Localized Realized Volatility Modelling

Ying Chen*
Wolfgang Karl Härdle**
Uta Pigorsch***

BERLIN

ECONOMIC RISK

SFB 649

*National University of Singapore
**Humboldt-Universität zu Berlin, Germany
*** Universität Mannheim, Germany

Localized Realized Volatility Modelling[*]

Ying Chen[†], Wolfgang Karl Härdle[‡] and Uta Pigorsch[§]

January 28, 2009

Abstract

With the recent availability of high-frequency financial data the long range dependence of volatility regained researchers' interest and has lead to the consideration of long memory models for realized volatility. The long range diagnosis of volatility, however, is usually stated for long sample periods, while for small sample sizes, such as e.g. one year, the volatility dynamics appears to be better described by short-memory processes. The ensemble of these seemingly contradictory phenomena point towards short memory models of volatility with nonstationarities, such as structural breaks or regime switches, that spuriously generate a long memory pattern (see e.g. Diebold and Inoue, 2001; Mikosch and Stărică, 2004b). In this paper we adopt this view on the dependence structure of volatility and propose a localized procedure for modeling realized volatility. That is at each point in time we determine a past interval over which volatility is approximated by a local linear process. Using S&P500 data we find that our local approach outperforms long memory type models in terms of predictability.

JEL codes: G17, C14, C51

Keywords: Localized Autoregressive Modeling, Realized Volatility, Adaptive Procedure

[†]Department of Statistics & Applied Probability, National University of Singapore, 6 Science Drive 2, Singapore 117546, email: stacheny@nus.edu.sg

[‡]CASE - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, School of Business and Economics, Spandauerstr. 1, 10178 Berlin, Germany

[§]Universität Mannheim, Department of Economics, L7, 3-5, 68131 Mannheim, Germany

1 Introduction

One of the key elements in the modeling of the stochastic dynamic behavior of financial assets is the volatility. It is not only a measure of uncertainty about future returns but also an important input parameter in derivative pricing, hedging and portfolio selection. Accurate volatility modeling is therefore in the focus of the financial econometrics and quantitative finance research. Among the many possible volatility measures the (square root) of quadratic variation has emerged as the most frequently used one. With the availability of high-frequency data, so–called realized volatility estimators (sums of squared high-frequency returns) have been proposed and have been shown to provide better volatility forecasts than the concurrent volatility estimators based on a coarser, e.g. daily, sampling frequency (see e.g. Andersen, Bollerslev, Diebold and Labys, 2001b).

Realized volatility together with other volatility measures exhibit significant autocorrelation which is the basis for the statistical predictability of volatility. In fact, the sample autocorrelation function has typically a hyperbolically decaying shape, also known as "long memory". Therefore, a strand of literature (see e.g. the literature on autoregressive fractional integrated moving average, ARFIMA, and fractional integrated generalized autoregressive conditional heteroscedaticity, FIGARCH, models) focused on this kind of correlation phenomenon. The long memory "diagnosis" is usually stated for long sample periods such as typically three to ten years. The diagnosis can, however, also be generated by a simple model with change inside this rather long interval: the possibility of such intermediate changes provides an alternative view on the described phenomenon. Like in the physical sciences, where one uses wave and particle theory to explain the emission of light, we have here a duality of theories for the emission of volatility. It is the object of this paper to investigate this dual view

2

on volatility phenomenon.

In doing so, we need to determine the time-varying (local) structure of volatility. This is conveniently done via adaptive statistical techniques, that allow us to find for each time point a past time interval, where a local volatility model is a good approximator. This adaptively chosen time interval varies with the time point in consideration. We, thus, refer to this procedure as localized volatility modeling.

We investigate localized realized volatility (LRV) modeling based on autoregressive processes. We apply the LRV to S&P500 data and compare it to (approximate) long memory techniques, such as the ARFIMA and heterogenous autoregressive (HAR) models. We find that the LRV technique provides improved volatility forecasts.

In the literature of the long memory view of volatility, fractionally integrated $I(d)$ processes have frequently been under consideration due to their hyperbolically decaying shock propagation for $0 < d < 1$. These long memory processes have been proposed by e.g. Granger (1980), Granger and Joyeux (1980) and Hosking (1981), and can be opposed to the extreme case of short memory (i.e. $I(0)$ processes) and those of infinite memory (i.e. $I(1)$ processes). When applied to volatility they seem to provide a better description and predictability than short memory models estimated over (the same) long sample periods. A typical example is the empirically better performance of the FIGARCH model of Baillie, Bollerslev and Mikkelsen (1996) as opposed to a standard GARCH model. For realized volatility, the ARFIMA process emerged as a standard model (see e.g. Andersen, Bollerslev, Diebold and Labys, 2003; Pong, Shackleton, Taylor and Xu, 2004). An alternative and quite popular model, that does not belong to the class of fractionally integrated processes but approximates the long range dependence by a sum of several multi-period volatility components, is the HAR model proposed by Corsi (2008).

3

The question on the true source of the long memory diagnosis, however, still remains. In fact, the theoretical results provided in Diebold and Inoue (2001) and Granger and Hyung (2004) show that this phenomenon can also be spuriously generated by a short-memory model with structural breaks or regime-shifts. More generally, Mikosch and Stărică (2004b) even argue independently of any particular model assumptions that nonstationarities in the data, such as changes in the unconditional mean or variance, can lead to the diagnosis of long range dependencies. Thus, parameter changes in the volatility equation of GARCH models may induce the observed long memory in the absolute or squared daily returns (see also Čížek, Härdle and Spokoiny, 2009; Mikosch and Stărică, 2004a), who find that a locally adaptive GARCH model outperforms the stationary, i.e. constant (in parameters) GARCH model). For realized volatility, which is usually modeled directly, we expect from these arguments that a dynamic short memory model with changing parameters may be the driving source of volatility. This motivates our choice of a local AR(1) opposed to the long memory approaches.

The remainder of the paper is structured as follows. The next section describes the construction and empirical properties of our S&P500 index futures realized volatility measure. Section 3 presents in detail the LRV modeling approach along with some simulation experiments. Section 4 briefly reviews the standard long memory models and Section 5 empirically compares these dual views within a forecasting exercise. Section 6 concludes.

2 Realized volatility

The recent literature usually refers to realized volatility as a measure of the quadratic variation of the (logarithmic) price of a financial asset that is based on high-frequency,

i.e. intradaily, returns. Most commonly it is the daily quadratic variation that is of main interest. For the ease of exposition, we, thus, normalize the daily time interval to unity and assume that $M + 1$ intraday prices are observed at time points $n_0, \ldots, n_M$. The continuously compounded j-th within-day return of day t is denoted by:

$$r_{t,j} = p_{t,n_j} - p_{t,n_{j-1}}, \quad j = 1, \ldots, M, \tag{1}$$

where p_{t,n_j} is the logarithmic price observed at time point n_j of trading day t. *Daily realized volatility* is then defined by

$$\widetilde{RV}_t = \sum_{j=1}^{M} r_{t,j}^2. \tag{2}$$

If the logarithmic price process follows a continuous-time semimartingale, this quantity converges to the quadratic variation for $M \rightarrow \infty$ (see e.g. Andersen and Bollerslev, 1998; Barndorff-Nielsen and Shephard, 2002b). Importantly, if the price is given by a diffusion process, then realized volatility converges to the daily integrated volatility, which is the main object of interest. Consistency and asymptotic distribution of realized volatility as an estimator of the integrated volatility are derived in Barndorff-Nielsen and Shephard (2002a).

2.1 Market microstructure effects

The theoretical results on realized volatility build on the notion of an infinite sampling frequency. In practice, however, the very high-frequency prices are contaminated by market microstructure effects, such as bid-and-ask bounce effects, price discreteness etc., leading to biases in realized volatility (see e.g. Andersen, Bollerslev, Diebold and Ebens, 2001a; Barndorff-Nielsen and Shephard, 2002a). A common approach to

reduce these effects is to simply construct realized volatility based on lower frequency returns, such as 10 to 30 minutes. However, such a procedure comes at the cost of a less precise volatility estimate. Various alternative methods have been proposed to solve this bias-variance trade-off, such as selecting an in a mean-square-error-sense optimal sampling frequency (see e.g. Bandi and Russell, 2005; Zhang, Mykland and Aït-Sahalia, 2005), or to employ subsampling procedures, also called multiscale estimators, that average over the realized volatility based on all price observations and various realized volatilities each being calculated from different subsamples of high-frequency returns (see e.g. Zhang et al., 2005). The use of kernel-based estimators was originally proposed by Zhou (1996) and further extended in Hansen and Lunde (2006). However, both of these estimators are inconsistent. Barndorff-Nielsen, Hansen, Lunde and Shephard (2008) instead developed consistent, so–called realized kernel estimators for realized volatility, that have very attractive properties. In particular, for an appropriate choice of the kernel the estimator is efficient (relative to a parametric maximum-likelihood estimator), the asymptotic variance of the estimator is even smaller than that of the multiscale estimator and the realized kernels are robust to a host of market microstructure frictions including a particular type of endogenous noise.

