

# Recursive estimation of piecewise constant volatilities<sup>1</sup>

by

**Christian Höhenrieder**

Deutsche Bundesbank, Berliner Allee 14  
D-40212 Düsseldorf, Germany,

**Laurie Davies**

Fakultät Mathematik, Universität Duisburg-Essen  
D-45117 Essen, Germany  
Laurie.Davies@uni-essen.de

and

**Walter Krämer**

Fakultät Statistik, Universität Dortmund  
D-44221 Dortmund, Germany  
walterk@statistik.uni-dortmund.de

Version April 2009

## Summary

Returns of risky assets are often modelled as the product of a volatility function times standard Gaussian noise. This paper proposes a piecewise constant volatility function and shows how to construct such functions so that (i) the number of intervals of constant volatilities is minimized, and that (ii) these constant volatilities are equal to the root mean squared returns.

---

<sup>1</sup>Research supported by Deutsche Forschungsgemeinschaft (DFG). The algorithms suggested here were programmed in R and are available from the authors upon request.

# 1 The model and notation

Let  $r(t)$  be the excess-return of some risky asset in period  $t$ . For stocks with end of period price  $p_t$ ,  $r(t) = \log(p_t/p_{t-1})$ . In empirical finance,  $r(t)$  is often modelled as

$$R(t) = \Sigma(t) \cdot Z(t), \quad (t = 1, \dots, n) \quad (1)$$

where  $Z(t) \sim \text{i.i.d. } N(0, 1)$ . In the ARCH-class of parametric models  $\Sigma(t)$  depends on past values of the  $R(t)^2$  as typified by the GARCH(1,1) model

$$\Sigma(t)^2 = \alpha_0 + \alpha_1 R(t-1)^2 + \beta_1 \Sigma(t-1)^2.$$

The present paper in contrast follows a nonparametric approach to the modelling of  $\Sigma(t)$  based on Davies and Kovac (2001), Davies (2005, 2006) and Hönenrieder (2008). Related nonparametric approaches are those of Mercurio and Spokoiny (2004) and Granger and Stărică (2005).

The exact manner of the approximation can be carried out in various ways. Here we follow Davies (2005, 2006) and exploit the fact that, under the model (1) we have

$$\frac{R(t)^2}{\Sigma(t)^2} = \sum_{t \in I} Z(t)^2 \sim \chi_{|I|}^2 \quad (2)$$

for any nonempty interval  $I \subset \{1, \dots, n\}$  where  $|I|$  denotes the number of elements of  $I$  and  $\chi_k^2$  denotes the  $\chi^2$ -distribution with  $k$  degrees of freedom. This implies that, for all  $\alpha \in (0, 1)$ , there is some  $\alpha_n \in (0, 1)$  such that

$$P \left( \chi_{|I|, \frac{1-\alpha_n}{2}}^2 \leq \sum_{t \in I} Z(t)^2 \leq \chi_{|I|, \frac{1+\alpha_n}{2}}^2, \forall I \subset \{1, \dots, n\} \right) = \alpha \quad (3)$$

where  $\chi_{k, \beta}^2$  denotes the  $\beta$ -quantile of the  $\chi^2$ -distribution with  $k$  degrees of freedom. Simple approximations for  $\alpha_n$  for  $\alpha = 0.9$  and  $0.95$  and  $100 \leq n \leq 20000$  are given in Section 5. We define

$$\mathcal{A}_n = \left\{ \sigma : \sigma : \{1, \dots, n\} \rightarrow (0, \infty) \right. \\ \left. \chi_{|I|, \frac{1-\alpha_n}{2}}^2 \leq \sum_{t \in I} \frac{R(t)^2}{\sigma(t)^2} \leq \chi_{|I|, \frac{1+\alpha_n}{2}}^2, \forall I \subset \{1, \dots, n\} \right\}. \quad (4)$$

For data generated under (1) it follows that  $P(\Sigma \in \mathcal{A}_n) = \alpha$  so  $\mathcal{A}_n$  is a universal, exact  $\alpha$ -confidence region for the volatility  $\Sigma$  (see Davies et al (2009)). For real data  $r(t)$  which may or may not have been generated

under the model we refer to  $\mathcal{A}_n$  with the  $R(t)$  replaced by  $r(t)$  as an  $\alpha$ -approximation region. It may be checked that, for  $\alpha_n > 0.5$  (which will always be the case), we have

$$\chi_{|I|, \frac{1-\alpha_n}{2}}^2 < |I| < \chi_{|I|, \frac{1+\alpha_n}{2}}^2, \quad (5)$$

for all non-empty intervals  $I \subset \{1, \dots, n\}$ . This implies that  $\tilde{\sigma}(t) := |r_t|$  lies in  $\mathcal{A}_n$ , which is consequently a nonempty set whatever the data. The problem then becomes one of specifying one or more functions  $\sigma \in \mathcal{A}_n$  which reflect aspects of the data which are of interest. This is done by regularization.

