + \exp\left(-\gamma \prod_{l=1}^n (s_t - c_l)\right)\right)^{-1}, \gamma > 0, c_1 \leq \dots \leq c_n.$$

To simplify the calculation of the derivatives the logistic function is rewritten as:

$$H_n(s_t; \gamma, \mathbf{c}) = \frac{1}{2 \exp\left(-\frac{\gamma}{2} \prod_{l=1}^n (s_t - c_l)\right) \cosh\left(\frac{\gamma}{2} \prod_{l=1}^n (s_t - c_l)\right)}.$$

Let $\xi_t(s_t; \gamma, c_1, \dots, c_n) = \frac{\gamma}{2} \prod_{l=1}^n (s_t - c_l)$, drop the arguments and write the function as:

$$H_n(s_t; \gamma, \mathbf{c}) = \frac{e^{\xi_t}}{2 \cosh \xi_t}.$$

B.4.1. First order partial derivative of $H_n(s_t; \gamma, \mathbf{c})$

All of the first order derivatives have the same structure; as an example we consider the derivative with respect to φ .

$$\frac{\partial H_n(s_t)}{\partial \varphi'} = \frac{1}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \varphi}$$

Let c_i denote the i^{th} element of the location vector \mathbf{c} ; then the derivatives of ξ_t with respect to the parameters of the model are

$$\begin{aligned}\frac{\partial \xi_t}{\partial \varphi} &= \frac{\delta}{2} \frac{\partial s_t}{\partial \varphi} \times Sum \\ \frac{\partial \xi_t}{\partial \theta} &= \frac{\delta}{2} \frac{\partial s_t}{\partial \theta} \times Sum \\ \frac{\partial \xi_t}{\partial \delta} &= \frac{1}{2} \prod_{l=1}^n (s_t - c_l) \\ \frac{\partial \xi_t}{\partial c_i} &= -\frac{\delta}{2} \prod_{l=1, l \neq i}^n (s_t - c_l)\end{aligned}$$

where $Sum = \sum_{j=1}^n \left(\prod_{l=1, l \neq j}^n (s_t - c_l) \right)$ if $n > 1$, otherwise $Sum = 1$. Note that $\frac{\partial \xi_t}{\partial \varphi} = \mathbf{0}$ and $\frac{\partial \xi_t}{\partial \theta} = \mathbf{0}$ in the conditional mean case due to the fact that $s_t = y_{t-d}$ which does not depend on either φ nor θ .

C. Tables and Figures

	OMX index	JPY/USD
Characteristics:		
<i>Min</i>	-0.083	-0.049
<i>Max</i>	0.091	0.039
<i>Mean</i>	0	0
<i>Variance</i>	1.6×10^{-4}	4.7×10^{-5}
<i>Skewness</i>	0.056	-0.38
<i>Kurtosis</i>	8.7	6.4
Non-parametric test of IID (BDS)		
<i>statistic</i>	12.4	7.8
LM-test against nonlinearity of STAR type:		
num. df	21	21
denom. df	2656	3259
	<i>F</i> -statistic	<i>F</i> -statistic
$d = 1$	7.3	2.2
$d = 2$	5.2	2.5
$d = 3$	7.0	2.4
$d = 4$	4.0	2.9
$d = 5$	5.8	1.3

Table C.1: Certain characteristics, the BDS test and linearity tests for the OMX-index and the JPY/USD exchange rate. The characteristics and the BDS test are computed from the residuals of a linear model (a constant, daily dummy variable and an AR polynomial). The BDS statistic is asymptotically normally distributed with zero expectation and unit variance. The LM test against nonlinearity of LSTAR(3) type uses a lag length 7 for the autoregressive part. The test is computed, assuming constant conditional variance, against the alternative $1 \leq d \leq 5$. The null hypothesis is given in the table.



Figure C.1: The daily Swedish OMX index, December 30, 1983 to September 30, 1998. The dashed vertical line corresponds to October 5, 1994. Observations preceding this date are used for estimation and the remaining ones for one-step-ahead forecasting.

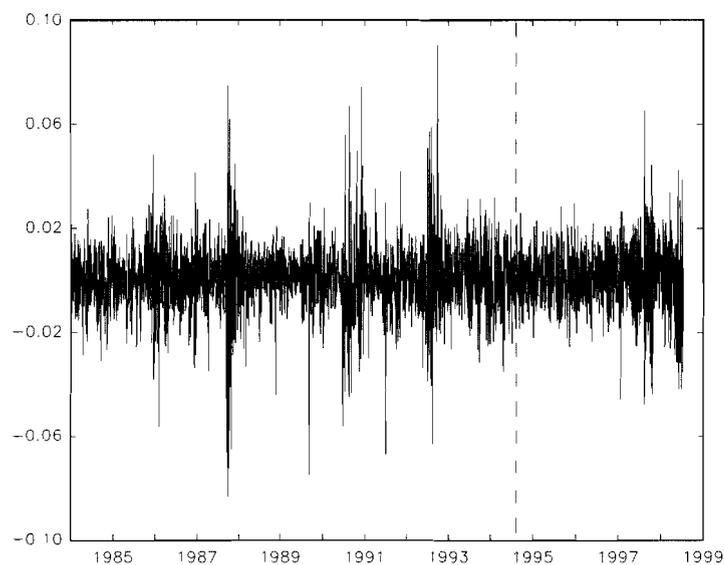


Figure C.2: Returns of the daily Swedish OMX index (first differences), December 31, 1983 to September 30, 1998. The dashed line corresponds to October 5, 1994. Observations preceding this date are used for estimation and the remaining ones for one-step-ahead forecasting.

Parameter	OMX index	JPY/USD
Conditional mean model		
D_{Mo}	-0.0033 (0.00052)	-0.00092 (0.00034)
D_{Tu}	-0.0023 (0.00050)	-0.00070 (0.00035)
D_{Wc}	-0.00048 (0.00050)	-0.0010 (0.00035)
D_{Th}	-0.00083 (0.00052)	-0.0013 (0.00035)
D_{Hol}	-0.00073 (0.0028)	-0.0015 (0.0013)
φ_0	.	0.00082 (0.00025)
φ_2	.	0.029 (0.019)
φ_3	.	0.042 (0.019)
φ_4	0.072 (0.028)	.
ϕ_0	0.0055 (0.00063)	-0.011 (0.0043)
ϕ_2	.	-0.55 (0.32)
ϕ_3	.	-0.43 (0.30)
ϕ_4	-0.11 (0.049)	.
d	1	4
γ	6.1 (2.1)	4.2 (3.7)
c	0.0039 (0.00087)	0.019 (0.0024)
Conditional variance model		
α_0	7.1×10^{-6} (1.2×10^{-6})	3.7×10^{-6} (6.4×10^{-7})
α_{11}	0.16 (0.021)	0.094 (0.013)
β_1	0.84 (0.018)	0.83 (0.022)
α_{01}	.	.
α_{21}	-0.10 (0.021)	.
δ	7.7 (5.5)	.
k	.	.

Table C.2: Parameter estimates of the STAR-STGARCH models (standard deviations in parentheses) for the OMX-index and the JPY/USD exchange rate.

	OMX index	JPY/USD
Characteristics:		
<i>Min</i>	-9.2	-9.1
<i>Max</i>	5.6	4.2
<i>Mean</i>	-0.018	-0.0089
<i>Variance</i>	1.0	1.0
<i>Skewness</i>	-0.44	-0.63
<i>Kurtosis</i>	7.7	7.9
Nonparametric test of IID (BDS)		
<i>BDS statistic</i>	0.28	1.1
<i>p</i> -value (asymptotic)	0.78	0.29
<i>p</i> -value (bootstrap)	0.78	0.30

Table C.3: Characteristics of the standardized residuals of the STAR-STGARCH model and the BDS test of independence for the standardized errors. For the BDS test a bootstrapped probability value based on 1000 resampled series is reported as well.

	OMX index	JPY/USD
Remaining autocorrelation (<i>p</i> -values)		
<i>l</i> = 1	0.27	0.11
<i>l</i> = 2	0.54	0.037
<i>l</i> = 3	0.69	0.086
<i>l</i> = 4	0.81	0.15
<i>l</i> = 5	0.90	0.13
Parameter constancy (<i>p</i> -values)		
<i>All</i>	0.076	0.13
<i>Dummies</i>	0.0080	0.80
<i>Linear</i>	0.73	0.033
<i>Non-linear</i>	0.82	0.57
Remaining nonlinearity (<i>p</i> -values)		
<i>d</i> = 1	0.11	0.49
<i>d</i> = 2	0.44	0.36
<i>d</i> = 3	0.38	0.019
<i>d</i> = 4	0.17	0.36
<i>d</i> = 5	0.35	0.72

Table C.4: *p*-values of specification tests for the conditional mean for the estimated STAR-STGARCH model. LM tests for the conditional mean. The test of no remaining autocorrelation is computed against the alternative of remaining autocorrelation up to the given lag, *l*. The test of parameter constancy is computed against the alternative of time-dependence given by an LSTAR(3) parametrization with time as the transition variable. The test against nonlinearity of LSTAR(3) type uses a lag length 7 for the autoregressive part. The test is computed separately against the alternatives $1 \leq d \leq 5$.

OMX index	
Remaining nonlinearity of STGARCH type (<i>p</i> -values)	
<i>All parameters</i>	1.8×10^{-8}
<i>Constant intercept</i>	9.9×10^{-7}
Remaining nonlinearity with a fixed delay (<i>p</i> -values)	
<i>d</i> = 1	1.8×10^{-8}
<i>d</i> = 2	2.3×10^{-5}
<i>d</i> = 3	0.017

Table C.5: *p*-values of the test against nonlinearity in the conditional variance for a STAR-GARCH model estimated for the OMX index. The tests against remaining nonlinearity make use of a third order Taylor approximation of the transition function. One test is against a STGARCH structure and the other one against a nonlinearity with a fixed delay. The latter is computed separately against $1 \leq d \leq 3$.

	OMX index	JPY/USD
Remaining ARCH (<i>p</i> -values)		
<i>l</i> = 1	0.15	0.17
<i>l</i> = 2	0.32	0.18
<i>l</i> = 3	0.51	0.0013
<i>l</i> = 10	0.45	0.022
Test of the functional form (<i>p</i> -values)		
<i>l</i> = 1	0.60	0.49
<i>l</i> = 2	0.44	0.52
<i>l</i> = 3	0.45	0.15
<i>l</i> = 10	0.11	0.33
Parameter constancy (<i>p</i> -values)		
<i>All</i>	0.012	0.053
<i>Intercept</i>	0.060	0.090
<i>Alfa</i>	0.13	0.74
<i>Beta</i>	0.055	0.38
<i>Smooth</i>	0.25	.
Remaining nonlinearity of STGARCH type (<i>p</i> -values)		
<i>All parameters</i>	0.0021	0.080
<i>Constant intercept</i>	0.0013	0.69
Remaining nonlinearity with a fixed delay (<i>p</i> -values)		
<i>d</i> = 1	0.0021	0.080
<i>d</i> = 2	0.60	0.53
<i>d</i> = 3	0.13	0.21

Table C.6: *p*-values of specification tests for the conditional variance of the estimated STAR-STGARCH model. LM tests for the conditional variance. The tests of no remaining serial dependence in the squared and standardized residuals are computed against the alternative of remaining dependence up to the given lag, *l*. The test of parameter constancy is computed against the alternative of time-dependence given by an LSTAR(2) parametrization with time as the transition variable. The tests against remaining nonlinearity make use of a third order Taylor approximation of the transition function. One test is against a STGARCH structure and the other one against a nonlinearity with a fixed delay. The latter is computed separately against the alternatives $1 \leq d \leq 3$.

Asset	RMSE		Test p-value
	Our model	Competitor	
OMX index	0.32	0.39	0
JPY/USD	0.27	0.27	0.99

Table C.7: Root mean square errors (RMSE) of the conditional mean part of the STAR-STGARCH model and its linear competitor. Also the p -value of the test of the null hypothesis $H_0 : r = 0$ against $H_1 : r > 0$ is given. For the Swedish OMX index we use a linear AR(7) model as the competitor. For the JPY/USD exchange rate we use the deviation of the actual observations from zero as the competitor.

		The model			
		-1	0	1	Total
The observed data	-1	108	281	3	392
	0	27	117	3	147
	1	1	126	334	461
	Total	136	524	340	1000

Table C.8: Cross-tabulation of the ordered actual observations and one-step-ahead forecasts of the Swedish OMX index. A value between -0.002 and 0.002 is denoted 0, otherwise positive and negative values are represented by 1 and -1.

		The model			
		-1	0	1	Total
The observed data	-1	89	453	104	646
	0	18	89	18	125
	1	95	471	123	689
	Total	202	1013	245	1460

Table C.9: Cross-tabulation of the ordered actual observations and one-step-ahead forecasts of the JPY/USD exchange rate. A value between -0.0005 and 0.0005 is denoted 0, otherwise positive and negative values are represented by 1 and -1.

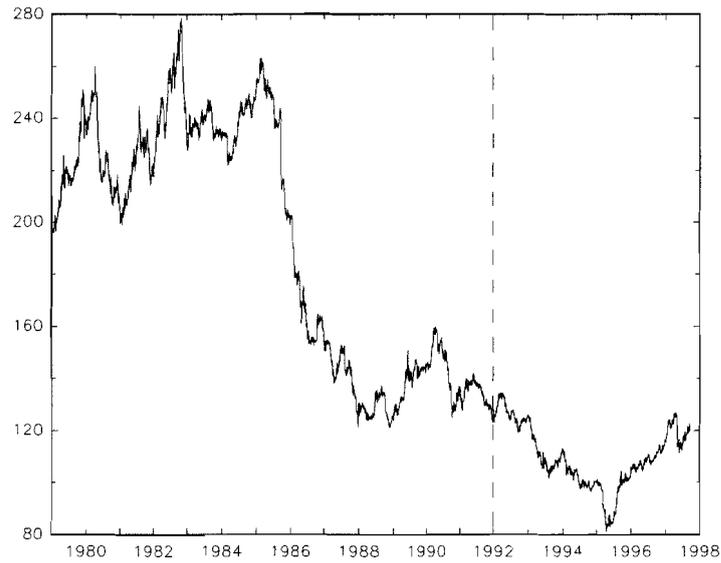


Figure C.3: The JPY/USD daily exchange rate, December 28, 1978 to September 30, 1997. The dashed vertical line corresponds to January 1, 1992. Observations preceding this date are used for estimation and the remaining ones for one-step-ahead forecasting.

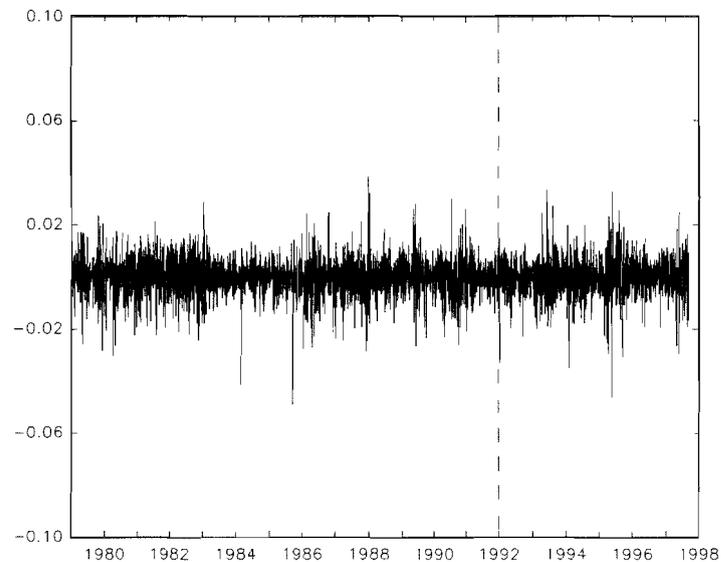


Figure C.4: Returns of the JPY/USD daily exchange rate (first differences), December 29, 1978 to September 30 1997. The dashed line corresponds to January 1, 1992. Observations preceding this date are used for estimation and the remaining ones for one-step-ahead forecasting.

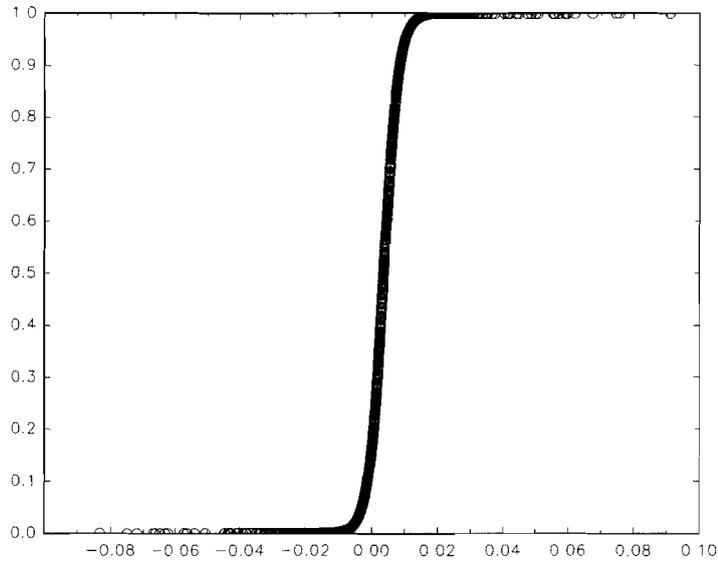


Figure C.5: Values of the transition function for the conditional mean part of model (2.4) for the Swedish OMX index return series. Each circle indicates an observation.

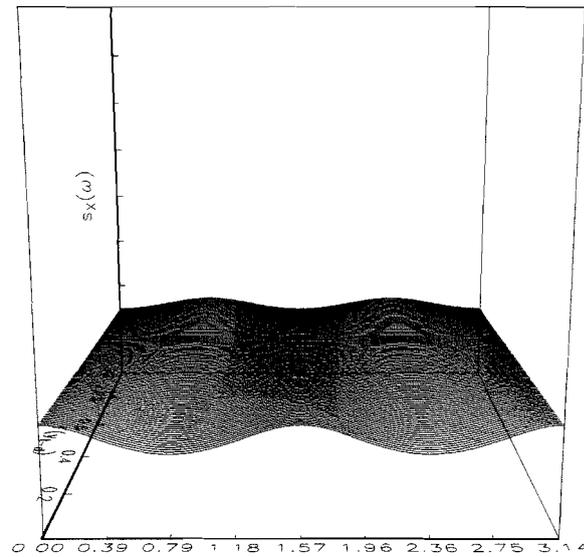


Figure C.6: The estimated 'sliced' spectrum of model (2.4) for the Swedish OMX index return series. The x-axis gives the frequency and the y-axis gives the value of the transition function. A slice (solid curve) represents at least one observed value of the transition function.

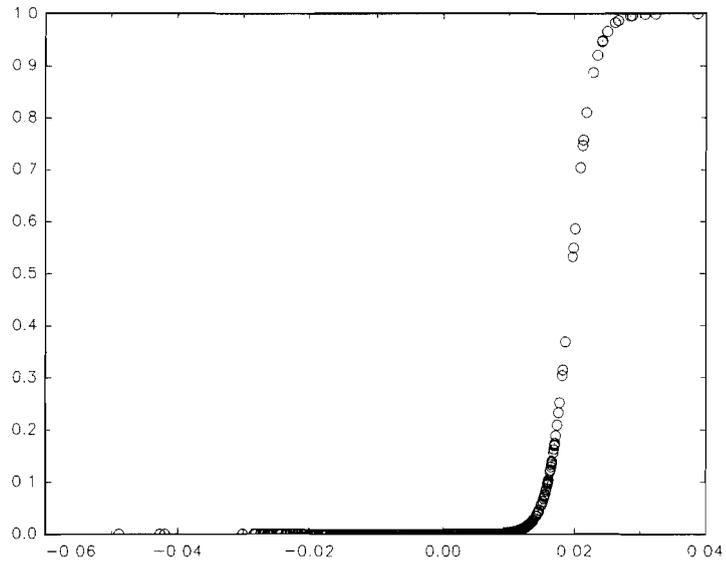


Figure C.7: Values of the transition function for the conditional mean part of model (2.4) for the JPY/USD exchange rate return series. Each circle indicates an observation.

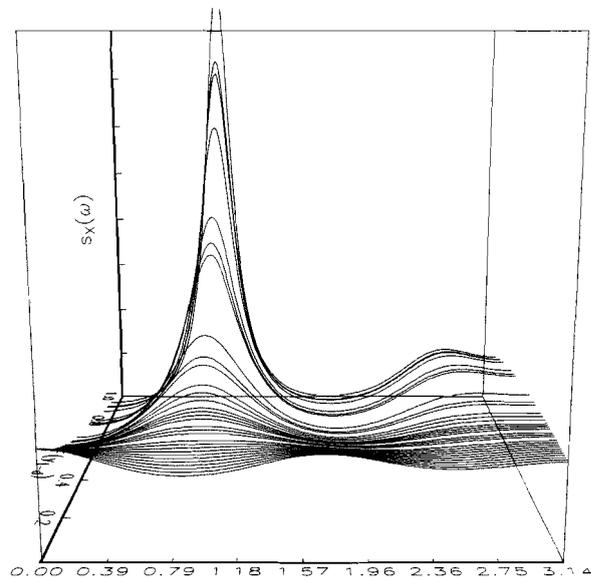


Figure C.8: The estimated 'sliced' spectrum of model (2.4) for the JPY/USD exchange rate return series. The x-axis gives the frequency and the y-axis gives the value of the transition function. A slice (solid curve) represents at least one observed value of the transition function.

Chapter 3

Evaluating GARCH models

1. Introduction

When modelling the conditional mean, at least when it is a linear function of parameters, the estimated model is regularly subjected to a battery of misspecification tests to check its adequacy. The hypothesis of no (conditional) heteroskedasticity, no error autocorrelation, linearity, and parameter constancy, to name a few, are tested using various methods. In models of conditional variance, such as the popular GARCH model, testing the adequacy of the estimated model has been much less common in practice. But then, misspecification tests do exist in the literature for GARCH models also. For example, Bollerslev (1986) already suggested a score or Lagrange Multiplier (LM) test for testing a GARCH model of a given order against a higher-order GARCH model. Li and Mak (1994) derived a portmanteau type test for testing the adequacy of a GARCH model. Engle and Ng (1993) considered testing the GARCH specification against asymmetry using the so-called sign-bias test. Chu (1995) derived a test of parameter constancy against a single structural break. This test has a nonstandard asymptotic null distribution, but Chu provided tables for critical values.

In this chapter we provide a unified framework for misspecification testing in GARCH models. The framework covers the most common alternative hypotheses. The idea is to make misspecification testing easy without sacrificing power. We

suggest tests for testing the null of no ARCH in the standardized errors, a general test for misspecification of the functional form, testing symmetry against a smooth transition GARCH, and a test of parameter constancy against smoothly changing parameters. A single structural break is nested in the alternative hypothesis of the parameter constancy test. A two-regime asymmetric GARCH model such as the so-called GJR model (Glosten, Jagannathan and Runkle, 1993) is nested in the alternative of smooth transition GARCH. Note that the test of Bollerslev (1986) fits well into our framework. Furthermore, we show that the portmanteau test of Li and Mak (1994) is asymptotically equivalent to our test of no remaining ARCH. All our tests are LM-tests and require only standard asymptotic distribution theory. They may be obtained from the same "root" by merely changing the definitions of the elements of the score vector corresponding to the alternative hypothesis. This makes testing easy as the sample counterparts of the analytical first and second order derivatives of the logarithmic likelihood function may be computed without difficulty using the results in Fiorentini, Calzolari and Panattoni (1996). Our Monte Carlo simulations show that the tests we propose have reasonable power, that is, they compare favourably with the tests currently available in the literature.

The plan of the chapter is as follows. In Section 2 we define the model. In Section 3 we discuss testing the null of no ARCH in the standardized errors and compare our LM-test with the portmanteau test of Li and Mak (1994). Section 4 considers testing the functional form, symmetry and parameter constancy. Section 5 contains results of a simulation experiment in which our tests are compared with other tests proposed in the literature and Section 6 concludes.

2. The model

Consider a conditionally heteroskedastic model where the conditional mean has the following structure:

$$y_t = f(\mathbf{w}_t; \boldsymbol{\varphi}) + u_t \quad (2.1)$$

where f is at least twice continuously differentiable with respect to $\boldsymbol{\varphi} \in \Phi$, for all $\mathbf{w}_t \in \mathcal{R}^k$ everywhere in Φ . The conditional variance is parameterized as:

$$u_t = \xi_t \sqrt{h(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta})} \quad (2.2)$$

where $\{\xi_t\}$ is a sequence of independent standard normal variables. The normality assumption is made just for the purpose of defining the likelihood function but is not needed for the asymptotic results. Existence of a number of moments has to be assumed, however, for each of the cases considered below. We assume that $h_t = \boldsymbol{\eta}'\mathbf{z}_t$, that is, a linear function of the parameters $\boldsymbol{\eta}$. The standard GARCH(p, q) model where $\boldsymbol{\eta} = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ and the observation vector $\mathbf{z}_t = (1, u_{t-1}^2, \dots, u_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$ constitutes an example. Furthermore, $u_t = y_t - f(\mathbf{w}_t; \boldsymbol{\varphi})$ and $\boldsymbol{\varphi}$ is assumed not to depend on $\boldsymbol{\eta}$. This guarantees that $E[u_t] = 0$ and $E[u_t u_{t-j}] = 0$, $j \neq 0$. We assume that the maximum likelihood estimators of the parameters of the GARCH process are consistent and asymptotically normal. Lumsdaine (1996) gave the required assumptions for the GARCH(1,1) process. For more general GARCH processes, see Bollerslev and Wooldridge (1992). Since the standard GARCH(p, q) model is symmetric and satisfies the usual regularity conditions, see Engle (1982), the information matrix is block-diagonal in $\boldsymbol{\varphi}$ and $\boldsymbol{\eta}$. The restrictions $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \dots, q-1$, $\alpha_q > 0$, $\beta_i \geq 0$, ensure nonnegative conditional variance with probability one but they can be relaxed as in Nelson and Cao (1992); see also He and Teräsvirta (1999c). In what follows we shall mainly focus on the conditional variance and do not consider the conditional mean. This we do for simplicity, and in cases where the information matrix of the log-likelihood is block diagonal this approach is justified. Joint modelling of the conditional mean and the conditional variance is discussed, for example, in Lee and Li (1998) and Chapter 2.

