Statistical inference for time-inhomogeneous volatility models

Čížek, Pavel *
Humboldt-Universität zu Berlin,
Spandauerstrasse 1,
D-10178 Berlin, Germany
cizek@wiwi.hu-berlin.de

Härdle, Wolfgang †
Humboldt-Universität zu Berlin,
Spandauerstrasse 1,
D-10178 Berlin, Germany
haerdle@wiwi.hu-berlin.de

Spokoiny, Vladimir
Weierstrass-Institute,
Mohrenstr. 39,
10117 Berlin, Germany
spokoiny@wias-berlin.de

July 2, 2003

Keywords: stochastic volatility model, adaptive estimation, local homogeneity

AMS 2000 Subject Classification: 62M10; Secondary 62P20

Abstract

This paper offers a new method for estimation (filtering) and forecasting of the volatility of financial time series when the stationarity assumption is violated. We consider a varying coefficient parametric (GARCH) model whose coefficient may arbitrary vary with time. This includes global parametric, smooth transition and change-point models as special cases. The method is based on pointwise adaptive selection of the largest interval of homogeneity with the given right end-point. We construct a *locally adaptive volatility estimate (LAVE)* which can perform this task and investigate it both from the theoretical point of view and through Monte Carlo simulations. Finally the LAVE procedure is applied to a data set of nine exchange rates and a comparison with a standard GARCH model is also provided.

*Financial support by the Deutsche Forschungsgemeinschaft, SFB 373 "Simulation and Quantification of Economic Processes" at Humboldt University, is gratefully acknowledged.

†Financial support by the Deutsche Forschungsgemeinschaft, SFB 373 "Simulation and Quantification of Economic Processes" at Humboldt University, is gratefully acknowledged.

1 Introduction

The aim of this paper is to offer a new perspective for the estimation (filtering) and forecasting of the volatility of financial asset returns such as stocks and exchange rate returns.

A remarkable amount of statistical research is devoted to financial time series, and in particular to the volatility of asset returns, where the term volatility indicates a measure of dispersion, usually the variance or the standard deviation. The interest in this topic is motivated by the needs of the financial industry, which regards the volatility as one of the main reference number for risk management and derivative pricing.

Let S_t denote the observed asset process, then the corresponding (log) returns $R_t = \log(S_t/S_{t-1})$ are usually modelled by the following conditional heteroscedasticity equation

$$R_t = \sigma_t \xi_t$$

where ξ_t are standard Gaussian independent innovations and σ_t is a time-varying *volatility* coefficient. It is often assumed that σ_t is measurable w.r.t. the σ-field generated by the preceding returns $R_1, \ldots, R_{t-1}$. For modeling this volatility process parametric assumptions are usually used. The main model classes are the ARCH and GARCH family (**?**), and the stochastic volatility (**?**). A great amount of papers has followed the first publications on this topic, and the original models have been extended in order to provide better explanations. For example models which take into account asymmetries in volatility have been proposed, such as EGARCH (**?**), QGARCH (**?**) and GJR (**?**); furthermore the research on integrated processes has produced integrated (**?**) and fractal integrated versions of the GARCH model.

The availability of very large samples of financial data has given the possibility of constructing models which display quite complicated parameterizations in order to explain all the observed stylized facts. Obviously those models rely on the assumption that the parametric structure of the process remains constant through the whole sample. This is a nontrivial and possibly dangerous assumption in particular as far as forecasting is concerned (**?**).

Financial models with time dependent parameters have been considered by a number of authors. We refer to the overview by Fan (2003) for a comprehensive list of references.

Our modeling approach is similar to varying coefficient idea from **?**, see also **?** and **?**. **?** discussed applications of this method to stock price volatility modeling. The proposed procedure is based on the assumption that the model parameters smoothly vary with time and can be locally approximated by a linear function of time. This approach has drawback of not allowing to incorporate structural breaks in the model.

Change-point modeling with applications to financial time series was considered in **?**. However, their approach involves a special step of identifying all pieces of stationarity which requires certain prior information like the number of changes or the minimal time interval between two neighbor changes. In addition, the method is designed towards the piecewise homogeneous models and its applications to the models with smoothly varying parameters is questionable. **?** applied non-Gaussian random walk modeling with heavy tails as the prior for the piecewise constant mean for one-step-ahead prediction of nonstationary time series. However, the mentioned approaches require some essential amount of prior information about the frequency of change-points and their size.

This paper proposes another approach which allows to proceed in a unified way with the change point and varying coefficients models. The idea of the method goes back to the pointwise adaptive procedure from Spokoiny (1998) and it extends Mercurio and Spokoiny (2003) and Härdle, Herwartz and Spokoiny (2003) from the local constant to local parametric modeling. However, the assumption of a constant volatility is too restrictive and it is not fulfilled for the most applications. Therefore, the local constant modeling can be applied only for a short time intervals and it cannot be used for forecasting over a large time horizon. In opposite, parametric models like GARCH mimic the most of stylized facts of financial time series and allow to reasonably fit the data for many practical situations. Possibility for time dependence of model parameters offers much more flexibility in modeling the real life financial time series, and parametric models with time depend coefficients build a very useful class which includes the global parametric models as a special case. However, an extension of the procedure from Mercurio and Spokoiny (2003) to the local parametric situation is a nontrivial problem and it requires to develop completely new tools. The main idea of the method is to describe in a data-driven way, for every time point t, the largest interval of homogeneity in which the parametric assumptions fits well the observations. The proposed procedure involves two basic steps: local (interval) parameter estimation and testing the parametric hypothesis against a change point alternative. The procedure starts from a small interval. At every step this interval is increased to the left and tested on homogeneity. The procedure is terminated if a change-point is detected or if the largest possible interval is achieved. The method can be directly applied to forecasting by modeling from the latest estimated model.

The rest of the paper is organized as follows. In the next section we discuss the two main ingredients of the method: local parameter estimation and local test of homogeneity. Section 3 introduces the adaptive estimation procedure and discusses the choice of its parameters and some numerical issues. Theoretical properties of the method are discussed in the general situation and for a change-point model in Section 4. A simulation study illustrates the performances of the new methodology with respect to the change-point model in Section 5 where some applications to real data are also presented.

2 Parametric volatility modeling

Let S_t be an observed asset process in discrete time, $t = 1, 2, \ldots, t^*$ and R_t are the corresponding returns: $R_t = \log(S_t/S_{t-1})$. We model this process via the *conditional heteroscedasticity* assumption

$$R_t = \sigma_t \xi_t \,, \tag{2.1}$$

where ξ_t, $t \geq 1$, is a sequence of independent standard Gaussian random variables and σ_t is the *volatility* process which is in general a predictable random process, that is, $\sigma_t \sim \mathcal{F}_{t-1}$ with $\mathcal{F}_{t-1} = \sigma(R_1, \ldots, R_{t-1})$ (σ-field generated by the first $t-1$ observations).

A standard way of modelling the volatility process is based on one or another parametric assumption. A large class of parametric models is built of the so called ARCH and GARCH models. An ARCH (p) models is described by the equation

$$\sigma_t^2 = \omega + \alpha_1 R_{t-1}^2 + \ldots + \alpha_p R_{t-p}^2.$$

The main disadvantage of this model is that the order p of this autoregressive type equation should be large to be able to capture the main stylized facts of the financial time series.