The form the regularization takes is dictated by the problem under investigation and also by the practicability of being able to carry it out. In many nonparametric problems the regularization is done in terms of shape or smoothness but neither of these is appropriate for modelling risky returns. In line with Davies (2005, 2006) we take a sparsity approach and look for functions  $\sigma(t)$  which are piecewise constant on intervals and such that the number of intervals is minimized. For real data  $r(t)$  we now have to minimize the number of intervals of constancy of  $\sigma(t)$  subject to  $\sigma \in \mathcal{A}_n$  that is

$$\chi_{|I|, \frac{1-\alpha_n}{2}}^2 \leq \sum_{t \in I} \frac{r(t)^2}{\sigma(t)^2} \leq \chi_{|I|, \frac{1+\alpha_n}{2}}^2, \quad \forall I \subset \{1, \dots, n\}. \quad (6)$$

As it stands this problem is computationally too difficult. We show below how a modified version can be solved and how additional constraints can be placed on the solution. We illustrate the method using 19260 daily returns of the Standard and Poor's index from 1928 to 2000 and 9569 daily returns of the German DAX stock indices from 4.1.1960 - 30.4.1998.

## 2 Minimizing the number of intervals

### 2.1 Local adequacy

As mentioned above it is computationally not feasible to solve (6). We therefore consider the following modified version. Let  $I_1, \dots, I_k \subset \{1, \dots, n\}$ ,  $I_\nu \cap I_\mu = \emptyset$  and  $I_1 \cup \dots \cup I_k = \{1, \dots, n\}$  be the intervals where  $\sigma(t)$  is constant, with value  $\sigma_{I_\nu}$ , ( $\sigma_{I_\nu} > 0$ ). The inequalities in equation (6) imply

$$\frac{\sum_{t \in J} r(t)^2}{\chi_{|J|, \frac{1+\alpha_n}{2}}^2} \leq \sigma_{I_\nu}^2 \leq \frac{\sum_{t \in J} r(t)^2}{\chi_{|J|, \frac{1-\alpha_n}{2}}^2}, \quad \forall J \subset I_\nu, \nu = 1, \dots, k. \quad (7)$$

This lead to the modified approximation region  $\mathcal{A}_n^*$  given by

$$\mathcal{A}_n^* = \left\{ \sigma : \sigma : \{1, \dots, n\} \rightarrow (0, \infty) \chi_{|I|, \frac{1-\alpha n}{2}}^2 \leq \sum_{t \in I} \frac{R(t)^2}{\sigma(t)^2} \leq \chi_{|I|, \frac{1+\alpha n}{2}}^2 \right. \\ \left. \forall I \subset I_\nu, I_\nu \text{ a constancy interval of } \sigma \right\}. \quad (8)$$

Clearly local adequacy is a weaker condition than (6). It is seen that  $P(\Sigma \in \mathcal{A}_n) \geq \alpha$  so  $\mathcal{A}_n^*$  is a universal, honest (Li (1989))  $\alpha$ -confidence region for the volatility  $\Sigma$ . A volatility function  $\sigma \in \mathcal{A}_n^*$  will be called locally adequate. It turns out that we can minimize the number of intervals for a locally adequate volatility function.

## 2.2 Algorithm 0

It follows from (7) that

$$\sigma_l^2(I_\nu) := \max_{J \subset I_\nu} \frac{\sum_{t \in J} R(t)^2}{\chi_{|J|, \frac{1+\alpha n}{2}}^2} \leq \sigma_{I_\nu}^2 \leq \min_{J \subset I_\nu} \frac{\sum_{t \in J} R(t)^2}{\chi_{|J|, \frac{1-\alpha n}{2}}^2} =: \sigma_u^2(I_\nu). \quad (9)$$

Given the left endpoint  $s_v$  of  $I_\nu$ , the lower bound  $\sigma_l^2(I_\nu)$  is an increasing function of the right endpoint  $t_v$ , and the upper bound  $\sigma_u^2(I_\nu)$  is a decreasing function of the right endpoint  $t_v$  of  $I_\nu$ . This suggests the following algorithm to obtain a locally adequate volatility function.