3. Testing the null of no ARCH in standardized errors

3.1. LM-Test

We consider the situation where we have estimated a GARCH(p, q) model under the assumption that the standardized errors $\xi_t = u_t h_t^{-1/2}$ of this model are independent normal. We want to test this hypothesis against the alternative that these errors follow an ARCH(m) process. Consider (2.2) but assume that

$$\xi_t = \varepsilon_t \sqrt{g(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})} \quad (3.1)$$

where $\{\varepsilon_t\}$ is a sequence of independent standard normal variables. The alternative of higher-order dependence in (3.1) is parametrized as $g_t = 1 + \boldsymbol{\pi}'\mathbf{v}_t$ where $\boldsymbol{\pi} = (\pi_1, \dots, \pi_m)'$ and $\mathbf{v}_t = (\xi_{t-1}^2, \dots, \xi_{t-m}^2)'$ so that $E[\xi_t^2 \xi_{t-j}^2] \neq 0$, $j \neq 0$. This implies that $\{\xi_t\}$ follows an ARCH(m) model. The null hypothesis of no ARCH in $\{\xi_t\}$ is

$H_0 : \boldsymbol{\pi} = \mathbf{0}$ which in turn is equivalent to $g_t \equiv 1$. Under the alternative $\boldsymbol{\pi} \neq \mathbf{0}$ the standard GARCH(p, q) model is misspecified because $\{\xi_t\}$ is no longer a sequence of independent variables. For simplicity, rewrite (2.2) as

$$u_t = \varepsilon_t \sqrt{h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})} \quad (3.2)$$

where $h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = (\boldsymbol{\eta}' \mathbf{z}_t) g(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$. Let $\boldsymbol{\omega} = (\boldsymbol{\varphi}, \boldsymbol{\eta})'$ denote the parameters of the standard GARCH model with the conditional mean specified according to (2.1). In that case, $h(\mathbf{z}_t; \boldsymbol{\omega}, \mathbf{a}) = (\boldsymbol{\eta}' \mathbf{z}_t)(1 + \boldsymbol{\pi}' \mathbf{v}_t)$. The Lagrange multiplier (or score) test statistic is defined as

$$\text{LM}_\pi = T \begin{pmatrix} 0 \\ \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t}{\partial \boldsymbol{\pi}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \end{pmatrix}' I(\widehat{\boldsymbol{\omega}}, \boldsymbol{\pi} |_{\boldsymbol{\pi}=\mathbf{0}})^{-1} \begin{pmatrix} 0 \\ \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t}{\partial \boldsymbol{\pi}} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \end{pmatrix} \quad (3.3)$$

where T is the number of observations. The information matrix $I(\widehat{\boldsymbol{\omega}}, \boldsymbol{\pi} |_{\boldsymbol{\pi}=\mathbf{0}})$ is estimated by the estimated negative expectation of the Hessian. Estimation of the GARCH model using analytical derivatives, Fiorentini et al. (1996), yields as a by-product numerically reliable estimates for the elements of the inverse of the information matrix and can therefore be recommended.

The first derivative of the log-likelihood of observation t with respect to $\boldsymbol{\pi}$ evaluated under H_0 has the form

$$\begin{aligned} \frac{\partial l_t}{\partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} &= \frac{1}{2} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} \\ &= \frac{1}{2} \left(\frac{u_t^2}{h_t} - 1 \right) \mathbf{v}_t \end{aligned}$$

where $\mathbf{v}_t = (\xi_{t-1}^2, \dots, \xi_{t-m}^2)' = (\frac{u_{t-1}^2}{h_{t-1}}, \dots, \frac{u_{t-m}^2}{h_{t-m}})' = \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}}$. Under the null hypothesis, the information matrix is block-diagonal in $\boldsymbol{\varphi}$ and $\boldsymbol{\eta}$. The negative (conditional) expectations of the relevant second order derivatives are

$$-\mathbf{E} \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2} \mathbf{x}_t \mathbf{x}_t' \quad -\mathbf{E} \frac{\partial^2 l_t}{\partial \boldsymbol{\pi} \partial \boldsymbol{\eta}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2} \mathbf{v}_t \mathbf{x}_t' \quad -\mathbf{E} \frac{\partial^2 l_t}{\partial \boldsymbol{\pi} \partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{2} \mathbf{v}_t \mathbf{v}_t'$$

where $\mathbf{x}_t = \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\eta}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}}$. The test statistic (3.3) may then be written as

$$\text{LM}_\pi = \frac{1}{4T} \sum \left(\left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right) \widehat{\mathbf{v}}_t' \right) V_{\text{LM}}(\widehat{\boldsymbol{\eta}})^{-1} \sum \left(\widehat{\mathbf{v}}_t \left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right) \right) \quad (3.4)$$

where $V_{\text{LM}}(\widehat{\boldsymbol{\eta}}) = \frac{1}{2T} (\sum \widehat{\mathbf{v}}_t \widehat{\mathbf{v}}_t' - \sum \widehat{\mathbf{v}}_t \widehat{\mathbf{x}}_t' (\sum \widehat{\mathbf{x}}_t \widehat{\mathbf{x}}_t')^{-1} \sum \widehat{\mathbf{x}}_t \widehat{\mathbf{v}}_t')$. Furthermore, $\widehat{\mathbf{v}}_t$ and

$\widehat{\mathbf{x}}_t$ are the sample counterparts of the corresponding derivatives under the null hypothesis and they may be computed iteratively; see Fiorentini et al. (1996) for details. Assuming that the relevant moments, including the fourth moment of ξ_t , exist, (3.4) has an asymptotic χ^2 distribution with m degrees of freedom when H_0 holds.

This test statistic may also be computed by using an artificial regression. The F -version of the test is then carried out as follows.

1. Estimate the parameters of the conditional variance model under the null, compute $\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1$ and the sum of squared residuals, $SSR_0 = \sum_{t=1}^T (\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)^2$.
2. Regress $(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)$ on $\widehat{\mathbf{x}}_t'$, $\widehat{\mathbf{v}}_t'$ and compute the sum of squared residuals, $SSR_1 = \sum_{t=1}^T \widehat{\varepsilon}_t^2$.
3. Compute the F -version of the test statistic as $F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - \dim(\boldsymbol{\omega}) - m)}$.

For the sample sizes relevant in GARCH modelling, there is no essential difference between the properties of the F -test and its asymptotically correct χ^2 counterpart $\chi^2 = T(SSR_0 - SSR_1)/SSR_0$.

3.2. Comparison with a portmanteau test

The test in the previous section was explicitly derived as an LM-test. Li and Mak (1994) recently introduced a portmanteau statistic for testing the adequacy of the standard GARCH(p, q) model. The null hypothesis is that the squared and standardized error process is not autocorrelated. In practice, one tests this hypothesis for the first m autocorrelations. Let $\mathbf{r} = (r_1, \dots, r_m)'$ be the $m \times 1$ vector of the first m autocorrelations so that $H_0 : \mathbf{r} = \mathbf{0}$. Li and Mak (1994) showed that under this hypothesis $\sqrt{T}\widehat{\mathbf{r}}$ is asymptotically normally distributed where T is the number of observations. The vector of autocorrelations is estimated by

$$\widehat{\mathbf{r}} = \frac{1}{T} \sum_{t=1}^T \left(\left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right) \widehat{\mathbf{v}}_t^* \right) / \left(\frac{1}{T} \sum_{t=1}^T \left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right)^2 \right)$$

where $\widehat{u}_t = y_t - f(\mathbf{w}_t; \widehat{\boldsymbol{\varphi}})$, $\widehat{h}_t = h_t(\mathbf{w}_t; \widehat{\boldsymbol{\omega}})$ and $\widehat{\mathbf{v}}_t^* = \widehat{\mathbf{v}}_t - \mathbf{1}_m = (\frac{\widehat{u}_{t-1}^2}{\widehat{h}_{t-1}} - 1, \dots, \frac{\widehat{u}_{t-m}^2}{\widehat{h}_{t-m}} - 1)'$ with $\mathbf{1}_m = (1, \dots, 1)'$ an $m \times 1$ vector. Note that under normality $\frac{1}{T} \sum_{t=1}^T (\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1)^2 \rightarrow 2$ in probability as $T \rightarrow \infty$. Thus

$$\widehat{\mathbf{r}} = \frac{1}{2T} \sum_{t=1}^T \left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right) \widehat{\mathbf{v}}_t^*$$

is asymptotically equivalent to $\widehat{\mathbf{r}}$. Under the null hypothesis, the asymptotic covariance matrix of $\sqrt{T}\widehat{\mathbf{r}}$ is block diagonal in $\boldsymbol{\varphi}$ and $\boldsymbol{\eta}$ and therefore estimated by

$$\mathbf{V}_r(\widehat{\boldsymbol{\eta}}) = \mathbf{I}_m - \mathbf{X}_r(\widehat{\boldsymbol{\eta}})' \mathbf{G}^{-1}(\widehat{\boldsymbol{\eta}}) \mathbf{X}_r(\widehat{\boldsymbol{\eta}})$$

since $\frac{1}{T} \sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{v}}_t^{*'} \rightarrow 2\mathbf{I}_m$ under H_0 as $T \rightarrow \infty$ and

$$\mathbf{X}_r(\widehat{\boldsymbol{\eta}}) = -\frac{1}{2T} \sum_{t=1}^T \left(\frac{1}{\widehat{h}_t} \frac{\partial \widehat{h}_t}{\partial \widehat{\boldsymbol{\eta}}} \widehat{\mathbf{v}}_t^{*'} \right) = -\frac{1}{2T} \sum \widehat{\mathbf{x}}_t \widehat{\mathbf{v}}_t^{*'}.$$

Furthermore, $\mathbf{G}^{-1}(\widehat{\boldsymbol{\eta}})$ is some consistent estimator of the relevant block of the information matrix, evaluated at $\boldsymbol{\eta} = \widehat{\boldsymbol{\eta}}$. The portmanteau statistic becomes

$$Q(m) = T\widehat{\mathbf{r}}' \mathbf{V}_r(\widehat{\boldsymbol{\eta}})^{-1} \widehat{\mathbf{r}} \quad (3.5)$$

which is asymptotically χ^2 -distributed with m degrees of freedom under the null hypothesis. We may now also define

$$Q(m)^* = \frac{1}{4T} \left(\sum \left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right) \widehat{\mathbf{v}}_t^{*'} \right) \mathbf{V}_r^*(\widehat{\boldsymbol{\eta}})^{-1} \left(\sum \widehat{\mathbf{v}}_t^* \left(\frac{\widehat{u}_t^2}{\widehat{h}_t} - 1 \right) \right) \quad (3.6)$$

where $\mathbf{V}_r^*(\widehat{\boldsymbol{\eta}}) = \frac{1}{2T} \left(\sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{v}}_t^{*'} - \sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{x}}_t' \left(\sum \widehat{\mathbf{x}}_t \widehat{\mathbf{x}}_t' \right)^{-1} \sum \widehat{\mathbf{x}}_t \widehat{\mathbf{v}}_t^{*'} \right)$. The only difference between (3.5) and (3.6) is the choice of the consistent estimator of the covariance matrix. $\mathbf{V}_r^*(\widehat{\boldsymbol{\eta}})$ makes use of $\frac{1}{T} \sum \widehat{\mathbf{v}}_t^* \widehat{\mathbf{v}}_t^{*'}$ whereas its expectation $2\mathbf{I}_m$ appears in $\mathbf{V}_r(\widehat{\boldsymbol{\eta}})$. The two covariance matrices are thus asymptotically equal. As (3.6) is identical to (3.4), the apparent difference in the expressions being due to centring, it follows that the Li and Mak (1994) portmanteau statistic and the LM-test (3.4) are asymptotically equivalent. This means that the portmanteau test of model adequacy is in fact an LM-test of no ARCH in the standardized errors against ARCH(m). This is analogous to the McLeod and Li (1983) portmanteau test being asymptotically equivalent to the classic LM-test of no ARCH of Engle (1982); see, for example, Luukkonen, Saikkonen and Teräsvirta (1988b) for a discussion. As a matter of fact, when $h_t \equiv \alpha_0$, our test and that of Li and Mak (1994) reduce to the Engle (1982) and McLeod and Li (1983) tests, respectively.

4. Misspecification of structure

In this section we present three different misspecification tests for an estimated conditional variance model. The first one can be interpreted as a test of the functional form. The second one is a test against nonlinearity or, in some cases, asymmetry. It is a modification of a test in Hagerud (1997). Finally we propose a test of parameter constancy against smooth continuous change in parameters. All three tests may be viewed as conditional variance counterparts of the tests for the nonlinear conditional mean in Eitrheim and Teräsvirta (1996). To describe the common features in these tests we first introduce a general structure and thereafter consider each test separately.

4.1. General structure

Consider (2.2) and define

$$h(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) = \boldsymbol{\eta}' \mathbf{z}_t + G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) \quad (4.1)$$

We assume that $h(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$ satisfies the regularity conditions mentioned in Section 2, and that $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$ is at least twice differentiable for all $\boldsymbol{\pi}$ everywhere in its sample space. Furthermore, let $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \mathbf{0}) \equiv \mathbf{0}$ which does not affect the generality of the argument. We also assume that the necessary moments of $\{u_t\}$ exist. Let $\boldsymbol{\omega} = (\boldsymbol{\varphi}', \boldsymbol{\eta}')$ denote the parameters of the standard GARCH(p, q) model with the conditional mean specified according to (2.1). In that case, $h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \boldsymbol{\eta}' \mathbf{z}_t$, and the null hypothesis of no additional structure in $h(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})$ becomes $H_0 : \boldsymbol{\pi} = \mathbf{0}$. The Lagrange multiplier (or score) test statistic is again (3.3) which is asymptotically χ^2 -distributed with $\dim(\boldsymbol{\pi})$ degrees of freedom under the null hypothesis and the required regularity conditions. Due to the fact that the information matrix is block diagonal under the null hypothesis it has the expression (3.4), where $\mathbf{v}_t = \frac{\partial l_t}{\partial \boldsymbol{\pi}'} |_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\pi}'} |_{\boldsymbol{\pi}=\mathbf{0}}$ and $\mathbf{x}_t = \frac{\partial l_t}{\partial \boldsymbol{\eta}'} |_{\boldsymbol{\pi}=\mathbf{0}} = \frac{1}{h_t} \frac{\partial h_t}{\partial \boldsymbol{\eta}'} |_{\boldsymbol{\pi}=\mathbf{0}}$. Furthermore $\widehat{\mathbf{v}}_t$ and $\widehat{\mathbf{x}}_t$ are the sample counterparts of the corresponding derivatives under H_0 . Thus (3.4) may be computed by using an artificial regression as described in the previous section. The partial derivatives required to construct the Hessian needed for the estimation of the information matrix can be found in Appendix A.

4.2. Testing the functional form

Another way of testing the null hypothesis of no error autocorrelation in the squared residuals is to lag the linear combination $\boldsymbol{\eta}'\mathbf{z}_t$ and enter it in the conditional variance process, h_t , under the alternative. This may be viewed as a general but possibly parsimonious misspecification test along the lines in Bollerslev (1986). The test is obtained by defining $G(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \boldsymbol{\pi}'\mathbf{v}_t$ where $\boldsymbol{\pi} = (\pi_1, \dots, \pi_r)'$ and $\mathbf{v}_t = (\boldsymbol{\eta}'\mathbf{z}_{t-1}, \dots, \boldsymbol{\eta}'\mathbf{z}_{t-r})'$, in (4.1). The null hypothesis of no remaining serial dependence in the squared residuals or no model misspecification is $H_0 : \boldsymbol{\pi} = \mathbf{0}$. The moment condition $Eu_t^4 < \infty$ must hold for the asymptotic theory to go through. Under the null hypothesis, the LM-statistic (3.3) is asymptotically χ^2 -distributed with $\dim(\boldsymbol{\pi})$ degrees of freedom. On the other hand, Bollerslev (1986) suggested another test which is obtained by defining $G(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \boldsymbol{\eta}'\mathbf{z}_t^*$ where $\mathbf{z}_t^* = (u_{t-q-1}^2, \dots, u_{t-q-m}^2)'$ or $\mathbf{z}_t^* = (h_{t-p-1}, \dots, h_{t-p-n})'$. The alternative is thus a higher-order GARCH model. Note that it is not possible to test a GARCH(p, q) against a GARCH($p+n, q+m$) model using standard procedures when both m and n are assumed positive.

4.3. Testing linearity (symmetry)

The above test of the functional form offers few hints about what may be wrong when the null hypothesis is rejected. Bollerslev's (1986) test is more explicit about the alternative, which is a higher-order GARCH model. Nevertheless, it may also be useful to consider other parametric alternatives to the symmetric GARCH(p, q) model. In some cases we may expect the response to be a function not only of the size of the shock but also of its direction. Engle and Ng (1993), see also references therein, considered this possibility. We call such a shock response asymmetric and parameterize it by generalizing the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). This is done in three ways. First, we make the transition between the extreme regimes smooth. Second, we incorporate a nonlinear version of the quadratic GARCH model of Sentana (1995) in our alternative. Finally, while the GJR-GARCH model is asymmetric, the present generalization may also retain the symmetry although the model becomes nonlinear. For smooth transition GARCH, see also Hagerud (1997) and González-Rivera (1998). Let

$$H_n(s_t; \gamma, \mathbf{c}) = \left(1 + \exp\left(-\gamma \prod_{l=1}^n (s_t - c_l)\right) \right)^{-1}, \gamma > 0, c_1 \leq \dots \leq c_n \quad (4.2)$$

where s_t is the transition variable at time t , γ is a slope parameter, and \mathbf{c} a location vector. When $\gamma = 0$, $H_n(s_t; \gamma, \mathbf{c}) \equiv 1/2$ and when $\gamma \rightarrow \infty$, $H_n(s_t; \gamma, \mathbf{c})$ becomes a step function. The logistic function (4.2) is used for parameterizing the maintained model, and we assume $n \leq 2$. We define $\bar{H}_n = H_n - \frac{1}{2}$, subtracting $1/2$ from H_n is just a notational convenience in deriving the test and does not affect the generality of the argument. The alternative may now be written as

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_{0j} \bar{H}_n(u_{t-j}; \gamma, \mathbf{c}) + \sum_{j=1}^q \{ \alpha_{1j} + \alpha_{2j} \bar{H}_n(u_{t-j}; \gamma, \mathbf{c}) \} u_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (4.3)$$

where $\alpha_0 + \sum_{j=1}^q \alpha_{0j} \bar{H}_n > 0$ and $\alpha_{1j} + \alpha_{2j} \bar{H}_n \geq 0$, $j = 1, \dots, q$, for $0 \leq \bar{H}_n \leq 1$. By letting $\boldsymbol{\pi} = (\alpha_{01}, \dots, \alpha_{0q}, \alpha_{21}, \dots, \alpha_{2q}, \gamma, \mathbf{c}')'$ we can write

$$G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) = \sum_{j=1}^q \alpha_{0j} \bar{H}_n(u_{t-j}; \gamma, \mathbf{c}) + \sum_{j=1}^q \alpha_{2j} \bar{H}_n(u_{t-j}; \gamma, \mathbf{c}) u_{t-j}^2 \quad (4.4)$$

where $n = 1$ or 2 . Assuming that the first sum on the right-hand side of (4.4) is identical equal to zero and letting $\gamma \rightarrow \infty$ yields the GJR-GARCH model. The test of the standard GARCH model against nonlinear GARCH in Hagerud (1997) may be viewed as a special case of this specification with $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) = \sum_{j=1}^q \alpha_{2j} \bar{H}_n(u_{t-j}; \gamma, \mathbf{c}) u_{t-j}^2$ where $\boldsymbol{\pi} = (\alpha_{21}, \dots, \alpha_{2q}, \gamma, \mathbf{c}')'$.

Another way of parameterizing the alternative is to assume that the transition variable has a fixed delay. This assumption results in the following conditional variance model

$$h_t = \alpha_0 + \alpha_{0d} \bar{H}_n(u_{t-d}; \gamma, \mathbf{c}) + \sum_{j=1}^q (\alpha_{1j} + \alpha_{2j} \bar{H}_n(u_{t-d}; \gamma, \mathbf{c})) u_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (4.5)$$

which gives $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) = \alpha_{1d} \bar{H}_n(u_{t-d}; \gamma, \mathbf{c}) + \sum_{j=1}^q \alpha_{2j} \bar{H}_n(u_{t-d}; \gamma, \mathbf{c}) u_{t-j}^2$ where $\boldsymbol{\pi} = (\alpha_{1d}, \alpha_{21}, \dots, \alpha_{2q}, \gamma, \mathbf{c}')'$. This specification is very much in the spirit of Teräsvirta (1994) for the STAR-type conditional mean. Probably the most common case in practice is $d = q = p = 1$ so that (4.3) and (4.5) are equal. We do not impose any nonlinear structure on the h_{t-j} , $j = 1, \dots, p$, as the alternative structure is already very flexible even without such an extension.

These smooth transition alternatives pose an identification problem. The null

hypothesis can be expressed as $H_0 : \gamma = 0$ in (4.2). It is seen that the parameters in (4.3) or (4.5) assuming a logistic transition function (4.2) are only identified under the alternative. Thus the classical test procedures do not work; see, for example, Hansen (1996) for discussion. We circumvent the identification problem by following Luukkonen, Saikkonen and Teräsvirta (1988a). This is done by expanding the transition function \bar{H}_n into a Taylor series around $\gamma = 0$, replacing the transition function with this Taylor approximation in (4.4) or (4.5) and rearranging terms. This results in

$$h_t = \boldsymbol{\eta}' \mathbf{z}_t + \tilde{\boldsymbol{\pi}}' \mathbf{v}_t + R_1(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})$$

where it can be seen that $\tilde{\boldsymbol{\pi}} = \gamma \bar{\boldsymbol{\pi}}$ with $\bar{\boldsymbol{\pi}} \neq \mathbf{0}$. This being the case, the new null hypothesis $H'_0 : \tilde{\boldsymbol{\pi}} = \mathbf{0}$. Thus $G(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \tilde{\boldsymbol{\pi}}' \mathbf{v}_t + R_1(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})$ in (4.1). Note that under H_0 we have $R_1(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) \equiv 0$ so that the remainder does not affect the distribution theory. For the smooth GJR-like alternative (4.3) we have $\tilde{\boldsymbol{\pi}} = (\tilde{\pi}_1, \dots, \tilde{\pi}_{(1+n)q})'$ and $\mathbf{v}_t = (u_{t-1}, u_{t-1}^3, \dots, u_{t-1}^{n+2}, \dots, u_{t-q}, u_{t-q}^3, \dots, u_{t-q}^{n+2})'$. For the alternative with a fixed delay (4.5) there are two possibilities, depending on the delay. If the delay is to be found within q then $\tilde{\boldsymbol{\pi}} = (\tilde{\pi}_1, \dots, \tilde{\pi}_{1+nq})'$ and $\mathbf{v}_t = (u_{t-d}, u_{t-d} u_{t-1}^2, \dots, u_{t-d}^n u_{t-1}^2, \dots, u_{t-d} u_{t-q}^2, \dots, u_{t-d}^n u_{t-q}^2)'$, otherwise $\tilde{\boldsymbol{\pi}} = (\tilde{\pi}_1, \dots, \tilde{\pi}_{3+nq})'$ and $\mathbf{v}_t = (u_{t-d}, u_{t-d}^2, u_{t-d}^3, u_{t-d} u_{t-1}^2, \dots, u_{t-d}^n u_{t-1}^2)'$. For both models the moment condition of $E u_t^{2n+2} < \infty$ must hold for the asymptotic theory to go through. Under the null hypothesis the LM-type teststatistic (3.3) is asymptotically χ^2 -distributed with $\dim(\tilde{\boldsymbol{\pi}})$ degrees of freedom. If d in (4.5) is assumed unknown a priori, the test may be generalized to that situation along the lines in Luukkonen et al. (1988a). Note that if $n = 2$ and $c_1 = -c_2$ the alternative is symmetric in that a positive shock and a negative one of the same magnitude still have the same (mirror) effect on the conditional variance. The response to the shock, however, is a nonlinear function of lags of u_t^2 .

4.4. Testing parameter constancy

Testing parameter constancy is important in its own right but also because non-constancy may manifest itself as an apparent lack of weak stationarity (IGARCH); see, for example, Lamoureux and Lastrapes (1990). In this chapter we assume that the alternative to constant parameters in the conditional variance is that the parameters, or a subset of them, change smoothly over time. Lin and Teräsvirta

(1994) applied this idea to testing parameter constancy in the conditional mean. We postulate that the time-varying parameter is $\boldsymbol{\eta} = \boldsymbol{\eta}^* + \boldsymbol{\lambda} \overline{H}_n(t; \gamma, \mathbf{c})$. If the null hypothesis only concerns a subset of parameters then only the corresponding elements in $\boldsymbol{\lambda}$ are assumed to be nonzero a priori. Again we define $\overline{H}_n = H_n - \frac{1}{2}$ and the transition function $H_n(t; \gamma, \mathbf{c})$ is assumed to be a logistic function of order n defined in (4.2) with $s_t \equiv t$. If $\gamma \rightarrow \infty$, $\overline{H}_1(t; \gamma, \mathbf{c})$ becomes a step-function and characterizes a single structural break in the model. Chu (1995) discussed testing parameter constancy against this alternative. The null hypothesis of parameter constancy becomes $H_0 : \gamma = 0$ under which $\boldsymbol{\eta}^* = \boldsymbol{\eta}$. By setting $\boldsymbol{\pi} = (\boldsymbol{\lambda}, \gamma, \mathbf{c})'$ we can write this as a special case of (4.1) with $G(\mathbf{z}_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi}) = (\boldsymbol{\lambda}' \mathbf{z}_t) \overline{H}_n(t; \gamma, \mathbf{c})$. We can again circumvent the lack of identification under the null hypothesis by a Taylor approximation of the transition function. A first-order Taylor-expansion of $\overline{H}_n(t; \gamma, \mathbf{c})$ around $\gamma = 0$ yields, after a reparameterization,

$$h_t = \boldsymbol{\beta}'_0 \mathbf{z}_t + \tilde{\boldsymbol{\pi}}' \mathbf{v}_t + R_2(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})$$

where $\tilde{\boldsymbol{\pi}} = (\boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_n)' = \gamma \boldsymbol{\pi}$ with $\boldsymbol{\pi} \neq \mathbf{0}$ and $\mathbf{v}_t = ((\mathbf{z}_t t)', \dots, (\mathbf{z}_t t^n)')'$. Thus, $G(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) = \tilde{\boldsymbol{\pi}}' \mathbf{v}_t + R_2(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi})$. Note, however, that $R_2(\mathbf{z}_t; \boldsymbol{\omega}, \boldsymbol{\pi}) \equiv 0$ under H_0 so that it does not affect the distribution theory. Our null hypothesis is $H_0 : \tilde{\boldsymbol{\pi}} = \mathbf{0}$. We note that components of \mathbf{w}_t are trending but modifying a corresponding proof in Lin and Teräsvirta (1994); the asymptotic null distribution of LM-type teststatistic (3.4) can be shown to be a χ^2 distribution with $\dim(\tilde{\boldsymbol{\pi}})$ degrees of freedom if the fourth moment of u_t exists.

An advantage of a parametric alternative to parameter constancy is that if the null hypothesis is rejected we can estimate the parameters of the alternative model. This helps us find out where in the sample the parameters under test seem to be changing and how rapid the change is. This is useful information if respecification of the model to achieve parameter constancy is attempted.

5. Simulation experiments

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

$$\begin{aligned} y_t &= u_t \\ u_t &= \varepsilon_t \sqrt{h_t}. \end{aligned} \tag{5.1}$$

The conditional variance h_t varies with the alternative hypothesis being testing against. Under the null hypothesis h_t is the conditional variance of the standard GARCH(1,1) model. The random numbers, ε_t , are generated by the random number generator in GAUSS 3.2.31. The random numbers sampled are all assumed to be normally distributed with expectation zero and unit variance. The first 200 observations of each generated series are always discarded to avoid initialization effects. For the tests against remaining structure, series of 2000 observations are generated. Series of 1000 observations are used for both the linearity test and the parameter constancy test. For each design a total of 1000 replications are performed.