A GARCH (p,q) model is described by an ARMA-type equation

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{2.2}$$

The main attractive feature of this model is that even with very few coefficients (two or three) one can model the most of stylized facts of the financial time series like e.g. volatility clustering or excessive curtosis. The linear GARCH equation might lead to negative values for the expression in the right hand side of (2.2) with the estimated GARCH-coefficients. As an alternative, one can use an EGARCH model defined as

$$\log \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{q} \beta_j \log \sigma_{t-j}^2. \tag{2.3}$$

where the Y_t's are some transformations of the R_t, e.g. $Y_t = (\log R_t^2 + \epsilon)$ or $Y_t = (\log\log e^{-R_t} + e + \epsilon)$ for some $\epsilon \ge 0$. All such models can be extended by adding one more volatility noise component ζ_t in the right hand side of (2.2) or (2.3) leading to *stochastic volatility models*. Asymmetric behavior of the volatility is usually modelled by the thresholding TGARCH-models. This list can be substantially extended.

All these models can be put into a common class of *generalized linear volatility models*. Those models can be described by

$$R_t \sim \mathcal{N}(0, \sigma_t^2), \qquad \sigma_t^2 = g(X_t), \qquad X_t = \omega + \sum_{i=1}^{p} \alpha_i h(R_{t-i}) + \sum_{j=1}^{q} \beta_j X_{t-j} \tag{2.4}$$

where g and h are some prescribed functions, and (X_t) is a latent process (structural variable) that models volatility coefficient σ_t via the transform g. For instance, in the GARCH-case we have $g(u) = u$ and $h(r) = r^2$. For the EGARCH-situation, $g(u) = e^u$ and $h(r) = \log(r^2 + \epsilon)$ or $Y_t = (\log\log e^{-R_t} + e + \epsilon)$. In what follows we denote $Y_t = h(R_t)$ and write the model equation in the linear form $X_t = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{q} \beta_j X_{t-j}$.

2.1 Parameter estimation

This section discusses the method of parameter estimation for the model (2.4) using the observations R_t from some time-interval $I = [t_0, t_1]$.

The volatility process σ_t^2 is modelled by the equation $\sigma_t^2 = g(X_t)$ where the structural process X_t fulfills the linear constraints from (2.4). Note that these linear constraints do not uniquely determined the process X_t. Apart the vector of parameters $\boldsymbol{\gamma} = (\omega, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$, one has to specify the boundary (initial) value

$\boldsymbol{\eta} = (X_{t_0-1}, \ldots, X_{t_0-q})$. However, the dependence on this parameter in typical situation is rather small and we simply set it to some prescribed value. We therefore use the notation $X_t = X_t(\boldsymbol{\theta})$. The structural linear equation for this process can be written as

$$X_t(\boldsymbol{\theta}) = \Psi_t(\boldsymbol{\theta})\boldsymbol{\gamma} = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{q} \beta_j X_{t-j}(\boldsymbol{\theta}), \qquad t = t_0, \ldots, t_1 \tag{2.5}$$

with the boundary (initial) conditions $X_{t_0-j}(\boldsymbol{\theta}) = \eta_j$ for $j = 1, \ldots, q$. We used here the notation $\Psi_t(\boldsymbol{\theta}) = \bigl(1, Y_{t-1}, \ldots, Y_{t-p}, X_{t-1}(\boldsymbol{\theta}), \ldots, X_{t-q}(\boldsymbol{\theta})\bigr)$. Using this linear equation one can recursively compute the process $X_t(\boldsymbol{\theta})$ starting from the initial values $\boldsymbol{\eta}$. Similarly we can compute the derivatives $\nabla X_t = \partial X_t(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$. Namely it holds

$$\frac{\partial X_t(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} = \Psi_t^{\top}(\boldsymbol{\theta}) + \sum_{j=1}^{q} \beta_j \frac{\partial X_{t-j}(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}}, \qquad t = t_0, \ldots, t_1 \tag{2.6}$$

with the boundary conditions $\partial X_{t_0-j}(\boldsymbol{\theta})/\partial \boldsymbol{\gamma} = 0$ for $j = 1, \ldots, q$.

For estimating the parameter $\boldsymbol{\theta}$, we apply the quasi maximum likelihood approach assuming the standard normal innovations $\{\varepsilon_t\}$. The quasi log-likelihood for the model (2.4) can be represented in the form

$$L_I(\boldsymbol{\theta}) = \sum_{t \in I} \ell(R_t, g(X_t(\boldsymbol{\theta})))$$

with $\ell(r, u) = -0.5\bigl(\log(u) + r_t^2/u\bigr)$. We define the quasi maximum likelihood estimate $\widehat{\boldsymbol{\theta}}_I$ of the parameter $\boldsymbol{\theta}$ by maximizing $L_I(\boldsymbol{\theta})$:

$$\widehat{\boldsymbol{\theta}}_I = \operatorname*{argsup}_{\boldsymbol{\theta}} L_I(\boldsymbol{\theta}) = \operatorname*{argsup}_{\boldsymbol{\theta}} \sum_{t \in I} \ell(R_t, g(X_t(\boldsymbol{\theta}))). \tag{2.7}$$

By $\widehat{L}_I$ we denote the corresponding maximum

$$\widehat{L}_I = L_I(\widehat{\boldsymbol{\theta}}_I) = \sup_{\boldsymbol{\theta}} \sum_{t \in I} \ell(R_t, g(X_t(\boldsymbol{\theta}))) \tag{2.8}$$

The MLE $\widehat{\boldsymbol{\theta}}_I$ fulfills the estimating equation $\partial L_I(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = 0$ leading to

$$\sum_{t \in I} \Bigl(R_t^2 - g(X_t(\boldsymbol{\theta}))\Bigr) V(X_t(\boldsymbol{\theta})) \, \frac{\partial X_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0 \tag{2.9}$$

where $V(X_t(\boldsymbol{\theta})) = g'(X_t(\boldsymbol{\theta}))/g^2(X_t(\boldsymbol{\theta}))$.

A closed form solution of the optimization problem (2.9) is known only for very special situations, for instance, if $q = 0$. In the general situation, an iterative Newton-Raphson iterative procedure can be applied. Let some initial values $\boldsymbol{\theta}^{(0)}$ be fixed. The idea is to recurrently update this initial value using the linear expansion of the estimating equation w.r.t. the parameter $\boldsymbol{\theta}$. At the kth step of the Newton-Raphson algorithm for $k = 1, 2, \ldots$, for a given previous step estimate $\boldsymbol{\theta}^{(k-1)}$, one has to compute the process $X_t^{(k)} = X_t(\boldsymbol{\theta}^{(k-1)})$ by (2.5) and its gradient $\nabla X_t^{(k)} = \nabla X_t(\boldsymbol{\theta}^{(k-1)})$ by (2.6). Next, the approximation $g(X_t(\boldsymbol{\theta})) \approx g(X_t^{(k)}) + g'(X_t^{(k)})\nabla X_t^{(k)}(\boldsymbol{\theta} - \boldsymbol{\theta}^{(k-1)})$ leads to the following approximating estimating equation

$$\sum_{t \in I} \nabla X_t^{(k)} \bigl(\nabla X_t^{(k)}\bigr)^\top H(X_t^{(k)})(\boldsymbol{\theta} - \boldsymbol{\theta}^{(k-1)}) - \sum_{t \in I} \nabla X_t^{(k)} \Bigl(R_t^2 - g(X_t^{(k)})\Bigr) V(X_t^{(k)}) = 0$$

with $H(u) = g'(u)V(u) = |g'(u)|^2/g^2(u)$, and define the update $\boldsymbol{\theta}^{(k)}$ as $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + (B^{(k)})^{-1}S^{(k)}$ with

$$\begin{aligned} S^{(k)} &= \sum_{t \in I} \nabla X_t^{(k)} \left(R_s^2 - g(X_s^{(k)})\right) V(X_s^{(k)}), \\ B^{(k)} &= \sum_{t \in I} \nabla X_t^{(k)} \left(\nabla X_t^{(k)}\right)^\top H(X_s^{(k)})\,. \end{aligned}$$

The update $\boldsymbol{\theta}^{(k)}$ can be interpreted as gradient descent in direction of the estimated gradient of the log-likelihood. It is recommended to check that the this update really improves the likelihood, that is, $L(\boldsymbol{\theta}^{(k)}) < L(\boldsymbol{\theta}^{(k-1)})$. If this inequality does not hold, the step in the gradient direction should be taken smaller, $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} + \rho(B^{(k)})^{-1}S^{(k)}$ for some $\rho < 1$, e.g. $\rho = 1/2$.