For our empirical application we therefore construct realized volatility based on the kernel procedure. In particular, we use a flat-top realized kernel, i.e.

$$RV_t = \widetilde{RV}_t + \sum_{h=1}^{H^*} k\left(\frac{h-1}{H^*}\right)(\gamma_{t,h} + \gamma_{t,-h}) \tag{3}$$

with kernel weight function $k(x)$ being twice continuously differentiable on $[0, 1]$ and satisfying $k(0) = 1$, and $k(1) = k'(0) = k'(1) = 0$. The h-th realized autocovariance for day t is defined by $\gamma_{t,h} = \sum_{j=1}^{M} r_{t,j} r_{t,j-h}$. Note, that for each day t the number of autocovariances considered in the realized kernel is determined by $H = c\xi\sqrt{M}$,

where ξ denotes the noise-to-signal ratio, that relates the (daily) variance of the market microstructure noise to the (daily) integrated volatility, and c is a constant that depends (inter alia) on the specific kernel weight function. Since the asymptotic variance of the realized kernel, i.e. of the RV_t estimator, depends on c, this variance can be minimized by an appropriate choice of c^*, leading to the optimal number of lags H^*.

The computation of the realized kernel estimator requires the precise specification of the kernel weight function. For our empirical application we consider the modified Tukey-Hanning kernel with weight function $k(x) = \sin^2\left\{\frac{\pi}{2}(1-x)^p\right\}$, as it is most efficient among the finite lag kernels analyzed in Barndorff-Nielsen et al. (2008). Note that for increasing p the number of autocovariances considered in the realized kernel (3) is also increasing (due to an increasing value of c^*). Moreover, for increasing p the realized kernel approaches the (parametric) efficiency bound. As such, a large value of p might be preferable. In practice, however, the increasing number of lags imposes some limitations as it involves a large number of returns outside the daily time interval. We, thus, follow Barndorff-Nielsen et al. (2008) and choose $p = 2$ for our empirical application. Noteworthy, for this choice of p the realized kernel is still close to efficient. The optimal number of lags in the realized kernel is then determined for each day by

$$H^* = 5.74\hat{\xi}\sqrt{M} \tag{4}$$

with 5.74 being the optimal c^* for the Tukey-Hanning kernel with $p = 2$. We estimate the noise-to-signal ratio by $\hat{\xi} = \frac{\widetilde{RV}_{t,1/2M}}{\widetilde{RV}_{t,15}}$, where the numerator gives the estimator of the noise variance suggested by Bandi and Russell (2005), which is basically given by the conventional realized volatility estimator based on one minute returns (as is indicated here by the second subscript of $\widetilde{RV}$). The denominator estimates the daily integrated volatility at a low, i.e. 15 minutes, sampling frequency, at which market

microstructure effects should be negligible. Given H^*, realized volatility is finally computed according to (3). The market microstructure noise uncorrected realized volatility $\widetilde{RV}_t$ and the realized autocovariances $\gamma_{t,h}$ are based on one minute returns.

2.2 Data description

Our empirical analysis is based on realized volatility of the S&P500 index futures ranging from January 2, 1985 to February 4, 2005. From the various S&P500 Index futures with maturity dates in March, June, September and December, we consider only the most liquid contracts. We then construct realized volatility according to the realized kernel estimator. In particular, we compute realized volatility based on equation (3) using 1 minute returns. For the estimation of H^* we further consider 15 minute returns, as described above. The returns are constructed using the previous-tick method and by excluding overnight returns.

Table 1: Descriptive Statistics

Series	Mean	Std.Dev.	Skewness	Kurtosis	Ljung-Box(21)[1]
RV_t	1.0709	8.1691	59.0882	3861	1375
$\log(RV_t)$	-0.5139	0.8797	0.4335	4.9912	46809

[1] The critical value of the Ljung-Box test statistic of no autocorrelation up to approximately 1 month is 32.671.

The descriptive statistics of the resulting realized volatility series are presented in Table 1. In summary, the empirical characteristics of the series are in line with the findings reported in the earlier literature on realized volatility. In particular, realized volatility is strongly skewed and fat-tailed, while its logarithmic version is much closer to Gaussianity. This is also confirmed by the kernel density estimate of logarithmic realized volatility, which is presented in Figure 1 along with the kernel density esti-
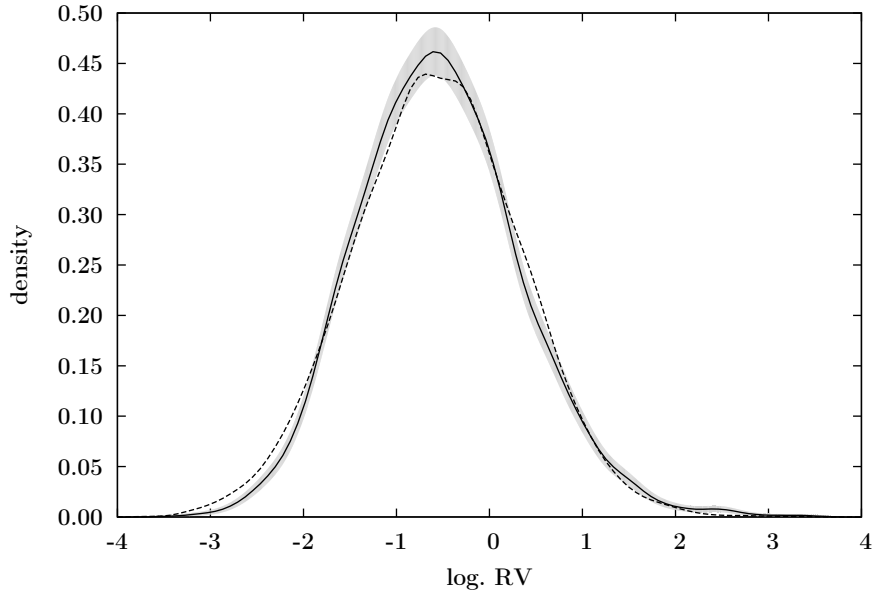
Figure 1: Kernel density estimate of logarithmic realized volatility of the S&P500 index futures (solid line). The shaded area corresponds to the pointwise 95% confidence intervals and the dashed line represents the kernel density estimate of i.i.d. random variables simulated from the fitted normal distribution.

mate of i.i.d. random variables simulated from the fitted normal distribution (with a sample size corresponding to the empirical one). Moreover, the sample autocorrelation function of (logarithmic) realized volatility (depicted in the lower panel of Figure 2) exhibits the aforementioned hyperbolic decay. We evaluate this long memory diagnosis in more detail in the empirical application. In the following, however, we first introduce our localized approach to realized volatility modeling.