5.1. Test of the functional form

First we consider the test of no ARCH in the standardized errors and the functional misspecification test of Section 4.2. We define a DGP such that the conditional variance either (a) follows a GARCH(1, 2) process or (b) follows a GARCH(2, 1) process. Thus,

$$h_t = 0.5 + 0.05u_{t-1}^2 + \alpha u_{t-2}^2 + 0.9h_{t-1} \quad (5.2a)$$

$$h_t = 0.5 + 0.05u_{t-1}^2 + 0.9h_{t-1} + \beta h_{t-2} \quad (5.2b)$$

In the experiment the values of α and β vary within limits such that the conditional variance of the process remains positive with probability 1, see Nelson and Cao (1992), and the condition for covariance stationarity holds. For (a) this is the case when $-0.045 < \alpha < 0.05$ and for (b) when $-0.2025 < \beta < 0.1$. For $\alpha = 0$ in (a) and $\beta = 0$ in (b) the DGP reduces to a standard GARCH(1,1) model. The results for the test of functional misspecification and for the test of no ARCH in the standardized errors for (a) and (b) are reported in Figures B.1 and B.2. The nominal significance level equals 0.1, and we use the test of Bollerslev (1986) as a benchmark. In the simulations these tests were all computed with a single parameter in the alternative, that is, with either $\hat{v}_t = \hat{u}_{t-1}^2/h_{t-1}$ (no ARCH) or with $\hat{v}_t = \hat{\eta}\mathbf{z}_{t-1}$ where $\mathbf{z}_{t-1} = (1, \hat{u}_{t-1}^2, \hat{h}_{t-1})$ (functional form). We also used Bollerslev's test such that the alternative was in the DGP. In practice one does not know if the unconditional fourth moment of $\{u_t\}$ exists, although the estimated model does contain information about this. It is therefore of interest to investigate how the tests behave when the unconditional fourth moment does not exist; the existence condition is given in He and Teräsvirta (1999a). This is the case for parameter

values on the right-hand side of the dashed vertical line in Figures B.1 and B.2.

In the experiment where the conditional variance of the DGP is (5.2a) the power of the functional form test is lower than that of Bollerslev's test. On the other hand, for positive values of β when the DGP is (5.2b) the test of the functional form performs better than Bollerslev's test. The latter test has no power against the GARCH(2,1) alternative when the unconditional fourth moment exists. When the existence condition for the unconditional variance is violated, the power of the functional form test first sharply increases and then decreases as a function of β . The power of Bollerslev's test increases as a function of β without dropping again. The power of our test against remaining ARCH(1), which is asymptotically equivalent to the test of Li and Mak (1994), is found to be equal to that of Bollerslev's test in both cases (a) and (b).

It may be pointed out already that these tests have no power when the DGP is a nonlinear GJR-GARCH(1,1) model (5.3) or contains a structural break (5.4). Simulations of other tests against such alternatives will be discussed in the next two sections.

5.2. Testing linearity (symmetry)

In this section we consider the small-sample performance of the linearity test. The test can be expected to be powerful against smooth transition alternatives for which it is designed. Considering its performance against, say, the GJR-GARCH model would constitute a tougher trial for our test and make it possible to compare the performance of the tests with that of the sign-bias test of Engle and Ng (1993). The DGP is (2.2) with

$$h_t = 0.005 + 0.28 [|u_{t-1}| + \omega u_{t-1}]^2 + 0.7h_{t-1} \quad (5.3)$$

where u_t is assumed conditionally normal. For (5.3), $|\omega| < 0.267$ is required for covariance stationarity. In this special case the unconditional fourth moment also exists under this restriction; the relevant moment condition appeared in He and Teräsvirta (1999b). The joint sign-bias test of Engle and Ng (1993) mentioned above is designed for detecting asymmetry in the conditional variance. We compute the values of the sign-bias test statistic by using (3.3); the difference between this test and the TR^2 version, suggested by Engle and Ng (1993), is negligible at our sample size. Engle and Ng (1993) used (2.2) and (5.3) with $\omega = -0.23$ as the DGP in their

evaluation of the sign-bias test. Note that for $\omega = 0$ the DGP reduces to a standard GARCH(1,1) model. The power simulations are reported in Figure B.3 for $\omega \leq 0$. For positive values of ω the results are similar and therefore not reported. Our smooth transition GARCH test was computed by assuming $n = 3$ in (4.2). The power of the test compares very favourably with that of the sign-bias test. This accords with the results in Hagerud (1997)

5.3. Testing parameter constancy

We consider two cases of parameter nonconstancy: the DGP is a GARCH(1,1) model with either (a) a single or (b) a double structural break. We did not simulate smooth parameter change because our test can be expected to perform well against such an alternative. Our choice of alternative also gives us an opportunity to compare our test against that of Chu (1995), which is a test against a single structural break. If T is the total number of observations, the single structural break parametrization corresponding to alternative (a) is assumed to have a change at time ηT where η lies between 0 to 1. The double structural break parametrization corresponding to (b) first postulates a change at time $\eta_1 T$ and a return to the original parameters at $\eta_2 T$, $0 \leq \eta_1 < \eta_2 \leq 1$. The test of Chu (1995) should outperform ours in case (a). We use the version of Chu's test that assumes normal errors. Our test is computed with $n = 1$, case (a), and $n = 2$, case (b), where n is the order of the logistic function in (4.2).

We consider the following model for a change in the constant term:

$$\begin{aligned} h_t &= 0.5 + 0.1u_{t-1}^2 + 0.8h_{t-1}, \quad (\text{a}) \ t < \eta T, \quad (\text{b}) \ t < \eta_1 T, t > \eta_2 T \\ h_t &= 0.5(1 + \Delta) + 0.1u_{t-1}^2 + 0.8h_{t-1}, \quad (\text{a}) \ t \geq \eta T, \quad (\text{b}) \ \eta_1 T \leq t \leq \eta_2 T \end{aligned} \quad (5.4)$$

where $\Delta = 0.4, 0.8$. Chu (1995) used (a) in (5.4) as the DGP in his own simulation experiments. The power simulations for the DGPs in (5.4) with a single structural break at η for $\Delta = 0.4$ and 0.8 appear in Figures B.4 and B.5. The values $\eta = 0, 1$ correspond to the null hypothesis.

For $0.3 < \eta < 0.8$ our test has the same as or higher power than the test of Chu (1995), but otherwise the relationship is the opposite. This occurs for both $\Delta = 0.4$ and $\Delta = 0.8$, which includes these DGPs. Thus the Chu test does not dominate ours as one might expect. We can see that a small change, $\Delta = 0.4$, is difficult to detect. As Δ doubles to 0.8 , the change is detected more easily. For comparison, we also simulated another version of the test in which we assumed that only the constant

term is time-varying under the alternative. In that case our test outperforms that of Chu (1995) for almost all η . In yet another experiment we allowed the coefficient of u_{t-1}^2 to change once within the sample period. The behaviour of the tests was similar to the previous case and the details are not reported here.

We turn to the case of a double structural break. The DGP for the experiment is such that $\eta_2 = \eta_1 + 0.3$ where η_1 is varied from 0 to 0.7. Thus for $\eta_1 = 0$ and $\eta_1 = 0.7$ the DGP has only a single structural break. The power simulations of this experiment for $\Delta = 0.4$ and 0.8 can be found in Figures B.6 and B.7. In this case the test of Chu (1995) cannot be expected to be very powerful because the design of the experiment does not favour it, and our test does have superior power for all double break points considered. The power of the Chu test is high for η_1 close to zero and 0.7 because the test is designed for detecting a single structural break.

6. Conclusions

In this chapter we have derived a unified framework for testing the adequacy of an estimated GARCH model. Our selection contains a number of new tests while some existing ones fit into this framework as well. Nothing more complicated than standard asymptotic distribution theory is required. As a result, misspecification of a GARCH model may be detected quite easily at low computational cost. Because the tests of symmetry and parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This helps the model builder find out the weaknesses of the estimated specification and may give useful ideas of how the current specification could be improved.

We also show that our test of no ARCH in the standardized error process is asymptotically equivalent to a portmanteau test of Li and Mak (1994). This links the work of these authors to our framework and indicates that the null hypothesis of no remaining ARCH can be tested in different ways while the asymptotic theory remains the same. This chapter does not contain any applications of these new tests. Empirical examples can be found instead in the Chapters 2,4 and 5.

References

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**: 307–327.
- Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews* **11**: 143–172.
- Chu, C.-S. J. (1995). Detecting parameter shift in GARCH models, *Econometric Reviews* **14**: 241–266.
- Eitrheim, Ø. and Teräsvirta, T. (1996). Testing the adequacy of smooth transition autoregressive models, *Journal of Econometrics* **74**: 59–75.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**: 987–1007.
- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *Journal of Finance* **48**: 1749–1777.
- Fiorentini, G., Calzolari, G. and Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics* **11**: 399–417.
- Glosten, L., Jagannathan, R. and Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**: 1779–1801.
- González-Rivera, G. (1998). Smooth transition GARCH models, *Studies in Non-linear Dynamics and Econometrics* **3**: 61–78.
- Hagerud, G. E. (1997). *A New Non-Linear GARCH Model*, EFI, Stockholm School of Economics.
- Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* **64**: 413–430.
- He, C. and Teräsvirta, T. (1999a). Fourth moment structure of the GARCH(p, q) process, *Econometric Theory* (forthcoming) .
- He, C. and Teräsvirta, T. (1999b). Properties of moments of a family of GARCH processes, *Journal of Econometrics* (forthcoming) .

- He, C. and Teräsvirta, T. (1999c). Properties of the autocorrelation function of squared observations for second order GARCH processes under two sets of parameter constraints, *Journal of Time Series Analysis* **20**: 23–30.
- Lamoureux, C. G. and Lastrapes, W. G. (1990). Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**: 225–234.
- Lee, Y. N. and Li, W. K. (1998). On smooth transition double threshold models. Department of Statistics, University of Hong Kong, Research Report No. 198.
- Li, W. K. and Mak, T. K. (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity, *Journal of Time Series Analysis* **15**: 627–636.
- Lin, C.-F. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**: 211–228.
- Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, *Econometrica* **64**: 575–596.
- Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988a). Testing linearity against smooth transition autoregressive models, *Biometrika* **75**: 491–499.
- Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988b). Testing linearity in univariate time series models, *Scandinavian Journal of Statistics* **15**: 161–175.
- McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations, *Journal of Time Series Analysis* **4**: 269–273.
- Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics* **10**: 229–235.
- Sentana, E. (1995). Quadratic ARCH models, *Review of Economic Studies* **62**: 639–661.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**: 208–218.

A. Information matrix

Consider the model defined by (2.1) and (2.2):

$$\begin{aligned} y_t &= \boldsymbol{\varphi}' \mathbf{w}_t + u_t \\ u_t &= \varepsilon_t \sqrt{h(z_t; \boldsymbol{\eta}, \boldsymbol{\pi})} \end{aligned}$$

where $h_t = h(z_t; \boldsymbol{\varphi}, \boldsymbol{\eta}, \boldsymbol{\pi})$ is the parametrization of the conditional variance including the alternative. The assumed null hypothesis is $H_0 : \boldsymbol{\pi} = \mathbf{0}$. If we assume that $\{\varepsilon_t\}$ is a sequence of independent standard normal errors, the log-likelihood function at time t is given by:

$$l_t = \text{const} - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{u_t^2}{h_t}$$

where $u_t = y_t - \boldsymbol{\varphi}' \mathbf{w}_t$ as $\boldsymbol{\varphi}$ is not assumed to depend on $\boldsymbol{\eta}$ or $\boldsymbol{\pi}$.

A.1. Partial derivative of l_t

The first-order partial derivative (the gradient) of the log likelihood function at time t is

$$G_t = \begin{pmatrix} \frac{\partial l_t}{\partial \boldsymbol{\varphi}'} & \frac{\partial l_t}{\partial \boldsymbol{\eta}'} & \frac{\partial l_t}{\partial \boldsymbol{\pi}'} \end{pmatrix}$$

where the corresponding elements are

$$\begin{aligned} \frac{\partial l_t}{\partial \boldsymbol{\varphi}'} &= \frac{u_t}{h_t} \mathbf{w}_t' + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \\ \frac{\partial l_t}{\partial \boldsymbol{\eta}'} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\eta}'} \\ \frac{\partial l_t}{\partial \boldsymbol{\pi}'} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\pi}'} \end{aligned}$$

The second-order partial derivative (the Hessian) of the log-likelihood function at time t is

$$H_t = \begin{pmatrix} \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\pi}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\eta} \partial \boldsymbol{\pi}'} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\pi} \partial \boldsymbol{\varphi}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\pi} \partial \boldsymbol{\eta}'} & \frac{\partial^2 l_t}{\partial \boldsymbol{\pi} \partial \boldsymbol{\pi}'} \end{pmatrix}. \quad (\text{A.1})$$

If H_0 holds, the expectations of the elements in (A.1) are

$$\begin{aligned} E \left[\frac{\partial^2 l_t}{\partial \varphi \partial \varphi'} \middle| \pi=0 \right] &= -E \left[\frac{1}{h_t} \mathbf{w}_t \mathbf{w}_t' + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \varphi'} \right] & E \left[\frac{\partial^2 l_t}{\partial \varphi \partial \eta'} \middle| \pi=0 \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \eta'} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \eta \partial \eta'} \middle| \pi=0 \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \eta} \frac{\partial h_t}{\partial \eta'} \right] & E \left[\frac{\partial^2 l_t}{\partial \varphi \partial \pi'} \middle| \pi=0 \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \varphi} \frac{\partial h_t}{\partial \pi'} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \pi \partial \pi'} \middle| \pi=0 \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \pi} \frac{\partial h_t}{\partial \pi'} \right] & E \left[\frac{\partial^2 l_t}{\partial \pi \partial \eta'} \middle| \pi=0 \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \pi} \frac{\partial h_t}{\partial \eta'} \right]. \end{aligned}$$

A.2. Partial derivative of the multiplicative conditional variance h_t .

Assume that the conditional variance under the alternative hypothesis is parameterized as in (3.2). The conditional variance is then

$$h_t = (\boldsymbol{\eta}' \mathbf{z}_t)(1 + \boldsymbol{\pi}' \mathbf{v}_t) = \left(\alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \right) (1 + \boldsymbol{\pi}' \mathbf{v}_t)$$

which reduces into the standard GARCH model under H_0 . To initialize the iterative computation of h_t under null hypothesis the conditional variance is estimated with the unconditional variance (sample variance) in the pre-sample case. This is done for all $t \leq 0$ by setting $h_t = u_t^2 = \frac{1}{T} \sum_{s=1}^T u_s^2$ where $u_s = y_s - \boldsymbol{\varphi}' \mathbf{w}_s$.

To compute the test statistic (3.4) we need the first-order partial derivatives of the conditional variance h_t under the null hypothesis.

First-order derivative

$$\frac{\partial h_t}{\partial \varphi'} \middle| \pi=0 = -2 \sum_{i=1}^q \alpha_i u_{t-i} \mathbf{w}'_{t-i} + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \varphi'} \middle| \pi=0$$

$$\frac{\partial h_t}{\partial \eta'} \middle| \pi=0 = \mathbf{z}'_t + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \eta'} \middle| \pi=0$$

$$\frac{\partial h_t}{\partial \pi'} \middle| \pi=0 = (\boldsymbol{\eta}' \mathbf{z}_t) \mathbf{v}'_t$$

Pre-sample values ($t \leq 0$)

$$-2 \mathbf{u}_t \mathbf{w}'_t \middle| \pi=0 = \frac{\partial h_t}{\partial \varphi'} \middle| \pi=0 = -\frac{2}{T} \sum_{s=1}^T u_s \mathbf{w}'_s$$

$$\frac{\partial h_t}{\partial \eta'} \middle| \pi=0 = 0$$

$$\frac{\partial h_t}{\partial \pi'} \middle| \pi=0 = 0$$

A.3. Partial derivative of the additive conditional variance h_t .

Assume that the conditional variance under the alternative hypothesis is parameterized as in (4.1) The conditional variance is then

$$h_t = \boldsymbol{\eta}' \mathbf{z}_t + \boldsymbol{\pi}' \mathbf{v}_t = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} + \boldsymbol{\pi}' \mathbf{v}_t$$

which reduces to the standard GARCH model under H_0 . The iterative computation of h_t under null hypothesis is initialized in the same way as for the multiplicative

conditional variance model. The first-order derivatives of the conditional variance h_t under the null hypothesis are required to compute the test statistic (3.3). These derivatives are given as follows:

First-order derivative

$$\frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = -2 \sum_{i=1}^q \alpha_i u_{t-i} \mathbf{w}'_{t-i} + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\varphi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}}$$

$$\frac{\partial h_t}{\partial \boldsymbol{\eta}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{z}'_t + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\eta}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}}$$

$$\frac{\partial h_t}{\partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{v}'_t + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}}$$

Pre-sample values ($t \leq 0$)

$$-2u_t \mathbf{w}'_t \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \frac{\partial h_t}{\partial \boldsymbol{\varphi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = -\frac{2}{T} \sum_{s=1}^T u_s \mathbf{w}'_t$$

$$\frac{\partial h_t}{\partial \boldsymbol{\eta}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{0}$$

$$\frac{\partial h_t}{\partial \boldsymbol{\pi}'} \Big|_{\boldsymbol{\pi}=\mathbf{0}} = \mathbf{0}$$

B. Figures

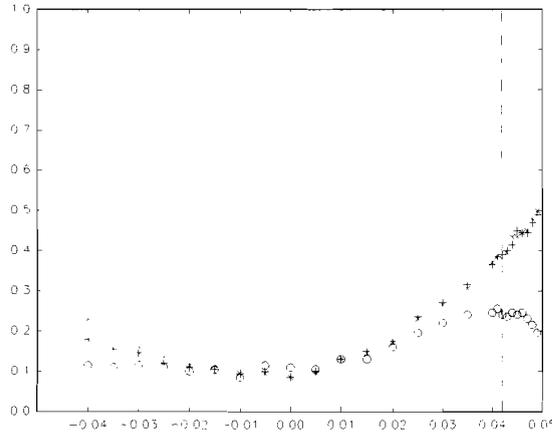


Figure B.1: Power simulations at significance level 0.1 for the no ARCH test (cross), the functional form test (circle) and Bollerslev's test (plus). The DGP is a GARCH(1,2). The value of α in (5.2a) is given on the x-axis. The null hypothesis is the GARCH(1,1) model.

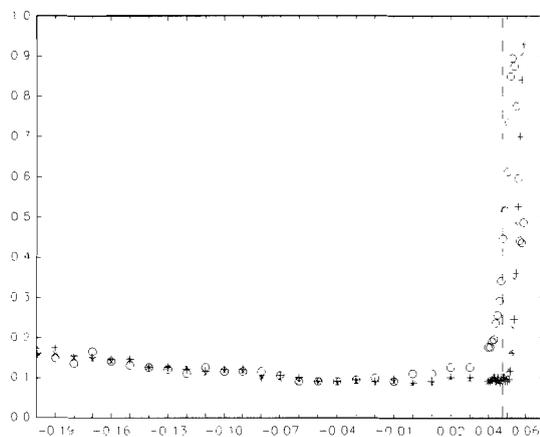


Figure B.2: Power simulations at significance level 0.1 for the no ARCH test (cross), the functional form test (circle) and Bollerslev's test (plus). The DGP is a GARCH(2,1). The value of β in (5.2b) is given on the x-axis. The null hypothesis is the GARCH(1,1) model.

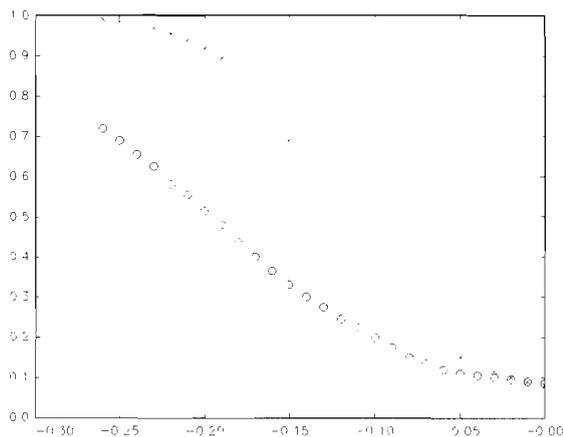


Figure B.3: Power simulations at significance level 0.1 for the linearity (symmetry) test (cross) and the sign-bias test (circle). The DGP is a GJR-GARCH. The value of ω in (5.3) is given on the x-axis. The null hypothesis is the GARCH(1,1) model.

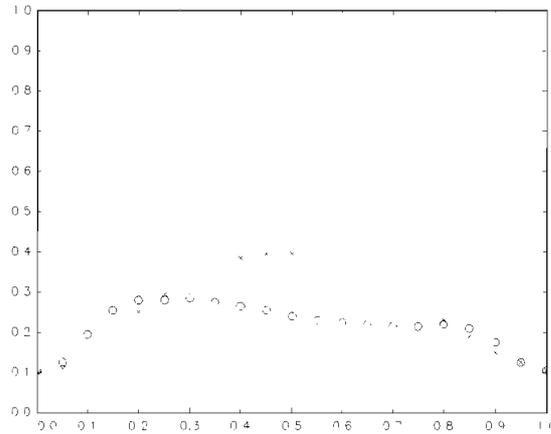


Figure B.4: Power simulations at significance level 0.1 for the parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) with a single structural shift of size $\Delta = 0.4$ at η . In the figure η is given on the x-axis. The null hypothesis is a constant parameter GARCH(1,1) model.

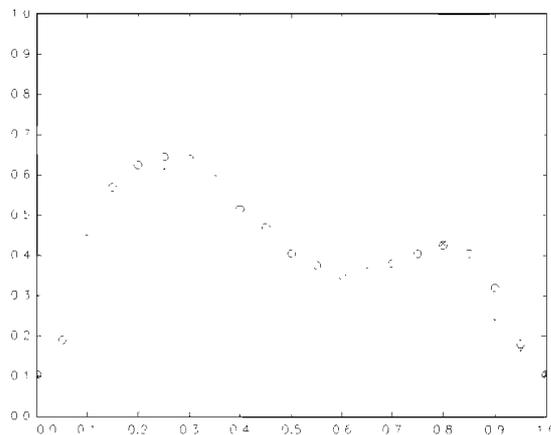


Figure B.5: Power simulations at significance level 0.1 for the parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) with a single structural shift of size $\Delta = 0.8$ at η . In the figure η is given on the x-axis. The null hypothesis is a constant parameter GARCH(1,1) model.

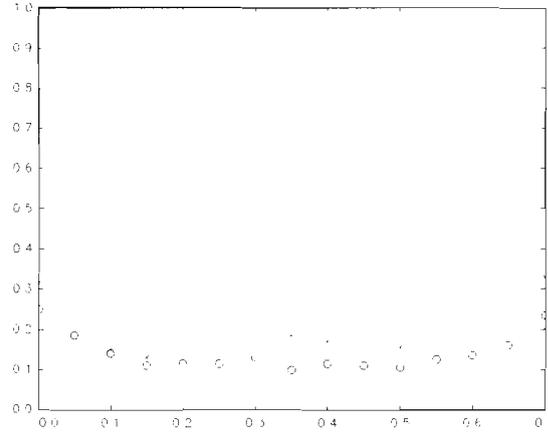


Figure B.6: Power simulations at significance level 0.1 for our parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) process with a double structural shift in the constant of size $\Delta = 0.4$. The first shift appearing begins at η_1 and the return to original parameter values occurs at η_2 . In the figure the value of η_1 is given on the x-axis and defined such that $\eta_2 = \eta_1 + 0.3$. The null hypothesis is a constant parameter GARCH(1,1) model.

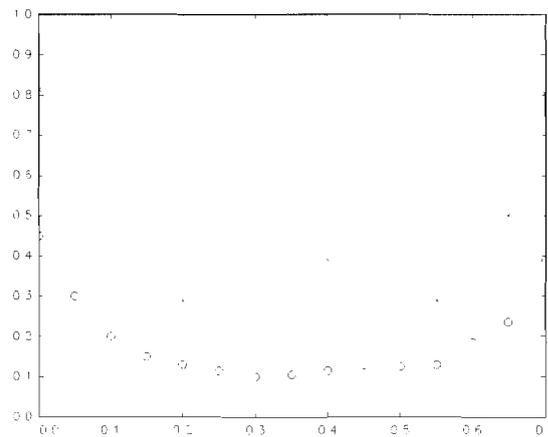


Figure B.7: Power simulations at significance level 0.1 for our parameter constancy test (cross) and the test of Chu (circle). The DGP is a GARCH(1,1) process with a double structural shift in the constant of size $\Delta = 0.8$. The first shift appearing begins at η_1 and the return to original parameter values occurs at η_2 . In the figure the value of η_1 is given on the x-axis and defined such that $\eta_2 = \eta_1 + 0.3$. The null hypothesis is a constant parameter GARCH(1,1) model.



Chapter 4

A GARCH model with time-varying parameters

1. Introduction

When modelling the conditional mean of a univariate process, it may be difficult to distinguish between an autoregressive process with a structural break and a random walk (unit root) process, see for example Hendry and Neale (1991). This observation is of importance when modelling empirical series, since a regime shift may generate "unit root like" behaviour in piecewise stationary autoregressive time series. This may put the modeller in a situation where it is necessary to choose between unit roots or stationarity with structural breaks. A vast literature has emerged on this topic, starting with Perron (1989). A similar problem is apparent when modelling the conditional variance. Parameter nonconstancy may manifest itself as an apparent lack of weak stationarity (IGARCH); see, for example, Diebold (1986) and Lamoureux and Lastrapes (1990). In this paper we attempt to model a structural shift in the conditional variance. A reasonable assumption is that the transition between the regimes from one structure to the other is smooth. A model capable of describing this kind of behaviour is the smooth transition GARCH (STGARCH) where the transition is a deterministic function of time. We suggest a modelling cycle for the time-varying GARCH model, consisting of specification, estimation and evaluation. A parameter constancy test designed against a smooth

structural shift but also having power against structural shifts was suggested in Chapter 3 and will be applied here.

The plan of the paper is as follows. We define the model in Section 2, discuss its specification and estimation in Section 3 and consider the evaluation in Section 4. Section 5 contains a small simulation study that illustrates what may happen if one attempts to estimate a standard constant parameter GARCH model when the parameters in fact change over time. To demonstrate these ideas and the applicability of our model we apply the time-varying GARCH model to Affärsvärdens share index and the USD/GBP exchange rate in Section 6. Section 7 concludes.

2. The Model

For simplicity we assume that the conditional mean of the process equals zero. This restriction can easily be relaxed to allow for structure in the conditional mean. See Chapter 2 for an example where the conditional mean is parameterized by a smooth transition autoregressive model. We consider a GARCH(p, q) parameterization in which certain parameters are allowed to vary smoothly over time. This may be viewed as a special case of the following additive model in which the conditional mean has no parametric structure:

$$\begin{aligned} y_t &= u_t \\ u_t &= \varepsilon_t \sqrt{h_t} \end{aligned} \tag{2.1}$$

where $\{\varepsilon_t\} \sim \text{nid}(0, 1)$ and $h_t = \boldsymbol{\delta}_t(\boldsymbol{\zeta})' \mathbf{z}_t$ is the conditional variance which is stochastically independent of ε_t and positive with probability one. Parameter $\boldsymbol{\delta}_t$ is assumed to vary deterministically over time and the stochastic vector \mathbf{z}_t will be defined later. We also assume that $\boldsymbol{\delta}_t(\boldsymbol{\zeta})' \mathbf{z}_t$ is at least twice continuously differentiable for $\boldsymbol{\zeta} \in \Xi$ everywhere in the sample space. The normality assumption of errors $\{\varepsilon_t\}$ is not necessary but is retained for deriving the asymptotic theory of inference. It is assumed that the parameters are subject to restrictions such that the process is weakly stationary and ergodic, which requires the existence of at least fourth-order moments. Furthermore, it is assumed that the moments necessary for the asymptotic theory to go through exist.