2.2 Test of homogeneity against a change-point alternative

Let $I = [t_0, t_1]$ be a time interval. Here we discuss a test of the null hypothesis that the observed returns R_t follow within I the parametric model (2.4). We consider a likelihood ratio test against a change-point alternative within I.

The null hypothesis for I means that the observations R_t for $t \in I$ follow the parametric model (2.4) with the parameter $\boldsymbol{\theta}$. This yields the parametric MLE $\widehat{\boldsymbol{\theta}}_I$ and the corresponding fitted log-likelihood $\widehat{L}_I$ due to (2.7) and (2.8).

Now we describe the change-point alternative. Let $\mathcal{T}(I)$ be a set of internal points within I. We consider the change-point alternative with location of the change at points $\mathcal{T}(I)$. Every point $\tau \in \mathcal{T}(I)$ splits the interval I in two subintervals J and $J^c = I \setminus J$. A change-point alternative with location at $\tau \in \mathcal{T}(I)$ means that $\boldsymbol{\theta}_t = \boldsymbol{\theta}_1$ for $t \in J$ and $\boldsymbol{\theta}_t = \boldsymbol{\theta}_2$ for $t \in J^c$ with two different values $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. Under such an alternative, the observations $(R_t, t \in I)$ are described by the log-likelihood $L_J(\boldsymbol{\theta}_1) + L_{J^c}(\boldsymbol{\theta}_2)$.

The likelihood ratio test statistic is of the form

$$\begin{aligned} T_{I,\tau} &= \max_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2} \{L_J(\boldsymbol{\theta}_1) + L_{J^c}(\boldsymbol{\theta}_2)\} - \max_{\boldsymbol{\theta}} L_I(\boldsymbol{\theta}) \\ &= \widehat{L}_J + \widehat{L}_{J^c} - \widehat{L}_I\,. \end{aligned}$$

The change-point test for the interval I is defined as maximum of such defined test statistics over $\tau \in \mathcal{T}(I)$:

$$T_I = \max_{\tau \in \mathcal{T}(I)} T_{I,\tau}\,.$$

The estimated location of the change point is given by the point of maximum:

$$t_{\text{cp}} = \operatorname*{argmax}_{\tau \in \mathcal{T}(I)} T_{I,\tau}.$$

The local likelihood analysis is based on the nonasymptotic version of the Wilks theorem: under homogeneous parametric assumption $2\widehat{L}_I$ is approximately χ^2-distributed, see e.g. ???.

3 Adaptive estimation under local time-homogeneity

An obvious feature of the model (2.4) is that the parametric structure of the process is assumed constant over the whole sample. This assumption has drawback of not allowing to incorporate changes and structural breaks in the model. A natural generalization is given by models whose coefficient may vary with time. Following to Fan et a. (???) we consider in this paper the following *varying coefficient model*

$$R_t \sim \mathcal{N}(0, \sigma_t^2), \quad \sigma_t^2 = g(X_t), \quad X_t = \omega(t) + \sum_{i=1}^{p} \alpha_i(t) Y_{t-i} + \sum_{j=1}^{q} \beta_j(t) X_{t-j} \tag{3.1}$$

where $\alpha_i(t)$ and $\beta_j(t)$ are some functions of time which have to be estimated from the observations R_t. Of course, this is only possible under some additional assumptions on these functions. We mention two typical examples of such assumptions: (1) varying coefficients are smooth functions of time; (2) they are piecewise constant (piecewise smooth) functions.

Our approach is based on the following structural assumption: for every point t there exists an *interval of homogeneity* $I = I(t)$ in which the model parameters are nearly constant and the data R_s, $s \in I$, follow the local parametric model with a parameter $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$ corresponding to the right end-point of I. The main question under this approach is how the local interval of homogeneity is defined. Fan et al. (???) base this choice on the smoothness properties of the function $\boldsymbol{\theta}(t)$ while Mikosch and Starica (2001) suggest to first identify all the changes and then to apply the parametric model between every two neighbor changes.

We apply here the pointwise adaptive choice, that is, for every point t the corresponding interval of homogeneity is selected independently. This allows to naturally incorporate the both mentioned special cases in a unified way.

The idea of the method is to find the closest from the left to t change point t_{cp} and to define the interval of homogeneity as $[t_{\mathrm{cp}}, t]$. The search of a change point is done by successive testing interval I on homogeneity against a change-point alternative. If the hypothesis of homogeneity is not rejected then a larger interval is taken and tested again. The main difference between this method and that of Mikosch and Starica (2001) is that we do not aim to detect all the change points for the given time series but focus on the local change point analysis near the point of estimation. This particularly allows restricting to local alternatives with only one change.

3.1 Adaptive choice of the interval of homogeneity

This section proposes a *pointwise adaptive* method of selecting the interval of homogeneity for the conditional heteroscedastic model (2.1). Let t^* be the point of interest and let $\mathcal{T}$ be the ordered set of time points $1 = t_1 < t_2 < \ldots < t_{m^*} = t^*$. Specific proposals for the choice of this set will be discussed in Section 3.2. We successively test every interval $I_m = [t_m, t^*]$ on homogeneity where m decreases from $m^* - 2$ to one. Rejection of this hypothesis means a detection of a change point t_{cp} which will be taken as the left end-point of the interval of homogeneity.

The procedure reads as follows.

Initialization Set $m = m^* - 2$.

Iteration Select the interval $I = [t_m, t^*]$.

Testing homogeneity Test the hypothesis of homogeneity within I against a change-point alternative.

Loop If a change point is detected with location at $t_{\mathrm{cp}} \in I$, then set $\widehat{I} = [t_{\mathrm{cp}}, t^*]$. Otherwise, decrease m by one, $m := m - 1$ and continue with the iteration step. If $m = 0$ is achieved, then set $\widehat{I} = [1, t^*]$.

The estimate $\widehat{\boldsymbol{\theta}}(t^*)$ of $\boldsymbol{\theta}(t^*)$ is defined as the MLE for the selected interval $\widehat{I}$, that is, $\widehat{\boldsymbol{\theta}}(t^*) = \widehat{\boldsymbol{\theta}}_{\widehat{I}}$.

3.2 Implementation and numerical issues

To run the proposed procedure, a number of parameters have to be specified. Here we briefly discuss their choice and some questions of numerical implementation.

3.2.1 Set of intervals

An important element of the method is the set $\mathcal{I}$ of intervals candidates and for every such interval I the set of testing intervals J or equivalently, the set of break points $\mathcal{T}(I)$. Note that all these sets depends on the reference end-point t^*. Similarly to HHS and Mercurio and Spokoiny (2003) we suggest to fix some arithmetic grid $\mathcal{T} = \{t_m, m = 1, 2, \ldots\}$ with the step m_0 and to run the procedure only for the points t^* which belong to this set $\mathcal{T}$. The intervals I and the testing intervals J can also be taken with the end-points at $\mathcal{T}$. Moreover, it is recommended to consider only those testing intervals J for which the relative length $|J|/|I|$ is bounded away from zero and one, that is, $\rho \le |J|/|I| \le 1 - \rho$ for some $\rho > 0$. e.g. $\rho = 1/4$. The reason is that the likelihood ratio test statistic behaves in a very irregular way when $|J|/|I|$ is small or close to one, and the intervals J corresponding to small values of $|J|/|I|$ have been already checked on change points at the previous steps of the algorithm.