### Algorithm 0

*Starting with  $s_1 = 1, t_1 = 1$ , we let  $t_1$  increase until the upper bound  $\sigma_u^2(I_1)$  becomes smaller than the lower bound  $\sigma_l^2(I_1)$  at  $t_1 + 1$ . Setting  $s_2 = t_1 + 1$ , we then repeat the procedure until we reach the end of the sample at  $n$ .*

The algorithm requires the calculation of the  $\sigma_l^2(I_\nu)$  and  $\sigma_u^2(I_\nu)$  of (9) which can be done efficiently as follows. For  $I = \{s, \dots, t\}$  we write

$$\sigma_l^2(s, t) = \sigma_l^2(I), \quad \sigma_u^2(s, t) = \sigma_u^2(I). \quad (10)$$

We have

$$\sigma_l^2(s, t) = \max \left\{ \sigma_l^2(s+1, t), \sigma_l^2(s, t-1), \frac{\sum_{i=s}^t r(i)^2}{\chi_{t-s+1, \frac{1+\alpha n}{2}}^2} \right\} \quad (11)$$

and

$$\sigma_u^2(s, t) = \min \left\{ \sigma_u^2(s+1, t), \sigma_u^2(s, t-1), \frac{\sum_{i=s}^t r(i)^2}{\chi_{t-s+1, \frac{1-\alpha n}{2}}^2} \right\}. \quad (12)$$

**Remark 2.1.** *The complexity of this algorithm is  $O(n^2)$ . The actual running time depends on the data, the worst case being when the method results in a single interval. For daily returns over several years this is never the case so that actual running times are much shorter than the complexity would suggest; for the Standard and Poor's returns with  $n = 19260$  observations it takes about three seconds.*

The algorithm does not specify a volatility function except for the unlikely case that the upper and lower bounds are equal at the end of each interval. Because of this  $\sigma_{sp,n}^0$  we be interpreted as any volatility function which lies between the bounds. In this sense we have the following theorem.

**Theorem 2.2.** *Any volatility function  $\sigma_{sp,n}^0$  constructed above is locally adequate and has the minimum number of intervals of constancy amongst all locally adequate volatility functions.*

**Proof:** Assume that there exists another locally adequate volatility function  $\tilde{\sigma}$  with corresponding partition  $\tilde{I}_1, \dots, \tilde{I}_{\tilde{k}}$  with  $\tilde{k} < k$ . It is clear that  $\tilde{I}_1 \subset I_1$  and by induction it follows that  $\cup_{j=1}^i \tilde{I}_j \subset \cup_{j=1}^i I_j$ . Consequently

$$\{1, \dots, n\} = \cup_{j=1}^{\tilde{k}} \tilde{I}_j \subset \cup_{j=1}^{\tilde{k}} I_j \stackrel{\neq}{\subset} \{1, \dots, n\} \quad (13)$$

which is a contradiction.  $\square$

Figure 1 shows the first 280 absolute values of the Standard and Poor's returns together with the lower and upper bounds calculated with  $\alpha = 0.9999991$  which is the default value (see Section 5) for a data set of length 19260. At observation 277 the value of the lower bound is 0.008924 and that of the upper bound is 0.009062. At observation 278 the lower bound increases to 0.009162 and so is higher than the upper bound which remains constant. The first interval is therefore  $[1, 277]$ . Figure 2 shows the absolute daily returns of the German stock price index DAX with a volatility function  $\sigma_{sp,n}^{(0)}$  as follows:  $n = 9589$ ,  $\alpha_n = 0.9999975$  and the value of the volatility on the interval  $I_\nu$  set to  $\sqrt{(\sigma_l^2(I_\nu) + \sigma_u^2(I_\nu))/2}$ . There are 36 intervals of constant volatility.

The partitioning  $I_1 \cup I_2 \cup \dots \cup I_k = \{1, \dots, n\}$  obtained via the above procedure is in general not the only one with the smallest number of intervals of constancy. We could have started from the endpoint and moved to the left; for most data sets this would yield a different solution.



Figure 1: The first 280 absolute values of the Standard and Poor's returns together with the lower and upper bounds for the volatility.

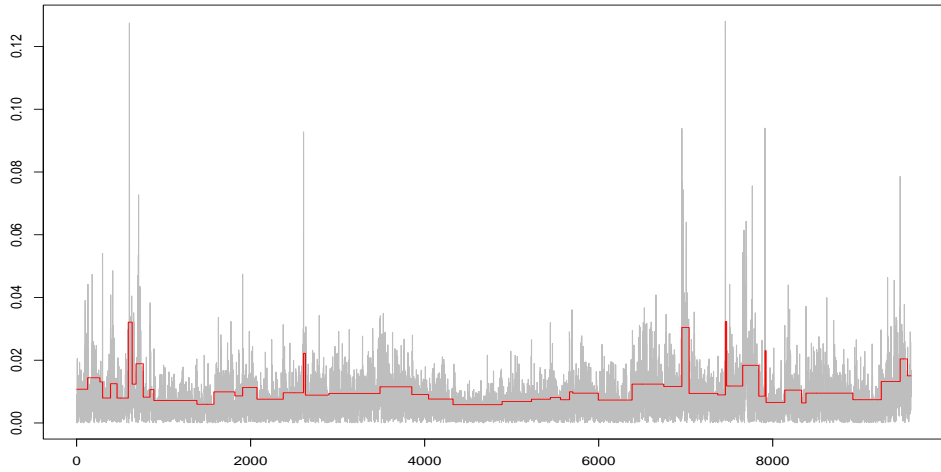


Figure 2: 9589 absolute daily returns of the German DAX-index together with the locally adequate piecewise constant volatility  $\sigma_{sp,n}^{(0)}$  described in the text with 36 intervals of constancy.

### 3 Empirical volatility and dynamic programming

#### 3.1 Empirical volatility

As already mentioned Algorithm 0 does not specify a value for the volatility function. For simplicity of interpretation there are grounds for requiring that its value is the empirical volatility on that interval

$$\sigma_{n,\{s,\dots,t\}}^2 = \sigma_{n,I}^2 := \frac{1}{t-s+1} \sum_{t=s}^t r(t)^2, \quad I = \{s, \dots, t\}. \quad (14)$$

If the end of period prices  $p(t)$  are generated by an underlying exponential Brownian Motion on  $[s, t]$ , and no micro-structure noise is present,  $\sigma_{n,I}^2$  will, after suitable normalization, tend in probability to the innovation variance of this Brownian Motion as the number of (equally spaced) observations in the  $[s, t]$ -interval increases. This follows immediately from the finite quadratic variation of Brownian motion and has engendered an enormous literature in empirical finance recently (see Anderson and Benzoni (2008) for an overview).

At first glance it may seem that we can simply modify Algorithm 0 of Section 2.2 as follows. Starting with  $t = 1$  we increase the interval by one point at a time until the empirical volatility no longer lies between the upper and lower bounds. We stop at the immediately preceding point and start with a new interval. Unfortunately this will not always yield the minimum number of intervals. Suppose that the first time that the empirical volatility does not lie between the two bounds is because it is too small. If we include the next observation this may be sufficiently large to increase the empirical volatility so that it now does lie between the bounds. This is not just a theoretical possibility as it occurs for the Standard and Poor's returns: the empirical volatility moves outside the bounds for the first time at time  $t = 233$ , stays outside the bounds until time  $t = 274$  when it moves back in again. The values of the empirical volatility and the lower and upper bounds at  $t = 274$  are 0.008380, 0.008139 and 0.009063 respectively. The optimization problem can be solved by adapting the dynamic programming algorithm of Friedrich et al. (2008). We describe the algorithm in the following section.

#### 3.2 Dynamic programming

The idea of the algorithm is the following. Given  $t$  suppose we have the optimal partition of each interval  $\{1, \dots, s\}$ ,  $1 \leq s \leq t$ . That is, for each  $s$  we have a partition of the interval  $\{1, \dots, s\}$  into the minimum number

disjoint intervals such that the empirical volatility on each interval  $I_\nu^{(s)}$  of the partition lies between the bounds  $\sigma_l^2(I_\nu^{(s)})$  and  $\sigma_u^2(I_\nu^{(s)})$  for that interval as defined by (9). Given this we show how it is possible to extend these optimal partitions from  $t$  to  $t + 1$ .

The empirical volatility is locally adequate on the interval  $\{s, \dots, t\}$  if

$$\sigma_l^2(s, t) \leq \frac{1}{t - s + 1} \sum_{\ell=s}^t r(\ell)^2 \leq \sigma_u^2(s, t). \quad (15)$$

We put

$$J_t = \{s \in \{1, \dots, t\} : (15) \text{ is satisfied}\} \quad (16)$$

$$\mathfrak{P}_t = \left\{ \{I_1, \dots, I_k\} : \sum_{\nu=1}^k I_\nu = \{1, \dots, t\}, \right. \\ \left. I_\nu \text{ satisfies (15), } \nu = 1, \dots, k \right\} \quad (17)$$

$$L_0 = 0 \quad (18)$$

$$L_t = \min_{|\mathcal{P}| \in \mathfrak{P}_t} |\mathcal{P}| \quad (19)$$

where  $\sum$  in (17) denotes a disjoint union and  $|\mathcal{P}|$  in (19) denotes the number of elements in  $\mathcal{P}$ .

**Theorem 3.1.** *If  $\alpha_n \geq 0.5$  then*

$$t \in J_t \quad (20)$$

$$\{\{1\}, \dots, \{t\}\} \in \mathfrak{P}_t \quad (21)$$

$$L_t = \min_{s \in J_t} L_{s-1} + 1 \quad (22)$$

**Proof:** The claims (20) and (21) follow from

$$\chi_{1, \frac{1-\alpha_n}{2}}^2 < 1 < \chi_{1, \frac{1+\alpha_n}{2}}^2$$

for  $\alpha_n \geq 0.5$  and (22) follows from

$$\mathfrak{P}_t = \{\mathcal{P} \cup \{\{s, \dots, t\}\} : s \in J_t, \mathcal{P} \in \mathfrak{P}_{s-1}\}.$$

□

**Theorem 3.2.** *Let  $p_1, \dots, p_n$  satisfy*

$$p_t \in \arg \min_{s \in J_t} L_{s-1}, \quad 1 \leq t \leq n \quad (23)$$



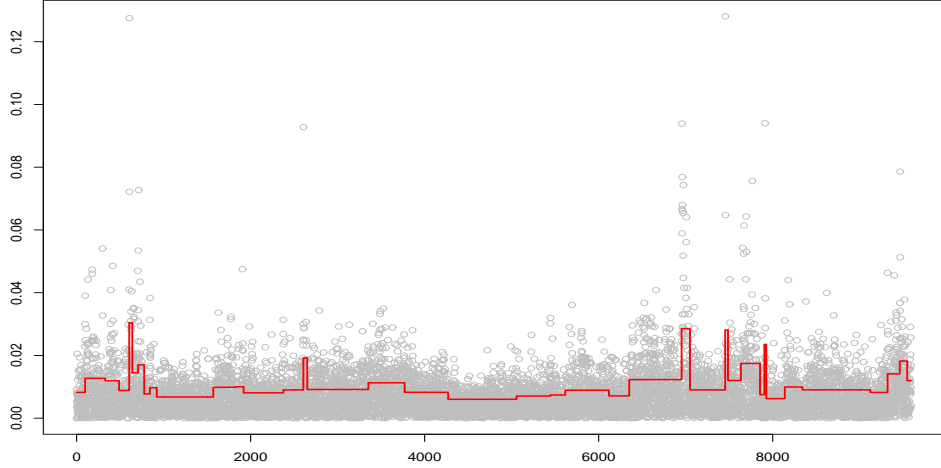


Figure 3: The absolute daily returns of the German DAX-index together with the locally adequate piecewise constant empirical volatility function  $\sigma_{sp,n}^{(1)}$  with 38 intervals of constancy.

and put recursively

$$t_{L_n} := n, s_{L_n} := p_{t_{L_n}}, \quad t_{\nu-1} := s_{\nu} - 1, s_{\nu-1} := p_{t_{\nu-1}}.$$

Then the partition  $I_{\nu} = \{s_{\nu}, \dots, t_{\nu}\}, \nu = 1, \dots, L_n$  is a partition which satisfies (15) and the resulting volatility function  $\sigma_{sp,n}^{(1)}$  has the minimum number of intervals amongst all locally adequate empirical volatility functions.

**Proof:** This is clear from the construction.  $\square$

As we have included an extra restriction, the minimum number of intervals cannot now be less than the number obtained from Algorithm 0 of Section 2.2. For the DAX-data of Figure 2 the minimum number of intervals is now 38 against the 36 obtained from Algorithm 0. The result is shown in Figure 3.

### 3.3 Algorithm 1

We now give an algorithm to calculate the partition of Theorem 3.2 following Höhenrieder (2008). It requires the efficient calculation of  $\sigma_t^2(s, t)$  and  $\sigma_t^2(s, t)$  as given by (11). Remark 2.1 also applies to this algorithm.

---

**Algorithm 1:** Calculation of a piecewise constant locally adequate empirical volatility function with the minimum number of intervals of constancy.

---

**Input:** Sample size  $n \in \mathbb{N}$ , quadratic increments  $r(1)^2, \dots, r(n)^2 \in \mathbb{R}$ ,  
quantiles  $\chi_{k, \frac{1-\alpha_n}{2}}^2, \chi_{k, \frac{1+\alpha_n}{2}}^2 \in \mathbb{R}_+$  for  $k = 1, \dots, n$  (with given  $\alpha_n \in [0.5, 1)$ )

**Output:** Piecewise constant locally adequate empirical volatility function with  
minimum number of intervals  $\sigma_n(1), \dots, \sigma_n(n) \in \mathbb{R}$

**local:** Interval end points  $s, t \in \mathbb{N}$ , number of intervals  $L_0, \dots, L_n \in \mathbb{N}_0$ ,  
left end points  $p_1, \dots, p_n \in \mathbb{N}$ ,  $\sigma_{n,l}(s, t)^2, \sigma_{n,u}(s, t)^2 \in \mathbb{R}$  as in (10),  
 $\sigma_{n, \{s, \dots, t\}}^2 \in \mathbb{R}$  as in (14), index  $i \in \mathbb{N}$

```

begin
    /* Calculation of  $L_1, \dots, L_n$  and  $p_1, \dots, p_n$  */
     $L_0 \leftarrow 0$ ;
     $L_1 \leftarrow 1$ ;                                     /*  $\sigma_{n,l}(1, 1)^2 \leq \sigma_{n, \{1\}}^2 \leq \sigma_{n,u}(1, 1)^2$  (15) */
     $p_1 \leftarrow 1$ ;
    for  $t \leftarrow 2$  to  $n$  do
         $L_t \leftarrow L_{t-1} + 1$ ;                         /*  $\sigma_{n,l}(t, t)^2 \leq \sigma_{n, \{t\}}^2 \leq \sigma_{n,u}(t, t)^2$  (15) */
         $p_t \leftarrow t$ ;
        for  $s \leftarrow t-1$  to  $1$  do
            if  $\sigma_{n,l}(s, t)^2 \leq \sigma_{n, \{s, \dots, t\}}^2 \leq \sigma_{n,u}(s, t)^2$  then
                if  $L_{s-1} + 1 < L_t$  then
                     $L_t \leftarrow L_{s-1} + 1$ ;
                     $p_t \leftarrow s$ ;
                end
            end
        end
    end
    /* Calculation of  $\sigma_n(1), \dots, \sigma_n(n)$  */
     $t \leftarrow n$ ;
    while  $t > 0$  do
         $s \leftarrow p_t$ ;
        for  $i \leftarrow s$  to  $t$  do
             $\sigma_n(i) \leftarrow \sigma_{n, \{s, \dots, t\}}$ ;
        end
         $t \leftarrow s-1$ ;
    end
end

```

---

## 4 Minimizing the empirical quadratic deviations

### 4.1 The empirical quadratic deviations

In general there will not be a unique solution to the problem of minimizing the number of intervals for piecewise constrained empirical volatility. We therefore look for a solution  $\sigma_{sp,n}^{(2)}$  which is closest to the data in that it minimizes the sum of the empirical quadratic deviations

$$\sum_{t=1}^n (r(t)^2 - \sigma(t)_{sp,n}^2)^2 \quad (24)$$

amongst all empirical volatility functions  $\sigma_{sp,n}$  with the minimum number of intervals. This does not involve any new principles but it necessitates an additional search amongst the admissible partitions to find those which minimize the sum of the quadratic deviations. As the volatility function is constant on each interval of the partition and its value on each interval independent of the other intervals, the search can be considerably reduced. With  $\sigma_{n,I}$  as defined by (14) we write for  $I = \{s, \dots, t\}$

$$\text{abw}_{\{s,\dots,t\}} := \text{abw}_I := \sum_{i \in I} (r(i)^2 - \sigma_{n,I}^2)^2 \quad (25)$$

and

$$LS_0 = 0 \quad (26)$$

$$LS_t = \min_{\mathcal{P} \in \mathfrak{P}_t, |\mathcal{P}|=L_t} \sum_{I \in \mathcal{P}} \text{abw}_I. \quad (27)$$

Corresponding to (22) we have

$$LS_t = \min_{s \in J_t, L_{s-1}+1=L_t} (LS_{s-1} + \text{abw}_{\{s,\dots,t\}}) \quad (28)$$

which gives rise to

**Theorem 4.1.** *Let  $p_1, \dots, p_n$  satisfy*

$$p_t \in \text{argmin}_{s \in J_t, L_{s-1}+1=L_t} (LS_{s-1} + \text{abw}_{\{s,\dots,t\}}) \quad (29)$$

*and put recursively*

$$t_{L_n} := n, s_{L_n} := p_{t_{L_n}}, \quad t_{\nu-1} := s_{\nu} - 1, s_{\nu-1} := p_{t_{\nu-1}}.$$

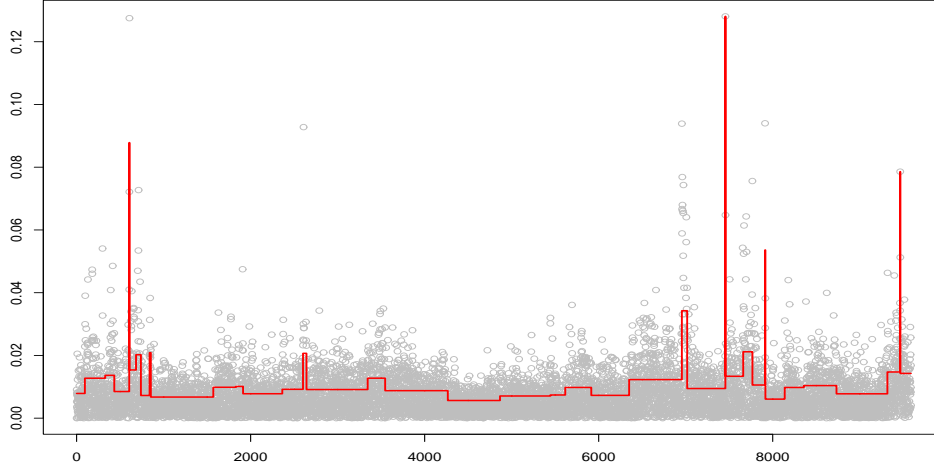


Figure 4: The absolute daily returns of the German DAX-index together with the locally adequate piecewise constant empirical volatility function  $\sigma_{sp,n}^{(2)}$  with 38 intervals of constancy which minimizes the sum of the empirical quadratic deviations (24).

*Then the partition  $I_\nu = \{s_\nu, \dots, t_\nu\}, \nu = 1, \dots, L_n$  satisfies (15) defines a locally adequate empirical volatility function which has the minimum number of intervals and minimizes the sum of the empirical quadratic deviations (24) amongst all such functions.*

**Proof:** Again this is clear from the construction. □

Figure 4 shows the result of applying this procedure to the DAX-data of Figure 2. It should be compared with Figure 3 which has the same number of intervals, namely 38. The sums of the empirical quadratic deviations are 0.001308 for Figure 3 and 0.000813 for Figure 4. The advantage of the minimization of the quadratic deviations is apparent: the solution is closer to the data with the consequence that structural breaks are not smoothed away. Of the 38 intervals of constancy the shortest is of length one (observation 7455 – ‘Black Monday’ 1987) and the longest is of length 720 (observations 3546-4265).

## 4.2 Algorithm 2

We now give an algorithm to calculate the partition of Theorem 4.1 following Höhenrieder (2008). If the data have been generated under the model (1) then the solution is unique with probability one. For real data of finite precision it is possible that there is more than one solution but we take this to be highly improbable. The complexity of the algorithm is  $O(n^3)$  because of the required search amongst the adequate empirical volatility functions with the minimum number of intervals. Again, the worst case in terms of running time is when the solution is exactly one interval. The running time for the Standard and Poor's returns with  $n = 19260$  and 71 intervals is about four seconds.

## 5 Choice of $\alpha_n$

The use of local adequacy as the criterion for accepting a model has as a consequence that the number of intervals for which the model is tested depends on the data. To overcome this the method can be calibrated for data generated as an exponential Brownian motion by taking  $\alpha_n m$  to be defined by (3) for a given  $\alpha$ . For such data the volatility is constant and we choose  $\alpha_n$  so that the method returns exactly one interval with probability  $\alpha$ . The values of  $\alpha_n$  may be determined by simulations. Figure 5 shows a plot of  $-\log(1 - \alpha_n)$  against  $\log n$  for  $n = 100, 250, 500, 1000, 2500, 5000, 10000, 20000$  and with  $\alpha = 0.90$ . The points were fitted by linear regression to the linear fit

$$\log(1 - \alpha_n) - \log n = a_0 + a_1 \log \log n$$

with the maximum absolute deviation being 0.0532 for  $\alpha = 0.9$  and 0.0384 for  $\alpha = 0.95$ . The resulting simple functions for  $\alpha_n$  are

$$\alpha_n = 1 - 0.0343 \exp(-0.286 \log \log(n))/n, \quad \alpha = 0.90, \quad (30)$$

and

$$\alpha_n = 1 - 0.0175 \exp(-0.329 \log \log(n))/n, \quad \alpha = 0.95. \quad (31)$$

The default value for  $\alpha_n$  we use is (30). Unfortunately it is a non-trivial problem to derive the correct asymptotic behaviour of  $\alpha_n$ . Even in the case of Gaussian white noise where the suitably normalized sums over intervals  $\sum_{t \in I} Z(t)/\sqrt{|I|}$  are all  $N(0, 1)$ , the correct asymptotic behaviour of the maximum is not easy to derive. Surprisingly it is not sufficient to embed the  $Z_i$  in a Brownian motion (see Kabluchko (2007)). Nevertheless the accuracy of the above approximations in the range  $n = 100$  to  $n = 20000$  gives hope that they continue to hold for much larger values of  $n$ .



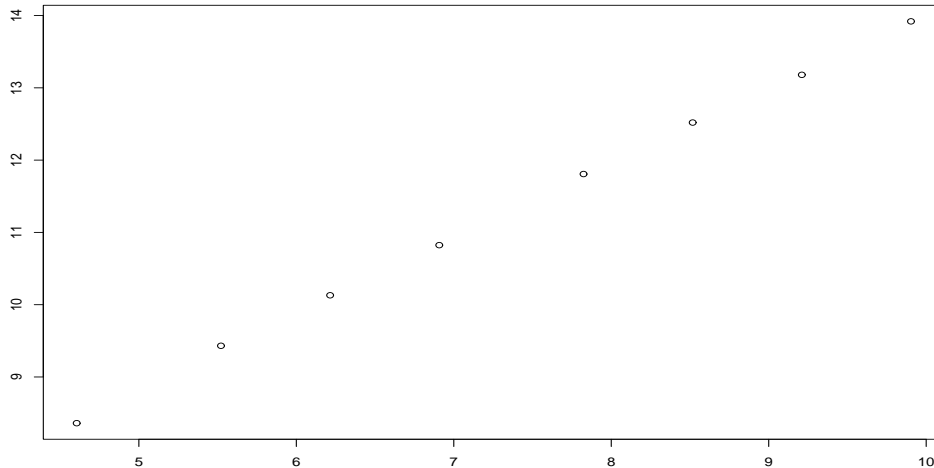


Figure 5: The plot of  $\log(1 - \alpha_n)$  against  $\log n$  for  $\alpha = 0.90$ .

## 6 Possible extensions

This paper took (1) as its basic model for the daily returns of a financial asset. One obvious way of extending the model is to allow for other distributions for the  $Z(t)$  rather than the standard normal. Obvious candidates are the family of  $t$ -distributions. First results suggest that a better fit can be obtained with a  $t$ -distribution with 10 degrees of freedom but that the gains are in no manner substantial. Allowing for non-Gaussian  $Z(t)$  requires a more complex algorithm as there are no simple algorithms for calculating the quantiles of sums of squares of  $t$ -distributed random variables. For many data sets it is clear that  $Z(t)$  cannot be modelled as a symmetric random variable but again no substantial gains can be expected by allowing for asymmetric  $Z(t)$ . This does not cause an increase in algorithmic complexity. Another possibility is to minimize the number of local extreme values of  $\sigma(t)$  (see Davies and Kovac (2001)) but this is probably less interesting for financial data where one major concern is the identification of structural breaks which may have external causes. The one advantage of minimizing the number of local extreme values is that use can be made of the taut string algorithm which is very fast and efficacious.

## References

- T. G. Anderson and L. Benzoni (2008):** "Realized volatility." Federal Reserve Bank of Chicago working paper, **87**,3, 2008-2014.
- P. L. Davies (2005):** "Universal Principles, Approximation and Model Choice." Invited talk, European Meeting of Statisticians, Oslo.
- P. L. Davies (2006):** "Long range financial data and model choice." Technical report 21/2006. SFB 475. Universität Dortmund.
- P. L. Davies and A. Kovac (2001):** "Local extremes, runs, strings and multiresolution. (with discussion)" *Annals of Statistics* **29**, 1–65.
- P. L. Davies, A. Kovac and M. Meise (2009)** "Nonparametric Regression, Confidence Regions and Regularization." *Annals of Statistics*. to appear.
- F. Friedrich, A. Kempe, V. Liebscher and G. Winkler (2008):** "Complexity penalized M-estimation: Fast computations." *Journal of Computational and Graphical Statistics* **17**, 201-224.
- C. Granger and C. Stărică (2005):** "Nonstationarities in Stock Returns." *The Review of Economics & Statistics* **87**,3, 523-538.
- C. Höhenrieder (2008)** "Nichtparametrische Volatilitäts- und Trendapproximation von Finanzdaten". Ph.D thesis, Department of Mathematics, University Duisburg-Essen, Germany.
- Z. Kabluchko (2007)** "Extreme-value analysis of standardized Gaussian increments." arXiv:0706.1849.
- K. C. Li (1989):** "Honset confidence regions for nonparametric regression." *Annals of Statistics* **17**, 1001–1008.
- D. Mercurio and V. Spokoiny (2004):** "Statistical inference for time-inhomogeneous volatility models." *Annals of Statistics* **32**, 577-602.