In order to define the vector of time-varying parameters, $\delta_t(\zeta)$, let

$$H_n(t; \gamma, \mathbf{c}) = \left(1 + \exp(-\gamma \prod_{l=1}^n (t - c_l)) \right)^{-1}, \gamma > 0, c_1 \leq \dots \leq c_n \quad (2.2)$$

where the time, t , is the deterministic transition variable, γ the slope parameter and $\mathbf{c} = (c_1, \dots, c_n)'$ the location vector. The parameter restrictions $\gamma > 0$ and $c_1 \leq \dots \leq c_n$ are identifying restrictions. The value of the logistic function (2.2), which is bounded between a and 1, $0 \leq a \leq 1/2$, depends on the transition variable t . If $\gamma \rightarrow \infty$ then $a \rightarrow 0$. In this paper we rescale t between $1/T$ and 1 where T is the number of observations. Note that for $\gamma = 0$, $H_n(t; \gamma, \mathbf{c}) \equiv 1/2$ and when $\gamma \rightarrow \infty$, and $n = 1$, $H_n(t; \gamma, \mathbf{c})$ becomes a step function. It becomes a "multistep" function as $\gamma \rightarrow \infty$, if $n > 1$. The cases of a single structural break and, more generally, that in which the process switches between two parameterizations are thus included in this specification as special cases.

The time-varying parameter GARCH model is defined as follows. Set $\zeta = (\boldsymbol{\eta}', \boldsymbol{\lambda}', \gamma, \mathbf{c}')'$. This allows us to write the time-varying conditional variance as

$$h_t = (\boldsymbol{\eta} + \boldsymbol{\lambda}H_n(t; \gamma, \mathbf{c}))' \mathbf{z}_t \quad (2.3)$$

where $\boldsymbol{\eta} = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$, $\boldsymbol{\lambda} = (\alpha_{20}, \alpha_{21}, \dots, \alpha_{2q}, \beta_{21}, \dots, \beta_{2p})'$ and $\mathbf{z}_t = (1, u_{t-1}^2, \dots, u_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$. The time-varying GARCH model collapses into the standard GARCH model when $\gamma = 0$ and $h_t = \boldsymbol{\eta}' \mathbf{z}_t > 0$ for all t . Sufficient conditions for positivity of the conditional variance are $\alpha_0 > 0$, $\alpha_0 + \sum_{j=1}^q \alpha_{0j} > 0$, $\alpha_j \geq 0$, $\alpha_j + \alpha_{2j} \geq 0$, $j = 1, \dots, q$; $\beta_j \geq 0$, $\beta_j + \beta_{2j} \geq 0$, $j = 1, \dots, p$. Nelson and Cao (1992) showed how these conditions that are necessary for $j = 1$ may be relaxed for the standard GARCH model. Similar relaxation is obviously possible for the time-varying GARCH model but is not discussed in detail here. In the following section we consider the specification and estimation of this model.

3. Specification and estimation

When modelling time series with time-varying GARCH models we have to select the dimensions of p and q for the TVGARCH(n, p, q) process. The order ($n = 1, 2, \dots, n_0$) for the logistic transition function of the TVGARCH has to be specified from the data as well because usually it cannot be assumed to be known in advance. All this requires a coherent specification strategy of the kind intro-

duced, for example, in Box and Jenkins (1970), Li and Li (1996), Tsay (1989, 1998), Teräsvirta (1994, 1998) and Chapter 2. We propose the following strategy.

1. First test the null hypothesis of no ARCH against high-order ARCH. This serves as a test against GARCH.
2. If the null hypothesis is rejected, estimate the parameters of the standard GARCH(1,1) model and evaluate the model by appropriate misspecification tests. Expand to a higher order GARCH model, if necessary. Test the final model against parameter nonconstancy. If parameter constancy is rejected, estimate a TVGARCH model.
3. For the TVGARCH estimation we have to choose the order of the logistic function n . A test sequence such as the one in Teräsvirta (1994) may be applied. On the other hand, with the computational power that an average modeller has at his/her disposal today it is quite easy to estimate the parameters for $n = 1, 2, \dots, n_0$ and select the best model thereafter.
4. If the selected model passes the misspecification tests, tentatively accept it. In the opposite case try another specification search or choose another family of models.

Time is standardized, as mentioned before, such that $0 \leq t^* \leq 1$ where $t^* = t/T$. In estimating the parameters at stage 3 we obtain information about the optimal n . For example, if \hat{c}_i lies so far (depending on the slope $\hat{\gamma}$) outside the range of t^* that it does not affect the values of the transition function within $0 \leq t^* \leq 1$, we can discard either $c_i \leq \dots \leq c_1$ or $c_i \geq c_{i+1} \geq \dots \geq c_n$ depending on whether $\hat{c}_i < 1/T$ or $\hat{c}_i > 1$ and use a lower order transition function instead.

All parameter estimates are obtained by maximizing the log-likelihood under the assumption that $\{\varepsilon_t\}$ is a sequence of independent standard normal errors. In that case we write the log-likelihood function at time t as

$$l_t = \text{const} - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{u_t^2}{h_t} \quad (3.1)$$

where u_t and h_t are defined in (2.1). We assume that the model under consideration satisfies the necessary regularity conditions for stationarity and ergodicity needed for consistency and asymptotic normality of the estimates. In the following we consider the specification stages in detail.

3.1. A test of parameter constancy in GARCH models

After estimating a suitable GARCH(p, q) model it is straightforward to construct a Lagrange multiplier (or score) type test for parameter constancy against smooth structural change following Lin and Teräsvirta (1994) and Chapter 3. It is of importance that the squared standardized errors are not autocorrelated because remaining serial correlation may cause size distortion in the test and thus be mistaken for parameter nonconstancy.

The model comprising the alternative of smoothly time-varying parameters is defined in (2.1) with

$$h_t = (\boldsymbol{\eta} + \boldsymbol{\lambda} \bar{H}_n(t, \gamma, \mathbf{c}))' \mathbf{z}_t \quad (3.2)$$

where $\{\varepsilon_t\}$ is a sequence of independent standard normal variables. We define $\bar{H}_n = H_n - \frac{1}{2}$; subtracting $1/2$ from H_n is just a notational convenience in deriving the test and does not affect the generality of the argument. The null hypothesis of parameter constancy is $H_0 : \gamma = 0$ whereas the alternative is $H_1 : \gamma > 0$. If the null hypothesis concerns a subset of parameters then only the corresponding elements in $\boldsymbol{\lambda}$ are assumed to be nonzero a priori. The transition function H_n is assumed to be a logistic function of order n defined in (2.2) so that under the alternative the parameters, or a subset of them, change smoothly over time. We can circumvent the lack of identification under the null hypothesis by a Taylor approximation of the logistic transition function, which in practice means trading off information about the structure of the alternative against a standard asymptotic distribution theory. A first-order Taylor-expansion of \bar{H}_n around $\gamma = 0$ yields, after a reparameterization,

$$h_t = \boldsymbol{\beta}'_0 \mathbf{z}_t + \tilde{\boldsymbol{\pi}}' \mathbf{v}_t + R_1(\mathbf{z}_t; \boldsymbol{\eta}, \boldsymbol{\lambda}, \gamma, \mathbf{c})$$

where it can be seen that $\tilde{\boldsymbol{\pi}} = (\boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_n)' = \gamma(\bar{\boldsymbol{\beta}}'_1, \dots, \bar{\boldsymbol{\beta}}'_n)' = \gamma \bar{\boldsymbol{\pi}}$ with $\bar{\boldsymbol{\pi}} \neq \mathbf{0}$ and that $\mathbf{v}_t = ((\mathbf{z}_t t)', \dots, (\mathbf{z}_t t^n)')'$. Note, however, that $R_1(\mathbf{z}_t; \boldsymbol{\eta}, \boldsymbol{\lambda}, \gamma, \mathbf{c}) \equiv 0$ under H_0 so that the remainder does not affect the asymptotic distribution theory. Our new null hypothesis is $H_0 : \tilde{\boldsymbol{\pi}} = \mathbf{0}$, under which $\boldsymbol{\beta}_0 = \boldsymbol{\eta}$. See Chapter 3 for a similar development. The Lagrange multiplier (or score) type test statistic is then defined as

$$\text{LM} = T \left(\begin{array}{c} 0 \\ \frac{1}{T} \sum \frac{\partial \hat{h}_t}{\partial \tilde{\boldsymbol{\pi}}} |_{H_0} \end{array} \right)' \hat{\mathbf{I}}(\hat{\boldsymbol{\eta}}, \tilde{\boldsymbol{\pi}})^{-1} |_{H_0} \left(\begin{array}{c} 0 \\ \frac{1}{T} \sum \frac{\partial \hat{h}_t}{\partial \tilde{\boldsymbol{\pi}}} |_{H_0} \end{array} \right) \quad (3.3)$$

where $\frac{1}{T} \sum \frac{\partial \hat{h}_t}{\partial \tilde{\boldsymbol{\pi}}} |_{H_0}$ is the average score evaluated under H_0 and $\hat{\mathbf{I}}$ is a consistent estimator of the population information matrix under the null hypothesis. We use

the expectation of the negative average Hessian as our estimator for the population information matrix. The partial derivatives forming the Hessian may be found in Appendix A. Under the null hypothesis the test statistic (3.3) is asymptotically χ^2 -distributed with $\dim(\tilde{\pi})$ degrees of freedom. Since the GARCH(p,q) model is symmetric in the sense that the response of the process to a shock is similar regardless of the sign of the shock, the parameter constancy test may be computed simply by an artificial regression.

4. Misspecification of structure

In this section we present three different misspecification tests for an estimated time-varying GARCH model. These tests are modifications of misspecification tests in Chapter 3. The first test can be interpreted as a test of the functional form. The second one is a test against nonlinearity or, in some cases, asymmetry. The last one is a test against remaining parameter nonconstancy. To describe the common features in these tests we first introduce a general structure and thereafter briefly consider each test separately. We also need a general test of independence of the (standardized) errors. The BDS statistic (Brock, Dechert, Scheinkman and LeBaron, 1996) appears to be a suitable large-sample test for this purpose, see Chapter 2.

4.1. General structure

Consider the conditional variance model (3.2) and define

$$h_t = \boldsymbol{\delta}_t(\boldsymbol{\zeta})' \mathbf{z}_t + G(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi}). \quad (4.1)$$

We assume that h_t satisfies the regularity conditions mentioned in Section 3.1, and that $G(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi})$ is at least twice continuously differentiable for all elements of $\boldsymbol{\pi} \in \Pi$ everywhere in the corresponding sample space. We also assume that the necessary moments of $\{u_t\}$ exist. By assuming without loss of generality that $G(\mathbf{z}_t; \boldsymbol{\zeta}, \mathbf{0}) \equiv \mathbf{0}$ the null hypothesis of no additional structure in h_t becomes $H_0 : \boldsymbol{\pi} = \mathbf{0}$, under which $h_t = \boldsymbol{\delta}_t(\boldsymbol{\zeta})' \mathbf{z}_t$. The Lagrange multiplier (or score) test statistic is again (3.3) which, due to the boundedness of H_n , is asymptotically χ^2 -distributed with $\dim(\boldsymbol{\pi})$ degrees of freedom under the null hypothesis and the required regularity conditions. The partial derivatives in the Hessian needed for the estimation of the information matrix can be found in Appendix A.

4.2. Testing the functional form

A way of testing the null hypothesis of no error autocorrelation in the squared residuals is to lag $\delta_t(\zeta)'z_t$ and enter it in the conditional variance process, h_t , under the alternative. The test is obtained by defining $G(z_t; \zeta, \pi) = \pi'v_t$ where $\pi = (\pi_1, \dots, \pi_r)'$ and $v_t = (\delta_t(\zeta)'z_{t-1}, \dots, \delta_t(\zeta)'z_{t-r})'$, in (4.1). The null hypothesis of no serial dependence in the squared standardized errors or no model misspecification is $H_0 : \pi = 0$. Under the null hypothesis and the assumption $Eu_t^4 < \infty$, the LM-statistic (3.3) is asymptotically χ^2 -distributed with $\dim(\pi)$ degrees of freedom.

4.3. Testing linearity (symmetry)

When modelling return series it may be interesting to see if the estimated linear (symmetric) TVGARCH model provides an adequate representation of the data, because it has been argued (see for example, Engle and Ng (1993); Hagerud (1997)) that the responses to negative and positive shocks to returns may not be symmetric. The alternative to symmetry in this paper is the so-called smooth transition GARCH, see Hagerud (1997), González-Rivera (1998) and Chapter 3. By letting $\pi = (\alpha_{301}, \alpha_{31}, \dots, \alpha_{30q}, \alpha_{3q}, \gamma^*, c^*)'$ this alternative may be written as a special case of (4.1) with $G(z_t; \zeta, \pi) = \sum_{j=1}^q \alpha_{30j} \bar{H}_{n^*}(u_{t-j}; \gamma^*, c^*) + \sum_{j=1}^q \{\alpha_{3j} \bar{H}_{n^*}(u_{t-j}; \gamma^*, c^*)\} u_{t-j}^2$. Again we define $\bar{H}_{n^*} = (H_{n^*} - \frac{1}{2})$ as in Section 3.1, but with the difference that we now use u_{t-j} instead of t as transition variable. No nonlinear structure is imposed on the h_{t-j} , $j = 1, \dots, p$, as the alternative structure is already very flexible even without such an extension. The null hypothesis of linearity is $H_0 : \gamma^* = 0$ under which $\bar{H}_{n^*}(u_{t-j}; \gamma^*, c^*) \equiv 0$.

The identification problem discussed earlier (the model is not identified under H_0) is circumvented by expanding the transition function \bar{H}_{n^*} into a Taylor series around $\gamma^* = 0$, and rearranging terms. This results in the extended model with $G(z_t; \zeta, \pi) = \tilde{\pi}'v_t + R(z_t; \zeta, \pi)$ where $\tilde{\pi} = (\tilde{\pi}_1, \dots, \tilde{\pi}_{(1+n^*)q})'$ and $v_t = (u_{t-1}, u_{t-1}^3, \dots, u_{t-1}^{n^*+2}, \dots, u_{t-q}, u_{t-q}^3, \dots, u_{t-q}^{n^*+2})'$. It can be seen that $\tilde{\pi} = \gamma^* \bar{\pi}$ with $\bar{\pi} \neq \mathbf{0}$. This being the case, the new null hypothesis is $H_0 : \tilde{\pi} = 0$. Note that if the null hypothesis holds, the remainder $R(z_t; \zeta, \pi) \equiv 0$ does not affect the distribution theory. The moment condition of $Eu_t^{2n^*+2} < \infty$ must hold for the asymptotic theory to go through. Under the null hypothesis the LM-type test statistic (3.3) is asymptotically χ^2 -distributed with $\dim(\tilde{\pi})$ degrees of freedom.

4.4. Testing parameter constancy

In this section we consider the possibility that there is remaining parameter non-constancy of smooth transition type in the conditional variance. Our alternative to parameter constancy is that the parameters of the estimated model change deterministically over time. In our notation $\boldsymbol{\eta}$ and $\boldsymbol{\lambda}$ in (2.3) are replaced with $\boldsymbol{\eta}(t^*) = \boldsymbol{\eta}^* + \mathbf{a}\overline{H}_{n_a}(t^*; \gamma_a, \mathbf{c}_a)$ and $\boldsymbol{\lambda}(t^*) = \boldsymbol{\lambda}^* + \mathbf{b}\overline{H}_{n_b}(t^*; \gamma_b, \mathbf{c}_b)$. The vectors $\boldsymbol{\eta}^*, \mathbf{a}$ and $\boldsymbol{\lambda}^*, \mathbf{b}$ are of the same size as $\boldsymbol{\eta}$ and $\boldsymbol{\lambda}$ respectively. Again we define $\overline{H}_{n_i} = H_{n_i} - \frac{1}{2}$, $i = a, b$. We thus have a "double nonconstancy" compared to a constant parameter GARCH model. The transition function $H_{n_i}(t^*; \gamma_i, \mathbf{c}_i)$ is assumed to be a logistic function of order n_i , defined in (2.2). The null hypothesis of no additional parameter nonconstancy is $H_0 : \gamma_a = \gamma_b = 0$, under which $\boldsymbol{\eta}(t^*) = \boldsymbol{\eta}^*$ and $\boldsymbol{\lambda}(t^*) = \boldsymbol{\lambda}^*$. By setting $\boldsymbol{\pi} = (\mathbf{a}, \gamma_a, \mathbf{c}_a, \mathbf{b}, \gamma_b, \mathbf{c}_b)'$ we can write this as a special case of (4.1) with $G(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi}) = (\mathbf{a}\overline{H}_{n_a}(t^*; \gamma_a, \mathbf{c}_a) + \mathbf{b}\overline{H}_{n_b}(t^*; \gamma_b, \mathbf{c}_b)H_n(t^*; \gamma, \mathbf{c}))' \mathbf{z}_t$.

The identification problem under the null hypothesis is again circumvented by Taylor expanding $\overline{H}_{n_i}(t^*; \gamma_i, \mathbf{c}_i)$ around the null hypothesis $\gamma_i = 0$. It is a realistic assumption in a test situation to assume that the order of the logistic function (2.2) is $n_a = l_a$ and $n_b = l_b$ respectively. After a reparameterization we obtain the extended model (4.1) with $G(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi}) = \tilde{\boldsymbol{\pi}}' \mathbf{v}_t + R(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi})$, where $\tilde{\boldsymbol{\pi}} = (\boldsymbol{\tau}'_1, \dots, \boldsymbol{\tau}'_{l_a}, \boldsymbol{\tau}'_{21}, \dots, \boldsymbol{\tau}'_{2l_b})' = (\gamma_a(\overline{\boldsymbol{\tau}}'_1, \dots, \overline{\boldsymbol{\tau}}'_{l_a}), \gamma_b(\overline{\boldsymbol{\tau}}'_{21}, \dots, \overline{\boldsymbol{\tau}}'_{2l_b}))'$ and $\mathbf{v}_t = ((\mathbf{z}_t t^*)', \dots, (\mathbf{z}_t t^{*l_a})', ((\mathbf{z}_t t^*)', \dots, (\mathbf{z}_t t^{*l_b})')H_n(t^*; \gamma, \mathbf{c}))'$ with all $\overline{\boldsymbol{\tau}}'_i \neq 0$ and $\overline{\boldsymbol{\tau}}'_{2i} \neq 0$. Note that $R(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi}) \equiv 0$ under H_0 so that it does not affect the distribution theory. Our new null hypothesis is $H_0 : \tilde{\boldsymbol{\pi}} = \mathbf{0}$. The asymptotic null distribution of LM-type test statistic (3.3) can be shown to have a χ^2 distribution with $\dim(\tilde{\boldsymbol{\pi}})$ degrees of freedom if the fourth moment of u_t exists. Note that if $H_n(t^*; \gamma, \mathbf{c})$ in (2.3) is constant then the test collapses into the parameter constancy test of Section 3.1.

5. Simulations of time varying parameters

It is important to know that a structural shift in some of the parameters for a standard GARCH(1, 1) process, when ignored, affects the estimates of the parameters of this model. To investigate this a simple simulation study is conducted. The data generating process (DGP) is

$$\begin{aligned} y_t &= u_t \\ u_t &= \varepsilon_t \sqrt{h_t} \end{aligned} \tag{5.1}$$

where the conditional variance follows a TVGARCH model. At least one of the parameters of the model changes smoothly over time, that is,

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} + (\alpha_{20} + \alpha_{21} u_{t-1}^2 + \beta_{21} h_{t-1})(1 + \exp(-\gamma(t^* - c)))^{-1} \quad (5.2)$$

For all simulations $\alpha_0 = 0.1$, $\alpha_1 = 0.05$, and $\beta_1 = 0.9$ are fixed. Parameter values for the experiments are to be found in Table 5.1. For all experiments the magnitude Δ of the change in parameters of the DGPs is varied from 0 to 0.6 in steps of 0.05. When $\Delta = 0$ we have the standard GARCH(1,1) model with constant parameters. A missing value in Table 5.1 means that the corresponding parameter is set to zero and in the case when $\gamma = \infty$ we have a step change at c .

Simulation design					
Experiment	α_{20}	α_{21}	β_{21}	γ	c
1	Δ	.	.	∞	0.5
2	Δ	.	.	25	0.5
3	.	Δ	$-\Delta$	25	0.5

Table 5.1: Simulation design; in each experiment Δ is changed in the DGPs from 0 to 0.6 in steps of 0.05.

For each of the three experiments in the simulation study we define a change in parameters such that mid-point of the change lies at the mid-sample ($t^* = 1/2$). In Experiment 1 the location parameter equals infinity so that the change is abrupt. In Experiment 2 we allow the transition from one set of parameters to the other to be smooth by setting $\gamma = 25$. This implies that it takes about half the sample for the whole change to take place. In Experiment 3 we consider a model where the sum of α_1 , α_{21} , β_1 and β_{21} is held constant over time. This is achieved by letting $\alpha_{21} = -\beta_{21} = \Delta$. With this choice of parameters we retain the shape of the autocorrelation structure in the squared errors in the sense that the exponential decay rate remains unchanged in each "local model".

The random numbers, ε_t , used in the simulation study were generated by the random number generator in GAUSS 3.2.31. The random numbers sampled were all assumed to be normally distributed with expectation zero and unit variance. The first 200 observations of each generated series were always discarded to avoid initialization effects. Series of 1000 observations were used and a total of 10000 replications were performed. In each of the simulation experiments we computed the arithmetic mean of the parameter estimates. The averages of the sum of $\hat{\alpha}_1$ and $\hat{\beta}_1$ for the estimated standard GARCH(1,1) model for each experiment are graphed

in Figure B.1. In the first two experiments the average of $\hat{\alpha}_1 + \hat{\beta}_1$ approaches 1 as Δ increases. There is no apparent difference between the results from these two experiments; in both cases when no change is present, i.e. $\Delta = 0$, then $\hat{\alpha}_1 + \hat{\beta}_1 = 0.93$ which indicates that the sum of the two parameters is underestimated on average. For a change of magnitude $\Delta = 0.15$ in both Experiments 2 and 3 the sum $\hat{\alpha}_1 + \hat{\beta}_1$ increases to 0.99. In Experiment 3 the sum $\hat{\alpha}_1 + \hat{\beta}_1$ increases almost linearly as a function of Δ , reaches 1 at $\Delta = 0.55$ and continues to increase even thereafter.

The results thus demonstrate the fact that ignoring parameter changes may lead to erroneous conclusions concerning the conditional variance and the existence of the unconditional second moments of the process. It is clear that a (smooth) shift in parameters may in fact be mistaken for an integrated GARCH when a standard fixed parameter GARCH model is fitted to the series.

6. Empirical example

In this section we apply the time-varying GARCH model to the Affärsvärldens share index and the USD/GBP exchange rate. Both series are transformed to percentage changes in continuously compounded rates. This is done by differencing the logarithms of the original series. In both examples we employed linear models for the conditional mean. That is, $y_t = u_t$ in (2.1) is replaced with $y_t = \varphi' \mathbf{w}_t + u_t$ where $\varphi = (\varphi_0, \varphi_1, \dots, \varphi_s)'$ is the parameter vector for the autoregressive part and $\mathbf{w}_t = (1, y_{t-1}, \dots, y_{t-s})'$ the corresponding intercept-lag vector.

6.1. Affärsvärldens share index (ASI)

We consider the monthly observations of the Affärsvärldens share index, from January 1919 until June 1995, a total of 919 observations, see Figure B.2. An unbroken return series was constructed by Frennberg and Hansson (1992) where cash flows, such as dividends and new issues, were accounted for. Frennberg and Hansson later extended the return series to June 1995.

The parameter estimates of the standard fixed parameter GARCH model can be found in Table B.1. It is seen that the sum $\hat{\alpha}_1 + \hat{\beta}_1 = 0.98$. This may lead us to believe that the true model is an integrated GARCH one. The results of the parameter constancy test in Table B.2 indicate that the parameters (or at least the intercepts) vary over time. This standard GARCH model is extended by allowing for parameter nonconstancy of TVGARCH type. The parameters of the estimated TVGARCH model can be found in Table B.3. When the time-variation in the

intercept parameter is accounted for, the sum of $\hat{\alpha}_1$ and $\hat{\beta}_1$ is reduced to 0.91.

No evidence of misspecification is found for the conditional mean, see Table B.5. In the conditional variance the TVGARCH model passes all misspecification tests except the one for nonlinearity, see Table B.6. Nevertheless, we tentatively accept the current specification. The BDS test in Table B.4 does not reject the IID hypothesis of the standardized errors. The time variation only affects the intercept parameter, $\alpha_0^*(t) = \alpha_0 + \alpha_{20}H_n(t^*; \gamma, \mathbf{c})$, and the estimated TVGARCH model makes use of a logistic function of order $n = 3$. The values of $\alpha_0^*(t)$ are graphed in Figure B.4. It is seen that the Affärsvärldens share index has been less volatile from the mid 1920s to the early 1970s than during the remaining observation period. However, volatility seems to be decreasing again in the 1990s. Thus taking account of the time variation in volatility properly offers quite a different view of the data, compared to the standard fixed parameter GARCH model.

6.2. USD/GBP exchange rate

We consider the daily observations of the USD/GBP exchange rate from November 21, 1986 until September 30, 1997, a total of 2756 observations. The data are plotted in Figure B.3.

The parameter estimates of the standard GARCH model can be found in Table B.1. It is seen that $\hat{\alpha}_1 + \hat{\beta}_1 = 0.99$, and this may again lead us to believe that the true model is an integrated GARCH one. The null hypothesis of constant parameters in Section 3.1 is rejected in favour of parameter nonconstancy of TVGARCH type in the intercept, see Table B.2. A TVGARCH model is estimated and its parameter estimates can be found in Table B.3. After the time-varying intercept parameter has been accounted for by the TVGARCH model, $\hat{\alpha}_1 + \hat{\beta}_1 = 0.94$.

No evidence of misspecification is found in either the conditional mean or the conditional variance, see Tables B.5 and B.6 respectively. The BDS test in Table B.4 does not reject the IID hypothesis of the standardized errors. Again the time variation only affects the intercept parameter, $\alpha_0^*(t) = \alpha_0 + \alpha_{20}H_n(t^*; \gamma, \mathbf{c})$. The selected TVGARCH model has $n = 1$ and the estimated $\alpha_0^*(t)$ is graphed in Figure B.4. The value of $\alpha_0^*(t)$ decreases by two thirds in September 1993, implying that the unconditional variance of the exchange rate decreases by that amount at that point in time. A crisis in the European exchange rate mechanism (ERM) resulted in the withdrawal of the British Pound from the ERM, in September 1992. The turbulence in the ERM continued for one year and culminated in August 1993 when the exchange rate bands for the participating currencies were widened to $\pm 15\%$ from

parity. This seems to have reduced the uncertainty about the value of the pound sterling as well. Thus the TVGARCH model gives a picture rather different from the one given by the fixed parameter standard GARCH model.

7. Conclusions

Several authors have noticed that a GARCH process with time-varying parameters can be mistaken for an integrated GARCH process if the parameter variation is ignored in the estimation of the parameters. In this paper we propose a GARCH model whose parameters are allowed to change deterministically over time. The difference compared to previous models and considerations is that the parameter variation can be continuous and thus smooth. A modelling strategy consisting of specification, estimation and evaluation is constructed and applied to data. The proposed misspecification tests make use of standard asymptotic theory and are therefore easy to compute.