3.2.2 Choice of the parameters p, q and transforms g and h

For practical applications of the method one has to specify the transforms g, h and values p, q. ¿From numerical reasons we suggest the EGARCH model corresponding to $g(u) = e^u$ and $h(u) = \log(R_t^2 + \epsilon)$ or $h(u) = \log\log(e^{-R_t} + e + \epsilon)$. The latter proposal allows to mimic the asymmetric dependence of the volatility on the past price fluctuations.

It is well known that even GARCH(1,1) or EGARCH(1,1) model delivers a reasonable fit for many financial time series and allows to mimic the most of stylized facts of the financial time series. Our varying coefficient model is much more flexible due to possibility for the model parameter to depend on time. So, it is reasonable to apply $p, q = 1$. Such a choice allows to effectively estimate the model parameter $\boldsymbol{\theta}$ even from a small time interval of observations and it is flexible enough for modeling changes in time.

3.2.3 Numerical issues

The procedure can be easily carried out for a every point $t^* \in \mathcal{T}$ as soon as the values $\widehat{L}_J$ are available for every intervals J with the end-points at $\mathcal{T}$. Then the algorithm can be run by a few arithmetic operations. Indeed, every test of homogeneity for the interval $I = [t_{m'}, t_m]$ for $m' < m$ requires only about $C|m - m'|$ operations for some $C \leq 3$.

In many practical applications the procedure has to be carried through in online mode, that is, it is run once again as soon as next time point t_{m^*} from $\mathcal{T}$ is achieved. Since the values L_J for all intervals J with the end-points at $t_m < t_{m^*}$ have been computed previously, to perform the procedure at t_{m^*} the only values $\widehat{L}_I$ for the intervals $I = [t_m, t_{m^*}]$ should be computed for all $t_m \in \mathcal{T}$, $m < m^*$. We recommend to compute the estimate $\widehat{\boldsymbol{\theta}}_I$ for every such I by the Newton-Raphson algorithm as described in Section 2.1 using the previously computed estimate $\widehat{\boldsymbol{\theta}}_{I'}$ for the smaller interval $I' = [t_m, t_{m^*-1}]$ as starting value $\boldsymbol{\theta}_I^{(0)}$.

3.2.4 Choice of the critical value for the test of homogeneity

The proposed adaptive choice of the interval of homogeneity is based on the multiple testing procedure. A natural requirement to the critical value λ is that the first type error probability (which means the rejection of the hypothesis when it is true) is bounded above by some prescribed value. A standard way of describing the proper critical value relies to asymptotic distribution of the test statistics. Asymptotic properties of the likelihood ratio test statistic T_I from Section 2.2 is well known, see e.g. Darchovskii and Brodskii (???). In particular, the limit distribution is parameter free and does not depend on $\boldsymbol{\theta}$. Unfortunately, the most of established results is applicable only if the length of the interval I and all the considered test-intervals J tends to infinity. The proposed procedure has to be applied to a relatively small intervals I. To justify a practically relevant choice of the critical value, we have to establish a non-asymptotic bound for the deviation probability of the test statistic T_I over a threshold λ.

To be specific, we assume that $I = [t_m, t_{m^*}]$ and denote by N_I the number of all testing intervals $J = [t_{m_1}, t_{m_2}]$ with $m \leq m_1 < m_2 \leq m^*$.

Theorem 3.1. *Let (A1)–(A?) hold true. There exists a constant C depending on the constants from these conditions such that it holds for every $\alpha > 0$*

$$\boldsymbol{P}\left(T_I > C \log(N_I/\alpha)\right) \leq \alpha.$$

Table 1: GARCH(1,1)-models used for critical values simulations, the corresponding simulated 90% quantiles $q_{0.9}(|I|)$ of the test statistics T_I at different sample sizes $|I|$ and estimated C_r for every model.

Model	ω	α	β	$q_{0.9}(50)$	$q_{0.9}(100)$	$q_{0.9}(150)$	C_r
1	0.0001	0.05	0.57	2.44	3.72	4.43	0.52
2	0.0001	0.07	0.85	4.30	6.37	7.70	0.87
3	0.0001	0.15	0.82	6.41	9.37	10.57	1.24
4	0.0010	0.03	0.95	7.05	9.40	10.53	1.42
5	0.0010	0.06	0.82	6.46	7.95	9.02	1.19
6	0.0010	0.08	0.90	7.80	11.13	12.90	1.58
7	0.0010	0.10	0.89	8.77	11.55	14.32	1.69
8	0.0010	0.13	0.77	7.26	9.26	10.13	1.37
9	0.0100	0.04	0.90	9.23	10.67	12.10	1.70
10	0.1000	0.05	0.90	11.41	13.22	14.64	1.97

This statement ensures that a logarithmic threshold $\lambda = C \log(N_I/\alpha)$ provides a prescribed test level. This suggests to select the critical value λ proportional to $\log(N_I/\alpha)$. The factor C can be chosen by Monte-Carlo simulation from some parametric models with parameters fitted from the considered time series.

4 Theoretic properties

In this section we collect some results describing the quality of the proposed adaptive procedure.

First we present some technical conditions required for establishing the result.

(A1)

(A2)

(A3)

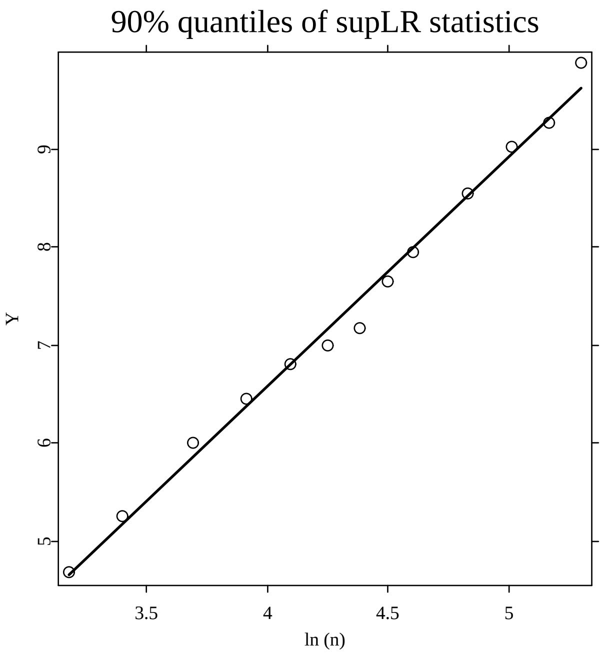
5 Simulation study and aplications

In this section, we first find, by means of Monte Carlo simulations, the finite-sample critical values for the test of homogeneity and then study the performance of the proposed adaptive estimation procedure discussed in Section 3 in simulated samples and real data for GARCH(1,1) models (see equation (2.2)).

5.1 Finite-sample critical values for the test of homogeneity

A practical application of the proposed adaptive procedure requires critical values for the test of local homogeneity of a time series. Theorem 3.1 establishes an upper bound for

Figure 1: 90% quantiles of the test of homogeneity for GARCH(1,1) model.



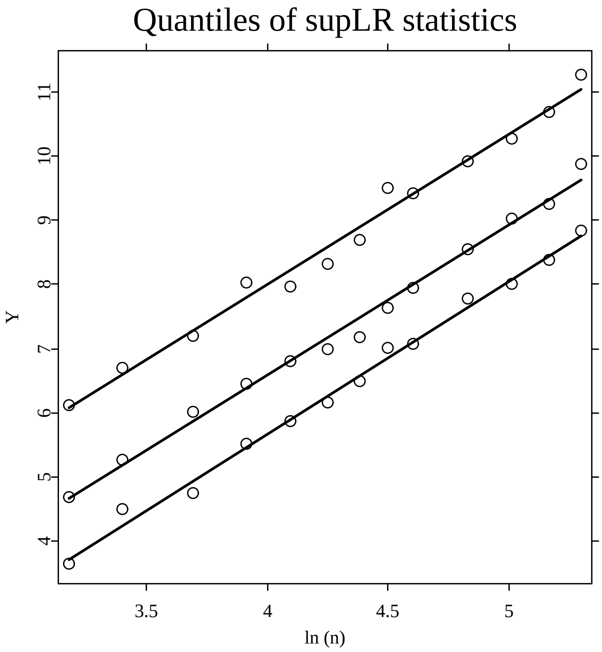
these critical values, which is proportional to $\log(N_I/\alpha)$ and some unknown constant C. To find the magnitude of C, we simulated 1000 samples for several time-homogeneous GARCH(1,1) models with parameters typical for high-frequency financial time series: volatility level $\omega(t)$ very close to zero, ARCH parameter $\alpha(t) \in (0, 0.25)$, and GARCH parameter $\beta(t) \in (0.75, 1)$. Such values of ARCH and GARCH coefficients were found, for example, by GARCH(1,1) modelling of the returns of the German stock index DAX within periods of 3 to 6 monthes in years 1990–2002. The list of simulated models is in Table 1.

Using the models in Table 1, we simulated 1000 times samples of sizes $|I|$ ranging from 20 to 200 for each model (under the null hypothesis of time-homogeneity), performed the test of homogeneity for each sample, and analyzed the values of the resulting test statistics T_I. The relative number of tested subintervals $|N_I|/|I|$ was kept in all simulation equal to 0.5. Given a significance level α (e.g., 0.10), the $1-\alpha$ quantile of the test statistics (at a given sample size $|I|$) can be computed; see Table 1 for selected model-specific results.

Based on the simulated quantiles, there are two important empirical observations one can make. The first observation follows from the results of simulated 90% quantiles for model 5 depicted in Figure 1. There is a least-squares regression line fitting the simulated critical values to $\ln(N_I/\alpha)$. Apparently, the simulated values follow very well a linear trend with respect to the logarithm of sample size $\log(|I|)$.

Second, plotting different quantiles, for example 85%, 90%, and 95% ones, of simulated values of the test statistics T_I shows that the behavior is very much the same for all of them—all exhibit a linear patterns wrt. $\log(|I|)$ and $\ln(N_i/\alpha)$ and they have almost

Figure 2: 85%, 90%, and 95% quantiles of the test of homogeneity for GARCH(1,1) model.



the same slope (see Figure 2). We found larger, but still relatively small variation in slopes of these regression lines also across different models from Table 1.

These findings offer two ways to proceed in constructing the critical values for the test of interval homogeneity. First, one can follow Theorem 3.1, take all simulations of all models, compute $1-\alpha$ quantiles $q_{1-\alpha}(|I|)$ of the test statistics T_I and estimate regressions $q_{1-\alpha}(|I|) = C_r \log(N_I/\alpha) + \varepsilon$. Then the constant C from Theorem 3.1 should be the maximum over all estimated C_r. Using GARCH(1,1) with models 1–10, the value of C_r never exceeded $C = 2$ and rarely went over $C = 1.7$ (see Table 1), which suggests that $C = 2$ could be a sufficiently large choice in this context.

Second, since the simulated quantiles of test statistics seems to increase linearly in $\log(N_I)$, it is possible to consider a linear regression model with intercept

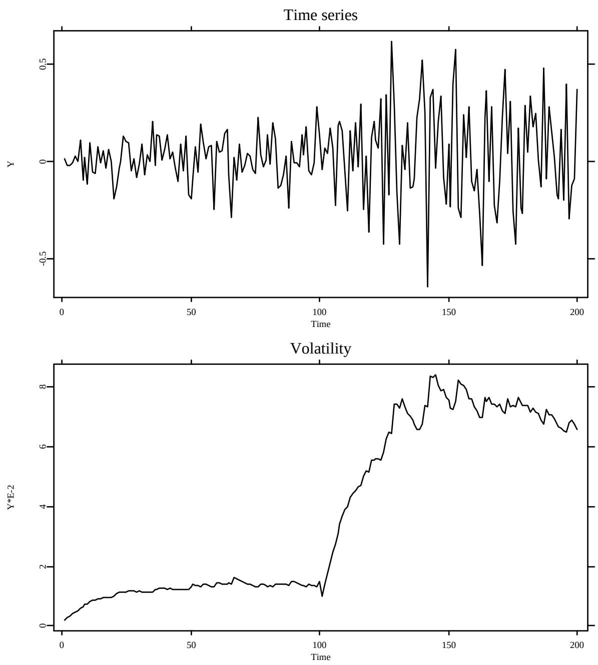
$$q_{1-\alpha}(|I|) = B + C_r \log(N_I/\alpha) + \varepsilon, \tag{5.1}$$

which describes lines in Figure 2, for instance. This way, one can obtain critical values tailored for a specific data set, which would typically be smaller than the "worst-case" scenario employing the maximum value of the constant C: as the values of $q_{1-\alpha}(|I|)$ in Table 1 document, the critical values in small samples increase as the variance increases (larger ω) and as the sum of ARCH and GARCH coefficients approaches 1 ($\alpha + \beta \to 1$). Thus given a data set, if GARCH(1,1) is estimated locally for small subsets of data to find out the level of volatility within the data, it is possible by simulating critical values for the respective models found in the data obtain smaller critical values than by taking the

Table 2: Critical values at 10% significance level for the test of homogeneity based on simulations from model 5.

Sample size $\lvert I \rvert$	30	50	100	150	200
Theorem 3.1	10.02	11.04	12.42	13.24	13.82
Linear fit (5.1)	5.26	6.45	7.94	9.02	9.87

Figure 3: GARCH(1,1) time series and its volatility under a change in constant $\omega(t)$ at $t = 101$.



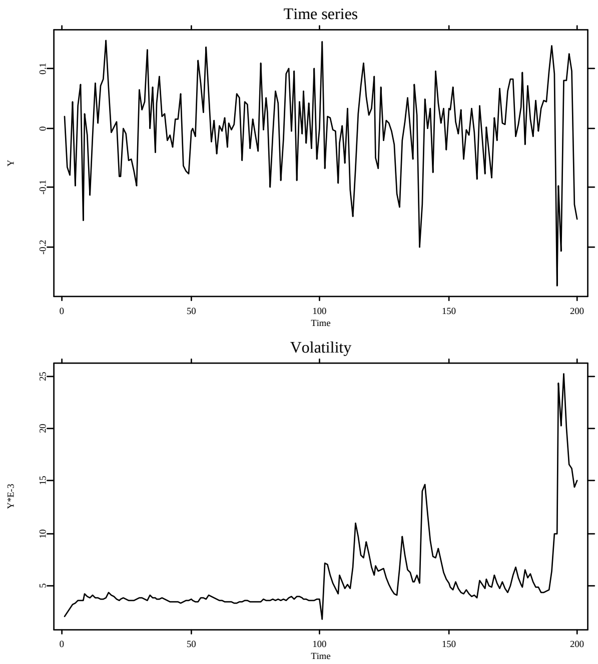
largest value of C found in any GARCH(1,1) model. The linear dependence on $\ln(N_I)$ simplifies then finding of the critical values since it suffices to simulate only for several sample sizes $|I|$. For example, fitting the regression model (5.1) for the simulated 90% quantiles of model 5 results in $B = -6.20$ and $C_r = 2.28$, which in turn produces smaller critical values (see Table 2) and thus a more sensitive procedure.

We will compare both approaches to constructing the critical values in the rest of this section.

5.2 Structural-break detection behavior

Since the GARCH models and their properties are well established in the literature, the main concern of this simulation study is to examine how well the proposed adaptive estimation method is able to detect structural breaks in time series. To find out what

Figure 4: GARCH(1,1) time series and its volatility under a change in ARCH parameter $\alpha(t)$ at $t = 101$.



kind of changes are difficult or easy to uncover, we performed three sets of simulations corresponding to a change in exactly one of the parameters of GARCH(1,1) model. In all cases, we used time series of length $T = 200$ observations, whereby a structural break in one of the parameters occurs at $t_b = 101$. That means that one model is valid for $t = 1, \ldots, 100$ and another model holds for $t = 101, \ldots, 200$. The structural breaks used throughout this section are:

1. $(\omega, \alpha, \beta) = (0.001, 0.03, 0.90)$ to $(\omega, \alpha, \beta) = (0.005, 0.03, 0.90)$ and vice versa,

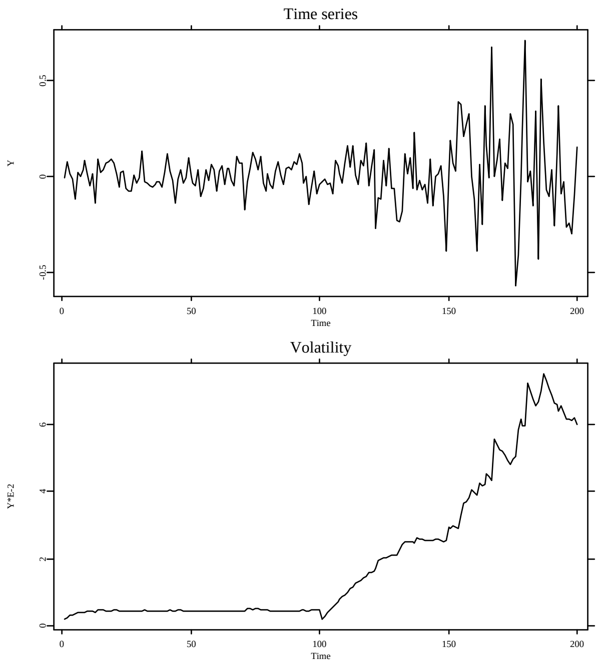
2. $(\omega, \alpha, \beta) = (0.001, 0.03, 0.75)$ to $(\omega, \alpha, \beta) = (0.005, 0.23, 0.75)$ and vice versa,

3. $(\omega, \alpha, \beta) = (0.001, 0.03, 0.75)$ to $(\omega, \alpha, \beta) = (0.005, 0.03, 0.95)$ and vice versa.

Pathes of simulated time series and their volatilities representing these changes in the parameters $\omega(t)$, $\alpha(t)$, and $\beta(t)$ in varying-coefficients GARCH(1,1) model (3.1) are depicted on Figures 3, 4, and 5, respectively.

The first task is to determine what significance level α should be used for the break detection. The aim is to use significance level, which should not exceed some reasonable level (e.g., 10% or 15%), which allows us to detect structural breaks in most cases, and finally, which does not often indicate some fictious breaks in time-homogenous data. We used for this decision the easiest-to-detect case 1 of change in the constant ω, and additionally, a GARCH(1,1) model $(\omega, \alpha, \beta) = (0.001, 0.13, 0.75)$ without any structural

Figure 5: GARCH(1,1) time series and its volatility under a change in GARCH parameter $\beta(t)$ at $t = 101$.



break. Moreover, we used at this point the critical values based on Theorem 3.1 and the maximum value of $C = 2$ determined in Section 5.1.

The results of these simulations based on 100 repetitions at three different significance levels $\alpha = 0.05, 0.10, 0.15$ are summarized in Table 3. The results are estimated only for the whole interval, that is, the adaptive procedure starts from $T = 200$ and tries to find the longest interval of stability. Each cell contains the median of the time, at which the interval of homogeneity starts, or in other words, at which a structural break was detected; the corresponding quartiles are in brackets. The results indicate that the significance level $\alpha = 0.10$ is best to use within the range $(0.05, 0.15)$: smaller value of significance level α (higher significance level of the test) makes the procedure insensitive

Table 3: Detected break times for three GARCH(1,1) models given by triplets (ω, α, β) using significance levels $\alpha = 0.05, 0.10, 0.15$.

Model 1 ($t = 1, \ldots, 100$):	$(0.001, 0.13, 0.75)$	$(0.005, 0.13, 0.75)$	$(0.001, 0.13, 0.80)$
Model 2 ($t = 101, \ldots, 200$):	$(0.005, 0.13, 0.75)$	$(0.001, 0.13, 0.75)$	$(0.001, 0.13, 0.80)$
$\alpha = 0.05$	110 (1 , 117)	94 (1 , 100)	1 (1 , 1)
$\alpha = 0.10$	111 (101 , 121)	97 (84 , 101)	1 (1 , 1)
$\alpha = 0.15$	112 (101 , 123)	98 (82 , 101)	1 (1 , 53)

Table 4: Detected break times for parameter β in GARCH(1,1) model using significance level $\alpha = 0.10$. Different starting points t^* are used.

Model 1 ($t = 1, \ldots, 100$):	$(0.001, 0.03, 0.75)$	$(0.001, 0.03, 0.95)$
Model 2 ($t = 101, \ldots, 200$):	$(0.001, 0.03, 0.95)$	$(0.001, 0.03, 0.75)$
$t^* = 105$	1 (1 , 1)	46 (1 , 96)
$t^* = 110$	1 (1 , 1)	98 (89 , 101)
$t^* = 115$	1 (1 , 1)	99 (95 , 101)
$t^* = 120$	1 (1 , 1)	99 (95 , 101)
$t^* = 130$	1 (1 , 84)	99 (95 , 101)
$t^* = 150$	1 (1 , 112)	100 (97 , 101)
$t^* = 200$	118 (112 , 125)	101 (97 , 110)

to some breaks (the first quartiles of break times for $\alpha = 0.05$ are 1 in all cases, so the break was not detected in at least 25% cases); on the other hand, higher values of α (lower significance of the test) detects non-existent breaks (the third quartile is 53 for the model without structural break and $\alpha = 0.15$).

Using the significance level $\alpha = 0.10$, let us now study the break-detection behavior as a function of distance from the break point $t = 101$. We simulated again GARCH(1,1) models with structural break following case 3 first, but we started the search for the longest interval of homogeneity not only from $t^* = T = 200$, but also from $t^* = 105, 110, 115, 120, 130, 150$. Thus, the simulation should indicate how fast the proposed procedure detects a break, measured by the time past the break. The simulation results for all starting points t^* are summarized in Table 4. Apparently, it takes a lot of time (100 observations) to detect change from a low volatility model $(\omega, \alpha, \beta) = (0.001, 0.03, 0.75)$ to a higher volatility model $(\omega, \alpha, \beta) = (0.001, 0.03, 0.95)$. On the other hand, the opposite change is uncovered already with 10 to 15 observations part the break time and the detected time is always very close to $t = 100$. This is a typical phenomenon: while it is possible to quickly detect shift from a high volatility to a low volatility model (column 2), it takes much more time to find out a change from low to high volatility model (column 1).

Next, let us perform the same set of simulations as presented in Table 4 for the change in the ARCH parameter. The results are summarized in Table 5. Again, one can see that the change from high to low volatility in the model can be detected with much smaller number of observations past the break than the opposite break. Moreover, contrary to the change in the GARCH parameter, a structural break is detected only in 25% of simulated time series. This may not be surprising once we have a look at one simulated path of a time series with such a structural break (Figure 4): visually, the path does not change very much after the break.

However, all simulation up to now used the critical values computed from Theorem 3.1 using the value of $C = 2$. An improvement in the performance of the break detection

Table 5: Detected break times for parameter α in GARCH(1,1) model using significance level $\alpha = 0.10$. Different starting points t^* are used.

Model 1 ($t = 1, \ldots, 100$):	$(0.001, 0.03, 0.75)$	$(0.001, 0.23, 0.75)$
Model 2 ($t = 101, \ldots, 200$):	$(0.001, 0.23, 0.75)$	$(0.001, 0.03, 0.75)$
$t^* = 105$	1 (1 , 1)	1 (1 , 1)
$t^* = 110$	1 (1 , 1)	1 (1 , 74)
$t^* = 115$	1 (1 , 1)	1 (1 , 86)
$t^* = 120$	1 (1 , 1)	1 (1 , 74)
$t^* = 130$	1 (1 , 1)	1 (1 , 92)
$t^* = 150$	1 (1 , 106)	1 (1 , 96)
$t^* = 200$	1 (1 , 116)	1 (1 , 99)

Table 6: Detected break times for parameter α in GARCH(1,1) model using significance level $\alpha = 0.10$ and critical values simulated for model 8 only. Different starting points t^* are used.

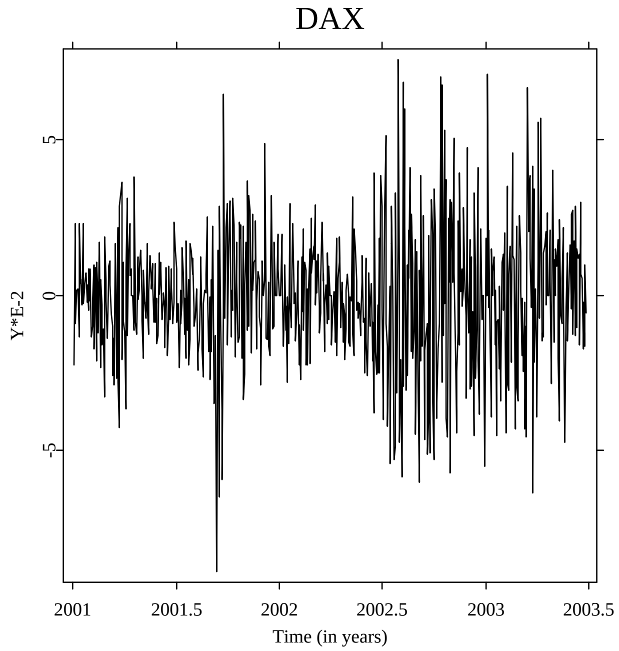
Model 1 ($t = 1, \ldots, 100$):	$(0.001, 0.03, 0.75)$	$(0.001, 0.23, 0.75)$
Model 2 ($t = 101, \ldots, 200$):	$(0.001, 0.23, 0.75)$	$(0.001, 0.03, 0.75)$
$t^* = 105$	1 (1 , 1)	82 (1 , 97)
$t^* = 110$	1 (1 , 1)	92 (1 , 98)
$t^* = 115$	1 (1 , 1)	90 (1 , 100)
$t^* = 120$	1 (1 , 1)	86 (1 , 99)
$t^* = 130$	1 (1 , 105)	93 (1 , 101)
$t^* = 150$	1 (1 , 120)	97 (59 , 101)
$t^* = 200$	136 (71 , 177)	98 (73 , 117)

might be improved by using critical values simulated for a specific model (see Section 5.1); for example, we can detect the structural break in case 2 by using critical values simulated only for model 8 (see Table 1). The corresponding simulation results analogous to Table 5 are presented in Table 6. An immediate impact of critical values specific for the models in the data is visible especially in column 2 of Table 6: the structural break is now detected in many more cases and the results are relatively close to $t = 100$, at least if there are enough data points observed after the break ($t^* \geq 150$). Despite this improvement, the break detection of the change from a lower to a higher value of α (column 1) does not seems to be very reliable.

5.3 Application to a stock index

The first application of the proposed adaptive estimation for GARCH(1,1) models concerns the German stock index DAX. Here, we are merely interested in seeing whether

Figure 6: DAX time series from January 1, 2001 to June 31, 2003.



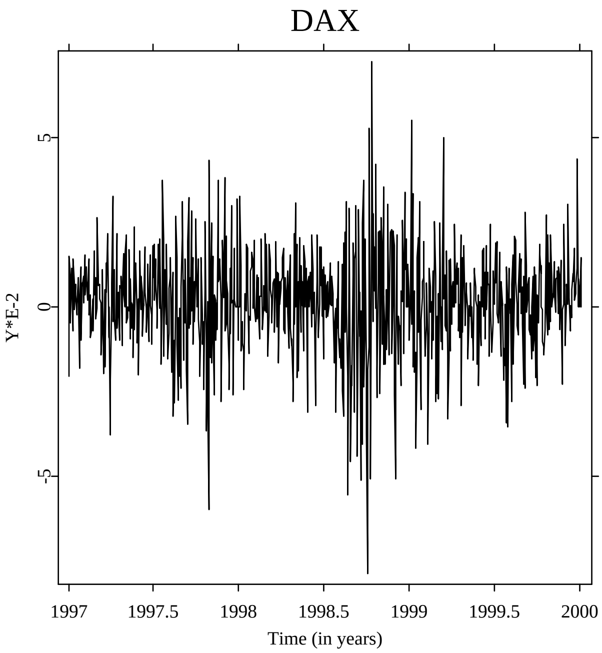
or not the adaptive procedure uncovers some structural breaks, which can possibly be supported by an economic intuition.

First of all, we took the log-returns of the DAX time series from January 1, 1990 to June 31, 2003 (see a part of the series in Figure 6), and having prediction based on historical volatilities in mind, started the search for the longest interval of homogeneity usable at the moment (June 31, 2003). The result indicated by the adaptive procedure is that the behavior of DAX follows a single GARCH(1,1) model for almost a year—from the beginning of July 2002 to the present date. Not surprisingly, starting the procedure from many other days between January and June 2003 produced the same result: the current window of homogeneity begins at the end of June 2003 or at the beginning of July 2003 (there was a slight fluctuation of the date within two weeks June 24 to July 4, 2002).

The change in the behavior of the DAX series around the middle of year 2002 can be confirmed visually, see Figure 6. Moreover, the period of higher volatility starting in July 2002 corresponds to a period of higher uncertainity on the stock market caused by the continuing economic decline of German economy without any hint on the actions (interest rate cuts) of the European Central Bank and increasing anxiety about the war in Iraq.

Second, we analyzed the same series from the perspective of year 1999, since it follows the Asian crisis, which started in the last quarter of year 1997 and most economic analysts agree that its impact lasted almost 6 quarters, that is, to the first quarter of 1999. Starting the proposed adaptive procedure from various time points during 1999 to find the longest

Figure 7: DAX time series from January 1, 1997 to December 31, 1999.



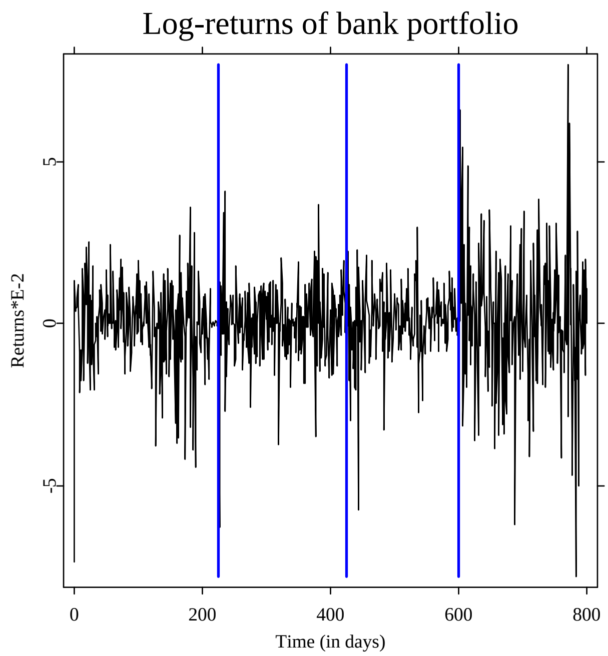
interval of homogeneity, March 15 was in all cases determined as the change point. This points ends the period of higher volatility in 1998 and the beginning of 1999 (Figure 7) and coincides very well with the assumed end of Asian crisis.

Finally, let us note that the adaptive GARCH estimation managed to detect in the case of DAX structural breaks both in the direction of change from low to high volatility (the first case, June/July 2002) and from high to low volatility (the second case, March 1999). This is a bit more optimistic result than the simulations would suggest, especially in the case of detecting change from a low to a high volatility model (smaller number of observations than in simulations was necessary). One has to realize however that the changes in the real data are usually more complex than in the simulations done in Section 5.2, which makes them easier to detect. On the other hand, the break detection in the DAX series resonates the simulation results in the sense that the change in March 1999 (high to low volatility) was detected (almost) always on the same day (March 15, 1999) no matter how far past this break we started, whereas the indication of a structural break in June/July 2002 (low to high volatility) varies in an interval of almost two weeks depending on the distance from the break.

5.4 VaR analysis

The second application aims at the so-called Value-at-Risk (VaR) analysis, which tries to predict the distribution of the value of a financial instrument or their portfolio at time $t+1$ given information available at time t, and more specifically, the quantiles of this

Figure 8: Three years of daily log-returns of a portfolio of a German bank.



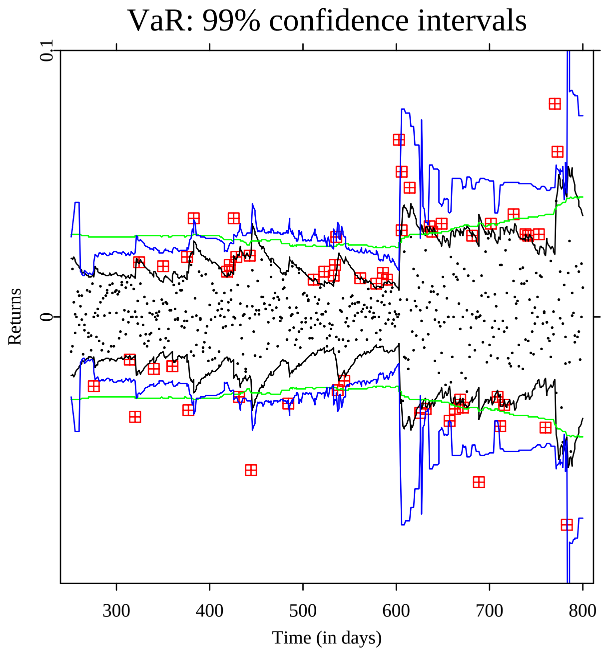
distribution and the corresponding confidence intervals. A usual assumption postulates that this distribution is Gaussian and a method to estimate just the variance (matrix) of the distribution is used. Typically, one takes the moving average (or exponentially weighted moving average) with window of some fixed length (e.g., one year). Alternatively, one can also use a time series models such as ARCH or GARCH to model the development of the series and its variance and then predict the volatility in the next time period. All these approaches suffer in an instable environment from the fixed number of past observations taken into account.

This is where the proposed adaptive procedure may bring in some improvement compared to the existing method, since it automatically detects how many observations can be used for prediction: either a rather small number if we are just after some structural break (likely to lead to consistent estimates with high variance), or a much larger number of observations than the moving average window if there is a long period of stability (likely to lead to reduction in variance compared to fixed-window procedures).

To verify this, we analyzed the returns of a portfolio of an unnamed German bank; the data are available daily for 4 years (see Figure 8). At every time point t we estimated the variance at $t+1$ by rectangular (RMA) and exponentially weighted (EMA) moving average as well as by the adaptive estimation procedure based on GARCH(1,1) (ADGARCH). Using the predicted variance by each method, we estimated 99% confidence intervals. If the estimation works correctly, about 1% of 750 observations (3 years) should lie outside of the constructed confidence intervals (the first year is omitted from the analysis, because non-adaptive methods use past observations within one year to construct the confidence intervals).

The estimation results by RMA (green line), EMA (black line), and ADGARCH (blue

Figure 9: 99% VaR estimates for daily log-returns of a portfolio of a German bank.



line) are in Figure 9. The black dots and red grids represent the time series itself, whereby the red color indicates that a point lies outside of the confidence interval based on EMA (exponential moving average is considered to be "more adaptive" because it gives more weight only to recent observations). Let us first have a look at the estimates after day 600, since both ADGARCH as well as the time series plot (Figure 8) indicate a period with an increased volatility. The Figure 9 shows that the adaptive GARCH estimation adapts very quickly to a regime change and produces a much larger confidence intervals than the other two methods. The ADGARCH confidence region is possibly too large, especially just after day 600, because it employes only small number of observations that lie after the break which increases the variance too. On the other hand, the RMA and EMA methods fail to adapt completely as the number of red points exceeding the EMA confidence regions in days 600–750 indicates.

Moreover, looking at the first part of Figure 9 (days 0–600), we can see that ADGARCH estimates can outperform RMA, although they are always bigger than EMA confidence intervals. The reason become clear from the graph of the time series (Figure 8), where the blue vertical lines indicate the point, which were mostly considered by ADGARCH as structural break points. Apparently, no window of homogeneity is longer than one year (250 working days), so ADGARCH cannot demonstrate its other advantage—taking a very long interval of stability if possible and thus reducing the variance of estimates.

Finally, let us note that there are just 12 points which lie outside of the confidence intervals determined by ADGARCH. One might considered this as still to much (given that optimally just 1% of the points, that is 7–8 observations, should lie outside). On the other hand, the ADGARCH confidence region is still much more reliable than those of EMA and RMA, where the number of exceedances is two to three times higher.

6 Conclusions and outlook

The local adaptive volatility estimate (LAVE) is described and analyzed in this paper. It provides a nonparametric way for for estimating and short term forecasting the volatility of financial returns.

It is assumed that a local constant approximation of the volatility process holds over some unknown interval. The issue of filtering this interval of time homogeneity out of the return time series is considered, and a nonparametric approach is presented. The estimate of the volatility process is then found by averaging over the interval of time homogeneity.

A theoretical analysis of the properties of the LAVE algorithm is provided and the problem of selecting the smoothing parameters is analyzed through Monte Carlo simulation. The estimation results on change-point models show that the method has a reasonable performance in practice. An empirical application to exchange rate returns and a comparison with a GARCH(1,1) also provides a good evidence that the new method is competitive and can even outperform the standard parametric models especially for forecasting with a short horizon.

An important feature of the proposed method is that it allows for a straightforward extension on the multivariate volatility estimation, see **?** for a detailed discussion.

Obviously, if the underlying conditional distribution is not normal, the estimated volatility can give only a partial information about the riskiness of the asset. Recent developments in the risk analysis tends to focus on the estimation of the quantiles of the distribution. In this direction the LAVE procedure can be used as a convenient tool for pre-whitening the returns and obtain a sample of "almost" identical and independently distributed returns, which do not display any more variance clustering. So that the usual techniques of quantile estimation could be applied in a static framework. We regard such a development as a topic for future research.