3   The localized realized volatility approach

An alternative view on the long memory phenomenon of volatility is given by a localization of the realized volatility dynamics. The idea of this localized approach for modeling realized volatility is as follows. It is assumed that at each point in
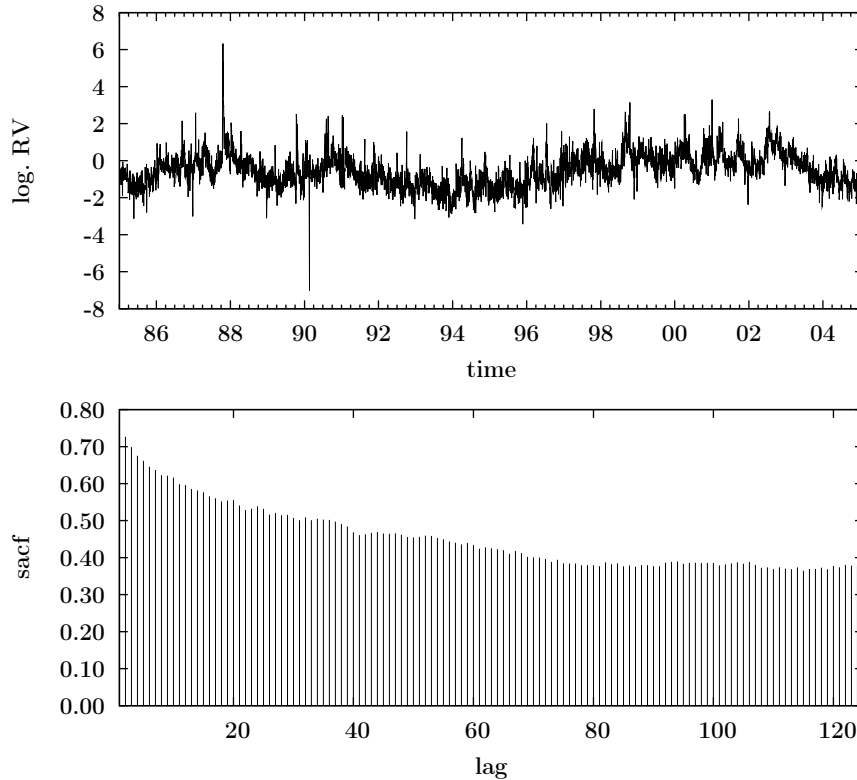
9

Figure 2: Time evolvement and sample autocorrelation function (sacf) of logarithmic realized volatility of the S&P500 index futures.

time there exists a past time interval over which volatility can be approximated by a local autoregressive model of order one (LAR(1)) with constant parameters. In contrast to fitting a global volatility model, this implies that we obtain at each point in time a potentially new set of parameters, which is estimated based on the so–called interval of homogeneity. This technique, thus, also involves the determination of the length of this interval over which the parameters of the local model are assumed to be constant. Once the interval length and the parameters are determined for a particular time period, the corresponding local model may be used for volatility predictions. This section describes the adaptive estimation and the implementation of the localized approach in more detail. The performance is also tested in a set of simulation experiments.

10

3.1   Adaptive estimation method

The local (time-varying) autoregressive scheme of order 1 is defined through a time-varying parameter set $\theta_t = (\theta_{1t}, \theta_{2t}, \theta_{3t})^\top$:

$$\log RV_t = \theta_{1t} + \theta_{2t} \log RV_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{N}(0, \theta_{3t}^2). \tag{5}$$

Here the Gaussian distributed innovations ε_t have a mean of zero and variance θ_{3t}^2. time-varying parameters at any point in time t are of course too flexible to really constitute a practical dynamic model. We therefore need to strike a balance between model flexibility and dimensionality. This is done by localizing a low dimensional time series dynamics in the high dimensional model (5).

The basic idea is to approximate (5) at a fixed time point τ by a parameter set $\theta_\tau = (\theta_{1\tau}, \theta_{2\tau}, \theta_{3\tau})^\top$, where θ_τ is constant over the interval $I_\tau = [\tau - s, \tau)$ with $0 < s < \tau$. The interval I_τ defines a "locally homogeneous" span of data and is called "interval of homogeneity". The question is how to find I_τ or the value of s over which the model parameters are estimated.

To this end, consider first the maximum likelihood (ML) estimator $\tilde{\theta}_\tau$ given an homogeneous interval I_τ:

$$
\begin{aligned}
\tilde{\theta}_\tau &= \operatorname{argmax}_{\theta \in \Theta} L(\log RV; I_\tau, \theta) \\
&= \operatorname{argmax}_{\theta \in \Theta} \left\{ -\frac{s}{2} \log 2\pi - s \log \theta_3 - \frac{1}{2\theta_3^2} \sum_{t=\tau-s}^{\tau-1} (\log RV_t - \theta_1 - \theta_2 \log RV_{t-1})^2 \right\}
\end{aligned}
$$

where Θ denotes the parameter space and $L(\log RV; I_\tau, \theta)$ the local conditional log-likelihood function. We, thus, refer to this estimator as the local ML estimator. For notational simplification, we also use the short notation $L(I_\tau, \theta)$ for the local conditional log-likelihood function.

Let the (logarithmic) realized volatility be exactly modelled by an AR(1) process with parameter θ_τ^* at time point τ, i.e θ_τ^* is constant over the Interval I_τ. The accuracy of estimation can then be measured by the log-likelihood ratio (LR):

$$LR(I_\tau, \tilde{\theta}_\tau, \theta_\tau^*) \;\; = \;\; L(I_\tau, \tilde{\theta}_\tau) - L(I_\tau, \theta_\tau^*). \tag{6}$$

Polzehl and Spokoiny (2006) have proved that LR and its power transformation $|LR(I_\tau, \tilde{\theta}_\tau, \theta_\tau^*)|^r$ with $r > 0$ are bounded for an i.i.d. sequence of Gaussian innovations (in our case this refers to the innovations of the local AR(1) process):

$$\mathsf{E}_{\theta_\tau^*} \left| LR(I_\tau, \tilde{\theta}_\tau, \theta_\tau^*) \right|^r \leq \xi_r \tag{7}$$

with $\xi_r = 2r \int_{\xi \geq 0} \xi^{r-1} e^{-\xi} d\xi = 2r\Gamma(r)$. This bound is non-asymptotic and allows to construct a confidence interval, which can be used to identify a homogeneous interval.

The number of possible interval candidates is large, e.g. the first interval may include just a few past observations and the intervals considered thereafter may be increased by just one observation in each step up to including all past observations. As this is computationally intensive, especially for large data sets, we consider only a finite set of intervals $\mathbf{I}_\tau = \{I_\tau^1, \ldots, I_\tau^K\}$ with a reasonable value of K, as proposed in Chen and Spokoiny (2009). The intervals are increasingly ordered according to their length, i.e. $I_\tau^1 \subset \ldots \subset I_\tau^K$. The first interval I_τ^1 should be short enough such that homogeneity can be assured. Note that to each interval there corresponds a local ML estimate, denoted by $\tilde{\theta}_\tau^k$ with $k = 1, \ldots, K$. In statistical learning theory these are called weak learners. The risk bound (7) is calculated under the hypothetical constant θ_τ^* parameter situation. By increasing the intervals in the nonparametric situation (5) we incur an increasing modeling bias. "Oracle" type of results as given in Belomestny and Spokoiny (2007) ensure that an optimal choice $\hat{I}_\tau$ of an interval of

12

homogeneity (striking a balance between bias and variation) can be obtained via an adaptive procedure. Details of the "oracle" results can be found in the cited literature.

The aim of this research is to put forward the mentioned dual view on the dependence structure of volatility. It is therefore appropriate here to concentrate on the construction details rather than on the theoretical technicalities. The main ingredient of the local model selection algorithm is based on a sequential testing procedure. The procedure starts from the shortest interval I_τ^1, where the homogeneity is assured and $\tilde{\theta}_\tau^1$ is automatically accepted as a homogeneous estimator: $\hat{\theta}_\tau^1 = \tilde{\theta}_\tau^1$. Sequentially at step $k \geq 2$, we test the hypothesis of homogeneity of the successive interval I_τ^k, see Figure 3. The test at step k is formulated as:

$$\left| LR(I_\tau^k, \tilde{\theta}_\tau^k, \hat{\theta}_\tau^{k-1}) \right|^r \leq \zeta_k \tag{8}$$

where ζ_k is the critical value (described in more detail below). The test is motivated by the bound (7). The likelihood ratio measures the difference of a new ML estimate $\tilde{\theta}_\tau^k$ over a "possibly" homogeneous interval I_τ^k and the previously accepted homogeneous estimate $\hat{\theta}_\tau^{k-1}$. If there is no significant difference between the two estimates, we accept the new one $\hat{\theta}_\tau^k = \tilde{\theta}_\tau^k$. The reason is that, compared to the former accepted estimate $\hat{\theta}_\tau^{k-1}$, the new estimate has a smaller variation as more observations are used in the estimation. On the other hand, if the difference is significant, it indicates that a structural shift rather than homogeneity is detected and the sequential testing terminates. The significance at each step is measured by a critical value. Therefore, a set of critical values $\{\zeta_k\}_{k=1}^K$ is required in the sequential testing. In the next section, we discuss the computation of the critical values and the choice of the involved parameters.

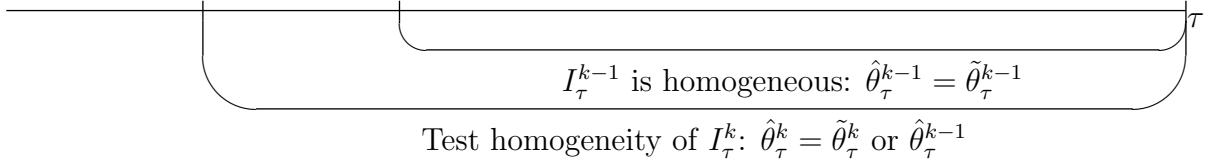The formal definition of the procedure is as follows:

13

Figure 3: Sequential test of homogeneity: the longer interval I_τ^k is tested after the hypothesis of homogeneity over the shorter interval I_τ^{k-1} has been accepted.

1. Initialization: $\hat{\theta}_\tau^1 = \tilde{\theta}_\tau^1$.

2. $k = 1$

   while $\left| LR(I_\tau, \tilde{\theta}_\tau^{k+1}, \hat{\theta}_\tau^k) \right|^r \leq \zeta_{k+1}$ and $k < K$,

$$
\begin{aligned}
k &= k + 1 \\
\hat{\theta}_\tau^k &= \tilde{\theta}_\tau^k
\end{aligned}
$$

3. Final estimate: $\hat{\theta}_\tau = \hat{\theta}_\tau^k$

3.2   Choice of parameters and implementation details

To run the proposed adaptive procedure, we need to determine the input parameters, i.e. the set of intervals, the power parameter r in (8) and the critical values. In the following we present our choices and the computation of the critical values using Monte Carlo simulation.

Set of intervals

We consider a finite set with $K = 13$ intervals. This set is composed of the following interval lengths:

$$\{1w, 1m, 3m, 6m, 1y, 1.5y, 2y, 2.5y, 3y, 3.5y, 4y, 4.5y, 5y\},$$

14

where w denotes a week (5 days), m refers to one month (21 days) and y to one year (252 days). In other words, $I_\tau^1 = [\tau - 1\mathrm{w}, \tau)$, $I_\tau^2 = [\tau - 1\mathrm{m}, \tau)$, ..., $I_\tau^{13} = [\tau - 5\mathrm{y}, \tau)$. Note that the same set $\mathbf{I} = \{I_\tau^k\}_{k=1}^{13}$ is used for each time point τ and for notational convenience we therefore drop the subscript in the following. Our interval choice is motivated by practical reason that investors are often concerned about special investment horizons. Using also different sets of intervals, we find that the procedure is insensitive to the interval choice. Nevertheless, it is important to assure homogeneity over the shortest interval.

Critical values

In the testing procedure, the critical values measure the significance of the ML estimate under the hypothesis of homogeneity. We here calculate the critical values under homogeneity, i.e. constant parameters in (5). In particular, we generate 100 000 AR(1) processes with $\theta_t = \theta^* = (\theta_1^*, \theta_2^*, \theta_3^*)^\top$ for all t: $y_t = \theta_1^* + \theta_2^* y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathrm{N}(0, \theta_3^{*2})$. The starting value was set to $y_0 = \theta_1^*/(1 - \theta_2^*)$. The sample size of each process is 1261 (corresponding to 5 years and 1 day) in accordance to the largest interval of $\mathbf{I}$. Clearly the ML estimate of the largest interval (I^{13}) is the choice, i.e. $\hat{\theta}_t = \hat{\theta}_t^K = \tilde{\theta}_t^K$. Remember that in the sequential testing the adaptive estimator at step k depends on the critical values $\{\zeta_1, \ldots, \zeta_k\}$ and we therefore use the notation $\hat{\theta}_{t(\zeta_1,\ldots,\zeta_k)}^k$ to emphasize the effect of the critical values. The estimate $\hat{\theta}_t^K$ is required to fulfill the following risk bound:

$$\mathsf{E}_{\theta^*} \left| LR \left(I^K, \tilde{\theta}_t^K, \hat{\theta}_{t(\zeta_1,\ldots,\zeta_K)}^K \right) \right|^r \leq \xi_r \tag{9}$$

Moreover, it has been discussed in the non- and semiparametric literature that an increase in sample size implies an increase of bias due to the increase in degrees of freedom, see e.g. Härdle, Müller, Sperlich and Werwatz (2004). To take this into

account, we introduce a weighting scheme for $k = 1, \ldots, K$ into the risk bound, i.e.:

$$\mathsf{E}_{\theta^*} \left| LR \left(I^k, \tilde{\theta}_t^k, \hat{\theta}_{t(\zeta_1, \ldots, \zeta_k)}^k \right) \right|^r \leq \frac{k-1}{K-1} \xi_r \qquad (10)$$

The weight $(k-1)/(K-1)$ reflects the nature of the bias increase. It is also worth mentioning that the above expressions deviate from (7), i.e. the adaptive estimator replaces the parameter set θ^*. The bias due to this replacement is controlled for by the critical values.

The sequential testing procedure is adopted to compute the critical values. At step $k = 1$, we set $\zeta_1 = \infty$ in agreement with the homogeneity of the shortest interval I^1, which delivers the result $\hat{\theta}_t^1 = \tilde{\theta}_t^1$. In selection of ζ_2, we set all the remaining $\zeta_k = \infty$ for $k \geq 3$ to specify the contribution of ζ_2. We choose the minimal value of ζ_2 satisfying the following risk function:

$$\mathsf{E}_{\theta^*} \left| LR \left(I^k, \tilde{\theta}_t^k, \hat{\theta}_{t(\zeta_1, \zeta_2)}^k \right) \right|^r \leq \frac{1}{K-1} \xi_r, \qquad k = 2, \ldots, K.$$

Consequently with $\zeta_1, \zeta_2, \ldots, \zeta_{k-1}$ fixed, we select the minimal value of ζ_k for $k = 3, \ldots, K$ which fulfills:

$$\mathsf{E}_{\theta^*} \left| LR \left(I^l, \tilde{\theta}_t^l, \hat{\theta}_{t(\zeta_1, \zeta_2, \ldots, \zeta_k)}^l \right) \right|^r \leq \frac{k-1}{K-1} \xi_r, \qquad l = k, \ldots, K.$$

Figure 4 depicts the critical values calculated for the simulated AR(1) processes with $\theta^* = (-0.1197, 0.7754, 0.5634)^\top$. The parameters correspond to our real data set, i.e. they are estimates of an AR(1) model fitted to the logarithm of realized volatility of the S&P500 index data.

The choice of critical values depends on the parameter r. Belomestny and Spokoiny (2007) suggest to choose $r = 1/2$ in order to provide a stable performance of simu-
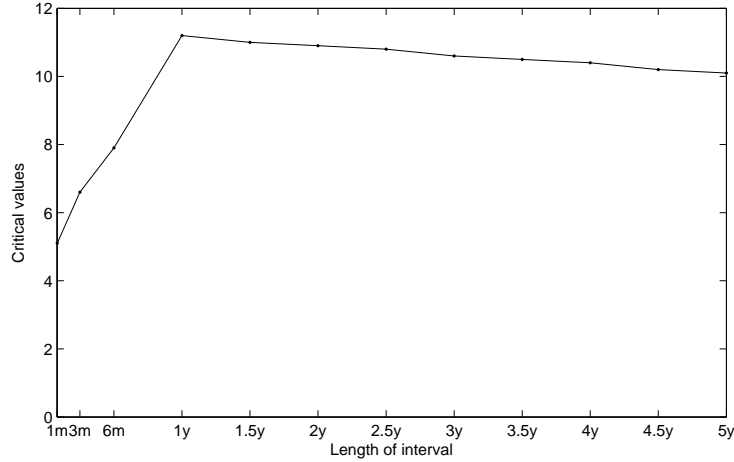
Figure 4: The set of critical values. They are based on $r = 1/2$ and on $\theta^* = (-0.1197, 0.7754, 0.5634)^\top$, which are calculated for the log realized volatility of the S&P500 index futures under the hypothesis of constant parameters in (5). The set of interval lengths is given on the X-axis.

lation, we here follow their recommendation. Note that the critical values also rely on the parameter θ^* in the simulation. In our study, we discuss two ways for selecting θ^*: in the first θ^* is estimated over the full sample period, whereas in the second θ^* is estimated at each time point using a rolling window with a fixed length. In general, a large rolling window size means that we put more attention to a time homogeneous situation. Such a choice leads to a rather conservative procedure with possibly low accuracy of estimation. On the contrary, a rolling window including fewer observations is more sensitive to structural shifts. The size of rolling window can be selected in a data driven way by minimizing some objective function, e.g., by minimizing the forecasting error. In Section 5 we report the prediction performance using both a constant set of critical values over all the observations and the time dependent sets with rolling windows including 1-month, 6-month, 1-year and 2.5-year observations. As expected, using the time dependent critical values increases the accuracy of prediction.

3.3 Simulation experiments

This section reports the performance of the adaptive estimation in a number of simulation experiments. The focus is the reaction of the new technique to a shift of parameters. The parameters estimated for the full S&P 500 realized volatility data serves again as a guidance. Several scenarios are studied here (see Table 2), by which one parameter varies over some time periods while the other two remain constant. Different size of shifts (big or small) followed by homogeneous intervals with different lengths (long and short) are examined. For example, case 1 involves a big shift in θ_{1t}, the intercept of the LAR(1) process, over a long homogeneous time span $[1501, 2800]$. For each case, we generated 500 LAR(1) processes with 3260 observations. The first 1260 observations, corresponding to $I^{13} = 5$ years, are used as training set. The

Table 2: Scenarios of the simulation study

1. $\theta_{1t}^* = \begin{cases} 1.1970 & t \in [1501, 2800] \\ -0.1197 & \text{otherwise} \end{cases}$	2. $\theta_{1t}^* = \begin{cases} 1.1970 & t \in [1501, 2000] \\ 0.3591 & t \in [2401, 2800] \\ -0.1197 & \text{otherwise} \end{cases}$	
3. $\theta_{2t}^* = \begin{cases} -0.7754 & t \in [1501, 2800] \\ 0.7754 & \text{otherwise} \end{cases}$	4. $\theta_{2t}^* = \begin{cases} -0.7754 & t \in [1501, 2000] \\ 0.6203 & t \in [2401, 2800] \\ 0.7754 & \text{otherwise} \end{cases}$	
5. $\theta_{3t}^* = \begin{cases} 0.1000 & t \in [1501, 2800] \\ 0.5634 & \text{otherwise} \end{cases}$	6. $\theta_{3t}^* = \begin{cases} 0.1000 & t \in [1501, 2000] \\ 0.4000 & t \in [2401, 2800] \\ 0.5634 & \text{otherwise} \end{cases}$	

Note: in each of the 6 scenarios only one parameter is changed either once or twice. The remaining ones are fixed to the values estimated from the full S&P500 realized volatility data, i.e. $(-0.1197, 0.7754, 0.5634)^\top$.

average value of the estimated parameters (solid line) and the pointwise 95% confidence intervals (shaded areas) are displayed in Figures 5 to 7. The true values of θ^* (dashed line) are depicted for judgement. It shows that in many cases the method reacts quickly to a big shift but slowly to a small shift. For example as the intercept
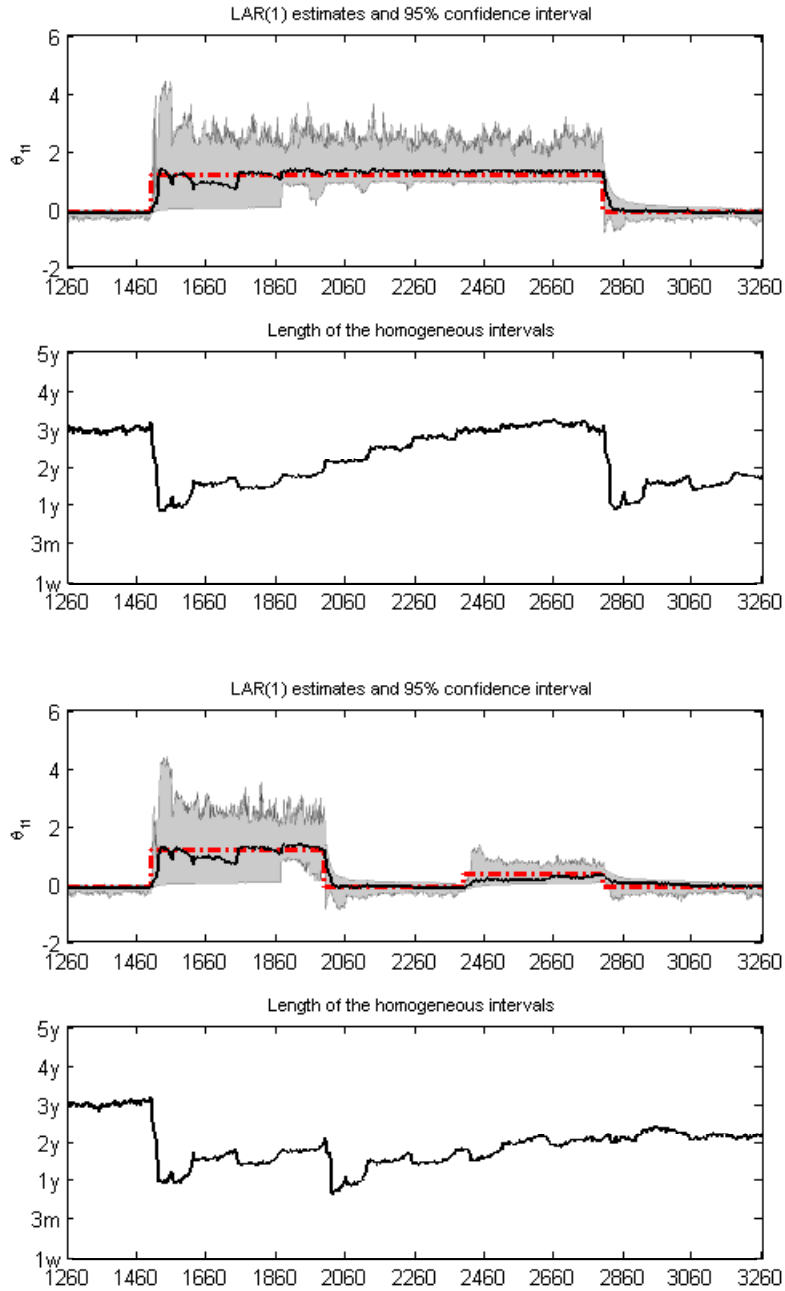
Figure 5: Simulation results for θ_{1t}: The red dashed line represents the process of the true time-varying parameter and the bold solid line is the average values of the estimated parameter over 500 simulations. The shaded area corresponds to the pointwise 95% confidence intervals. The average values of the selected homogeneous intervals for each time point are presented below each case of simulations.

Figure 6: Simulation results for θ_{2t}: The red dashed line represents the process of the true time-varying parameter and the bold solid line is the average values of the estimated parameter over 500 simulations. The shaded area corresponds to the pointwise 95% confidence intervals. The average values of the selected homogeneous intervals for each time point are presented below each case of simulations.
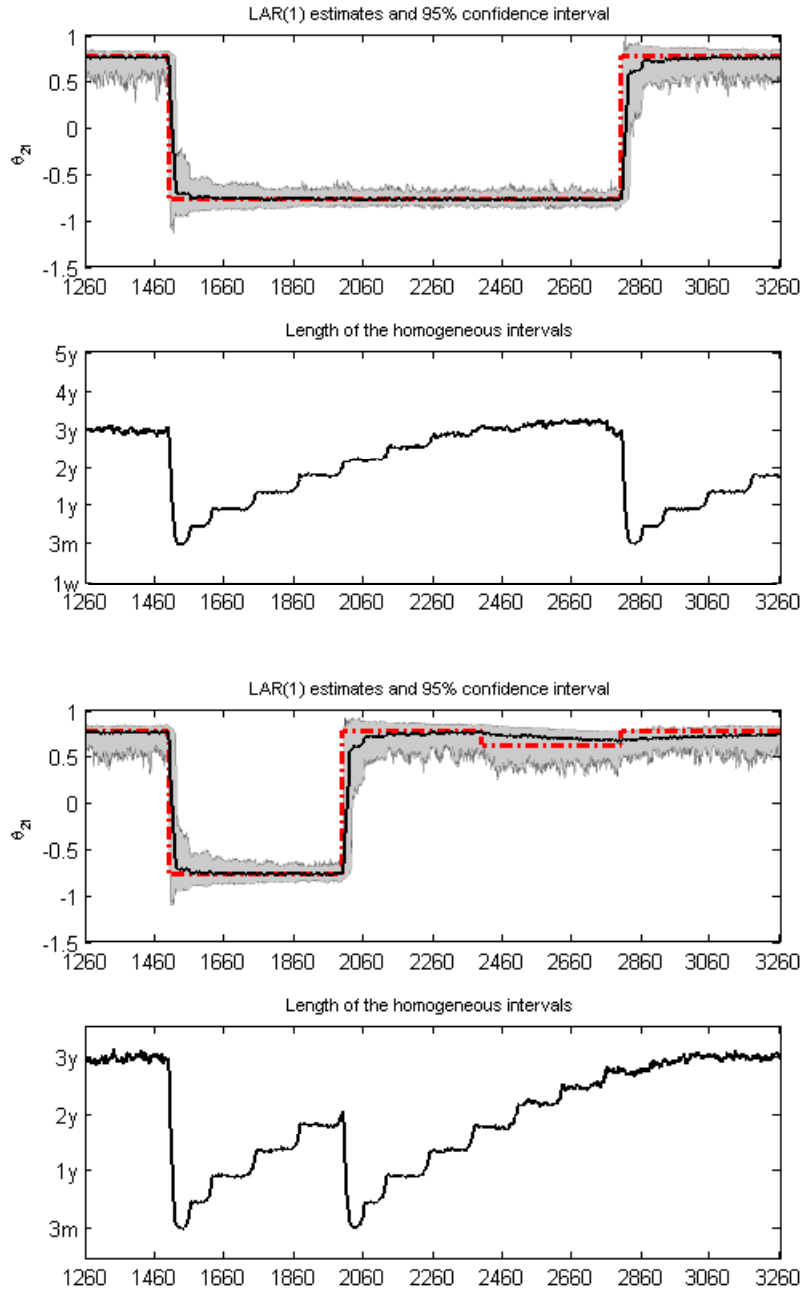
Figure 7: Simulation results for θ_{3t}: The red dashed line represents the process of the true time-varying parameter and the bold solid line is the average values of the estimated parameter over 500 simulations. The shaded area corresponds to the pointwise 95% confidence intervals. The average values of the selected homogeneous intervals for each time point are presented below each case of simulations.
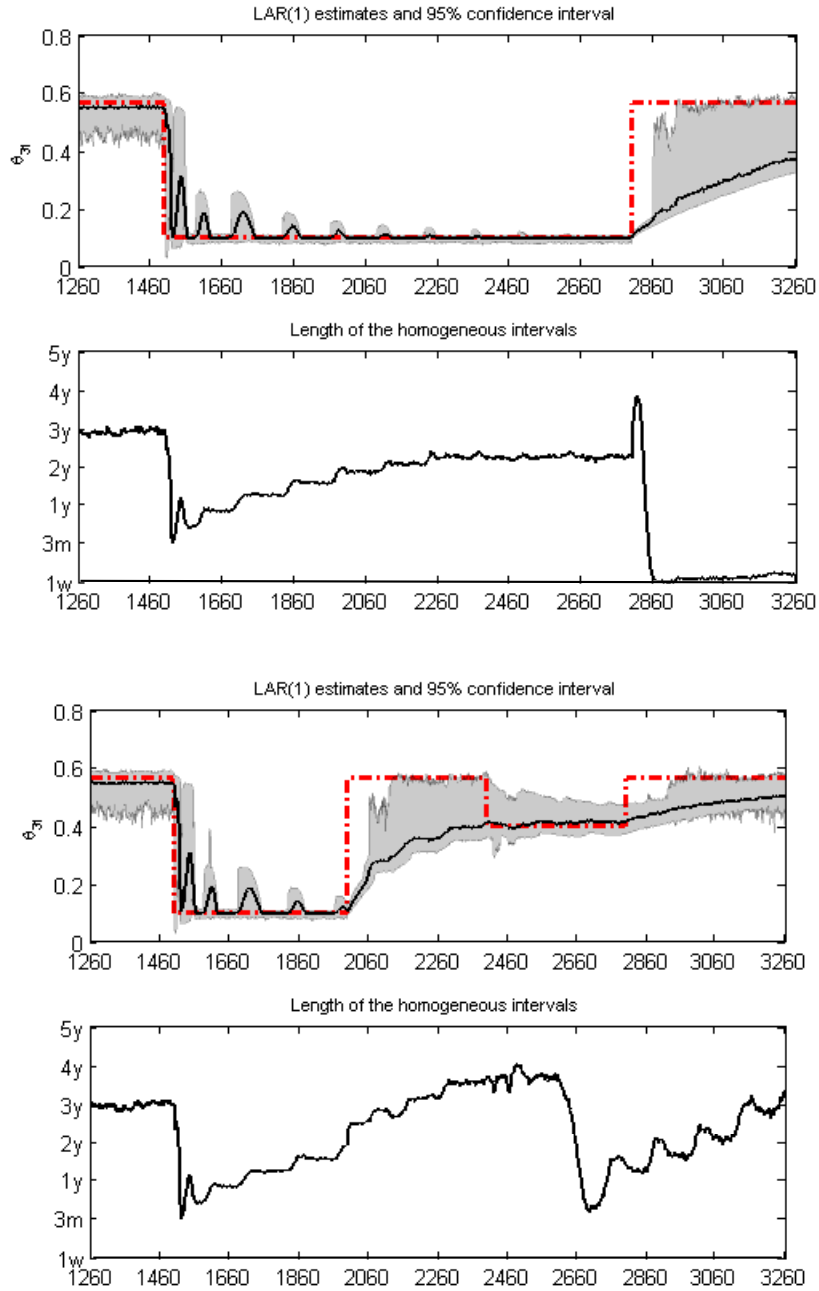
θ_{1t} jumps at $t = 1500$, it needs 20 steps to catch up 70% of the big shift. For the small shift at $t = 2400$, the procedure needs roughly 254 steps, see Figure 5. Similar results are obtained for the shifts in θ_{2t}. However, positive shifts in θ_{3t}, corresponding to an increase in the signal-to-noise ratio, are only slowly detected, see Figure 7.

The average values of the selected homogeneous intervals over the simulations for each time point are presented below each case, see Figures 5 to Figure 7. As expected, the homogeneous intervals are long when the parameters remain constant and become short sharply after a shift. It is also observed that the length of homogenous intervals seems to have no considerable influence on the adaptive estimation.

4 Long memory models

As we aim at a comparison to the long memory view of volatility, we briefly review here the most popular realized volatility models emanating from this view.

In contrast to the fractionally integrated GARCH models, in which volatility is a function of the daily squared return innovations that exhibits a hyperbolically decaying autocorrelation, the realized volatility literature applies fractionally integrated processes directly to the (more precise) realized volatility measure. Andersen et al. (2003), for example advocated the use of an ARFIMA(p, q) for modeling (logarithmic) realized volatility. The ARFIMA(p, q) model is given by

$$\phi(L)(1 - L)^d(\log RV_t - \mu) = \psi(L)u_t, \tag{11}$$

with $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$, $\psi(L) = 1 + \psi_1 L + \ldots + \psi_q L^q$, L denoting the lag operator, and $d \in (0, 0.5)$ is the fractional difference parameter. Given the empirical distributional properties of logarithmic realized volatility, u_t is usually assumed to

be a Gaussian white noise process, which facilitates the exact maximum-likelihood estimation of the model.

The HAR model aims at reproducing the observed volatility phenomenon. However, in contrast to the ARFIMA model, the HAR model is formally not a long memory model. Instead, the correlation structure is approximated by the sum of a few multi-period volatility components. The use of such components is motivated by the existence of heterogenous agents having different investment horizons (see Corsi, 2008; Müller, Dacorogna, Dav, Olsen, Pictet and von Weizsäcker, 1997). In particular, the HAR model put forward by Corsi (2008) builds on a daily, weekly and monthly component, which are defined by:

$$RV_{t+1-k:t} = \frac{1}{k}\sum_{j=1}^{k} RV_{t-j}. \tag{12}$$

with $k = 1, 5, 21$, respectively. The HAR model is then given by

$$\log RV_t = \alpha_0 + \alpha_d \log RV_{t-1} + \alpha_w \log RV_{t-5:t-1} + \alpha_m \log RV_{t-21:t-1} + u_t \tag{13}$$

with u_t typically being also Gaussian white noise. Maximum-likelihood estimation is straightforward. Interestingly, the HAR and ARFIMA models have been found to obtain a similar forecasting performance with both models outperforming the traditional volatility models based on daily returns (for the latter see e.g. Andersen, Bollerslev and Diebold, 2007; Koopman, Jungbacker and Hol, 2005).

5   Empirical analysis

We now turn to the empirical investigation of the dual views on the dynamics of volatility. We focus our analysis on realized volatility of the S&P500 index futures

from January 2, 1985 to February 4, 2005 (see Section 2). Like in the simulation exercise we use the first 5 years of our sample as a training set. For the local autoregressive procedure this means that January 2, 1990 is the first time point for which we estimate the LAR(1) model and that we allow the longest interval of homogeneity ($K = 13$) to be 5 years with the remaining set of subintervals given as in the simulation part, i.e. 1 week ($k = 1$), 1 month ($k = 2$), ..., 4.5 years ($k = 12$).

We estimate the LAR(1) model based on different sets of critical values. In particular, we first consider critical values obtained from a Monte Carlo simulation based on the parameter values of an AR(1) model being estimated over the full sample period. We refer to this as the global LAR(1) model. The other sets of critical values are obtained adaptively using a 1 month, 6 months, 1 year and 2.5 years sample period. Figure 8 shows the distribution of the lengths of the selected homogenous intervals over the evaluation period (January 2, 1990 to February 4, 2005) based on the global and the adaptive critical values, respectively. Obviously, the global LAR(1) model exhibits a slightly higher variation in the length of the selected intervals. Nevertheless, the average interval length is for nearly all LAR(1) models about 6 months, which indicates only a weak sensitivity of the interval selection procedure to the sample size used in the computation of the critical values. Interestingly, with the exception of the adaptive 1 month and 6 months LAR(1) models for which the median interval length is at $k = 3$, we find that the median is $k = 4$, which corresponds to 6 months of homogeneity.

We investigate the dual views by comparing the forecasting performance of the localized realized volatility procedure to ARFIMA and HAR models. To this end, we recursively compute one-step-ahead (logarithmic) realized volatility forecasts from all three model types over the evaluation period. We compute the ARFIMA and HAR forecasts based on a rolling window scheme, i.e. each forecast is based on an estimation
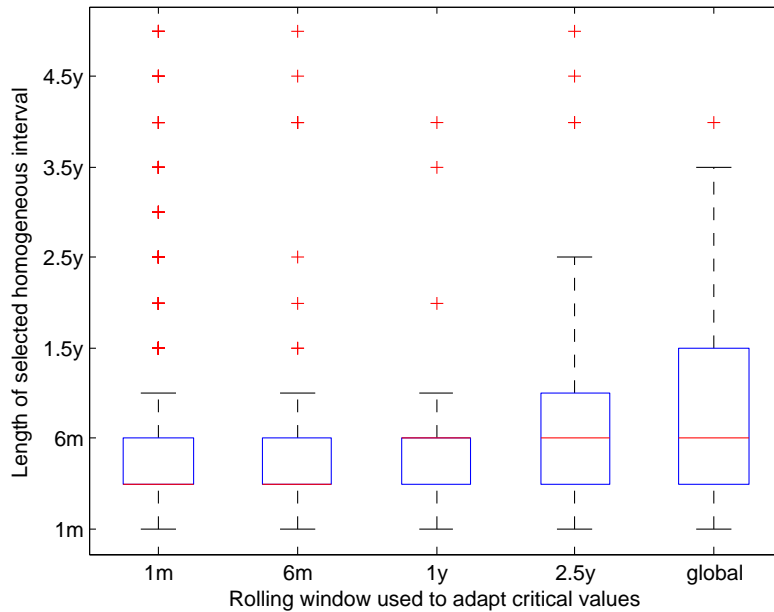
24

Figure 8: Boxplot of the homogenous intervals selected by the LAR(1) procedure with 1 month, 6 months, 1 year, 2.5 years adaptive critical values and the global LAR(1) procedure.

of the model over a constant number of observations. As it is sometimes argued, that both long memory and structural breaks are driving volatility, we attempt to account for this possibility by considering also smaller window sizes. Overall, we use 11 rolling window sizes ranging from 3 months to 5 years, which is broadly consistent with our choice of subintervals in the LAR(1) procedure. We additionally consider forecasts from constant AR(1) models based on the same rolling windows, as this allows for a direct evaluation of the relevance of the localization in our LAR(1) procedure. Recall, that the individual forecasts from the LAR(1) model are based on varying estimation periods, i.e. on the lengths of the homogenous time intervals that are selected at the forecast origins.

Table 3 shows the root mean square forecast errors (RMSFE) and the mean absolute forecast error (MAE) of the various models. Note, that the ARFIMA forecasts

25

Table 3: Forecast evaluation criteria

Sample period	RMSFE			MAE		
	LAR(1)			LAR(1)		
local adaptive 1m	0.4858	-	-	0.3667	-	-
local adaptive 6m	**0.4811**	-	-	**0.3654**	-	-
local adaptive 1y	0.4876	-	-	0.3704	-	-
local adaptive 2.5y	0.4916	-	-	0.3748	-	-
local global	0.5014	-	-	0.3824	-	-
	AR(1)	ARFIMA	HAR	AR(1)	ARFIMA	HAR
3m	**0.5149**	0.5328	0.5381	**0.3900**	0.3978	0.4025
6m	0.5288	0.5225	0.5240	0.3987	0.3902	0.3862
1y	0.5398	0.5178	0.5185	0.4057	0.3860	0.3857
1.5y	0.5462	0.5143	0.5172	0.4103	0.3836	0.3843
2y	0.5509	0.5133	0.5158	0.4136	0.3826	**0.3836**
2.5y	0.5555	0.5132	**0.5153**	0.4157	0.3816	0.3839
3y	0.5574	**0.5123**	0.5155	0.4177	**0.3814**	0.3843
3.5y	0.5607	0.5132	0.5164	0.4202	0.3820	0.3859
4y	0.5649	0.5129	0.5171	0.4238	0.3817	0.3851
4.5y	0.5686	0.5130	0.5173	0.4273	0.3821	0.3854
5y	0.5712	0.5129	0.5176	0.4300	0.3819	0.3858

The table reports the forecast evaluation criteria for 1-day-ahead forecast of the logarithmic realized variance of the S&P500 index futures based on the LAR(1), the constant AR(1), the ARFIMA and the HAR models. The first column refers either to the sample period used in the computation of the critical values in the LAR(1) procedure or to the rolling window sizes. Bold numbers indicate the minimum of the forecast evaluation criteria within each model class.

are based on an ARFIMA(2,d,0) specification, which was selected according to the Akaike as well as the Bayesian information criteria using the full sample period. Estimation and forecasting is carried out using the Ox ARFIMA 1.04 package, see Doornik and Ooms (2004), Doornik and Ooms (2006). Interestingly, our LAR(1) procedure provides the most accurate forecasts. This holds already for the forecasts based on the LAR(1) model with globally computed critical values. The performance can be further improved using adaptive critical values. More precisely, a reduction of

the sample period underlying the computation of the critical values introduces more flexibility into the procedure, which seems to result in an (albeit somewhat small) increase in forecast accuracy.

The direct comparison of the LAR(1) forecasts with those based on the constant AR(1) models also reveals, that the adaptive local selection of the homogenous intervals is indeed important. Obviously, the adaptive procedure, which determines at each time point the adequate length of the time interval over which the AR(1) model is appropriate, is superior. Noteworthy, for increasing window sizes the predictability of the constant AR(1) model worsens. This might be expected as for larger sample sizes, e.g. more than 2 years, the autocorrelation function of realized volatility exhibits more persistence and, thus, an AR(1) model tends to be misspecified.

For the same reason it is not surprising that the predictive performance of the long memory models increases when we consider larger rolling windows. Note also, that in accordance to the empirical results reported in the realized volatility literature so far, the HAR and ARFIMA models exhibit similar forecast accuracy with a slight tendency of the ARFIMA model to outperform the HAR model. Both models, however, are outperformed by the localized realized volatility method.

We further evaluate the predictive performance of the different realized volatility models on the grounds of the so–called Mincer–Zarnowitz regressions, i.e. by regressing the observed logarithmic realized volatility on the corresponding forecasts of model i:

$$\log RV_t = \alpha + \beta \log \widehat{RV}_{t,i} + \nu_t. \tag{14}$$

This allows us to test for the unbiasedness of the different forecasts. Table 4 reports the regression results along with the p-value of the F-test on unbiased forecasts, i.e. $H_0: \alpha = 0$ and $\beta = 1$. Note that for the ease of exposition, we solely focus on the

Table 4: Mincer-Zarnowitz regression results for log. realized volatility

Model	α	β	p-value	R^2
global LAR	-0.0130 (0.0125)	1.0128 (0.0142)	0.1007	0.6959
adaptive LAR, 1y	0.0025 (0.0123)	1.001 (0.0127)	0.9780	0.7117
1y AR(1)	-0.0010 (0.0144)	1.0117 (0.0158)	0.6002	0.4669
5y AR(1)	0.0221 (0.0162)	1.0367 (0.0213)	0.2216	0.6052
1y ARFIMA	0.0008 (0.0119)	1.001 (0.0132)	0.9962	0.6747
5y ARFIMA	0.0009 (0.0115)	1.0154 (0.0129)	0.4907	0.6811
1y HAR	-0.0076 (0.0128)	0.9907 (0.0128)	0.7509	0.6742
5y HAR	0.0145 (0.0119)	1.0237 (0.0133)	0.2036	0.6756

Reported are the estimation results of the Mincer-Zarnowitz regressions with heteroscedasticity and autocorrelation robust Newey-West standard errors given in parentheses. The third column reports the p-value of an F-test for $H_0: \alpha = 0$ and $\beta = 1$.

forecasts based on a moderately small sample (1 year) and a large sample (5 years). Correspondingly, we only consider the LAR(1) models based on the 1 year adaptively and on the globally computed critical values.

The results indicate that none of the forecasts is significantly biased at the 5% significance level. Noteworthy, the adaptive computation of the critical values seems to result in less systematic forecast errors. Similarly, the long memory and the constant AR(1) models exhibit smaller p-values for larger window sizes. The regression coefficients reported in Table 4 indicate a superior forecasting performance of the adaptive LAR(1) models. We investigate this result further and test for the significance of the observed differences in the forecast accuracies. In particular, we conduct a pairwise

28

test on the equality of the mean square forecast errors (MSFE) of the LAR(1) procedure and the other models (see Diebold and Mariano, 1995). To this end, we regress the difference between the squared forecast errors of the LAR(1) model and those of the competing model i, i.e. $e_{t,LAR}^2 - e_{t,i}^2$, on a constant μ. The null hypothesis of equal MSFEs is equivalent to $H_0 : \mu = 0$. Table 5 presents the test results. Obviously, the null hypothesis is always rejected in favor of a significant better forecasting performance of the adaptive LAR(1) model, as indicated by the significant negative estimate of μ. For the global LAR(1) model, however, we fail to reject the null.

6   Conclusion

This paper investigates a dual view on the long range dependence of realized volatility. While the current literature primarily advocates the use of long memory models to explain this phenomenon, we argue that volatility can alternatively be described by short memory models with structural breaks. To this end we propose the localized approach to realized volatility modeling where we consider the case of a dynamic short memory model. In particular, at each point in time we determine an interval of homogeneity over which the volatility is approximated by an AR(1) process. Our approach is based on local adaptive techniques developed in Belomestny and Spokoiny (2007), which make it flexible and allows for arbitrarily time-varying coefficients. This contrasts to smooth transition or regime switching models. Our procedure relies on parameters, that have to be predetermined but allow more flexibility. In particular, we show, that an adaptive view on intervals of homogeneity (and a decrease in the respective underlying sample size) is increasing the procedure's flexibility, yielding higher accuracy in estimation and a better forecasting performance. Furthermore, the choice of the underlying parameters can also be based upon criteria reflecting the

Table 5: Diebold Mariano test results

compared models	μ
(global LAR)-(1y AR(1))	-0.0400
	(0.0096)
(global LAR)-(5y AR(1))	-0.0141
	(0.0109)
(global LAR)-(1y ARFIMA)	-0.0168
	(0.0109)
(global LAR)-(5y ARFIMA)	-0.0118
	(0.0109)
(global LAR)-(1y HAR)	-0.0175
	(0.0104)
(global LAR)-(5y HAR)	-0.0165
	(0.0104)
(adaptive LAR, 1y)-(1y AR(1))	-0.0535
	(0.0097)
(adaptive LAR, 1y)-(5y AR(1))	-0.0546
	(0.0111)
(adaptive LAR, 1y)-(1y ARFIMA)	-0.0304
	(0.0108)
(adaptive LAR, 1y)-(5y ARFIMA)	-0.0253
	(0.0109)
(adaptive LAR, 1y)-(1y HAR)	-0.0310
	(0.0103)
(adaptive LAR, 1y)-(5y HAR)	-0.0258
	(0.0103)

Reported are test results for the Diebold Mariano test on equal forecast performance, i.e. $H_0 : \mu = 0$ in the regression $e_{t,LAR}^2 - e_{t,i}^2 = \mu + v_t$ with $e_{t,i}$ denoting the forecast error of model i. Heteroscedasticity and autocorrelation robust Newey-West standard errors are given in parentheses.

user's objective, such as in sample fit or forecasting criteria. Although we have refrained from doing so in our empirical application, we find that our adaptive localized realized volatility procedure provides accurate volatility forecasts and significantly outperforms the standard long memory realized volatility models. It seems that our alternative view on volatility is practical and realistic. Our forecast evaluation further suggests, that the adaptive localization is an important feature of our procedure, i.e. the locally adaptive selection of the homogenous intervals is superior to the specification of a short memory model that is assumed to be constant over a globally fixed period, such as e.g. an AR(1) model based on fixed rolling window sizes.

References

Andersen, T., Bollerslev, T., Diebold, F. and Ebens, H. (2001a). The distribution of realized stock return volatility, *Journal of Financial Economics* **61**: 43–76.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review* **39**: 885–905.

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility, *Review of Economics and Statistics* **89**: 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001b). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association* **96**: 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003). Modeling and forecasting realized volatility, *Econometrica* **71**: 579–625.

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **74**: 3–30.

Bandi, F. M. and Russell, J. R. (2005). Microstructure noise, realized volatility, and optimal sampling, *Review of Economic Studies* **75**: 339–369.

Barndorff-Nielsen, O. E. and Shephard, N. (2002a). Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society B* **64**: 253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2002b). Estimating quadratic variation using realized variance, *Journal of Applied Econometrics* **17**: 457–477.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica* **76**: 1481–1536.

Belomestny, D. and Spokoiny, V. (2007). Spatial aggregation of local likelihood estimates with applications to classification, *The Annals of Statistics* **35**: 2287–2311.

Chen, Y. and Spokoiny, V. (2009). Modeling and estimation for nonstationary time series with applications to robust risk management, *submitted*.

Čížek, P., Härdle, W. and Spokoiny, V. (2009). Statistical inference for time-inhomogeneous volatility models, *submitted*.

Corsi, F. (2008). A simple long memory model of realized volatility, *Journal of Financial Econometrics*. forthcoming.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy, *Journal of Business and Economic Statistics* **13**: 253–63.

Diebold, F. X. and Inoue, A. (2001). Long memory and regime switching, *Journal of Econometrics* **105**: 131–159.

Doornik, J. A. and Ooms, M. (2004). Inference and forecasting for arfima models, with an application to us and uk inflation, *Studies in Nonlinear Dynamics and Econometrics* **8**: Article 14.

Doornik, J. A. and Ooms, M. (2006). A package for estimating, forecasting and simulating arfima models: Arfima package 1.04 for ox, *http://www.doornik.com/download.html*.

Granger, C. W. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis* **1**: 5–39.

Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics* **14**: 227–238.

Granger, C. W. J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns, *Journal of Empirical Finance* **11**: 399–421.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise, *Journal of Business and Economic Statistics* **24**: 127–161.

Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*, Springer Verlag.

Hosking, J. R. M. (1981). Fractional differencing, *Biometrika* **68**: 165–176.

Koopman, S. J., Jungbacker, B. and Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements, *Journal of Empirical Finance* **12**: 445–475.

Mikosch, T. and Stărică, C. (2004a). Changes of structure in financial time series and the GARCH model, *REVSTAT Statistical Journal* **2**: 41–73.

Mikosch, T. and Stărică, C. (2004b). Non-stationarities in financial time series, the long range dependence and the IGARCH effects, *Review of Economics and Statistics* **86**: 378–390.

Müller, U. A., Dacorogna, M. M., Dav, R. D., Olsen, R. B., Pictet, O. V. and von Weizsäcker, J. E. (1997). Volatilities of different time resolutions - analyzing the dynamics of market components, *Journal of Empirical Finance* **4**: 213–239.

Polzehl, J. and Spokoiny, V. (2006). Propagation-separation approach for local likelihood estimation, *Probability Theory and Related Fields* **135**: 335–362.

Pong, S., Shackleton, M. B., Taylor, S. J. and Xu, X. (2004). Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models, *Journal of Banking & Finance* **28**: 2541–2563.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association* **100**: 1394–1411.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates, *Journal of Business and Economic Statistics* **14**: 45–52.

SFB 649 Discussion Paper Series 2009

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.