The model is illustrated by two applications, the Affärsvärldens share index and the USD/GBP exchange rate. In both examples $\hat{\alpha}_1 + \hat{\beta}_1$ is close to one when estimated from standard GARCH models. However, after the time variation in the intercept has been appropriately accounted for $\hat{\alpha}_1 + \hat{\beta}_1$ clearly decreases, implying that the series may not have been generated by an IGARCH model. In both examples, time variation is indeed found in the intercept, which implies that the unconditional variance of the process has changed over time.

References

- Box, G. and Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco.
- Brock, W. A., Dechert, W., Scheinkman, J. and LeBaron, B. (1996). A test for independence based on the correlation dimension, *Econometric Reviews* **15**: 197–235.
- Diebold, F. (1986). Modeling the persistence of conditional variances: A comment, *Econometric Reviews* **5**: 51–56.
- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *Journal of Finance* **48**: 1749–1777.
- Fiorentini, G., Calzolari, G. and Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics* **11**: 399–417.
- Frennberg, P. and Hansson, B. (1992). Computation of a monthly index for Swedish stock returns 1919–1989, *Scandinavian Economic History Review* **40**: 3–27.
- González-Rivera, G. (1998). Smooth transition GARCH models, *Studies in Non-linear Dynamics and Econometrics* **3**: 61–78.
- Hagerud, G. E. (1997). *A New Non-Linear GARCH Model*, EFI, Stockholm School of Economics.
- Hendry, D. F. and Neale, A. J. (1991). A monte carlo study of the effects of structural breaks on tests for unit roots, in P. Hackl and A. H. Westlund (eds), *Economic Structural Change, Analysis and Forecasting*, Springer-Verlag, Berlin, pp. 95–119.
- Lamoureux, C. G. and Lastrapes, W. G. (1990). Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**: 225–234.
- Li, C. W. and Li, W. K. (1996). On a double threshold autoregressive heteroscedastic time series model, *Journal of Applied Econometrics* **11**: 253–274.
- Lin, C.-F. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**: 211–228.

- Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics* **10**: 229–235.
- Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis, *Econometrica* **57**: 1361–1401.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**: 208–218.
- Teräsvirta, T. (1998). Modelling economic relationships with smooth transition regressions, in A. Ullah and D. E. A. Giles (eds), *Handbook of Applied Economic Statistics*, Dekker, New York, pp. 507–552.
- Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes, *Journal of the American Statistical Association* **84**: 231–240.
- Tsay, R. S. (1998). Testing and modeling multivariate threshold models, *Journal of the American Statistical Association* **93**: 1188–1202.

A. Analytical derivatives

In this section we consider the analytical derivatives of the suggested model. These first-order derivatives are used in the evaluation of an estimated model. On the other hand we also make use of second-order derivatives when estimating the model. These derivatives are not reported here, but they are straightforward to compute, see for example Fiorentini, Calzolari and Panattoni (1996) and Chapter 2. Consider the extended model defined by (2.1) and (4.1):

$$\begin{aligned} y_t &= u_t \\ u_t &= \varepsilon_t \sqrt{h_t} \end{aligned}$$

where $h_t = h(z_t; \delta_t(\zeta), \pi)$ is the parametrization of the conditional variance including the alternative. The assumed null hypothesis is $H_0 : \pi = \mathbf{0}$. If we assume that $\{\varepsilon_t\}$ is a sequence of independent standard normal errors, the log-likelihood function at time t is given by:

$$l_t = \text{const} - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{u_t^2}{h_t}.$$

A.1. Partial derivative of l_t

The first-order partial derivative (the gradient) of the log likelihood function at time t is

$$G_t = \begin{pmatrix} \frac{\partial l_t}{\partial \zeta'} & \frac{\partial l_t}{\partial \pi'} \end{pmatrix}$$

where the corresponding elements are

$$\begin{aligned} \frac{\partial l_t}{\partial \zeta'} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \zeta'} \\ \frac{\partial l_t}{\partial \pi'} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \pi'}. \end{aligned}$$

The second-order partial derivative (the Hessian) of the log-likelihood function at time t is

$$H_t = \begin{pmatrix} \frac{\partial^2 l_t}{\partial \zeta \partial \zeta'} & \frac{\partial^2 l_t}{\partial \zeta \partial \pi'} \\ \frac{\partial^2 l_t}{\partial \pi \partial \zeta'} & \frac{\partial^2 l_t}{\partial \pi \partial \pi'} \end{pmatrix}. \quad (\text{A.1})$$

If H_0 holds, the expectations of the elements in (A.1) are

$$\begin{aligned} E \left[\frac{\partial^2 l_t}{\partial \zeta \partial \zeta'} \mid \pi = \mathbf{0} \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \zeta} \frac{\partial h_t}{\partial \zeta'} \right] & E \left[\frac{\partial^2 l_t}{\partial \pi \partial \zeta'} \mid \pi = \mathbf{0} \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \pi} \frac{\partial h_t}{\partial \zeta'} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \pi \partial \pi'} \mid \pi = \mathbf{0} \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \pi} \frac{\partial h_t}{\partial \pi'} \right]. \end{aligned}$$

A.2. Partial derivative of the additive conditional variance h_t .

Assume that the conditional variance under the alternative hypothesis is parameterized as in (4.1) with $G(\mathbf{z}_t; \boldsymbol{\zeta}, \boldsymbol{\pi}) = \boldsymbol{\pi}'\mathbf{v}_t$. The conditional variance is then

$$\begin{aligned} h_t &= \boldsymbol{\delta}_t(\boldsymbol{\zeta}) + \boldsymbol{\pi}'\mathbf{v}_t = (\boldsymbol{\eta} + H_n(t; \gamma, \mathbf{c})\boldsymbol{\lambda})'\mathbf{z}_t + \boldsymbol{\pi}'\mathbf{v}_t = \\ &= (\alpha_0 + \alpha_{20}H_n(t; \gamma, \mathbf{c})) + \sum_{i=1}^q (\alpha_1 + \alpha_{21}H_n(t; \gamma, \mathbf{c}))u_{t-i}^2 + \\ &\quad \sum_{j=1}^p (\beta_1 + \beta_{21}H_n(t; \gamma, \mathbf{c}))h_{t-j} + \boldsymbol{\pi}'\mathbf{v}_t \end{aligned}$$

which reduces to the time-varying GARCH model under H_0 . To initialize the iterative computation of h_t under null hypothesis, the conditional variance is estimated with the unconditional variance (sample variance) in the pre-sample case. This is done for all $t \leq 0$ by setting $h_t = u_t^2 = \frac{1}{T} \sum_{s=1}^T u_s^2$ where $u_s = y_s$. The first-order derivatives of the conditional variance h_t under the null hypothesis are required to compute the test statistic (3.3). These derivatives are given as follows:

$$\begin{aligned} \left. \frac{\partial h_t}{\partial \boldsymbol{\zeta}'} \right|_{\boldsymbol{\pi}=\mathbf{0}} &= \frac{\partial \boldsymbol{\eta}'\mathbf{z}_t}{\partial \boldsymbol{\zeta}'} + \frac{\partial \boldsymbol{\lambda}'\mathbf{z}_t}{\partial \boldsymbol{\zeta}'} H_n(t; \gamma, \mathbf{c}) + \boldsymbol{\lambda}'\mathbf{z}_t \frac{\partial H_n(t; \gamma, \mathbf{c})}{\partial \boldsymbol{\zeta}'} \\ &\quad + \sum_{j=1}^p (\beta_j + \beta_{2j}H_n(t; \gamma, \mathbf{c})) \left. \frac{\partial h_{t-j}}{\partial \boldsymbol{\zeta}'} \right|_{\boldsymbol{\pi}=\mathbf{0}} \\ \left. \frac{\partial h_t}{\partial \boldsymbol{\zeta}'} \right|_{\boldsymbol{\pi}=\mathbf{0}} &= \mathbf{0}, \text{ pre-sample values, } t \leq 0 \\ \left. \frac{\partial h_t}{\partial \boldsymbol{\pi}'} \right|_{\boldsymbol{\pi}=\mathbf{0}} &= \mathbf{v}_t' + \sum_{j=1}^p (\beta_j + \beta_{2j}H_n(t; \gamma, \mathbf{c})) \left. \frac{\partial h_{t-j}}{\partial \boldsymbol{\pi}'} \right|_{\boldsymbol{\pi}=\mathbf{0}} \\ \left. \frac{\partial h_t}{\partial \boldsymbol{\pi}'} \right|_{\boldsymbol{\pi}=\mathbf{0}} &= \mathbf{0}, \text{ pre-sample values, } t \leq 0. \end{aligned}$$

A.3. Partial derivatives of the logistic function $H_n(t; \gamma, \mathbf{c})$.

Since $H_n(t; \gamma, \mathbf{c})$ is a function of the exogenous variable t , the partial derivatives are not affected by η^* and λ . Assume that the deterministic function $H_n(t; \gamma, \mathbf{c})$ is parameterized as in (2.2):

$$H_n(t; \gamma, \mathbf{c}) = \left(1 + \exp(-\gamma \prod_{l=1}^n (t - c_l)) \right)^{-1}.$$

To simplify the calculation of the derivatives the function $H_n(t; \gamma, \mathbf{c})$ is rewritten

as:

$$H_n(t; \gamma, \mathbf{c}) = \frac{1}{2 \exp(-\frac{\gamma}{2} \prod_{l=1}^n (t - c_l)) \cosh(\frac{\gamma}{2} \prod_{l=1}^n (t - c_l))}.$$

Let $\xi_t = \xi(t; \gamma, c_1, \dots, c_n) = \frac{\gamma}{2} \prod_{l=1}^n (t - c_l)$, drop the arguments and write the function as:

$$H_n(t; \gamma, \mathbf{c}) = \frac{e^{\xi_t}}{2 \cosh \xi_t}.$$

The first-order derivative of the logistic function with respect to the parameters of the slope and the location vector is

$$\frac{\partial H_n(t; \gamma, \mathbf{c})}{\partial \boldsymbol{\zeta}'} = \left(0, \dots, 0, \frac{1}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \gamma}, \frac{1}{2 \cosh^2 \xi_t} \frac{\partial \xi_t}{\partial \mathbf{c}'} \right).$$

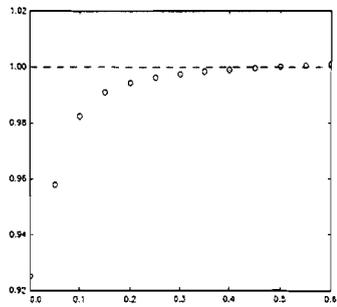
By letting c_i denote the i^{th} element in the location vector \mathbf{c} , we can write the first-order derivative of ξ_t as

$$\frac{\partial \xi_t}{\partial \gamma} = \frac{1}{2} \prod_{l=1}^n (t - c_l), \quad \frac{\partial \xi_t}{\partial c_i} = -\frac{\gamma}{2} \prod_{l=1, l \neq i}^n (t - c_l).$$

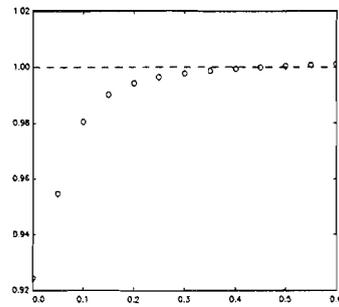
B. Tables and Figures

Parameter	ASI (1919-95)	GBP (1986-97)
φ_0	0.0077 (0.0013)	.
φ_1	0.15 (0.036)	.
φ_2	-0.067 (0.036)	0.054 (0.020)
φ_3	0.057 (0.036)	.
φ_6	0.10 (0.033)	.
α_0	0.000060 (0.000019)	4.9×10^{-7} (1.3×10^{-7})
α_1	0.18 (0.030)	0.071 (0.0085)
β_1	0.80 (0.028)	0.92 (0.0092)

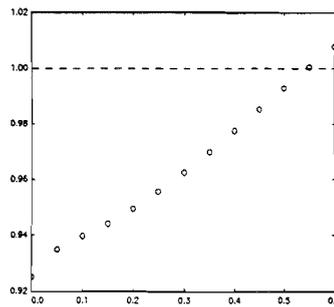
Table B.1: Parameter estimates of the standard GARCH models (standard deviation in parenthesis).



(a) Experiment 1

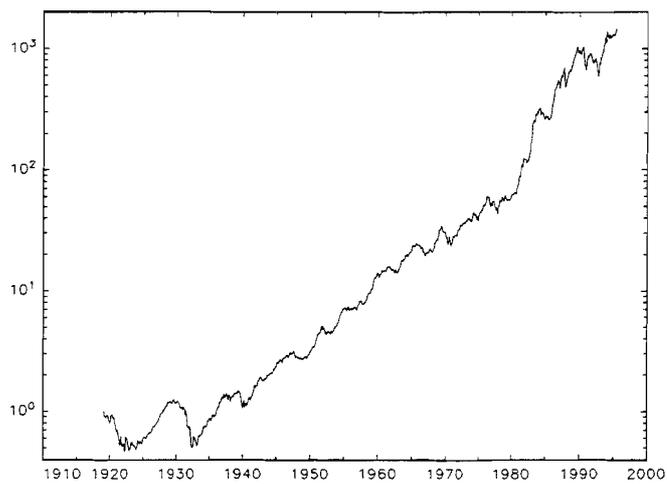


(b) Experiment 2

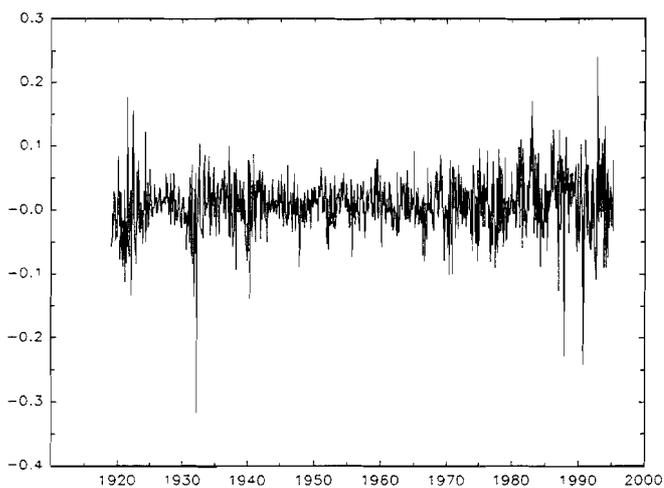


(c) Experiment 3

Figure B.1: The average sum of $\hat{\alpha}_1 + \hat{\beta}_1$ for the estimated standard GARCH(1,1) in Experiment 1-3 based on 10000 replications. In the figure the value of Δ is given on the x axis and the value of $\hat{\alpha}_1 + \hat{\beta}_1$ is given on the y axis. The horizontal dashed line corresponds to $\hat{\alpha}_1 + \hat{\beta}_1 = 1$

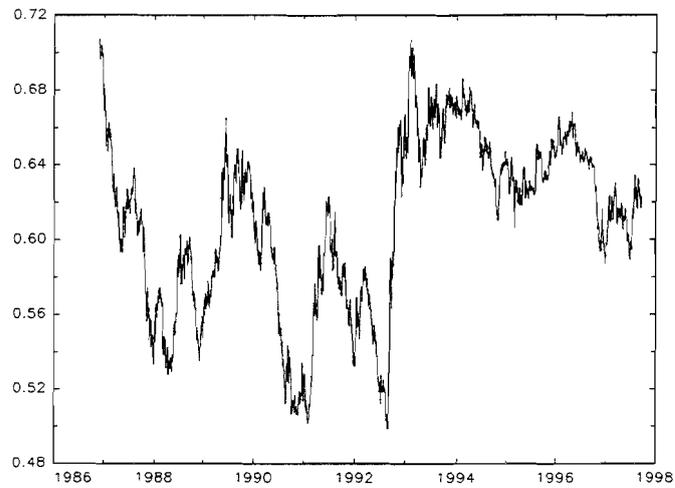


(a) Levels (logarithmic scale)

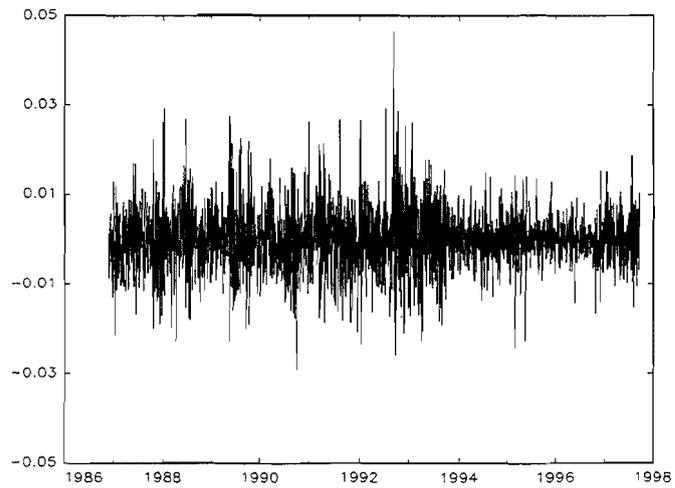


(b) Returns

Figure B.2: Affärsvärldens Share Index, monthly observations from January 1919 until July 1995.

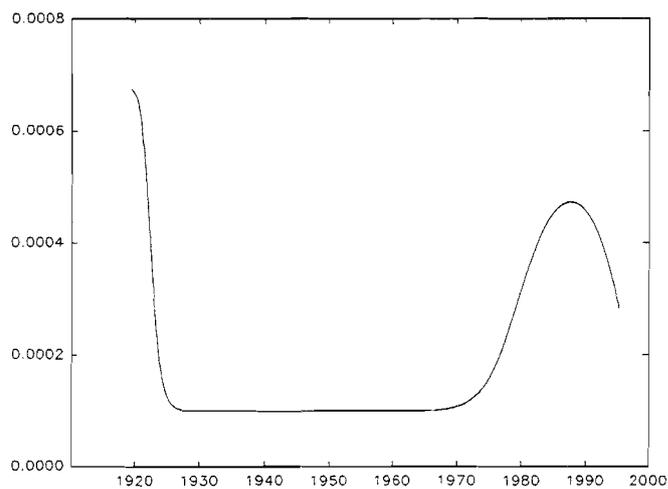


(a) Levels

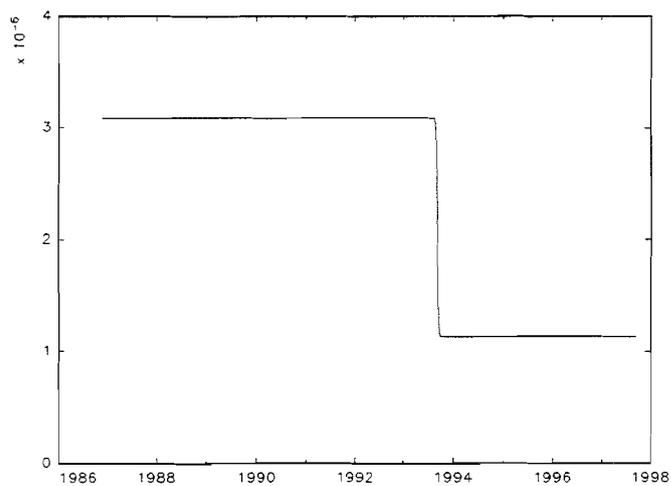


(b) Returns

Figure B.3: The daily exchange rate of USD/GBP from November 21, 1986 until September 30, 1997.



(a) Affärsvärdens share index



(b) The daily exchange rate of USD/GBP

Figure B.4: The value of the intercept parameter, $\alpha_0^*(t) = \alpha_0 + \alpha_{20}H_n(t^*; \gamma, \mathbf{c})$. In the figures the value of $\alpha_0^*(t)$ is plotted on the y axis against time on the x axis.

		ASI (1919-95)	GBP (1986-97)
All parameters	$n = 1$	0.029	0.0000043
	$n = 2$	0.0000018	0.0000043
Only intercept	$n = 1$	0.0030	0.0000023
	$n = 2$	0.0000028	0.0000062
Only alpha	$n = 1$	0.14	0.0052
	$n = 2$	0.0075	0.0034
Only beta	$n = 1$	0.077	0.0010
	$n = 2$	0.016	0.00034

Table B.2: p -values of the parameter constancy test. Under the null hypothesis the conditional variance follows a standard GARCH model. The LM-test is computed against the alternative of time dependence given by a logistic function of order $n = 1$ and $n = 2$ respectively using time as the transition variable. In practice $n = 1$ corresponds to a change at one point in time and $n = 2$ corresponds to a change at one point in time that is cancelled out at a later point in time.

Parameter	ASI (1919-95)	GBP (1986-97)
φ_0	0.0075 (0.0013)	.
φ_1	0.15 (0.036)	.
φ_2	-0.068 (0.036)	0.052 (0.020)
φ_3	0.062 (0.035)	.
φ_6	0.11 (0.033)	.
α_0	0.00068 (0.00042)	0.0000031 (0.0000074)
α_{20}	-0.00058 (0.00041)	-0.0000020 (0.0000050)
α_1	0.16 (0.034)	0.076 (0.012)
β_1	0.75 (0.049)	0.87 (0.021)
γ	140 (110)	1000 (fixed)
c_1	0.041 (0.026)	0.63 (0.010)
c_2	0.83 (0.11)	.
c_3	0.97 (0.099)	.

Table B.3: Parameter estimates of the TVGARCH models (standard deviation in parenthesis).

	ASI (1919-95)	GBP (1986-97)
Characteristics:		
<i>Min</i>	-4.8	-5.2
<i>Max</i>	3.3	4.3
<i>Mean</i>	-0.033	-0.016
<i>Variance</i>	1.0	1.0
<i>Skewness</i>	-0.23	0.078
<i>Kurtosis</i>	4.0	4.6
Nonparametric test of IID (BDS)		
<i>BDS statistic</i>	0.12	0.40
<i>p-value (asymptotic)</i>	0.90	0.69
<i>p-value (bootstrap)</i>	0.91	0.72

Table B.4: Characteristics of the standardized residuals of the TVGARCH model and the BDS test of independence for the standardized errors. For the BDS test a bootstrapped probability value based on 1000 resampled series is reported as well.

	ASI (1919-95)	GBP (1986-97)
Remaining autocorrelation (<i>p</i> -values)		
<i>l</i> = 1	0.12	0.73
<i>l</i> = 2	0.13	0.80
<i>l</i> = 3	0.041	0.85
<i>l</i> = 4	0.083	0.92
Parameter constancy (<i>p</i> -values)		
<i>n</i> = 1	0.79	0.67
<i>n</i> = 2	0.94	0.73
Remaining nonlinearity (<i>p</i> -values)		
<i>d</i> = 1	0.97	0.71
<i>d</i> = 2	0.84	0.26
<i>d</i> = 3	0.51	0.97
<i>d</i> = 4	0.96	0.61

Table B.5: *p*-values of specification tests for the conditional mean for the estimated TVGARCH models. LM tests for the conditional mean. The test of no remaining autocorrelation is computed against the alternative of remaining autocorrelation up to the given lag, *l*. The test of parameter constancy is computed against the alternative of time dependence given by a logistic function of order *n* using time as the transition variable. The test against nonlinearity of LSTAR(3) type uses a lag length 4 for the autoregressive part. The test is computed separately against the alternatives $1 \leq d \leq 4$.

	ASI (1919-95)	GBP (1986-97)
Test of the functional form (<i>p</i> -values)		
$l = 1$	0.95	0.044
$l = 2$	0.31	0.12
$l = 4$	0.12	0.33
$l = 8$	0.0054	0.63
Parameter constancy (<i>p</i> -values)		
$n = 1$	0.12	0.032
$n = 2$	0.035	0.020
Remaining nonlinearity of STGARCH type (<i>p</i> -values)		
<i>All parameters</i>	0.0023	0.044
<i>Without intercept</i>	0.00090	0.031

Table B.6: *p*-values of specification tests for the conditional variance of the estimated TVGARCH models. LM tests for the conditional variance. The tests of no remaining serial dependence in the squared and standardized residuals are computed against the alternative of remaining dependence up to the given lag, l . The test of parameter constancy is computed against the alternative of time dependence given by a logistic function of order n using time as the transition variable. The test against remaining nonlinearity is a test against an STGARCH(3) structure.

Chapter 5

A Smooth Transition Autoregressive Target Zone model

1. Introduction

Over the years target zones for exchange rates have been a reality in, for example, the Bretton Woods system, the gold standard, the Exchange rate Mechanism (ERM) of the European Monetary System (EMS) and the Nordic countries. There is a vast literature on target zones, both empirically and theoretically.

Empirically it has been found that a freely floating exchange rate seems to behave like a random walk with a fat-tailed distribution and clustering volatility; see Meese and Rogoff (1983), Baillie and Bollerslev (1989) and Diebold and Nerlove (1989). For currencies bounded within a target zone there exists empirical evidence of mean reversion; see, for example, Svensson (1993), Rose and Svensson (1994), Engle and Gau (1997) and Anthony and MacDonald (1998). This implies that if we tested the unit root hypothesis we would reject it and we should also be able to beat the naive (random walk) forecasts for the exchange rate series. However, in a simulation study, Froot and Obstfeld (1991) found that the mean reversion is often weak and hard to detect in small samples, especially for wide target zones. Furthermore, the same types of fat-tailed distribution and volatility clustering as for a freely floating exchange rate are found in exchange rates that fluctuate within a target zone.

In this paper we consider univariate modelling of an exchange rate within a target zone. Since the bands make the exchange rate process nonlinear we modify an AR-GARCH model to fit the target zone environment by allowing for nonlinearities of smooth transition type near the boundary of the exchange rate band. Nonlinearity is introduced in both the conditional mean and the conditional variance. Thus the model, which is called the Smooth Transition Autoregressive Target Zone (STARTZ) model, may be viewed as another variant of the STAR-STGARCH model in Chapter 2, which is a rather flexible nonlinear model.

To be able to apply the STARTZ model we propose a modelling cycle consisting of specification, estimation and evaluation stages. We apply the model to the exchange rates of the Nordic countries and find that the model is able to capture the dynamic behaviour of these currencies and display behaviour suggested by economic theory.

The plan of the paper is as follows. Theoretical target zone models are discussed in Section 2. The STARTZ model is defined in Section 3 and its specification and estimation are considered in Section 4. Section 5 contains misspecification tests for evaluation and Section 6 an application of the model to two Nordic exchange rates. Section 7 contains a small simulation study of the estimated STARTZ models. Finally, Section 8 concludes.

2. Target zone models

A vast literature has followed the basic target zone model published in Krugman (1991); for an excellent survey of target zone models see Svensson (1992). The Krugman model has three important empirical implications of the behaviour of the exchange rate movements within the band. First, the exchange rate movements are expected to behave nonlinearly (S-shaped) with respect to the underlying fundamentals, usually the money supply. By reducing (increasing) the money supply the central bank intervenes to strengthen (or weaken) the currency. Second, the model implies that the marginal density of the exchange rate inside the band is U-shaped, which means that the exchange rate spends most of the time near the edges of the band. Finally, the variance is expected to be \cap -shaped so that the conditional variance decreases when the rate approaches either of the boundaries of the target zone. These empirical implications of the theoretical model have been studied and rejected; see for example Bertola and Caballero (1992), Flood, Rose and Mathieson (1991) and Lindberg and Söderlind (1994b).

The empirical failure of the basic Krugman model has prompted some extensions. The model is constructed under two crucial assumptions. First, the credibility of the target zone is perfect and, second, possible interventions only take place at the margin of the exchange rate band. In the second generation of target zone models these assumptions are relaxed. In practice the exchange rate band may sometimes be realigned, which makes the assumption of perfect credibility less realistic. One extension of the target zone model has been to introduce imperfect credibility. The idea is to introduce a second fundamental, a time-varying realignment risk, see Bertola and Svensson (1993) and Bertola and Caballero (1992). Imperfect credibility implies that the \cap -shape of the variance is not so pronounced near the edges of the band. The other extension is to allow for intramarginal interventions. These are interventions that occur in the interior of the exchange rate band, see Froot and Obstfeld (1991) and Delgado and Dumas (1991). This implies that the exchange rate spends more time in the interior of the exchange rate band, and therefore the density has more probability mass in the interior than at the edges of the exchange rate band. Consequently such interventions give rise to a less pronounced (S-shaped) nonlinearity. It is worth noting that all these target zone models imply a mean reverting behaviour for the exchange rate.

3. The Model

A possible way of modelling the time series behaviour of an exchange rate within a target zone is to employ a parameterization similar to that in Chapter 2, where the idea of a smooth transition between two extreme regimes was used to model both the conditional mean and the conditional variance. In this paper we introduce bounds on the conditional mean to restrict the exchange rate, so that it remains within a band defined by the target zone. The model is a special case of the following additive nonlinear model

$$y_t = m_t(\boldsymbol{\varphi}, \gamma_a, \mu) + u_t \quad (3.1)$$

where y_t is the deviation of the exchange rate from the centre of the target zone. The function $m_t = m_t(\boldsymbol{\varphi}, \gamma_a, \mu)$ is assumed to be bounded and at least twice continuously differentiable for its parameters everywhere in the parameter space and the corresponding sample space. The error process of the model is parametrized as

$$u_t = \varepsilon_t \sqrt{h_t(\boldsymbol{\varphi}, \gamma_a, \mu, \boldsymbol{\eta}, \gamma_b, \delta)} \quad (3.2)$$

where $\{\varepsilon_t\} \sim \text{nid}(0, 1)$ and $h_t = h_t(\boldsymbol{\varphi}, \gamma_a, \mu, \boldsymbol{\eta}, \gamma_b, \delta)$ is a positive-valued function representing the conditional variance of u_t . Thus there is no autocorrelation in the error process $\{u_t\}$. Furthermore, $u_t = y_t - m_t$ such that $\boldsymbol{\varphi}$ is assumed not to depend on $\boldsymbol{\eta}$. Function h_t is at least twice continuously differentiable for the parameters everywhere in the corresponding parameter space. It is also assumed that the moments necessary for the inference exist and that the parameters are subject to restrictions such that the process defined by (3.1) and (3.2) is stationary and ergodic.

In order to define m_t , h_t and to consider the misspecification tests in Section 5, let

$$H_n(s_t; \gamma, \mathbf{c}) = \left(1 + \exp\left(-\gamma \prod_{l=1}^n (s_t - c_l)\right) \right)^{-1}, \gamma > 0, c_1 \leq \dots \leq c_n \quad (3.3)$$

where s_t is the transition variable, γ a slope parameter and $\mathbf{c} = (c_1, \dots, c_n)'$ a location vector. The parameter restrictions $\gamma > 0$ and $c_1 \leq \dots \leq c_n$ are identifying restrictions. The value of the logistic function (3.3), which is bounded between a and 1, $0 \leq a \leq 1/2$, depends on the transition variable s_t . Note that for $\gamma = 0$, $H_n(s_t; \gamma, \mathbf{c}) \equiv 1/2$ and when $\gamma \rightarrow \infty$, and $n = 1$, $H_n(s_t; \gamma, \mathbf{c})$ becomes a step function switching from 0 to 1 at c_1 . It becomes a "multistep" function with switches at c_1, \dots, c_n if $n > 1$.

According to theoretical target zone models, the conditional mean should follow a nonlinear function (S-shaped) of the underlying fundamentals with local nonlinearity emerging close to the band. This requirement is met by the following model specification

$$\begin{aligned} m_t = & \boldsymbol{\varphi}' \mathbf{w}_t + \\ & 2(\mu b^L - \boldsymbol{\varphi}' \mathbf{w}_t) H_1(-y_{t-1}; \gamma_a, -\mu b^L) + \\ & 2(\mu b^U - \boldsymbol{\varphi}' \mathbf{w}_t) H_1(y_{t-1}; \gamma_a, \mu b^U) \end{aligned} \quad (3.4)$$

where $\boldsymbol{\varphi} = (\varphi_0, \varphi_1, \dots, \varphi_s)'$ is the parameter vector for the autoregressive part of the model and $\mathbf{w}_t = (1, y_{t-1}, \dots, y_{t-s})'$ is the corresponding intercept-lag vector. The remaining structure in (3.4) characterizes the behaviour of the conditional mean close to the boundary of the target zone. Constants b^L and b^U represent the lower and the upper boundary respectively. For simplicity we assume that the local behaviour of the exchange rate is similar in the neighbourhood of both boundaries. Parameter $0 < \mu \leq 1$ adds flexibility to the specification and allows us, among other

things, to capture an implicit band inside the official one, when such a band exists. The slope parameter, $\gamma_a > 0$ tells us how pronounced the change is between the dynamic behaviour of the exchange rate near the centre of the target zone and close to either boundary.

This model may be interpreted as follows. The behaviour of the exchange rate is characterized by a linear combination of its lags, $\varphi' \mathbf{w}_t$, close to the centre of the band. Near both the upper and the lower boundary of the target zone the behaviour of the exchange rate becomes nonlinear. For example, in the case when the exchange rate approaches the upper boundary there is a smooth transition from the autoregressive behaviour represented by $\varphi' \mathbf{w}_t$ towards the upper boundary represented by a constant μb^U . This happens when H_1 in (3.4) approaches 1/2. Where and how quickly this happens depends on γ and c_1 .

According to theoretical target zone models the conditional variance specification should be \cap -shaped, as the conditional variance must decrease close to the band. We parameterize this requirement in a way similar to what was used for the conditional mean. Thus,

$$\begin{aligned} h_t = & \boldsymbol{\eta}' \mathbf{z}_t + & (3.5) \\ & 2(\delta - \boldsymbol{\eta}' \mathbf{z}_t) H_1(-y_{t-1}; \gamma_b, -\mu b^L) + \\ & 2(\delta - \boldsymbol{\eta}' \mathbf{z}_t) H_1(y_{t-1}; \gamma_b, \mu b^U) \end{aligned}$$

where constants b^L and b^U again represent the lower and the upper boundary. Parameter μ is defined as before. The slope parameter, $\gamma_b > 0$, determines the shape of the logistic functions that impose the upper and lower boundary restrictions. The logistic functions are defined in the same way as for the conditional mean. Setting $\boldsymbol{\eta} = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ and $\mathbf{z}_t = (1, u_{t-1}^2, \dots, u_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$, where $h_t > 0$ for all t with probability one, makes $\boldsymbol{\eta}' \mathbf{z}_t$ in (3.5) a standard GARCH(p, q) type specification. Assuming $\delta > 0$ together with the restrictions $\alpha_0 > 0$, $\alpha_j \geq 0$, $j = 1, \dots, q$; $\beta_j \geq 0$, $j = 1, \dots, p$; guarantees $h_t > 0$.

This parameterization implies that the conditional variance is nonlinear close to the boundaries. For example, in the case when the exchange rate approaches the upper boundary there is a smooth transition from a standard GARCH type behaviour represented by $\boldsymbol{\eta}' \mathbf{z}_t$ towards a small positive δ .

The parameterization of the STARTZ process is in principle only defined within the target zone. Nevertheless the STARTZ process thus defined is not bounded

such that the exchange rate would remain inside the target zone with probability 1. At the boundary the conditional variance of the STARTZ model is small, as $\delta > 0$ is close to zero. Thus, according to the model a shock such that the exchange rate breaks through the boundary of the zone does have a positive probability. On the other hand, in practice the rate may occasionally do that as there have been realignments in many exchange rate bands. The conditional mean model can be interpreted as follows: eventually a shock large enough to force a realignment of the target zone will arrive, but one does not know when this will occur or how large the realignment will be.

4. Specification and estimation

The nonlinear STARTZ model defined by (3.1-3.5) is our most general parameterization of the target zone model. Near the centre of the target zone a standard AR-GARCH model is assumed to adequately describe the behaviour of an exchange rate. The nonlinear behaviour imposed by the basic target zone model is only invoked close to the boundaries of the target zone. In this paper we propose a modelling strategy similar to the one suggested in Chapter 2. It can be described as follows.

1. Select an AR(s) model according to some suitable criterion such as the AIC (Akaike, 1974).
2. Estimate an AR(s)-GARCH(p, q) model to obtain initial values for estimating the STARTZ model.
3. Estimate the parameters of the full STARTZ model and test the adequacy of both the conditional mean and the conditional variance specification by appropriate misspecification tests.
4. If the model passes the tests, tentatively accept it. In the opposite case try another specification search or choose another family of models.

There is a notable difference between this modeling strategy and that proposed in Chapter 2: linearity is not tested against (3.1) with (3.4), nor is the linear GARCH model tested against (3.5). This is because we know that the boundary restrictions imply nonlinearity, as the movements of the exchange rate under consideration are at least most of the time restricted by the boundaries of the target zone. In practice this suggests that in step 3 we should start with a standard GARCH(1,1) model,

and if some remaining structure is detected by misspecification tests we extend the GARCH parameterization in the suggested direction.

All parameter estimates are obtained by maximizing the log-likelihood under the assumption that $\{\varepsilon_t\}$ is a sequence of independent standard normal errors. In that case we write the log-likelihood function at time t as

$$l_t = \text{const} - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{u_t^2}{h_t} \quad (4.1)$$

where u_t and h_t are defined in (3.5). We assume that the model under consideration satisfies the necessary regularity conditions for stationarity and ergodicity needed for the consistency and asymptotic normality of the estimators. In the following section we consider the specification, estimation and evaluation of this model.

5. Evaluation by misspecification tests

Once we have estimated a model it is important to investigate the validity of the assumptions used in the estimation. The STARTZ model is a modification of the STAR-STGARCH model. Thus we can modify the three different misspecification tests in Chapter 2 for the present situation. The tests in this Section only consider misspecification of the AR(s)-GARCH(p, q) parameterization within the target zone. The structure that constitutes the target zone is not considered since its parameterization is known from the definition of the STARTZ model. The first test can be interpreted as a test of the functional form. The second one is a test against nonlinearity or, in some cases, asymmetry. The last one is a test against parameter nonconstancy. To describe the common features in these tests we first introduce a general structure and thereafter briefly consider each test separately.

We also consider a general test of independence of the (standardized) errors. The BDS statistic (Brock, Dechert, Scheinkman and LeBaron, 1996) appears to be a suitable large-sample test for this purpose, see for example Chapter 2.

5.1. General

Consider the STARTZ model as defined in (3.1) and (3.2). An additive extension of the model may be written as

$$y_t = \varphi' \mathbf{w}_t + A(\mathbf{w}_t; \boldsymbol{\pi}_a) + 2(\mu b^L - \varphi' \mathbf{w}_t - A(\mathbf{w}_t; \boldsymbol{\pi}_a)) H_1(-y_{t-1}; \gamma_a, -\mu b^L) + \quad (5.1)$$

$$\begin{aligned}
& 2(\mu b^U - \boldsymbol{\varphi}' \mathbf{w}_t - A(\mathbf{w}_t; \boldsymbol{\pi}_a)) H_1(y_{t-1}; \gamma_a, \mu b^U) + u_t \\
u_t &= \varepsilon_t \sqrt{h_t} \\
h_t &= \boldsymbol{\eta}' \mathbf{z}_t + B(\mathbf{z}_t; \boldsymbol{\pi}_b) + \\
& 2(\delta - \boldsymbol{\eta}' \mathbf{z}_t - B(\mathbf{z}_t; \boldsymbol{\pi}_b)) H_1(-y_{t-1}; \gamma_b, -\mu b^L) + \\
& 2(\delta - \boldsymbol{\eta}' \mathbf{z}_t - B(\mathbf{z}_t; \boldsymbol{\pi}_b)) H_1(y_{t-1}; \gamma_b, \mu b^U)
\end{aligned}$$

where functions $A(\mathbf{w}_t; \boldsymbol{\pi}_a)$ and $B(\mathbf{z}_t; \boldsymbol{\pi}_b)$ are assumed twice continuously differentiable for all $\boldsymbol{\pi}_a$ and $\boldsymbol{\pi}_b$ everywhere in the corresponding sample spaces. For notational simplicity and without loss of generality we assume $A(\mathbf{w}_t; \boldsymbol{\pi}_a) \equiv 0$ for $\boldsymbol{\pi}_a = \mathbf{0}$ and $B(\mathbf{z}_t; \boldsymbol{\pi}_b)$ for $\boldsymbol{\pi}_b = \mathbf{0}$. Furthermore, $\boldsymbol{\eta}' \mathbf{z}_t + B(\mathbf{z}_t; \boldsymbol{\pi}_b)$ is assumed to be positive-valued for all t and $\{\varepsilon_t\}$ is a sequence of independent standard normal variables. Model (5.1) forms a unifying framework for our tests.

Let $\boldsymbol{\omega} = (\boldsymbol{\varphi}', \gamma_a, \mu, \boldsymbol{\eta}', \gamma_b, \delta)'$, which comprises all the parameters of the model. It is assumed that the maximum likelihood estimator of $\boldsymbol{\omega}$ is consistent and asymptotically normal under any null hypothesis to be considered, which requires $\{y_t\}$ to be stationary and ergodic and that the log-likelihood function satisfies the standard regularity conditions. Furthermore, the necessary moments of $\{u_t\}$ implied by the Hessian matrix and required for the asymptotic distribution theory to work are assumed to exist. The null hypothesis of no additional structure is $\mathbf{H}_0 : \boldsymbol{\pi}_a = \mathbf{0}$ and $\boldsymbol{\pi}_b = \mathbf{0}$. The Lagrange multiplier (or score) test statistic is defined as

$$\text{LM} = T \begin{pmatrix} \frac{1}{T} \sum \frac{\partial l_t}{\partial \boldsymbol{\pi}_a} |_{\mathbf{H}_0} \\ 0 \\ \frac{1}{T} \sum \frac{\partial l_t}{\partial \boldsymbol{\pi}_b} |_{\mathbf{H}_0} \end{pmatrix}' \widehat{\mathbf{I}}(\boldsymbol{\pi}_a, \widehat{\boldsymbol{\omega}}, \boldsymbol{\pi}_b)^{-1} |_{\mathbf{H}_0} \begin{pmatrix} \frac{1}{T} \sum \frac{\partial l_t}{\partial \boldsymbol{\pi}_a} |_{\mathbf{H}_0} \\ 0 \\ \frac{1}{T} \sum \frac{\partial l_t}{\partial \boldsymbol{\pi}_b} |_{\mathbf{H}_0} \end{pmatrix} \quad (5.2)$$

where $\widehat{\mathbf{I}}$ is a consistent estimator of the information matrix under the null hypothesis. We use the expectation of the estimated negative Hessian as our estimator of the information matrix. The partial derivatives forming the Hessian may be found in Appendix B.7. The test statistic (5.2) is asymptotically χ^2 -distributed with $\dim(\boldsymbol{\pi}_a) + \dim(\boldsymbol{\pi}_b)$ degrees of freedom under the null hypothesis.

5.2. Test against serial dependence

To test the joint null hypothesis of no serial dependence in either the conditional mean or the conditional variance or both, the alternative is stated as remaining serial dependence of order p_a in the ordinary error process and of order p_b in the squared (and standardized) errors. In the general case, this gives the extended model (5.1)

with $A(\mathbf{w}_t; \boldsymbol{\pi}_a) = \boldsymbol{\pi}'_a \mathbf{v}_t^a$ and $B(\mathbf{w}_t; \boldsymbol{\pi}_b) = \boldsymbol{\pi}'_b \mathbf{v}_t^b$ where $\boldsymbol{\pi}_a = (\pi_{a,1}, \dots, \pi_{a,p_a})'$, $\mathbf{v}_t^a = (u_{t-1}, \dots, u_{t-p_a})'$, $\boldsymbol{\pi}_b = (\pi_{b,1}, \dots, \pi_{b,p_b})'$, and $\mathbf{v}_t^b = (h_{t-1}, \dots, h_{t-p_b})'$; compare this with Chapter 2. The null hypothesis of no remaining serial dependence in either the conditional mean or in the conditional variance is equivalent to $\boldsymbol{\pi}_a = \mathbf{0}$ and $\boldsymbol{\pi}_b = \mathbf{0}$. Under the null hypothesis and assuming that at least all moments up to the fourth moment exist, the LM-statistic (5.2) is asymptotically χ^2 -distributed with $\dim(\boldsymbol{\pi}_a) + \dim(\boldsymbol{\pi}_b)$ degrees of freedom.

5.3. Test against remaining nonlinearity

To see if an estimated STARTZ model adequately characterizes all nonlinearity in the series, we should find out whether or not there is any nonlinearity left after fitting the model to the data. This can be done by testing the hypothesis of no remaining additive nonlinearity. The alternative to this null hypothesis is assumed to be an additive smooth transition component. This alternative is obtained as a special case of (5.1) with $A(\mathbf{w}_t; \boldsymbol{\phi}_a, \rho_a, \mathbf{c}_a) = \boldsymbol{\phi}'_a \mathbf{w}_t \overline{H}_{n_a}(y_{t-1}; \rho_a, \mathbf{c}_a)$ and $B(\mathbf{w}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b) = \boldsymbol{\psi}'_b \mathbf{z}_t \overline{H}_{n_b}(u_{t-1}; \rho_b, \mathbf{c}_b)$. We define $\overline{H}_{n_i} = H_{n_i} - \frac{1}{2}$ where $i = a, b$; subtracting $1/2$ from H_{n_i} is just a notational convenience in deriving the test and does not affect the generality of the argument. H_{n_i} is the logistic function (3.3) of order n_i with ρ_i as the slope parameter and \mathbf{c}_i as the location vector. The joint null hypothesis of no nonlinearity in both the conditional mean and the conditional variance can be written as $H_0 : \rho_a = \rho_b = 0$.

Parameters in $A(\mathbf{w}_t; \boldsymbol{\phi}_a, \rho_a, \mathbf{c}_a)$ and $B(\mathbf{w}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b)$ are not identified under the null hypothesis, see for example Teräsvirta (1994) for a discussion. We circumvent this identification problem under the null hypothesis by expanding functions $A(\mathbf{w}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b)$ and $B(\mathbf{w}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b)$ into a Taylor series around the null hypothesis. A realistic assumption in a testing situation is to assume that the orders of the logistic function (3.3) are $n_a = l_1$ and $n_b = l_2$ respectively. Using the first-order expansion this yields after reparameterization a transformed model with $A(\mathbf{w}_t; \boldsymbol{\phi}_a, \rho_a, \mathbf{c}_a) = \tilde{\boldsymbol{\pi}}'_a \mathbf{v}_t^a + R_1(\mathbf{w}_t; \boldsymbol{\phi}_a, \rho_a, \mathbf{c}_a, \boldsymbol{\omega})$ and $B(\mathbf{w}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b) = \tilde{\boldsymbol{\pi}}'_b \mathbf{v}_t^b + R_2(\mathbf{z}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b, \boldsymbol{\omega})$ where $\tilde{\boldsymbol{\pi}}_a = (\boldsymbol{\tau}'_{a,1}, \boldsymbol{\tau}'_{a,2}, \dots, \boldsymbol{\tau}'_{a,l_1+1})'$, $\mathbf{v}_t^a = ((\mathbf{w}_t^{lm})', (\mathbf{w}_t^{nc} \mathbf{y}_{t-c})', \dots, (\mathbf{w}_t^{nc} \mathbf{y}_{t-c}^{l_1})')'$, $\tilde{\boldsymbol{\pi}}_b = (\boldsymbol{\tau}_{b,1}, \dots, \boldsymbol{\tau}_{b,(l_1+l_2)q})'$ and $\mathbf{v}_t^b = (u_{t-1}, u_{t-1}^3, \dots, u_{t-1}^{l_2+2}, \dots, u_{t-q}, u_{t-q}^3, \dots, u_{t-q}^{l_2+2})'$. Any linear term, in the conditional mean, not included in the model is denoted by \mathbf{w}_t^{lin} , whereas $\mathbf{w}_t = (1, (\mathbf{w}_t^{nc})')'$. The two remainders of the Taylor expansions $R_1(\mathbf{w}_t; \boldsymbol{\phi}_a, \rho_a, \mathbf{c}_a, \boldsymbol{\omega})$ and $R_2(\mathbf{z}_t; \boldsymbol{\psi}_b, \rho_b, \mathbf{c}_b, \boldsymbol{\omega})$ do not affect the distribution theory because both are identically equal to zero under H_0 . The new null hypothesis is $\tilde{\boldsymbol{\pi}}_a = \mathbf{0}$ and $\tilde{\boldsymbol{\pi}}_b = \mathbf{0}$, under which the LM-type test statistic (5.2) is

asymptotically χ^2 - distributed with $\dim(\tilde{\pi}_a) + \dim(\tilde{\pi}_b)$ degrees of freedom.

In practice it is most often useful to divide this joint test into separate tests for the conditional mean and the conditional variance. This helps to locate the problem, if any, and thus makes it easier to remedy it.

5.4. Test against nonconstant parameters

We assume that the alternative to constant parameters in either the conditional mean or the conditional variance or both is that the parameters change smoothly over time, see Lin and Teräsvirta (1994) and Chapter 3. Our alternative is that the parameters change deterministically over time, that is, $\varphi = \varphi^* + \lambda_\varphi \bar{H}_{n_\varphi}(t; \rho_\varphi, \mathbf{c}_\varphi)$ and $\eta = \eta^* + \lambda_\eta \bar{H}_{n_\eta}(t; \rho_\eta, \mathbf{c}_\eta)$. The vectors φ^* , λ_φ , η^* and λ_η are of the same size as φ and η respectively. Again we define $\bar{H}_{n_i} = H_{n_i} - \frac{1}{2}$ where $i = \varphi, \eta$. H_{n_i} is a logistic function of order n_i defined in (3.3) with $s_t \equiv t$, $\gamma = \rho_i$ and $\mathbf{c} = \mathbf{c}_i$. The null hypothesis of parameter constancy can be stated as $H_0 : \rho_\varphi = \rho_\eta = 0$ under which $\varphi^* = \varphi$ and $\eta^* = \eta$. We can write this alternative as a special case of (5.1) by setting $A(\mathbf{w}_t; \lambda_\varphi, \rho_\varphi, \mathbf{c}_\varphi) = \lambda_\varphi \bar{H}_{n_\varphi}(t; \rho_\varphi, \mathbf{c}_\varphi)$ and $B(\mathbf{w}_t; \lambda_\eta, \rho_\eta, \mathbf{c}_\eta) = \lambda_\eta \bar{H}_{n_\eta}(t; \rho_\eta, \mathbf{c}_\eta)$. The identification problem under the null hypothesis is circumvented as before by expanding $\bar{H}_{n_i}(t; \rho_i, \mathbf{c}_i)$ into a Taylor series around the null hypothesis, $\rho_i = 0$. Again it is a realistic assumption in a test situation to assume that the orders of the logistic function (3.3) are $n_\varphi = l_1$ and $n_\eta = l_2$ respectively. Using the first-order expansion we obtain the extended model (5.1) with $A(\mathbf{w}_t; \rho_\varphi, \mathbf{c}_\varphi) = \tilde{\pi}'_a \mathbf{v}_t^a + R_1(\mathbf{w}_t; \rho_\varphi, \mathbf{c}_\varphi)$ and $B(\mathbf{w}_t; \rho_\eta, \mathbf{c}_\eta) = \tilde{\pi}'_b \mathbf{v}_t^b + R_2(\mathbf{z}_t; \rho_\eta, \mathbf{c}_\eta)$ where $\tilde{\pi}_a = (\boldsymbol{\tau}'_{a,1}, \dots, \boldsymbol{\tau}'_{a,l_1})'$, $\tilde{\pi}_b = (\boldsymbol{\tau}'_{b,1}, \dots, \boldsymbol{\tau}'_{b,l_2})'$, $\mathbf{v}_t^a = ((\mathbf{w}_t t)')', \dots, (\mathbf{w}_t t^{l_1})')'$ and $\mathbf{v}_t^b = ((\mathbf{z}_t t)')', \dots, (\mathbf{z}_t t^{l_2})')'$. The joint null hypothesis of parameter constancy in both the conditional mean and variance consists of the restrictions $\tilde{\pi}_a = \mathbf{0}$ and $\tilde{\pi}_b = \mathbf{0}$. Under the null hypothesis the two remainder terms of the Taylor expansions $R_1(\mathbf{w}_t; \rho_\varphi, \mathbf{c}_\varphi, \boldsymbol{\omega}) \equiv R_2(\mathbf{z}_t; \rho_\eta, \mathbf{c}_\eta, \boldsymbol{\omega}) \equiv 0$ so that they do not affect the asymptotic distribution theory. Under the null hypothesis, the LM-type test statistic (5.2) is asymptotically χ^2 - distributed with $\dim(\tilde{\pi}_a) + \dim(\tilde{\pi}_b)$ degrees of freedom.

6. Modelling two Nordic currencies

In this section the STARTZ model is applied to modelling Nordic currencies. In the second half of the 1980s the Nordic currencies had unilateral target zones against a trade weighted currency basket. We focus on periods with no realignments and no policy changes. The data for all currencies are daily data and we model the

deviation, in percent, of the exchange rate index from the central parity.

6.1. The Swedish case: 1985-1991

The data of the Swedish exchange rate index cover the period from July 1, 1985 until May 17, 1991 with a total of 1472 daily observations. The starting point of the series coincides with the introduction of an explicit band. In May 1991 the trade weighted currency basket was replaced by the ECU-index. For the period concerned the exchange rate index was allowed to vary within ± 1.5 percent from its central parity. The restrictions in the STARTZ model are symmetric in the sense that the behaviour of the exchange rate is assumed to be similar near both boundaries of the target zone. Note, however, that the Swedish exchange rate index has not been close to the upper boundary of the target zone at any time during the period concerned, see Figure B.1.

The estimated parameters of the STARTZ model for the Swedish exchange rate index can be found in Table B.1. The results from the misspecification tests of the conditional mean appear in Table B.3. There is weak evidence for parameter nonconstancy in the conditional mean model, but apart from this no evidence of misspecification is found. When evaluating the conditional variance part of the model it is found that the standard GARCH model does not adequately characterize the series. The misspecification tests suggest that evidence of asymmetry/nonlinearity is found for the constant term, α_0 , of the standard GARCH model. The conditional variance parameterization is therefore extended to the GQARCH model of Sentana (1995). This is done by setting $\boldsymbol{\eta} = (\alpha_0, \alpha_1^*, \dots, \alpha_q^*, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ and $\mathbf{z}_t = (1, u_{t-1}, \dots, u_{t-q}, u_{t-1}^2, \dots, u_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$ in (3.5). After this respecification the estimated STARTZ model passes all misspecification tests except the one for parameter constancy, see Table B.4. Nevertheless, we tentatively accept the current specification.

When considering the estimated STARTZ model we find in the conditional mean that the transition between the autoregressive part $\boldsymbol{\varphi}'\mathbf{w}_t$ and, for example, the lower boundary μb^L is not smooth but rather abrupt. The explicit boundaries serve as the actual boundaries of the target zone since $\hat{\mu} = 0.99$. Furthermore, the nonlinear mechanism is rarely active during the observation period. This implies that the conditional mean follows a locally linear autoregressive process that is close to a unit root ($\hat{\varphi}_1 + \hat{\varphi}_2 = 0.996$) almost everywhere within the target zone. This comes as no surprise, because during the observation period Sweden's central bank only intervened intramarginally. According to the theory this should result in a less

pronounced nonlinearity, see for example Svensson (1992). For a discussion of the Swedish intervention policy see Lindberg and Söderlind (1994a). The BDS test in Table B.2 does not reject the IID hypothesis of the errors.

The plot of the estimated conditional variance parameterization against the observed deviation (in percent) from the central parity shows that the conditional variance has a weak \cap -shape, see Figure B.3. This is in line with the results obtained if the agents do not consider the band fully credible, see Bertola and Svensson (1993) and Bertola and Caballero (1992).

6.2. The Norwegian case: 1986-1990

The Norwegian series covers the period from October 1, 1986 until October 21, 1990. A realignment of the Norwegian krone took place in May 1986 and the trade weighted currency basket was replaced by the ECU-index in October 22, 1990. Throughout the sample period the exchange rate index is allowed to vary within ± 2.25 percent from the central parity, see Figure B.2.

We divide the series into two subseries. The first one covers the period October 1, 1986 until June 17, 1988 with a total of 431 daily observations. The second subseries includes January 2, 1989 until October 21, 1990 with a total of 449 daily observations. A disadvantage of from splitting the data set is the low number of observations in the two subsets. There is a compelling reason for this decision, however. The intervention policy of the Central Bank of Norway was altered within the sample period; for a discussion see Lysebo and Mundaca (1997). From October 1986 until mid-June 1988 the interventions mainly occurred at the boundary of the target zone. In mid-June 1988 the authorities declared a change in their intervention policy and from July 1988 until October 1990 the interventions mostly took place inside the band. Late in 1988 the bank started to maintain an implicit target zone that was narrower than the official one, see Mundaca (1998).

We first consider the estimated STARTZ model for the period October 1, 1986 until June 17, 1988. Parameter estimates can be found in Table B.1. No evidence for misspecification of the conditional mean is found, see Table B.3. On the other hand, there is weak evidence against parameter constancy in the conditional variance specification, see Table B.4. Due to the small number of observations we decided not to extend the parameterization of the model as a response to the test result. The BDS test in Table B.2 rejects the IID hypothesis of the standardized errors but nevertheless we tentatively accept the current specification. The slope parameter, γ_a , in (3.4) for the conditional mean is estimated to 1.1, which implies that the

transition in the dynamic structure when the series approaches the boundary is very smooth, see Figure B.4 for an illustration. The explicit boundaries serve as the actual boundaries of the target zone since $\hat{\mu} = 0.99$. The estimated conditional variance has a pronounced \cap -shape, see Figure B.5. These results are in line with the behaviour implied by the basic target zone model of Krugman (1991).

Second, we consider the estimated STARTZ model for the period January 2, 1989 until October 21, 1990. The estimated parameters of the model appear in Table B.1. No evidence for misspecification is found in the conditional mean or the conditional variance, see Tables B.3 and B.4 respectively. The BDS test in Table B.2 does not reject the IID hypothesis of the standardized errors. We find that the estimated implicit band has a bandwidth of ± 0.5 percent around the central parity since $\hat{\mu} = 0.22$. The nonlinear mechanism in the conditional mean is only active very close to the boundary of the target zone and the process is mean reverting ($\hat{\varphi}_1 + \hat{\varphi}_2 = 0.96$) within the target zone. The estimated conditional variance does not have a \cap -shape, see Figure B.6. An implication of the model is that the boundaries of the implicit target zone are soft.

7. Reproducing stylized facts

In this section we demonstrate the properties of the estimated STARTZ models by simulation. We also want to see if the STARTZ model can reproduce the different shapes of the marginal density suggested by the target zone theories. This is done by small simulation studies. Data are generated from the estimated models. When doing this we force the conditional variance to be positive. The standard normal random numbers have been generated by the random number generator in GAUSS 3.2.31. The first 5000 observations of each generated series have been discarded to avoid initialization effects.

To examine if the STARTZ model is capable of reproducing the small sample properties we first simulate 10000 samples using the parameter values that were estimated and the same number of observations for both the Swedish krona and the Norwegian krone. For each of these 10000 samples a number of statistics are computed, see Table B.5. Kernel estimates of the density for the statistics are graphed in Figure B.10 for the Swedish krona and in Figures B.11 and B.12 for the Norwegian krone. For most statistics we find that empirical (observed) values lie within the simulated distribution.

To see if the STARTZ model can reproduce the marginal densities that were

found in the empirical data we simulate data using the parameter estimates for the daily Swedish krona and the daily Norwegian krone. In these simulations we only generate one sample of 100000 observations. For the model of the Swedish krona it is found that the marginal distribution is not U-shaped. The mass of the marginal distribution is located in the interior of the band instead, see Figure ?? . Only close to the boundaries of the target zone does the marginal distribution have a U-shape. Considering the model for the Norwegian krone in 1986-88, it is found that the marginal distribution has a pronounced U-shape, see Figure B.8. Thus the simulated observations are likely to fall close to either boundary of the zone. For the second period of the Norwegian krone, 1989-90, the marginal distribution has a \cap -shape, see Figure B.9. The mass of the distribution is located at central parity and there is no indication of a target zone behaviour at the boundary.

8. Conclusions

In the target zone literature, the emphasis has been on theoretical models. This paper proposes a rather flexible, empirical, time series model that is capable of capturing the behaviour implied by theoretical target zone models. The model also enables us to estimate the boundaries of an implicit band, should such a band exist. To model empirical data in a systematic way, having a coherent modelling strategy, is important and such a strategy is designed and applied to data here. A statistical advantage of the proposed strategy is that the misspecification tests we use only require standard asymptotic theory and are easy to perform.

The empirical examples indicate that there is structure in data that corresponds to theoretical target zone models. For the Swedish krona the behaviour within the target zone is in line with what theory suggests for a currency where the central bank intervenes intramarginally. In the case of the Norwegian krone 1986-88, where the Central Bank intervened only at the edges of the band, the behaviour of the estimated model is well in line with the results implied by the basic target zone model. For the other period, 1988-90, of the Norwegian krone the estimated model suggests that the implicit band maintained by the Central Bank during this period was not strictly enforced by the bank.

The STARTZ model may also be used to model other economic variables restricted by explicit or implicit boundaries, such as unemployment or interest rate series. The single-equation STARTZ model may also easily be made multivariate. In the present case this would allow the possibility of incorporating fundamentals

into the time series target zone model, a topic which is left for further research.

References

- Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions of Automatic Control* **19**: 716–723.
- Anthony, M. and MacDonald, R. (1998). On the mean-reverting properties of target zone exchange rates: Some evidence from the ERM, *European Economic Review* **42**: 1493–1523.
- Baillie, R. T. and Bollerslev, T. (1989). The message in daily exchange rates: A conditional-variance tale, *Journal of Business and Economic Statistics* **7**: 297–305.
- Bertola, G. and Caballero, R. J. (1992). Target zones and realignments, *Review of Economic Studies* **82**: 520–526.
- Bertola, G. and Svensson, L. E. O. (1993). Stochastic devaluation risk and the empirical fit of target-zone models, *Review of Economic Studies* **60**: 689–712.
- Brock, W. A., Dechert, W., Scheinkman, J. and LeBaron, B. (1996). A test for independence based on the correlation dimension, *Econometric Reviews* **15**: 197–235.
- Delgado, F. and Dumas, B. (1991). Target zones, broad and narrow, in P. Krugman and M. Miller (eds), *Exchange Rate Targets and Currency Bands*, Cambridge University Press, Cambridge, pp. 35–56.
- Diebold, F. and Nerlove (1989). The dynamics of exchange rate volatility: A multivariate latent factor ARCH model, *Journal of Applied Econometrics* **4**: 1–21.
- Engle, R. F. and Gau, Y.-F. (1997). Conditional volatility of exchange rates under a target zone. University of California, San Diego, Department of Economics, Discussion Paper No 97-06.
- Fiorentini, G., Calzolari, G. and Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates, *Journal of Applied Econometrics* **11**: 399–417.
- Flood, R. P., Rose, A. K. and Mathieson, D. J. (1991). An empirical exploration of exchange-rate target-zones, *Carnegie-Rochester Series on Public Policy* **35**: 7–65.

- Froot, K. A. and Obstfeld, M. (1991). Exchange-rate dynamics under stochastic regime shifts: A unified approach, *Journal of International Economics* **31**: 203–229.
- Krugman, P. R. (1991). Target zones and exchange rate dynamics, *Quarterly journal of economics* **106**: 669–682.
- Lin, C.-F. and Teräsvirta, T. (1994). Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* **62**: 211–228.
- Lindberg, H. and Söderlind, P. (1994a). Intervention policy and mean reversion in exchange rate target zones: The Swedish case, *Scandinavian Journal of Economics* **96**: 499–513.
- Lindberg, H. and Söderlind, P. (1994b). Testing the basic target zone model on Swedish data 1982–1990, *European Economic Review* **38**: 1441–1469.
- Lysebo, T. A. and Mundaca, B. G. (1997). Does the credibility of the exchange rate depend on the width of the target zone. Norges Bank.
- Meese, R. A. and Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample, *Journal of International Economics* **14**: 3–24.
- Mundaca, G. (1998). The effect of interventions on the probabilities of realignments. Working Paper, Research Department, Central Bank of Norway, Oslo, Norway.
- Rose, A. K. and Svensson, L. E. O. (1994). European exchange rate credibility before the fall, *European Economic Review* **38**: 1185–1216.
- Sentana, E. (1995). Quadratic ARCH models, *Review of Economic Studies* **62**: 639–661.
- Svensson, L. E. O. (1992). An interpretation of recent research on exchange rate target zones, *Journal of Economic Perspectives* **6**: 119–144.
- Svensson, L. E. O. (1993). Assessing target zone credibility, *European Economic Review* **37**: 763–802.
- Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* **89**: 208–218.

A. Analytical derivatives

In this section we consider the analytical derivatives of the suggested model. These first-order derivatives are used in the estimation and later on in the evaluation of the estimated model. The derivatives of the model are straightforward to compute, see for example Fiorentini, Calzolari and Panattoni (1996) and Chapter 2. Let $\boldsymbol{\omega}_a = (\boldsymbol{\varphi}', \mu, \gamma_a)'$ and $\boldsymbol{\omega}_b = (\boldsymbol{\eta}', \delta, \gamma_b)'$; consider the model defined by (3.1-3.5):

$$\begin{aligned} y_t &= m_t + u_t \\ u_t &= \varepsilon_t \sqrt{h_t} \end{aligned}$$

where $m_t = m_t(\boldsymbol{\omega}_a)$ and $h_t = h_t(\boldsymbol{\omega}_a, \boldsymbol{\omega}_b)$ are the parametrization of the conditional mean and the conditional variance. If we assume that $\{\varepsilon_t\}$ is a sequence of independent standard normal errors, the log-likelihood function at time t is given by:

$$l_t = \text{const} - \frac{1}{2} \ln h_t - \frac{1}{2} \frac{u_t^2}{h_t}$$

where $u_t = y_t - m_t$ as $\boldsymbol{\omega}_a$ is assumed not to depend on $\boldsymbol{\omega}_b$.

A.1. Partial derivative of l_t

The first-order partial derivative (the gradient) of the log likelihood function at time t is

$$G_t = \begin{pmatrix} \frac{\partial l_t}{\partial \boldsymbol{\omega}'_a} & \frac{\partial l_t}{\partial \boldsymbol{\omega}'_b} \end{pmatrix}$$

where the corresponding elements are

$$\begin{aligned} \frac{\partial l_t}{\partial \boldsymbol{\omega}'_a} &= \frac{u_t}{h_t} \frac{\partial m_t}{\partial \boldsymbol{\omega}'_a} + \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\omega}'_a} \\ \frac{\partial l_t}{\partial \boldsymbol{\omega}'_b} &= \frac{1}{2h_t} \left(\frac{u_t^2}{h_t} - 1 \right) \frac{\partial h_t}{\partial \boldsymbol{\omega}'_b}. \end{aligned}$$

The second-order partial derivative (the Hessian) of the log-likelihood function at time t is

$$H_t = \begin{pmatrix} \frac{\partial^2 l_t}{\partial \boldsymbol{\omega}_a \partial \boldsymbol{\omega}'_a} & \frac{\partial^2 l_t}{\partial \boldsymbol{\omega}_a \partial \boldsymbol{\omega}'_b} \\ \frac{\partial^2 l_t}{\partial \boldsymbol{\omega}_b \partial \boldsymbol{\omega}'_a} & \frac{\partial^2 l_t}{\partial \boldsymbol{\omega}_b \partial \boldsymbol{\omega}'_b} \end{pmatrix} \quad (\text{A.1})$$

and the expectations of the elements in (A.1) are

$$E \left[\frac{\partial^2 l_t}{\partial \boldsymbol{\omega}_a \partial \boldsymbol{\omega}'_a} \right] = -E \left[\frac{1}{h_t} \frac{\partial m_t}{\partial \boldsymbol{\omega}_a} \frac{\partial m_t}{\partial \boldsymbol{\omega}'_a} + \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \boldsymbol{\omega}_a} \frac{\partial h_t}{\partial \boldsymbol{\omega}'_a} \right]$$

$$\begin{aligned} E \left[\frac{\partial^2 l_t}{\partial \omega_b \partial \omega'_b} \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \omega_b} \frac{\partial h_t}{\partial \omega'_b} \right] \\ E \left[\frac{\partial^2 l_t}{\partial \omega_a \partial \omega'_b} \right] &= -\frac{1}{2} E \left[\frac{1}{h_t^2} \frac{\partial h_t}{\partial \omega_a} \frac{\partial h_t}{\partial \omega'_b} \right]. \end{aligned}$$

A.2. Partial derivative of the conditional mean m_t

The conditional mean is defined as in (3.1) and (3.4). For notational convenience also let $m_t^* = \boldsymbol{\varphi}' \mathbf{w}_t$, $\xi_a^L = 0.5(-y_{t-1} + \mu b^L) \gamma_a$ and $\xi_a^U = 0.5(y_{t-1} - \mu b^U) \gamma_a$. We can then write the conditional mean as

$$m_t = m_t^* + 2(\mu b^L - m_t^*) H_a^L + 2(\mu b^U - m_t^*) H_a^U$$

where $H_a^L = 0.5 \exp(\xi_a^L) / \cosh(\xi_a^L)$ and $H_a^U = 0.5 \exp(\xi_a^U) / \cosh(\xi_a^U)$. The first-order derivative is

$$\begin{aligned} \frac{\partial m_t}{\partial \boldsymbol{\varphi}'} &= (1 - 2H_a^L - 2H_a^U) \frac{\partial m_t^*}{\partial \boldsymbol{\varphi}'} \\ \frac{\partial m_t}{\partial \mu} &= 2b^L H_a^L + 2b^U H_a^U + \\ &\quad \frac{(\mu b^L - m_t^*)(\gamma_a b^L)}{2 \cosh^2(\xi_a^L)} + \frac{(\mu b^U - m_t^*)(-\gamma_a b^U)}{2 \cosh^2(\xi_a^U)} \\ \frac{\partial m_t}{\partial \gamma_a} &= \frac{(\mu b^L - m_t^*)(-y_{t-1} + \mu b^L)}{2 \cosh^2(\xi_a^L)} + \frac{(\mu b^U - m_t^*)(y_{t-1} - \mu b^U)}{2 \cosh^2(\xi_a^U)} \end{aligned}$$

where the necessary derivatives under the null hypothesis used in the evaluation tests are easily obtained by replacing m_t^* with $m_t^{alt} = \boldsymbol{\varphi}' \mathbf{w}_t + A(\mathbf{w}_t; \boldsymbol{\pi}_a)$.

A.3. Partial derivative of the conditional variance h_t

We assume that the conditional variance is parameterized as in (3.2) and (3.5). We rewrite the expression in the same way as we did for the conditional mean. Thus by setting $h_t^* = \boldsymbol{\eta}' \mathbf{z}_t$, $\xi_b^L = 0.5(-y_{t-1} + \mu b^L) \gamma_b$ and $\xi_b^U = 0.5(y_{t-1} - \mu b^U) \gamma_b$ the conditional variance may be written as

$$h_t = h_t^* + 2(\delta - h_t^*) H_b^L + 2(\delta - h_t^*) H_b^U$$

where $H_b^L = 0.5 \exp(\xi_b^L) / \cosh(\xi_b^L)$ and $H_b^U = 0.5 \exp(\xi_b^U) / \cosh(\xi_b^U)$. To initialize the iterative computation of h_t , the conditional variance is estimated with the sample (unconditional) variance in the pre-sample case. This is done for all $t \leq 0$ by setting $h_t = u_t^2 = \frac{1}{T} \sum_{i=1}^T u_i^2$ where $u_i = y_i - m_i$. The first-order derivatives may be computed iteratively by using the following expressions

$$\begin{aligned}
\frac{\partial h_t}{\partial \varphi'} &= (1 - 2H_b^L - 2H_b^U) \frac{\partial h_t^*}{\partial \varphi'} \\
\frac{\partial h_t}{\partial \mu} &= (1 - 2H_b^L - 2H_b^U) \frac{\partial h_t^*}{\partial \mu} + \\
&\quad \frac{(\delta - h_t^*)(\gamma_b b^L)}{2 \cosh^2(\xi_b^L)} + \frac{(\delta - h_t^*)(-\gamma_b b^U)}{2 \cosh^2(\xi_b^U)} \\
\frac{\partial h_t}{\partial \gamma_a} &= (1 - 2H_b^L - 2H_b^U) \frac{\partial h_t^*}{\partial \gamma_a} \\
\frac{\partial h_t}{\partial \eta'} &= (1 - 2H_b^L - 2H_b^U) \frac{\partial h_t^*}{\partial \eta'} \\
\frac{\partial h_t}{\partial \delta} &= (1 - 2H_b^L - 2H_b^U) \frac{\partial h_t^*}{\partial \delta} + (2H_b^L + 2H_b^U) \\
\frac{\partial h_t}{\partial \gamma_b} &= (1 - 2H_b^L - 2H_b^U) \frac{\partial h_t^*}{\partial \gamma_b} + \\
&\quad \frac{(\delta - h_t^*)(-y_{t-1} + \mu b^L)}{2 \cosh^2(\xi_b^L)} + \frac{(\delta - h_t^*)(y_{t-1} - \mu b^U)}{2 \cosh^2(\xi_b^U)}
\end{aligned}$$

where the necessary derivatives under the null hypothesis used in the evaluation tests are easily obtained by replacing h_t^* with $h_t^{alt} = \boldsymbol{\eta}' \mathbf{z}_t + B(\mathbf{w}_t; \boldsymbol{\pi}_b)$.

B. Tables and Figures

Parameter	SEK (85-91)	NOK (86-88)	NOK (89-90)
φ_1	0.85 (0.030)	0.40 (0.086)	0.67 (0.046)
φ_2	0.15 (0.030)	.	0.29 (0.044)
μ	0.99 (0.0034)	0.99 (0.093)	0.22 (0.0053)
γ_a	130 (120)	1.1 (0.26)	41 (28)
α_0	0.00091 (0.00024)	0.031 (0.014)	0.00028 (0.00025)
α_1^*	0.0091 (0.0025)	.	.
α_1	0.15 (0.026)	0.13 (0.066)	.
α_2	.	.	0.11 (0.053)
β_1	0.72 (0.047)	0.30 (0.25)	0.83 (0.091)
δ	0.00095 (0.00036)	0.0063 (0.018)	0.0068 (0.0015)
γ_b	36 (20)	6.1 (3.3)	26 (13)

Table B.1: Parameter estimates of the STARTZ models (standard deviation in parenthesis) for the Swedish krona (SEK) and the Norwegian krone (NOK).

	SEK (85-91)	NOK (86-88)	NOK (89-90)
Characteristics:			
<i>Min</i>	-5.1	-4.6	-3.1
<i>Max</i>	4.8	5.4	2.5
<i>Mean</i>	-0.018	-0.041	0.050
<i>Variance</i>	1.0	1.0	1.0
<i>Skewness</i>	0.15	0.54	-0.25
<i>Kurtosis</i>	4.9	6.4	3.1
Nonparametric test of IID (BDS)			
<i>BDS statistic</i>	-0.41	3.4	-1.6
<i>p-value (asymptotic)</i>	0.67	0.00057	0.11
<i>p-value (bootstrap)</i>	0.70	0.0	0.14

Table B.2: Characteristics of the standardized residuals of the STARTZ model and the BDS test of independence for the standardized errors. For the BDS test a bootstrapped probability value based on 1000 resampled series is reported as well.

	SEK (85-91)	NOK (86-88)	NOK (89-90)
Remaining autocorrelation (<i>p</i>-values)			
<i>l</i> = 1	0.94	0.94	0.61
<i>l</i> = 2	0.91	0.93	0.63
<i>l</i> = 3	0.97	0.97	0.50
<i>l</i> = 4	0.78	0.92	0.11
Parameter constancy (<i>p</i>-values)			
<i>n</i> = 1	0.99	0.98	0.74
<i>n</i> = 2	0.010	0.94	0.70
Remaining nonlinearity (<i>p</i>-values)			
<i>d</i> = 1	0.32	0.63	0.76
<i>d</i> = 2	0.13	0.79	0.69
<i>d</i> = 3	0.17	0.92	0.64
<i>d</i> = 4	0.012	0.95	0.42

Table B.3: *p*-values of specification tests for the conditional mean for the estimated STARTZ model. LM tests for the conditional mean. The test of no remaining autocorrelation is computed against the alternative of remaining autocorrelation up to the given lag, *l*. The test of parameter constancy is computed against an alternative of time dependence given by a logistic function of order *n* with time as the transition variable. The test against nonlinearity of LSTAR(3) type uses a lag length 4 for the autoregressive part. The test is computed separately against the alternatives $1 \leq d \leq 4$.

	SEK (85-91)	NOK (86-88)	NOK (89-90)
Test of the functional form (<i>p</i>-values)			
<i>l</i> = 1	0.20	0.036	0.16
<i>l</i> = 2	0.42	0.047	0.35
<i>l</i> = 4	0.55	0.024	0.72
<i>l</i> = 8	0.0020	0.050	0.31
Parameter constancy (<i>p</i>-values)			
<i>n</i> = 1	0.17	0.26	0.18
<i>n</i> = 2	0.0068	0.0017	0.094
Remaining nonlinearity of STGARCH type (<i>p</i>-values)			
<i>All parameters</i>	.	0.11	0.15
<i>Without intercept</i>	0.56	0.24	0.16

Table B.4: *p*-values of specification tests for the conditional variance of the estimated STARTZ model. LM tests for the conditional variance. The tests of no remaining serial dependence in the squared and standardized residuals are computed against the alternative of remaining dependence up to the given lag, *l*. The test of parameter constancy is computed against an alternative of time dependence given by a logistic function of order *n* with time as the transition variable. The test against remaining nonlinearity is a test against an STGARCH(3) structure.

	Simulated			Observed		
	SEK (85-91)	NOK (86-88)	NOK (89-90)	SEK (85-91)	NOK (86-88)	NOK (89-90)
<i>Min</i>	-1.4 (0.31)	-2.1 (0.69)	-0.55 (0.13)	-1.5	-2.2	-0.42
<i>Max</i>	1.4 (0.31)	2.1 (0.69)	0.55 (0.13)	0.82	2.3	0.59
<i>Mean</i>	-0.00094 (0.41)	-0.0028 (1.0)	0.0013 (0.088)	-0.53	-0.28	0.11
<i>Variance</i>	0.65 (0.14)	1.1 (0.35)	0.21 (0.038)	0.56	1.4	0.045
<i>Skewness</i>	0.036 (0.43)	0.0019 (0.81)	0.0011 (0.36)	0.31	0.15	-0.41
<i>Kurtosis</i>	2.4 (0.54)	2.6 (1.22)	2.8 (0.52)	2.1	1.8	2.9

Table B.5: Simulated and empirical characteristics for the STARTZ models (standard deviation in parentheses) for the Swedish krona and the Norwegian krone.

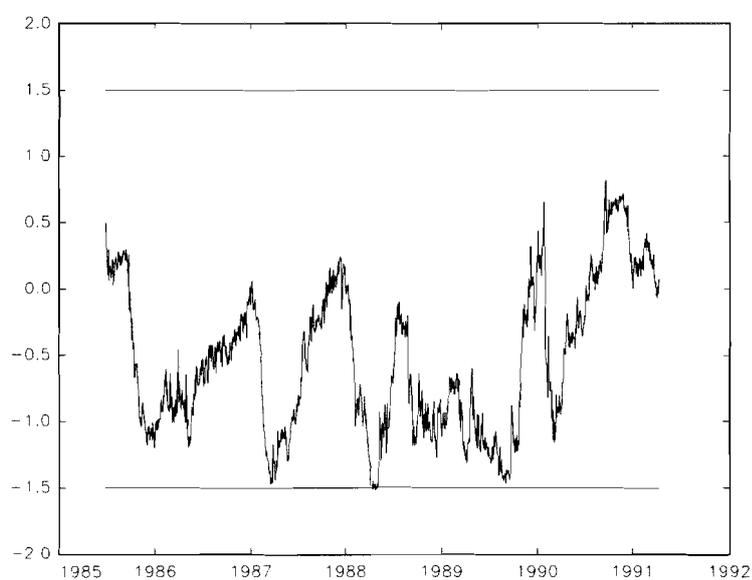


Figure B.1: The deviation (in percent) from central parity for the daily Swedish exchange rate index, July 1, 1985 to May 17, 1991. During this period the central parity of the target zone was 132 and the exchange rate index was allowed to vary within ± 1.5 percent from the central parity.

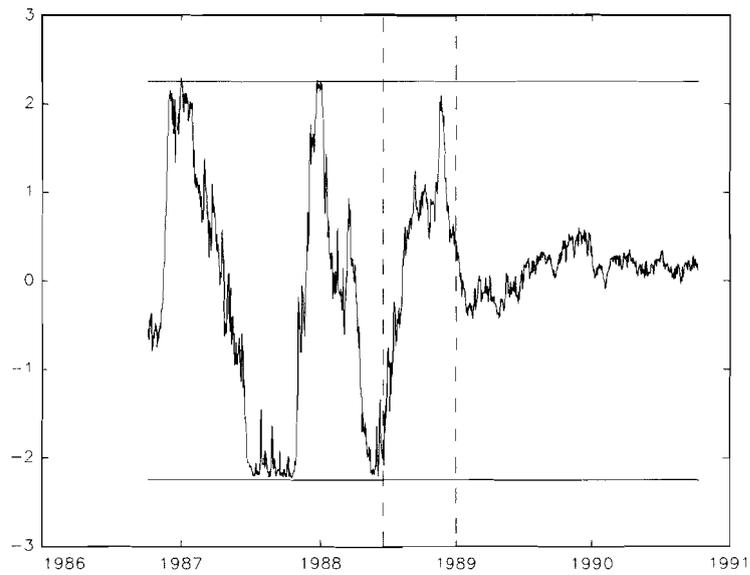


Figure B.2: The deviation (in percent) from the central parity for the daily Norwegian exchange rate index, October 1, 1986 to October 19, 1990. During this period the central parity of target zone was 112 and the exchange rate index was allowed to vary within ± 2.25 percent from the central parity. The first dashed line corresponds to June 17, 1988. At that date the authorities changed their intervention policy. The second dashed line corresponds to January 2, 1989.

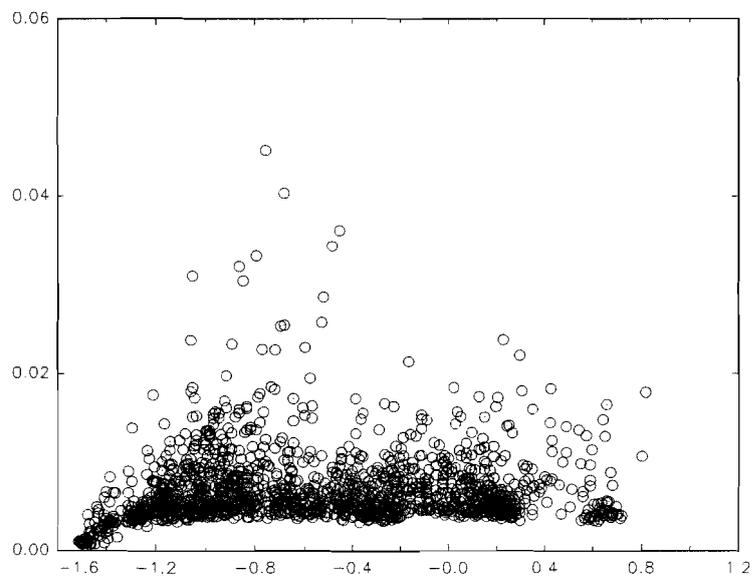


Figure B.3: The daily Swedish exchange rate index, July 1, 1985 to May 17, 1991. The parameterization of conditional variance, h_t , on the y-axis is plotted against the observed deviation from the central parity (in percent) on the x-axis.

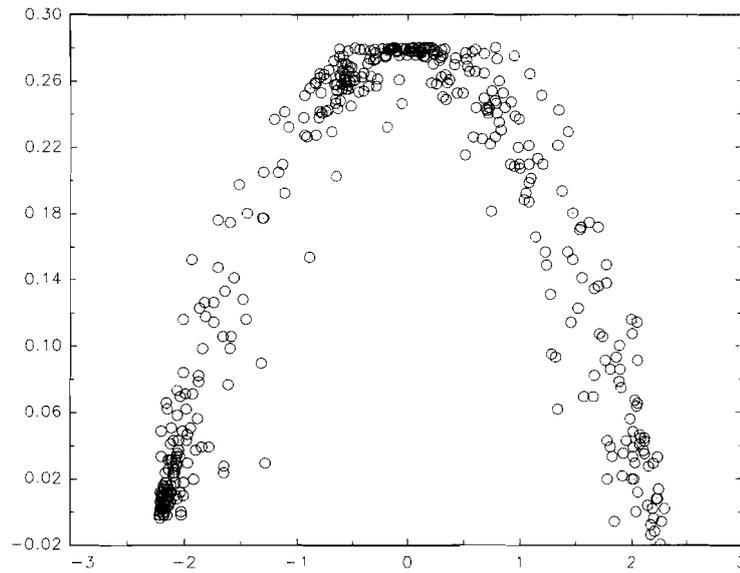


Figure B.4: The daily Norwegian exchange rate index, October 1, 1986 to June 17, 1988. The value of the restricted parameter in the conditional mean $[\varphi' - 2\varphi'H_1(-\varphi'w_t; \gamma_a, -b^L) - 2\varphi'H_1(\varphi'w_t; \gamma_a, b^U)]$ on the y-axis is plotted against the observed deviation from the central parity (in percent) on the x-axis.

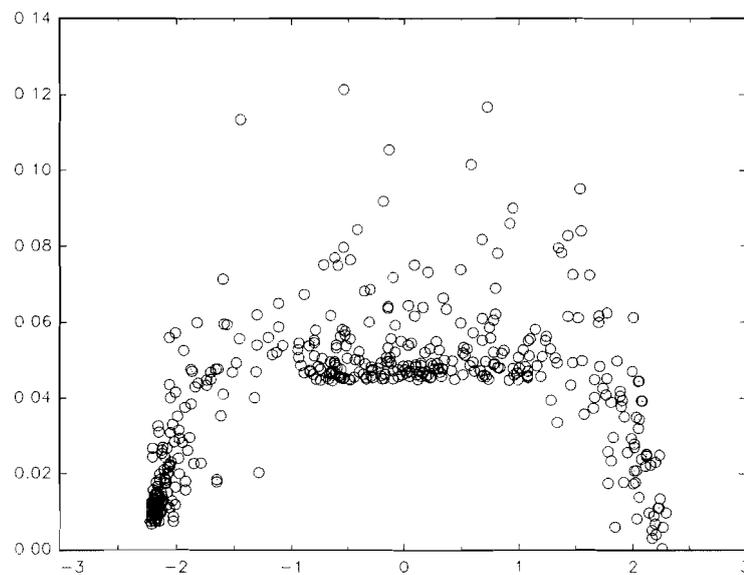


Figure B.5: The daily Norwegian exchange rate index, October 1, 1986 to June 17, 1988. The parameterization of conditional variance, h_t , on the y-axis is plotted against the observed deviation from the central parity (in percent) on the x-axis.

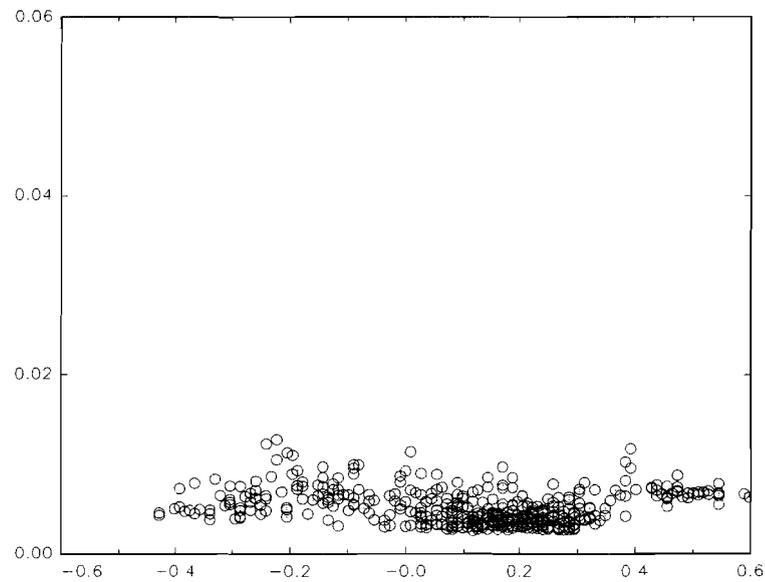


Figure B.6: The daily Norwegian exchange rate index, January 2, 1989 to October 21, 1990. The parameterization of conditional variance, h_t , on the y-axis is plotted against the observed deviation from the central parity (in percent) on the x-axis.

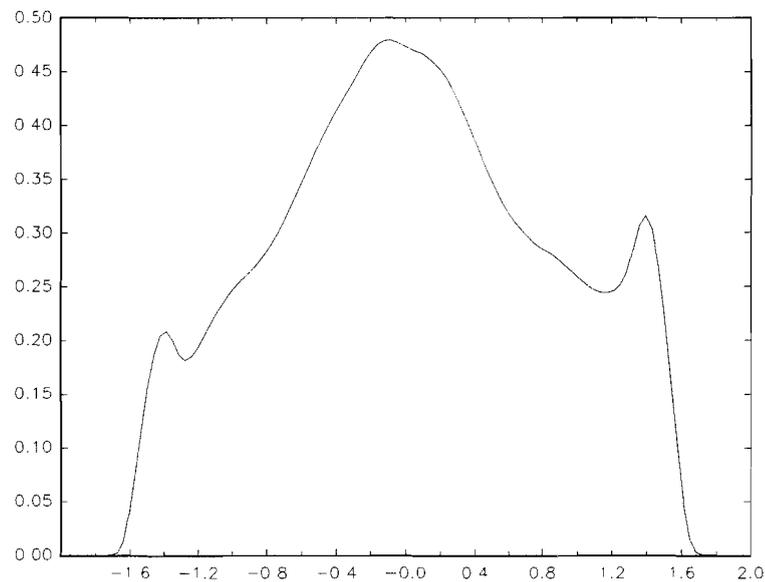


Figure B.7: Illustration of the marginal (mean) density for the STARTZ model. The parameter estimates obtained for the Swedish krona for the period July 1, 1985 to May 17, 1991 are used in this simulation. A kernel estimate of the marginal density based on 100000 generated data points is plotted in the figure.

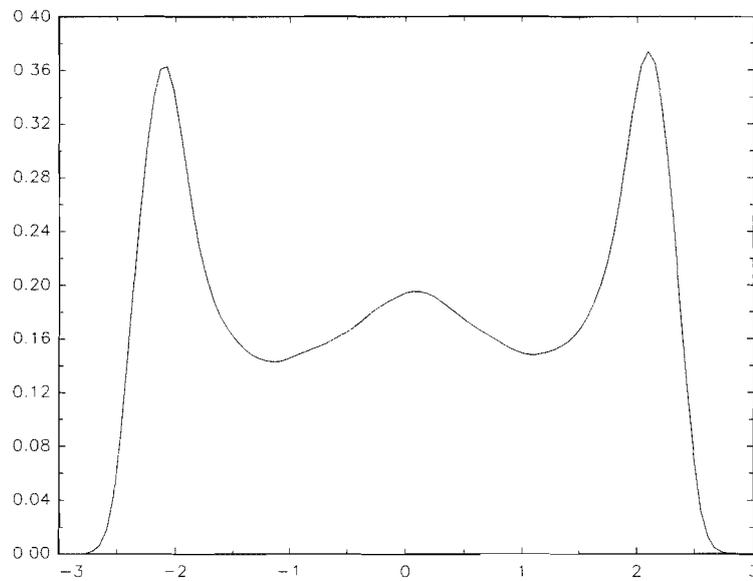


Figure B.8: Illustration of the marginal (mean) density for the STARTZ model. The parameter estimates obtained for the Norwegian krone in the period October 1, 1986 to June 17, 1988 are used in this simulation. A kernel estimate of the marginal density based on 100000 generated data points is plotted in the figure.

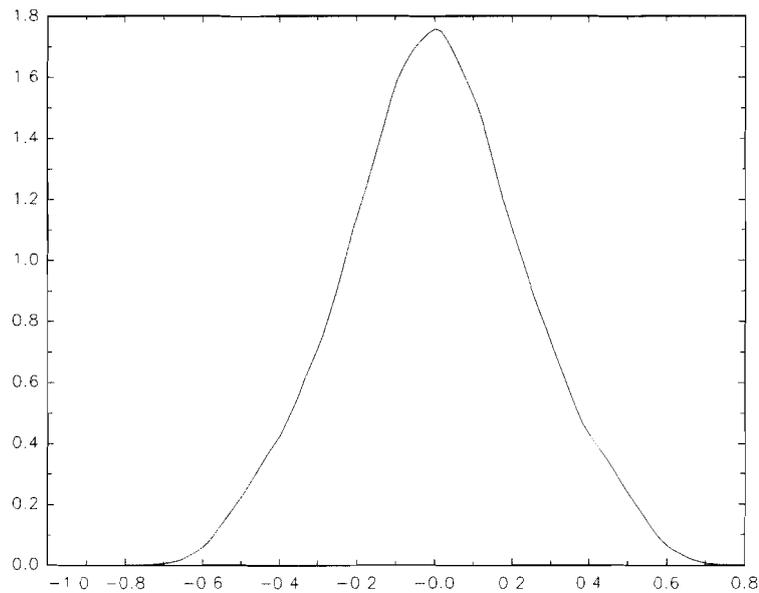
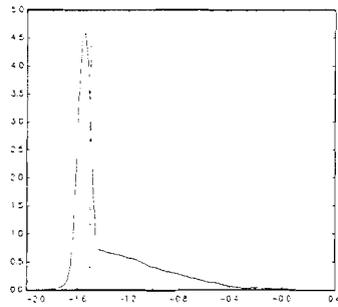
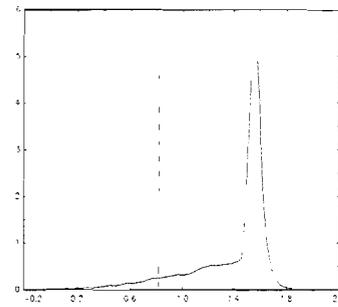


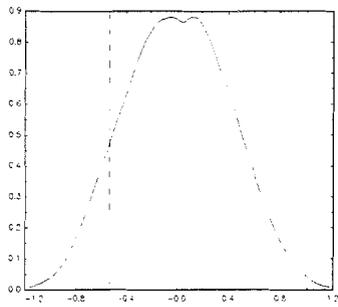
Figure B.9: Illustration of the marginal (mean) density for the STARTZ model. The parameter estimates obtained for the Norwegian krone for the period January 2, 1989 to October 21, 1990 are used in this simulation. A kernel estimate of the marginal density based on 100000 generated data points is plotted in the figure.



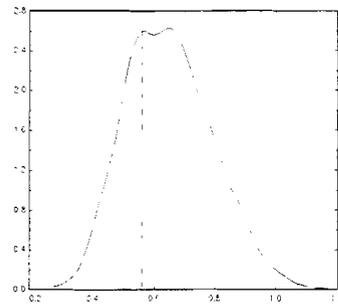
(a) Minimum



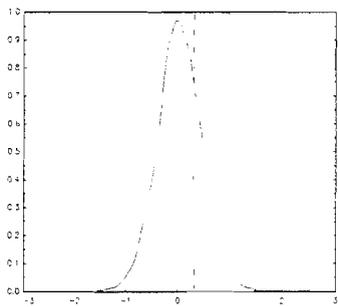
(b) Maximum



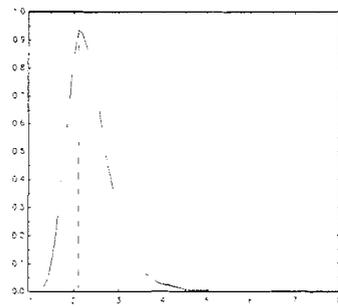
(c) Mean



(d) Variance

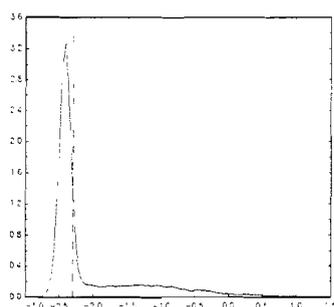


(e) Skewness

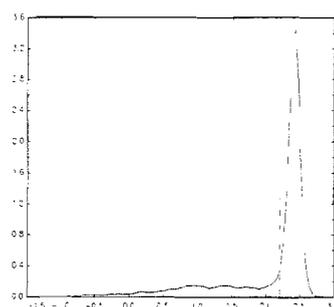


(f) Kurtosis

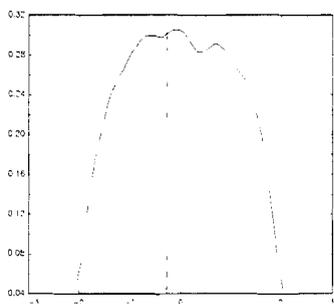
Figure B.10: Simulations of the small-sample properties of the model estimated for the Swedish krona during the period July 1, 1985 until May 17, 1991. Kernel estimates for a number of statistics based on 10000 replications are plotted in the figure. The dashed lines correspond to the observed values of the statistics.



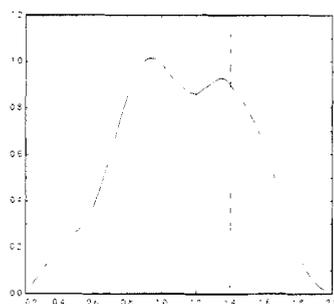
(a) Minimum



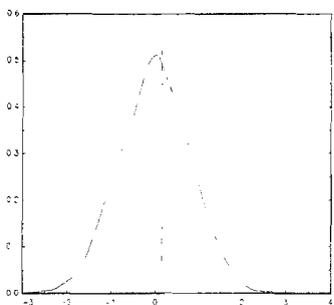
(b) Maximum



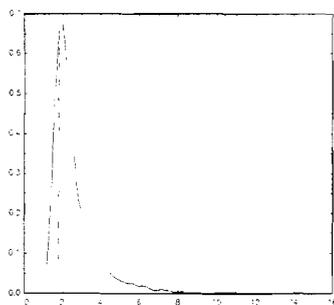
(c) Mean



(d) Variance

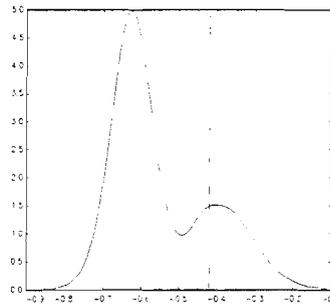


(e) Skewness

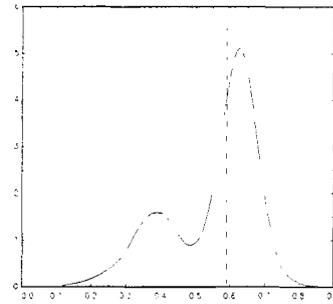


(f) Kurtosis

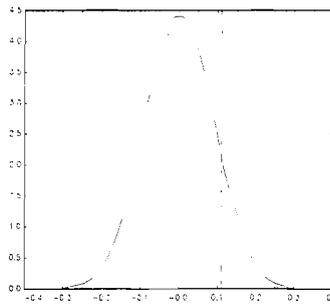
Figure B.11: Simulations of the small-sample properties of the model estimated for the Norwegian krone during the period October 1, 1986 until June 17, 1988. Kernel estimates for a number of statistics based on 10000 replications are plotted in the figure. The dashed lines correspond to the observed values of the statistics.



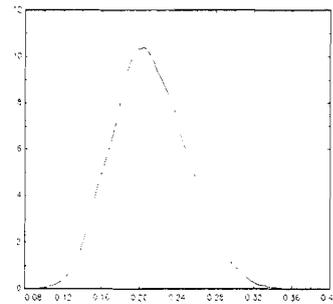
(a) Minimum



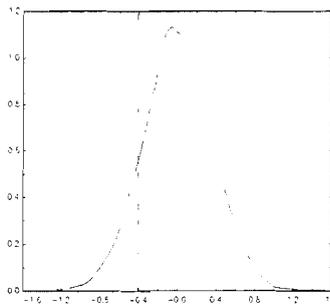
(b) Maximum



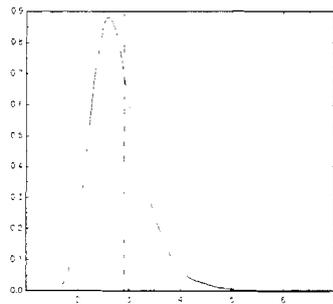
(c) Mean



(d) Variance



(e) Skewness



(f) Kurtosis

Figure B.12: Simulations of the small-sample properties of the model estimated for the Norwegian krone during the period January 2, 1989 until October 21, 1990. Kernel estimates for a number of statistics based on 10000 replications are plotted in the figure. The dashed lines correspond to the observed values of the statistics.

EFI

The Economic Research Institute

Reports since 1995

Published in the language indicated by the title

1999

- Andersson, P.**, Experto Credite: Three Papers on Experienced Decision Makers.
- Ekman, G.**, Från text till batong – Om poliser, busar och svennar.
- Gunnarsson, J.**, Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
- Hamrefors, S.**, Spontaneous Environmental Scanning.
- Japanese Production Management in Sunrise or Sunset.** Christer Karlsson (editor).
- Lindé, J.**, Essays on the Effects of Fiscal and Monetary Policy.
- Ljunggren, U.**, Indikatorer i grundskolan i Stockholms stad före stadsdelsnämndsreformen – en kartläggning
- Mägi, A.**, Store Loyalty? An Empirical Study of Grocery Shopping.
- Nilsson, K.**, Ledtider för ledningsinformation.
- Rognes, J.**, Telecommuting – Organisational Impact of Home Based – Telecommuting. 1999.
- Skalin, J.**, Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
- Spagnolo, G.**, Essays on Managerial Incentives and Product-Market Competition.
- Strauss, T.**, Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
- Svedberg Nilsson, K.**, Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
- Söderström, U.**, Monetary Policy under Uncertainty.

1998

- Andersson, M.**, On Testing and Forecasting in Fractionally Integrated Time Series Models.
- Berg-Suurwee, U.**, Styrning av kultur- och fritidsförvaltning innan stadsdelsnämndsreformen
- Berg-Suurwee, U.**, Nyckeltal avseende kultur- och fritidsförvaltning innan stadsdelsnämndsreformen.
- Bergström, F.**, Essays on the Political Economy of Industrial Policy.
- Bild, M.**, Valuation of Takeovers.
- Charpentier, C., Samuelson, L.A.**, Effekter av en sjukvårdsreform – en analys av Stockholmsmodellen.
- Eriksson-Skoog, G.**, The Soft Budget Constraint: The Emergence, Persistence and Logic of an Institution. The Case of Tanzania 1967-1992.
- Gredenhoff, M.**, Bootstrap Inference in Time Series Econometrics.
- Ioannidis, D.**, I nationens tjänst? Strategisk handling i politisk miljö – en nationell teleoperatörs interorganisatoriska, strategiska utveckling.
- Johansson, S.**, Savings Investment, and Economic Reforms in Developing Countries.
- Levin, J.**, Essays in Company Valuation.

Ljunggren, U., Styrning av grundskolan i Stockholms stad innan stadsdelsnämndsreformen.

Mattsson, S., Från stat till marknad – effekter på nätverksrelationer vid en bolagiseringsreform.

Nyberg, A., Innovation in Distribution Channels – An Evolutionary Approach.

Olsson, P., Studies in Company Valuation.

Reneby, J., Pricing Corporate Securities.

Roszbach, K., Essays on Banking Credit and Interest Rates.

Runsten, M., The Association Between Accounting Information and Stock Prices. Model development and empirical tests based on Swedish Data.

Segendorff, B., Essays on Bargaining and Delegation.

Sjöberg, L., Bagozzi, R., Ingvar, D.H., Will and Economic Behavior.

Sjögren, A., Perspectives on Human Capital: Economic Growth, Occupational Choice and Intergenerational Mobility.

Studier i kostnadsintäktsanalys, red Jennergren, P.

Söderholm, J., Målstyrning av decentraliserade organisationer. Styrning mot finansiella och icke-finansiella mål.

Thorburn, K., Cash Auction Bankruptcy and Corporate Restructuring

Wijkström, F., Different Faces of Civil Society.

Zethraeus, N., Essays on Economic Evaluation in Health Care. Evaluation of Hormone Replacement Therapy and Uncertainty in Economic Evaluations.

1997

Alexius, A., Essays on Exchange Rates, Prices and Interest Rates.

Andersson, B., Essays on the Swedish Electricity Market.

Berggren, N., Essays in Constitutional Economics.

Ericsson, J., Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.

Charpentier, C., Budgeteringens roller, aktörer och effekter. En studie av budgetprocesserna i en offentlig organisation.

De Geer, H., Silfverberg, G., Citizens' Trust and Authorities' Choices, A Report from The Fourth International Conference on Ethics in the Public Service, Stockholm June 15-18, 1994.

Friberg, R., Prices, Profits and Exchange Rates.

Från optionsprissättning till konkurslagstiftning. red. Bergström, C., Björk, T.

Hagerud, G.E., A New Non-Linear GARCH Model.

Haksar, A., Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka.

He, C., Statistical Properties of Garch Processes.

Holmgren, M., Datorbaserat kontrollrum inom processindustrin; erfarenheter i ett tidsperspektiv.

Jennergren, P., Tutorial on the McKinsey Model for Valuation of Companies.

Lagerlöf, J., Essays on Political Economy, Information, and Welfare.

Lange, F., Wahlund, R., Planerade och oplanerade köp - Konsumenternas planering och köp av dagligvaror.

Löthgren, M., Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.

Nilsson, H.E., Using Conceptual Data Modelling in Management Accounting: A Case Study.

Sjöberg, L., Ramsberg, J., En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.

Säfvenblad, P., Price Formation in Multi-Asset Securities Markets.

Sällström, S., On the Dynamics of Price Quality.
Södergren, B., På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.
Tambour, M., Essays on Performance Measurement in Health Care.
Thorén, B., Berg-Surwee, U., Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.
Zhang Gang., Chinese Rural Enterprises Between Plan and Market.
Åhlström, P., Sequences in the Process of Adopting Lean Production.
Åkesson, G., Företagsledning i strategiskt vakuum. Om aktörer och förändringsprocesser.
Åsbrink, S., Nonlinearities and Regime Shifts in Financial Time Series.

1996

Advancing your Business. People and Information Systems in Concert.
red. Lundeberg, M., Sundgren, B.
Att föra verksamheten framåt. Människor och informationssystem i samverkan.
red. Lundeberg, M., Sundgren, B.
Andersson, P., Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.
Andersson, P., The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.
Asplund, M., Essays in Industrial Economics.
Delmar, F., Entrepreneurial Behavior & Business Performance.
Edlund, L., The Marriage Market: How Do You Compare?
Gunnarsson, J., Three Studies of Financial Behavior. Research Report.
Hedborg, A., Studies of Framing, Judgment and Choice.
Holgersson, C., Höök, P., Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.
Holmberg, C., Stores and Consumers - Two Perspectives on Food Purchasing.
Håkansson, P., Wahlund, R., Varumärken. Från teori till praktik.
Karlsson, A., The Family Business as an Heirloom. Research Report.
Linghag, S., Man är handelsstudent. Research Report.
Molin, J., Essays on Corporate Finance and Governance.
Mägi, A., The French Food Retailing Industry - A Descriptive Study.
Mölleryd, B., Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk mobiltelefoni. Research Report.
Nielsen, S., Omkostningskalkulation for avancerede produktions-omgivelser - en sammenligning af stokastiske og deterministiske omkostningskalkulationsmodeller.
Normark, P., Danielsson, L., Larsson, A., Lundblad, P., Kooperativa nyckeltal. Research Report.
Sandin, R., Heterogeneity in Oligopoly: Theories and Tests.
Sandén-Håkansson, U., Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.
Stein, J., Söderlund, M., Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildningskvalitet. En empirisk studie av civil-ekonomer. Research Report.
Strömberg, P., Thorburn, K., An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.
Söderlund, M., Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995. Research Report.

Ulfsson, U., Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyoghurt 1982-1994.
Westelius, A., A Study of Patterns of Communication in Management Accounting and Control Projects.
Wijkström, F., Den svenska ideella sektorn och pengarna. Research Report.
Örtendahl, M., Health and Time - A Problem of Discounting.

1995

Becker, T., Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption.
Blomberg, J., Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Styrelsearbete i koncerner
Brännström, T., Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.
Ekonomisk politik i omvandling. red. Jonung, L.
Gunnarsson, J., Wahlund, R., Hushållens finansiella strategier. En explorativ studie.
Höök, P., Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.
Levin, J., Olsson, P., Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.
Mägi, A., Customer Satisfaction in a Store Performance Framework. Research Report.
Persson, P-G., Modeling the Impact of Sales Promotion on Store Profits.
Roman, L., Institutions in Transition. A Study of Vietnamese Banking.
Sandberg, J., Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.
Sandberg, J., How Do We Justify Knowledge Produced by Interpretative Approaches? Research Report.
Schuster, W., Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering.
Söderberg, K., Farmartjänsten - Ny Kooperation inom lantbruket. Research Report.
Söderqvist, T., Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation.
Thorén, B., Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor.