UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Center For International Economics

Working Paper Series

Working Paper No. 2013-08

Double-conditional smoothing of high-frequency volatility surface in a spatial multiplicative component GARCH with random effects

Yuanhua Feng

August 2013

Center For International Economics

Double-conditional smoothing of high-frequency volatility surface in a spatial multiplicative component GARCH with random effects

Yuanhua Feng[1]

Faculty of Business Administration and Economics
University of Paderborn, Germany

Abstract

This paper introduces a spatial framework for high-frequency returns and a faster double-conditional smoothing algorithm to carry out bivariate kernel estimation of the volatility surface. A spatial multiplicative component GARCH with random effects is proposed to deal with multiplicative random effects found from the data. It is shown that the probabilistic properties of the stochastic part and the asymptotic properties of the kernel volatility surface estimator are all strongly affected by the multiplicative random effects. Data example shows that the volatility surface before, during and after the 2008 financial crisis forms a *volatility saddle*.

Keywords: Spatial multiplicative component GARCH, high-frequency returns, double-conditional smoothing, multiplicative random effect, volatility arch, volatility saddle.

JEL Codes: C14, C51

[1]Corresponding Address: Yuanhua Feng, Department of Economics, University of Paderborn, Warburgerstr. 100, 33098 Paderborn, Germany
Tel. 0049 5251 60 3379; Fax: 0049 5251 60 5005; Email: yuanhua.feng@wiwi.upb.de

1 Introduction

Analysis of high-frequency returns can provide detailed features of market volatility. A widely used model in this context is introduced by Andersen and Bollerslev (1997, 1998) and Andersen et al. (2000), where the volatility is decomposed into a deterministic intraday component and a conditional daily component. This idea is extended in different ways. Feng and McNeil (2008) introduced a long-term deterministic component to deal with slowly changing volatility dynamics in high-frequency returns. Engle and Sokalska (2012), hereafter E&S, proposed a multiplicative component GARCH with an intraday GARCH (Engle, 1982 and Bollerslev, 1986) component. In this paper the idea of E&S will be extended in different ways for simultaneous modelling of deterministic long-term volatility dynamics, intraday variance patterns, daily and intraday conditional variances as well as possible multiplicative random effects in high-frequency returns, which we have found from the data.

The basic idea is to represent high-frequency returns under a spatial structure with the trading day defined as one dimension and the trading time on a day as the other. Although this proposal is just another illustration of the data, it provides a powerful framework for analyzing high-frequency volatility dynamics. Firstly, the deterministic component under this model forms a volatility surface, which shows an entire picture of long-term deterministic volatility dynamics and intraday volatility patterns and can be easily estimated. Note that the intraday seasonality considered in Andersen and Bollerslev (1997, 1998) and E&S can be thought of as the average curve of this volatility surface over all trading days, while the long-term deterministic volatility component in Feng and McNeil (2008) corresponds to the average curve of the volatility surface over all trading time points. Secondly, the proposed model also helps us to discover some detailed features of high-frequency returns, which cannot be found by the other models. Moreover, it is also possible to develop new GARCH models under this spatial framework.

In the literature, spatial models are usually applied to geological, ecological and social time series, because they focus on linkages between time series recorded at different locations. In recent years some spatial extensions of the GARCH model are proposed for modelling spatial dependence between financial time series. See e.g. Caporin and Paruolo (2006) and Borovkova and Lopuhaä (2009). Those GARCH models are indeed not intrinsic spatial models, because the definition of the spatial distance between different time series seems to be quite subjectively. However, the proposed framework for high-frequency data is a spatial model in its original sense. Now, the trading time points are considered as *locations*.

This paper focuses on estimating the deterministic volatility surface using bivariate kernel regression (Härdle, 1990, Scott, 1992 and Ruppert and Wand, 1994). Due to the huge number of observations, the traditional procedure runs very slowly. An equivalent but much quicker double-conditional smoothing technique is developed, where the data is first smoothed in one dimension. The intermediate smoothing results are then smoothed again in the other dimension. This idea adapts the double-smoothing for bandwidth selection in nonparametric regression (Müller, 1985, Heiler and Feng, 1998 and Feng and Heiler, 2009) to a dimension reduction technique. The smoothing results in the first stage also consist of useful information about some detailed features of the data, which cannot be found by the traditional procedure. In particular, these results indicate that high-frequency returns exhibit multiplicative random effects. Influence of this phenomenon is first discussed in detail under a simple specification of the error term with multiplicative random effects. To include possible conditional heteroskedasticity, a spatial multiplicative component GARCH with random effects is then proposed by introducing multiplicative random effects into the proposal of E&S. Properties of the stationary part and those of the conditional and unconditional sample variances of the stationary part are discussed in detail under this new model. Basic properties on the stationary part of the multiplicative component GARCH model of E&S are also obtained. It is shown that spatial stochastic processes (or random fields) with multiplicative random effects share the feature of a deterministic process. Now, the spatial autocovariances of the squared process do not decay to zero, if one of the two lags is zero. This in turn affects the variance of the sample variances. To discuss

The asymptotic variance, the rate of convergence and the optimal bandwidths of the kernel estimator of the volatility surface are hence all changed strongly. Now, the asymptotic variance of a kernel estimator converges much slower than that in the case without random effects. Furthermore, the smoothers in the first stage converge to the product of the true curves and a random scale and are hence inconsistent. Application to real data examples shows that the long-term volatility dynamics before, during and after the 2008 financial crisis at a given trading time point form a *volatility arch* (volatility bridge) with a very sharp peak, which together with the daily *volatility smiles* build a *volatility saddle*. Proposals in this paper also provide us new tools for analyzing market microstructures.

The spatial models with random effects are introduced in Section 2. The double-conditional smoothing is defined in Section 3. Properties of the error processes and the proposed kernel estimators are studied in Section 4. The application is reported in Section 5. Final remarks in Section 6 conclude the paper. Proofs of the results are given in the appendix.

2 The proposed models

This section introduces the multiplicative spatial model for equidistant high-frequency returns with different specifications on the multiplicative random effects and discusses some possible extensions of this model.

2.1 The spatial multiplicative component model

Let $r_{i,j}$ denote the high-frequency (log-)returns observed at n_t equidistant trading time points t_j, $j = 1, ..., n_t$, on the i-th trading day, where $i = 1, ..., n_x$. Although $r_{i,j}$ can be represented as a single time series with $n = n_x n_t$ observations, it is however more convenient to model $r_{i,j}$ directly based on a matrix form defined by the two indexes. Now, several features of high-frequency returns, which cannot be found under known models in the literature, can be easily discovered. Hence, we will introduce the following (lattice) spatial multiplicative component volatility model for $r_{i,j}$, where the lattice is defined by the trading days and the trading time points on a day:

$$r_{i,j} = n_t^{-1/2} \sigma(x_i, t_j) Y_{i,j}, \tag{1}$$

$i = 1, ..., n_x, j = 1, ..., n_t$, and

$$Y_{i,j} = \omega_i^{1/2} \lambda_j^{1/2} \varepsilon_{i,j}, \tag{2}$$

where $n_t^{-1/2}$ is a standardized factor to avoid the effect of the use of different frequencies, $x_i = (i - 0.5)/n_x$ is a re-scaled variable of the trading day, $\sigma^2(x, t) > 0$ is a slowly changing deterministic (unconditional) volatility surface, $Y_{i,j}$ denotes the stationary stochastic part, where $\omega_i > 0$ and $\lambda_j > 0$ are i.i.d. random variables with unit means and finite variances, and $\varepsilon_{i,j}$ are i.i.d. random variables with zero mean and unit variance. It is also assumed that ω_i, λ_j and $\varepsilon_{i,j}$ are mutually independent. We will see that $r_{i,j}$ defined in Model (1) are uncorrelated with each other. Throughout this paper we will define $\tilde{r}_{i,j} = n_t^{1/2} r_{i,j}$. The volatility surface illustrates the joint deterministic long-term and the intraday volatility dynamics in $\tilde{r}_{i,j}$. The two random variables ω_i and λ_j are introduced into the model to deal with possible multiplicative random effects found by our empirical analysis. It is easy to see that $Y_{i_1,j}$ and $Y_{i_2,j}$ are not independent. This is also true for Y_{i,j_1} and Y_{i,j_2}. We will see that the influence of the random effects are indeed even stronger than it is expected.

Models (1) and (2) provide a useful tool for modelling high-frequency returns observed over a long period, in particular for defining and estimating the high-frequency volatility

surface. After displaying the data in a 3D illustration we can immediately discover some interesting features of high-frequency returns, e.g. the effect of the financial crisis 2008 and the intraday volatility dynamics can already be discovered by eye. These features can be seen more clearly from the fitted volatility surface. See Figure 1 in Section 5, where the curves in the direction of the trading time indicate the estimated diurnal volatility patterns over all trading days. On the other hand, changes of the volatility surface over the observation period indicate the long-term volatility dynamics. Moreover, it was this model together with the double-conditional smoothing procedure defined in the next section, which helped us to discover the multiplicative random effects in high-frequency returns. See Figure 2.

2.2 Spatial multiplicative component GARCH with random effect

The specification in Model (2) allows us to discover the theoretical influences of the multiplicative random effects easily. However, possible conditional variance components in high-frequency data cannot be analyzed by this simple approach. Model (2) is hence extended by introducing corresponding random effects into that of E&S. This leads to

$$Y_{i,j} = \omega_i^{1/2} h_i^{1/2} \lambda_j^{1/2} q_{i,j}^{1/2} \varepsilon_{i,j}, \tag{3}$$

where h_i is a daily conditional variance component and $q_{i,j}$ are unit intraday GARCH components. Furthermore, we denote the intraday GARCH processes by $Z_{i,j} = q_{i,j}^{1/2} \varepsilon_{i,j}$. The daily volatility component h_i may be governed by a separate stochastic process. Models (1) and (3) together extend the multiplicative component GARCH of E&S in different ways. Firstly, the deterministic diurnal pattern, $\sigma(x_i, \cdot)$ for given i, is now allowed to change slowly over the trading days. And a long-term deterministic volatility trend, $\sigma(\cdot, t_j)$ for given j, is introduced, which is also allowed to change slowly over all trading time points. Secondly, multiplicative random effects in both dimensions are introduced into the stochastic part of their model. Following this approach, volatility in high-frequency returns is decomposed into different deterministic or stochastic components under the spatial framework. The model defined by (1) and (3) will hence be called *a spatial multiplicative component GARCH with random effects*. Note that the two GARCH components and the two random effect components in Model (3) play different roles, where ω_i and λ_j occur fully randomly, while h_i and $q_{i,j}$ obey some models and are predictable. If $h_i \equiv q_{i,j} \equiv 1$, Model (3) becomes (2). When $\omega_i \equiv \lambda_j \equiv 1$, Model (3) reduces to

$$Y_{i,j} = h_i^{1/2} q_{i,j}^{1/2} \varepsilon_{i,j}, \tag{4}$$

which is similar to that defined by Eq. (6) and (7) in E&S. Models (1) and (4) hence provide another simple spatial extension of their model with the intraday seasonality there being replaced by the entire nonparametric volatility surface, while keeping the stochastic part unchanged.

2.3 Possible extensions

Model (1) can be first extended by introducing a non-zero deterministic mean surface $\mu(x,t)$, which can be estimated similarly following the proposed procedure in the next section. It is well known that the error in $\hat{\mu}(x,t)$ usually does not affect the asymptotic properties of $\hat{\sigma}^2(x,t)$. The stochastic part $Y_{i,j}$ can also be specified in other ways. For instance, it may be of interest to introduce and study different daily GARCH components at different trading time points.

Although Model (1) is defined for equidistant high-frequency returns, it also applies to tick-by-tick high-frequency returns. Now, the double-conditional smoothing in the next section should be adapted properly and the intraday discrete GARCH components can e.g. be replaced by suitable continuous-time GARCH processes (Klüppelberg et al., 2004 and Brockwell et al., 2006). Furthermore, the proposals in this paper also apply to other kinds of high-frequency financial data, such as average transaction durations or trading volumes within time intervals with a given length, or non-financial data with suitable structure.

3 The double-conditional smoothing technique

In this paper we mainly focus on discussing the estimation of the volatility surface. Estimation of the parameters of the stochastic part will be studied elsewhere.

3.1 The proposed algorithm

Due to the huge number of observations the common bivariate kernel estimator of the volatility surface runs very slowly. An equivalent double-conditional smoothing technique is hence developed to reduce the computing time. This approach also provides us very useful additional information about the detailed features of high-frequency returns.

It is well known that $\sigma^2(x,t)$ can be thought of as a nonparametric regression in $\tilde{r}^2_{i,j}$, because Model (1) can be rewritten as follows

$$\begin{aligned} \tilde{r}^2_{i,j} &= \sigma^2(x_i,t_j) + \sigma^2(x_i,t_j)(Y^2_{i,j} - 1) \\ &= \sigma^2(x_i,t_j) + \sigma^2(x_i,t_j)\eta_{i,j}, \end{aligned} \tag{5}$$

where $\eta_{i,j} = Y^2_{i,j} - 1$ with $E(\eta_{i,j}) = 0$ and $\mathrm{var}\,(\eta_{i,j}) = \mathrm{var}\,(Y^2_{i,j})$. The autocorrelations of $\eta_{i,j}$ are also the same as those of $Y^2_{i,j}$. Model (5) is a nonparametric regression with heteroskedastic dependent errors, where $\sigma^2(x_i,t_j)$ is its regression and scale function at the same time. Another characteristic of (5) is that, given j, we have repeated observations at each point (x_i,t_j), and vice versa. And there is exactly one observation at each point.

For convenience, we assume that $t_j = (j-0.5)/n_t$ is the standardized trading time. Now, the spatial model is defined on an $n_x * n_t$ grid on $[0,1]\times[0,1]$. Assume that $\tilde{r}^2_{i,j}$ are arranged in the form of a single time series $\check{r}^2_l$ associated with the coordinates $(\check{x}_l, \check{t}_l)$, where $l = 1, ..., n$ and $n = n_x * n_t$ is the total number of observations. A well known approach for estimating $\sigma^2(x,t)$ is the bivariate kernel estimator (Härdle et al., 1998, and Feng and Heiler, 1998), which is defined by

$$\hat{\sigma}^2(x,t) = \sum_{l=1}^{n} \check{w}_l \check{r}^2_l, \tag{6}$$

where

$$\check{w}_l = K\left(\frac{\check{x}_l - x}{b_x}, \frac{\check{t}_l - t}{b_t}\right)\left[\sum_{l=1}^{n} K\left(\frac{\check{x}_l - x}{b_x}, \frac{\check{t}_l - t}{b_t}\right)\right]^{-1}, \tag{7}$$

where $K(u_x,u_t)$ is a bivariate kernel function, and b_x and b_t are the bandwidths for $\check{x}$ and $\check{t}$, respectively. The volatility surface is then given by $\hat{\sigma}(x,t) = \sqrt{\hat{\sigma}^2(x,t)}$. Under models (1) or (5), the kernel estimator can however be rewritten as

$$\hat{\sigma}^2(x,t) = \sum_{j=1}^{n_t}\sum_{i=1}^{n_x} w_{i,j}\tilde{r}^2_{i,j}, \tag{8}$$

where

$$w_{i,j} = K\left(\frac{x_i - x}{b_x}, \frac{t_j - t}{b_t}\right)\left[\sum_{r=1}^{n_x}\sum_{s=1}^{n_t} K\left(\frac{x_r - x}{b_x}, \frac{t_s - t}{b_t}\right)\right]^{-1}. \tag{9}$$

Assume that K is a product kernel $K(u_x,u_t) = K_1(u_x)K_2(u_t)$, then $\hat{\sigma}^2(x,t)$ can be further represented as

$$\hat{\sigma}^2(x,t) = \sum_{i=1}^{n_x}\sum_{j=1}^{n_t} w_{ix}w_{jt}\tilde{r}^2_{i,j}, \tag{10}$$

which is equivalent to

$$\hat{\sigma}^2(x,t) = \sum_{j=1}^{n_t} w_{jt}\hat{\sigma}^2(x|t_j) \tag{11}$$

or

$$\hat{\sigma}^2(x,t) = \sum_{i=1}^{n_x} w_{ix}\hat{\sigma}^2(t|x_i), \tag{12}$$

where

$$\hat{\sigma}^2(x|t_j) = \sum_{i=1}^{n_x} w_{ix}\tilde{r}_{ij}^2 \qquad \text{and} \qquad \hat{\sigma}^2(t|x_i) = \sum_{j=1}^{n_t} w_{jt}\tilde{r}_{ij}^2 \tag{13}$$

are two univariate kernel estimators with corresponding weights

$$w_{ix} = K_1\left(\frac{x_i - x}{b_x}\right)\left[\sum_{r=1}^{n_x} K_1\left(\frac{x_r - x}{b_x}\right)\right]^{-1},$$

$$w_{jt} = K_2\left(\frac{t_j - t}{b_t}\right)\left[\sum_{s=1}^{n_t} K_2\left(\frac{t_s - t}{b_t}\right)\right]^{-1}.$$

Note that $\hat{\sigma}^2(x|t_j)$ for given j is obtained with returns at the trading time point t_j over all trading days and $\hat{\sigma}^2(t|x_i)$ for given i is calculated with all intraday returns on the ith trading day, respectively. Each of the intermediate smoothers $\hat{\sigma}^2(x|t_j)$ and $\hat{\sigma}^2(t|x_i)$ obtained in the first stage consists of a panel of univariate kernel estimators obtained over one of the two explanatory variables conditioning on the other. The final estimators defined in (11) or (12) are hence called double-conditional kernel estimators, which provide two equivalent procedures of the double-conditional smoothing approach. It is obvious that the estimators defined by (6), (8) or (10) to (12) are all equivalent to each other. Note however that the first estimator applies to any bivariate kernel regression models but the others are only well defined under Model (1) or (5). Although the final estimate given by the double-conditional smoothing is exactly the same as that obtained by the standard bivariate kernel regression, the former exhibits a few important advantages.

3.2 Advantages and possible extensions

The obvious advantage of the double-conditional smoothing technique is that it is a dimension reduction technique, which transfers a two-dimensional smoothing into two one-dimensional smoothing procedures. It is easy to see that this approach runs much quicker than the standard one, in particular if the bandwidths are large and the estimation is carried out at all observation points. Detailed discussion on the computational advantage of the proposal is beyond the aim of the current paper.

Another more important advantage of the double-conditional smoothing technique is that the intermediate results provide useful information. For instance, $\hat{\sigma}^2(x|t_j)$ shows the long-term volatility trend at the time point t_j over all trading days and $\hat{\sigma}^2(t|x_i)$ stands for the estimated intraday volatility seasonality on the ith trading day. From Figure 2 in Section 5 we can see that $\hat{\sigma}^2(x|t_j)$ and $\hat{\sigma}^2(x|t_{j+1})$ differ to each other in a random way and the conditional smoothing of the intraday volatility patterns exhibits similar phenomenon. This was the motivation for the introduction of the multiplicative random effects. The fact that high-frequency returns may exhibit multiplicative random effects is almost impossible to find by means of the standard bivariate kernel approach.

The double-conditional smoothing can also be applied to estimate the volatility surface in tick-by-tick high-frequency returns. But now the first stage has to be carried out at n_t equidistant trading time points on each trading day. The second stage is the same as defined in (11). Some further possible extensions are for instance the adaptation of the algorithm to the use of non-product bivariate kernel functions, the definition of double-conditional local polynomial regression, the development of algorithms for further reduction of the computing time as well as the extension of this technique to high-dimensional data.

4 Main results

In this section properties of the processes defined in (2) and (3) and those of their sample variances are investigated. These results are then adapted to obtain the asymptotic variances of the proposed nonparametric estimators of the volatility surface.

4.1 Corresponding sample variances and necessary notations

Note that $Y_{i,j}$ has zero mean. The conditional and unconditional sample variances of $Y_{i,j}$ are hence the corresponding sample means of the squared observations $Y_{i,j}^2$. The sample variances conditioning on i and j, respectively, are given by

$$\hat{\sigma}_{i,\bullet}^2 = \frac{1}{n_t}\sum_{j=1}^{n_t} Y_{i,j}^2 \qquad \text{and} \qquad \hat{\sigma}_{\bullet,j}^2 = \frac{1}{n_x}\sum_{i=1}^{n_x} Y_{i,j}^2.$$

The unconditional sample variance is defined by

$$\hat{\sigma}_Y^2 = \frac{1}{n_x n_t}\sum_{i=1}^{n_x}\sum_{j=1}^{n_t} Y_{i,j}^2.$$

Moreover, we have

$$\hat{\sigma}_Y^2 = \frac{1}{n_x}\sum_{i=1}^{n_x}\hat{\sigma}_{i,\bullet}^2 = \frac{1}{n_t}\sum_{j=1}^{n_t}\hat{\sigma}_{\bullet,j}^2.$$

To describe the properties of these estimators, different nations, in particular those of the autocovariance functions (hereafter ACV) of different processes or components, are required. For instance, the ACV of the random fields $Y_{i,j}$, $Y_{i,j}^2$, $Z_{i,j}$ and $Z_{i,j}^2$ with lag k_1 and k_2 will be denoted by $\gamma_Y(k_1,k_2)$, $\gamma_{Y^2}(k_1,k_2)$, $\gamma_Z(k_1,k_2)$ and $\gamma_{Z^2}(k_1,k_2)$, respectively, and those of the components h_i, ω_i and λ_j with lag k by $\gamma_h(k)$, $\gamma_\omega(k)$ and $\gamma_\lambda(k)$, respectively. Sometimes, we still need the ACV of some products, e.g. that of $h_i\omega_i$ with lag k, which will be denoted by $\gamma_{h\omega}(k)$. Variances, second or fourth oder moments of the corresponding processes and components are e.g. denoted by σ_Y^2, $\sigma_{Y^2}^2$, σ_Z^2, $\sigma_{Z^2}^2$, σ_ω^2, σ_λ^2, m_2^h, m_2^ω, m_2^λ, m_2^q and m_4^ε etc.

4.2 Properties of $Y_{i,j}$ and the sample variances

Those properties will be first derived under Model (2). For this purpose, we need the following regularity assumptions.

A1. Assumed that $Y_{i,j}$ is defined by (2), where $\varepsilon_{i,j}$ are i.i.d. random variables with zero mean, unit variance and finite fourth moment $E(\varepsilon_{i,j}^4) = m_4^\varepsilon < \infty$.

A2. ω_i and λ_j are two independent series of positive i.i.d. random variables with unit mean and finite second moments $E(\omega_i^2) = m_2^\omega < \infty$ and $E(\lambda_j^2) = m_2^\lambda < \infty$, respectively. Furthermore, it is assumed that ω_i and λ_j are also independent of $\varepsilon_{i,j}$.

Assumptions A1 and A2 ensure that $Y_{i,j}$ is a stationary random field with finite fourth moments. Properties of $Y_{i,j}$, $Y_{i,j}^2$, $\hat{\sigma}_Y^2$, $\hat{\sigma}_{\bullet,j}^2$ and $\hat{\sigma}_{i,\bullet}^2$ are summarized in the following theorem.

Theorem 1. *Under assumptions A1 and A2 we have*

i) $Y_{i,j}$ *is a stationary random field with zero mean, unit variance and ACV* $\gamma_Y(k_1,k_2) = 0$, *if* $k_1 \neq 0$ *or* $k_2 \neq 0$.

ii) $Y_{i,j}^2$ *is also stationary with unit mean,* $\sigma_{Y^2}^2 = m_2^\omega m_2^\lambda m_4^\varepsilon - 1$, $\gamma_{Y^2}(0,k_2) = \sigma_\omega^2$ *for* $k_2 \neq 0$, $\gamma_{Y^2}(k_1,0) = \sigma_\lambda^2$ *for* $k_1 \neq 0$ *and* $\gamma_{Y^2}(k_1,k_2) = 0$, *if* $k_1 \neq 0$ *and* $k_2 \neq 0$.

iii) The mean and variance of $\hat{\sigma}_{i,\bullet}^2$ *conditioning on* i *are given by*

$$E(\hat{\sigma}_{i,\bullet}^2) = \omega_i \text{ and } var(\hat{\sigma}_{i,\bullet}^2) = \frac{\omega_i^2}{n_t} var(\lambda_j \varepsilon_{i,j}^2 | i). \tag{14}$$

iv) The mean and variance of $\hat{\sigma}^2_{\bullet,j}$ conditioning on j are given by

$$E(\hat{\sigma}^2_{\bullet,j}) = \lambda_j \text{ and } var(\hat{\sigma}^2_{\bullet,j}) = \frac{\lambda_j^2}{n_x} var(\omega_i \varepsilon^2_{i,j}|j). \tag{15}$$

v) The sample variance $\hat{\sigma}^2_Y$ is unbiased with

$$var(\hat{\sigma}^2_Y) = \frac{\sigma^2_{Y^2} - \sigma^2_\omega - \sigma^2_\lambda}{n_x n_t} + \frac{\sigma^2_\omega}{n_x} + \frac{\sigma^2_\lambda}{n_t}. \tag{16}$$

The proof of Theorem 1 is given in the appendix. Item *i*) shows that $Y_{i,j}$ is a zero mean uncorrelated random field. The formula of $\sigma^2_{Y^2}$ given in item *ii*) shows that each of the random effect components will increase the variance of $Y^2_{i,j}$, because $m^\omega_2 > 1$ and $m^\lambda_2 > 1$. The other results in this part indicate that, although $Y^2_{i,j}$ is stationary, $\gamma_{Y^2}(k_1, 0)$ and $\gamma_{Y^2}(0, k_2)$ are two positive constants and do not decay to zero. This is a feature of a deterministic process and leads to the fact given in items *iii*) and *iv*) that both $\hat{\sigma}^2_{i,\bullet}$ and $\hat{\sigma}^2_{\bullet,j}$ are inconsistent. On the other hand, when $Y_{i,j}$ are i.i.d or when $\gamma_{Y^2}(k_1, k_2)$ tend to zero exponentially, as $k_1 \to \infty$ or $k_2 \to \infty$. Now, both $\hat{\sigma}^2_{i,\bullet}$ and $\hat{\sigma}^2_{\bullet,j}$ are consistent estimators of σ^2 with rates of convergence determined by n_x and n_t, respectively. Consequentially, the variance of $\hat{\sigma}^2$ consists now of three terms. Where the first term corresponds to the variance of the sample variance in the i.i.d. case, the second and third are caused by random effects on a trading day and at a given trading time point, respectively. Now, the rate of convergence of $\hat{\sigma}^2$ is not determined by the first term on the rhs (right-hand-side) of (16), but by one of the last two terms and is of the order $\max[O(n_x^{-1/2}), O(n_t^{-1/2})]$.

The above results will be now extended to the spatial multiplicative component GARCH with random effects defied in (3). Now, the following regularity assumptions are required.

A1′. $Y_{i,j}$ is defined by (3) and fulfills the other conditions of A1.

A3. For given i, $Z_{i,j}$ follows a GARCH model with unit variance and finite fourth moment $m^Z_4 = m^q_2 m^\varepsilon_4$. It is assumed that the intraday GARCH processes on different trading days have the same coefficients. Those processes are however independent of each other.

A4. The daily conditional variance component h_i is independent of $\varepsilon_{i,j}$, stationary with unit mean and finite variance, whose ACV decays to zero exponentially.

A5. The components ω_i, λ_j, $Z_{i,j}$ and h_i are mutually independent.

Conditions in A3 on the existence of the fourth moments of a GARCH model are well known (see e.g. Bollerslev, 1986, and He and Teräsvirta, 1999a, b). For instance, if $Z_{i,j}$ follows a

GARCH(1, 1) with $N(0,1)$ innovations, this condition reduces to $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$. The further requirement that the GARCH coefficients are the same for all i is necessary for the stationarity of $Z_{i,j}$. A4 ensures that the daily volatility component h_i can be estimated and eliminated separately. The assumption that the ACV of h_i decays exponentially is made for convenience, which is true, if h_i follows a daily GARCH with finite fourth moment.

Properties of $Y_{i,j}$, $Y_{i,j}^2$, $\hat{\sigma}_Y^2$, $\hat{\sigma}_{\bullet,j}^2$ and $\hat{\sigma}_{i,\bullet}^2$ under Model (3) are given in the following theorem.

Theorem 2. *Under assumptions A1′ and A3 through A5 we have*

i) $Y_{i,j}$ *is a stationary random field with zero mean, unit variance and uncorrelated observations.*

ii) $Y_{i,j}^2$ *is stationary with* $E(Y_{i,j}^2) = 1$,

$$var(Y_{i,j}^2) = m_2^\omega m_2^\lambda m_2^h m_2^q m_4^\varepsilon - 1, \tag{17}$$

$$\begin{aligned} \gamma_{Y^2}(0,k_2) &= m_2^h m_2^\omega [\gamma_{Z^2}(0,k_2) + 1] - 1 \\ &\to \sigma_{h\omega}^2, \ \ as\ |k_2| \to \infty, \end{aligned} \tag{18}$$

where $\sigma_{h\omega}^2 = var(\omega_i h_i) = m_2^h m_2^\omega - 1$,

$$\begin{aligned} \gamma_{Y^2}(k_1,0) &= [\gamma_h(k_1) + 1] m_2^\lambda - 1 \\ &\to \sigma_\lambda^2, \ \ as\ |k_1| \to \infty, \end{aligned} \tag{19}$$

where $\sigma_\lambda^2 = m_2^\lambda - 1$ *is the variance of* λ_j, *and*

$$\gamma_{Y^2}(k_1,k_2) = \gamma_h(k_1) \tag{20}$$

for $k_1 \neq 0$ *and* $k_2 \neq 0$.

iii) The mean and variance of $\hat{\sigma}_{i,\bullet}^2$ *conditioning on* i *are given by*

$$E(\hat{\sigma}_{i,\bullet}^2) = h_i\omega_i \ \ and\ var(\hat{\sigma}_{i,\bullet}^2) \approx \frac{h_i^2\omega_i^2}{n_t} V_t. \tag{21}$$

iv) The mean and variance of $\hat{\sigma}_{\bullet,j}^2$ *conditioning on* j *are given by*

$$E(\hat{\sigma}_{\bullet,j}^2) = \lambda_j \ \ and\ var(\hat{\sigma}_{\bullet,j}^2) \approx \frac{\lambda_j^2}{n_x} V_x. \tag{22}$$

v) And the sample variance $\hat{\sigma}_Y^2$ is unbiased with

$$var(\hat{\sigma}_Y^2) \approx \frac{V}{n_x n_t} + \frac{m_2^h \sigma_\omega^2 + V_h}{n_x} + \frac{\sigma_\lambda^2}{n_t}. \tag{23}$$

Where V_x, V_t, V and V_h are constants defined in the appendix.

Proof of Theorem 2 is given in the appendix. As for a GARCH process, observations following the spatial multiplicative component GARCH are also uncorrelated. Item *ii*) shows that $q_{i,j}$ has a clear effect on $\text{var}(Y^2)$ but its effect on $\gamma_{Y^2}(0, k_2)$ is negligible, when $< k_2| \to \infty$. More detailed formulas of the ACV will be given during the proof, where we will see that, for fixed k_2, $\gamma_{Y^2}(0, k_2)$ does depend on $q_{i,j}$. Results in items *iii*) to *v*) show that the asymptotic variances of the sample variances are affected by the daily volatility component but not by the intraday volatility component. The three terms on the rhs of (23) correspond to those in (16). We see both the random effects and the daily GARCH component will cause the fact that sometimes the ACV of the process does not decay to zero. This will in turn affect the rate of convergence of the conditional and unconditional samples variances.

The above results provide the basis for deriving the asymptotic variances of the proposed nonparametric estimators, where we will see that the variances of the nonparametric estimators of the volatility surface also share similar features shown in Theorem 2 *iii*) to *v*).

Note that the model defined by Eq. (6) and (7) in E&S has the spatial representation (4) without the random effects ω_i and λ_j. We can hence obtain the spatial ACV of this process from Theorem 2 *ii*). The assumptions should now be adjusted accordingly.

A1″. Assumed that $Y_{i,j}$ is defined by (4) with i.i.d. $N(0, 1)$ random variables $\varepsilon_{i,j}$.

A3′. Let $Z_{i,j} = q_{i,j}^{1/2} \varepsilon_{i,j}$. For given i, we have $q_{i,j} = (1 - \alpha - \beta) + \alpha Z_{i,j-1}^2 + \beta q_{i,j-1}$, where $\alpha, \beta \geq 0$ and $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$.

Under A1″ and A3′ we have $E(\varepsilon_{i,j}^4) = 3$ and

$$m_2^q = \frac{[1 - (\alpha + \beta)]^2}{1 - 3\alpha^2 - 2\alpha\beta - \beta^2}. \tag{24}$$

Corollary 1. *Let A1″, A3′ and A4 hold. The squared multiplicative component GARCH of E&S, $Y_{i,j}^2$ say, is stationary with $E(Y_{i,j}^2) = 1$, $var(Y_{i,j}^2) = 3m_2^h m_2^q - 1$, $\gamma_{Y^2}(0, k_2) = [var(h_i) + 1][\gamma_q(0, k_2) + 1] \to var(h_i)$, as $|k_2| \to \infty$, and $\gamma_{Y^2}(k_1, 0) = \gamma_h(k_1)$ for $k_1 \neq 0$ and any k_2.*

We see, the exogenous daily volatility component in Model (4) will also cause the fact that the ACV of the squared process does not always decay to zero. This will in turn affect the asymptotic variance of the kernel estimators of the volatility surface.

4.3 Asymptotic properties of the nonparametric estimators

To investigate the asymptotic properties of the proposed kernel estimators of the volatility surface the following additional regularity conditions are required.

B1. $\mathcal{K}(u)$ is a product kernel $\mathcal{K}(u) = K_1(u_x)K_2(u_t)$. For simplicity assume that K_1 and K_2 are the same Lipschitz continuous symmetric density in the support $[-1, 1]$.

B2. $\sigma^2(x, t)$ is a smooth function with absolutely continuous second derivatives.

B3. The bandwidths b_x and b_t fulfill $b_x, b_t \to 0$, $n_x b_x, n_t b_t \to \infty$ as $n_x, n_t \to \infty$.

The above assumptions are suitable adaptations of the regularity conditions used in the literature for deriving the asymptotic properties of bivariate kernel regression estimators. The discussion of the effect of the autocovariances on the variance of a kernel estimator in a nonparametric regression with dependent errors is of general interest. In the following we will introduce a common tool, called the *ACV response function of a kernel function*, for investigating this, which will then be applied to obtain the asymptotic variances of the proposed estimators of the volatility surface in this paper.

Definition 1. *For a univariate kernel function $K(u)$ with support $[-1, 1]$, its ACV response function $\Gamma_K(u)$ is a non-negative symmetric function with support $[-2, 2]$:*

$$\Gamma_K(u) = \int_{-1}^{u+1} K(v)K(v-u)dv \tag{25}$$

for $u \in [-2, 0]$,

$$\Gamma_K(u) = \int_{u-1}^{1} K(v)K(v-u)dv \tag{26}$$

for $u \in [0, 2]$ and zero otherwise.

The ACV response function of a kernel function measures the asymptotic contribution of the ACV of certain lag to the variance of a kernel estimator. Consider a univariate kernel regression with n observations, bandwidth b and stationary errors with ACV $\gamma(k)$. Then the weight of $\gamma(k)$ in the variance of the kernel estimator is asymptotically $\Gamma_K(u_k)$ with $u_k = k/(nb)$. In particular, the weight of the innovation variance σ^2 is always asymptotically $\Gamma_K(0) = R(K)$. Hence, $\Gamma_K(u)$ provides a powerful tool for deriving the asymptotic variance of kernel estimators with dependent errors. For the product bivariate kernel function $\mathcal{K}(u) =$

$K(u_1)K(u_2)$ considered in this paper we define $\Gamma_{\mathcal{K}}(u) = \Gamma_K(u_1)\Gamma_K(u_2)$, which will help us to obtain the asymptotic variance of the proposed double-conditional kernel variance estimator.

Furthermore, define $\mu_2(K) = \int u^2 K_1(u)du$, $R(K) = \int K_1^2(u)du$ and $I(\Gamma_K) = \int \Gamma_K(u)du$. Our main findings on the double-conditional kernel estimator $\hat{\sigma}^2(x,t)$ and the two associate intermediate smoothers in the first stage are summarized in the following theorem.

Theorem 3. *Consider the estimation at an interior point (x,t) with $0 < x, t < 1$. Under the same assumptions of Theorem 2 and the additional assumptions B1 through B3 we have*

i) The mean and variance of the conditional smoother $\hat{\sigma}^2(t|x_i)$ are given by

$$E[\hat{\sigma}^2(t|x_i)] \approx h_i\omega_i \left\{ \sigma^2(x_i,t) + \frac{\mu_2(K)}{2} b_t^2 [\sigma^2(x,t)]''_t \right\}, \tag{27}$$

$$var\,[\hat{\sigma}^2(t|x_i)] \approx h_i^2\omega_i^2 \frac{\sigma^4(x_i,t)V_t}{n_x b_x} R(K). \tag{28}$$

ii) The mean and variance of the conditional smoother $\hat{\sigma}^2(x|t_j)$ are given by

$$E[\hat{\sigma}^2(x|t_j)] \approx \lambda_j \left\{ \sigma^2(x,t_j) + \frac{\mu_2(K)}{2} b_x^2 [\sigma^2(x,t)]''_x \right\}, \tag{29}$$

$$var\,[\hat{\sigma}^2(x|t_j)] \approx \lambda_j^2 \frac{\sigma^4(x,t_j)V_x}{n_x b_x} R(K). \tag{30}$$

iii) The bias and the variance of $\hat{\sigma}^2(x,t)$ are given by

$$B[\hat{\sigma}^2(x,t)] \approx \frac{\mu_2(K)}{2} \left\{ b_x^2 [\sigma^2(x,t)]''_x + b_t^2 [\sigma^2(x,t)]''_t \right\}, \tag{31}$$

$$var\,[\hat{\sigma}^2(x,t)] \approx \sigma^4(x,t) \left[\frac{VR^2(K)}{n_x b_x n_t b_t} + \left(\frac{m_2^h \sigma_\omega^2 + V_h}{n_x b_x} + \frac{\sigma_\lambda^2}{n_t b_t} \right) R(K) I(\Gamma_K) \right]. \tag{32}$$

Where V_x, V_t and V_1 are the same as defined in Theorem 2.

The proof of Theorem 3 is given in the appendix. Items *i*) and *ii*) show that the two intermediate smoothers converge to two random functions, respectively, and are inconsistent. Their asymptotic variances are also random variables. When the errors are i.i.d., the asymptotic variance of $\hat{\sigma}^2(x,t)$ is of the order $O(n_x b_x n_t b_t)^{-1}$, which is the order of the first term on the rhs of (32). But item *iii*) shows that the asymptotic variance of $\hat{\sigma}^2(x,t)$ is of the order $\max[O(n_x b_x)^{-1}, O(n_t b_t)^{-1}]$, which is much lower than the order $O(n_x b_x n_t b_t)^{-1}$.

The proof of item *iii*) in the appendix shows that $\text{var}\,[\hat{\sigma}^2(x,t)]$ is determined by different kinds of the ACV of $Y_{i,j}^2$. As far as we know, asymptotic results of bivariate kernel estimators

with dependent errors are not yet studied in the literature. Results in Theorem 3 *iii*) can be adapted to fill this blank. Assume for instance that Model (1) holds but $Y_{i,j}$ is defined without the random effects and the daily volatility component. Now, we can assume further that the ACV of $Y_{i,j}^2$ is absolutely summable. In this case the variance of $\hat{\sigma}^2(x,t)$ will be dominated by the first term on the rhs of (32) with V to be the sum of all $\gamma_{Y^2}(k_1,k_2)$. However, the asymptotic variances of $\hat{\sigma}^2(x,t)$ under Models (1) and (3) with multiplicative random effects and a daily GARCH component are quite different.

Let b_x^{A} and b_t^{A} denote the asymptotically optimal bandwidths which minimize the dominating part of the MSE (mean squared error) of $\hat{\sigma}^2(x,t)$. For simplicity, we assume that n_t is proportional to n_x such that $n_t = c_n n_x$ for any n_x, where $c_n > 0$ is a constant. Now, it is easy to show that both b_x^{A} and b_t^{A} are of the order $O(n_x^{-1/5}) = O(n^{-1/10})$. Note that the asymptotically optimal bandwidths for bivariate kernel regression with i.i.d. errors are of the order $O(n^{-1/6})$ (see e.g. Herrmann et al., 1995). That is bivariate kernel regression under multiplicative random effects requires much larger bandwidths. This fact will results in a much lower rate of convergence of the kernel estimator of the volatility surface. The explicit formulas of b_x^{A} and b_t^{A} are now also very complex. Detailed discussion on this and on the development of a plug-in bandwidth selector for double-conditional smoothing based on those formulas is beyond the aim of the current paper and will be studied elsewhere.

5 Application

The proposal is illustrated using 1-minute returns of Allianz AG from Jan. 2006 to Sep. 2011 with observations on $n_x = 1442$ trading days and $n_t = 510$ returns on each day. The data represented under the spatial model are illustrated in Figure 1 (left). From this figure the long-term volatility dynamics can be easily discovered by eye. In particular, the returns during 2008 and 2009 are much larger in absolute value indicating the effect of the 2008 financial crisis on volatility. If we look in the direction of the trading time exactly, it is even possible to discover a rough picture of the intraday volatility pattern. We see the spatial representation can help us to discover some interesting features of high-frequency returns.

The estimated volatility surface using bandwidths b_x and b_t such that $b_x * n_x = 200$ and $b_t * n_t = 100$ is shown in Figure 1 (right). In the dimension of the trading day we can see that the volatility before the financial crisis was very low. It became much higher during the financial crisis and reduced again to a low level thereafter. However, the volatility after the

financial crisis is still clearly higher than that before it. Furthermore, the form of the long-term volatility tendency is similar at all trading time points, which looks like the form of an "arch" (bridge). On the other hand, the daily volatility pattern is indicated by the curve on the volatility surface for a given trading day. On each day it looks like a volatility smile. The long-term volatility tendency and the daily volatility smiles together form a volatility saddle. Intermediate smoothing results conditioning on the trading time are displayed in Figure 2 (left). The randomness in these results can be seen clearly. Again, the long-term volatility patterns at different time points are quite similar. This indicates that the assumption of a random scale at each trading time point seems to be quite reasonable. The intermediate smoothing results conditioning on the trading day are displayed in Figure 2 (right). We see that these results also exhibit clear randomness. Comparing the right panels of Figures 1 and 2, we see that although the average daily volatility patterns show a smooth form, which change however from one day to another very strongly, if they are estimated using observations on a single trading day. This indicates that not only the scale but also the form of the daily volatility patterns might be affected by some random effect.

6 Concluding remarks

This paper introduced a spatial multiplicative component GARCH model for high-frequency returns with a nonparametric volatility surface. A double-conditional smoothing is developed for estimating the volatility surface, which leads to equivalent estimates as a common bivariate kernel estimator but runs much faster. This approach also helps us to find that high-frequency returns exhibit multiplicative random effects. Suitable models are introduced to deal with this phenomenon. Basic probabilistic properties of the proposed processes and asymptotic properties of the proposed estimators of the volatility surface are investigated in detail. It is in particular shown that in the presence of random effects, the rate of convergence of the nonparametric variance estimators will be strongly affected. There are still many open questions in this area. For instance, it will be worthwhile to study different extensions and further applications of the spatial multiplicative component GARCH model. Moreover, the proposals in this paper can also be employed for discussing market microstructures.

Acknowledgments. The graphics were generated by Mr Christian Peitz, University of Paderborn. We would like to thank Liudas Giraitis for helpful discussion and comments.

Appendix. Proofs of the results

Proof of Theorem 1.

i) The results $E(Y_{i,j}) = 0$ and $\text{var}\,(Y_{i,j}) = 1$ follow from the definition. Moreover,

$$\gamma_Y(k_1, k_2) = E(\sqrt{\omega_i \omega_{i+k_1} \lambda_j \lambda_{j+k_2}}) E(\varepsilon_{i,j} \varepsilon_{i+k_1,j+k_2}),$$

which is zero, if $k_1 \neq 0$ or $k_2 \neq 0$, because $\varepsilon_{i,j}$ are zero mean independent random variables.

ii) Note that $E(Y_{i,j}^2) = \text{var}\,(Y_{i,j}) = 1$. This leads to $\text{var}\,(Y_{i,j}^2) = E(Y_{i,j}^4) - 1$. Since $E(Y_{i,j}^4) = E(\omega_i^2) E(\lambda_j^2) E(\varepsilon_{i,j}^4)$, we have $\text{var}\,(Y_{i,j}^2) = m_2^\omega m_2^\lambda m_4^\varepsilon - 1$. For given i and $k_2 \neq 0$,

$$\begin{aligned}
\gamma_{Y^2}(0, k_2) &= E(Y_{i,j}^2 Y_{i,j+k_2}^2) - 1 \\
&= E(\omega_i \lambda_j \varepsilon_{i,j}^2 \omega_i \lambda_{j+k_2} \varepsilon_{i,j+k_2}^2) - 1 \\
&= E(\omega_i^2) E(\lambda_j \varepsilon_{i,j}^2 \lambda_{j+k_2} \varepsilon_{i,j+k_2}^2) - 1 \\
&= E(\omega_i^2) - 1 \\
&= \sigma_\omega^2. \qquad \text{(A.1)}
\end{aligned}$$

Similarly, it can be shown that, for given j and $k_1 \neq 0$, $\gamma_{Y^2}(k_1, 0) = \text{var}\,(\lambda_j)$. Furthermore, it can be shown that, for $k_1 \neq 0$ and $k_2 \neq 0$, $E(Y_{i,j}^2 Y_{i+k_1,j+k_2}^2) = 1$ and hence $\gamma_{Y^2}(k_1, k_2) = 0$.

iii) For the conditional sample variance given i, we have

$$\begin{aligned}
E(\hat{\sigma}_{i,\bullet}^2) &= E\left(\frac{1}{n_t} \sum_{j=1}^{n_t} \omega_i \lambda_j \varepsilon_{i,j}^2\right) \\
&= \frac{\omega_i}{n_t} \sum_{j=1}^{n_t} E(\lambda_j) E(\varepsilon_{i,j}^2) \\
&= \omega_i.
\end{aligned}$$

The variance of the conditional sample variance given i is

$$\begin{aligned}
\text{var}\,(\hat{\sigma}_{i,\bullet}^2) &= E\left[\left(\frac{1}{n_t} \sum_{j=1}^{n_t} \omega_i \lambda_j \varepsilon_{i,j}^2\right)^2\right] - \omega_i^2 \\
&= \frac{\omega_i^2}{n_t^2} \sum_{j_1=1}^{n_t} \sum_{j_2=1}^{n_t} [E(\lambda_{j_1} \lambda_{j_2} \varepsilon_{i,j_1}^2 \varepsilon_{i,j_2}^2) - 1].
\end{aligned}$$

Note that $E(\lambda_{j_1} \lambda_{j_2} \varepsilon_{i,j_1}^2 \varepsilon_{i,j_2}^2)$ is equal to $m_2^\lambda m_4^\varepsilon$, for $j_1 = j_2$, and 1, otherwise. Also note that

$\text{var}\,(\lambda_j \varepsilon_{i,j}^2 | i) = m_2^\lambda m_4^\varepsilon - 1$. We have

$$\begin{aligned}
\text{var}\,(\hat{\sigma}_{i,\bullet}^2) &= \frac{\omega_i^2}{n_t^2}[n_t m_2^\lambda m_4^\varepsilon + n_t(n_t-1) - n_t^2] \\
&= \frac{\omega_i^2}{n_t}[m_2^\lambda m_4^\varepsilon - 1] \\
&= \frac{\omega_i^2}{n_t}\text{var}\,(\lambda_j \varepsilon_{i,j}^2 | i).
\end{aligned}$$

iv) Proof of the results in this part is similar to that of *iii*) and is omitted.

v) It is easy to see that $E(\hat{\sigma}_Y^2) = 1$, i.e. $\hat{\sigma}_Y^2$ is unbiased. For the variance of $\hat{\sigma}_Y^2$ we have

$$\begin{aligned}
\text{var}\,(\hat{\sigma}_Y^2) &= \text{var}\left(\frac{1}{n_x n_t}\sum_{i=1}^{n_x}\sum_{j=1}^{n_t} Y_{i,j}^2\right) \\
&= \frac{1}{n_x^2 n_t^2}\left[\sum_{i_1=1}^{n_x}\sum_{j_1=1}^{n_t}\sum_{i_2=1}^{n_x}\sum_{j_2=1}^{n_t}\gamma_{Y^2}(i_2-i_1, j_2-j_1)\right] \\
&= \frac{1}{n_x^2 n_t^2}(T_1 + T_2 + T_3), \qquad \text{(A.2)}
\end{aligned}$$

where

$$T_1 = \sum_{i_1=1}^{n_x}\sum_{j_1=1}^{n_t}\sum_{i_2=1}^{n_x}\sum_{j_2=1}^{n_t}\gamma_{Y^2}^*(i_2-i_1, j_2-j_1) \qquad \text{(A.3)}$$

with $\gamma_{Y^2}^*(0,0) = \sigma_{Y^2}^2 - \sigma_\omega^2 - \sigma_\lambda^2$ and $\gamma_{Y^2}^*(k_1,k_2) = 0$, if $k_1 \neq 0$ or $k_2 \neq 0$,

$$T_2 = \sum_{i=1}^{n_x}\sum_{j_1=1}^{n_t}\left[\sigma_\omega^2 + \sum_{j_2 \neq j_1}\gamma_{Y^2}(0, j_2-j_1)\right] \qquad \text{(A.4)}$$

and

$$T_3 = \sum_{i_1=1}^{n_x}\sum_{j=1}^{n_t}\left[\sigma_\lambda^2 + \sum_{i_2 \neq i_1}\gamma_{Y^2}(i_2-i_1, 0)\right]. \qquad \text{(A.5)}$$

Here $\gamma_{Y^2}^*(k_1,k_2)$ are defined to show the performance of $\text{var}\,(\hat{\sigma}_Y^2)$, when $Y_{i,j}^2$ were uncorrelated. We have $T_1 = n_x n_t(\sigma_{Y^2}^2 - \sigma_\omega^2 - \sigma_\lambda^2)$. Following (A.1), it can be shown that $T_2 = n_x n_t^2 \sigma_\omega^2$. Similarly, we have $T_3 = n_x^2 n_t \sigma_\lambda^2$. Inserting these results into (A.2) leads to

$$\text{var}\,(\hat{\sigma}_Y^2) = \frac{\sigma_{Y^2}^2 - \sigma_\omega^2 - \sigma_\lambda^2}{n_x n_t} + \frac{\sigma_\omega^2}{n_x} + \frac{\sigma_\lambda^2}{n_t}\sigma_\lambda^2. \qquad \text{(A.6)}$$

Theorem 1 is proved. ◇

Remark A.1. *Although there is a close relationship between the results in items iii) and iv), and those in item v), it is however not easy to prove the results in v) based on those in iii) and iv), because results in the latter cases are obtained conditioning on i or j.*

Proof of Theorem 2.

Firstly, under the assumptions of Theorem 2, all of the autocovariances of h_i, q_j and $Z_{i,j}|i$ tend to zero exponentially, as the lag tends to infinite.

i) The proof of results in this part is straightforward and is omitted.

ii) Following the results in i) we have $E(Y_{i,j}^2) = \text{var}\,(Y_{i,j}) = 1$. Hence,

$$\begin{aligned}
\text{var}\,(Y_{i,j}^2) &= E(Y_{i,j}^4) - 1 \\
&= E(\omega_i^2 \lambda_j^2 h_i^2 q_{i,j}^2 \varepsilon_{i,j}^4) - 1 \\
&= m_2^h m_2^\omega m_2^\lambda m_2^q m_4^\varepsilon - 1.
\end{aligned}$$

For given i and $k_2 \neq 0$, we have

$$\begin{aligned}
\gamma_{Y^2}(0, k_2) &= E(h_i^2 \omega_i^2 \lambda_j \lambda_{j+k_2} Z_{i,j}^2 Z_{i,j+k_2}^2) - 1 \\
&= m_2^h m_2^\omega [\gamma_{Z^2}(0, k_2) + 1] - 1 \\
&\to m_2^h m_2^\omega - 1 = \sigma_{h\omega}^2, \qquad \text{(A.7)}
\end{aligned}$$

because λ_j and λ_{j+k_2} are independent, and $\gamma_{Z^2}(0, k_2) \to 0$, as $|k_2| \to \infty$. Similarly, for given j and $k_1 \neq 0$, we have

$$\begin{aligned}
\gamma_{Y^2}(k_1, 0) &= E(h_i h_{i+k_1} \omega_i \omega_{i+k_1} \lambda_j^2 Z_{i,j}^2 Z_{i+k_1,j}^2) - 1 \\
&= m_2^\lambda [\gamma_h(k_1) + 1] - 1 \\
&\to m_2^\lambda - 1 = \sigma_\lambda^2, \qquad \text{(A.8)}
\end{aligned}$$

because $Z_{i,j}^2$ and $Z_{i+k_1,j}^2$ are independent, and $\gamma_h(k_1) \to 0$, as $|k_1| \to \infty$. Furthermore, for $k_1 \neq 0$ and $k_2 \neq 0$, we have

$$\begin{aligned}
\gamma_{Y^2}(k_1, k_2) &= E(h_i h_{i+k_1} \omega_i \omega_{i+k_1} \lambda_j \lambda_{j+k_2} Z_{i,j}^2 Z_{i+k_1,j+k_2}^2) - 1 \\
&= E(h_i h_{i+k_1}) - 1 \\
&= \gamma_h(k_1). \qquad \text{(A.9)}
\end{aligned}$$

iii) The proof of $E(\hat{\sigma}_{i,\bullet}^2) = \omega_i h_i$ is straightforward and is omitted. For given i, the variance of $\hat{\sigma}_{i,\bullet}^2$ can be written as

$$\begin{aligned}
\text{var}\,(\hat{\sigma}_{i,\bullet}^2) &= E\left[\left(\frac{1}{n_t}\sum_{j=1}^{n_t} h_i \omega_i \lambda_j Z_{i,j}^2\right)^2\right] - h_i^2 \omega_i^2 \\
&= \frac{h_i^2 \omega_i^2}{n_t^2} \sum_{j_1=1}^{n_t} \sum_{j_2=1}^{n_t} [E(\lambda_{j_1} \lambda_{j_2} Z_{i,j_1}^2 Z_{i,j_2}^2) - 1].
\end{aligned}$$

It can be shown that $E(\lambda_{j_1}\lambda_{j_2}Z^2_{i,j_1}Z^2_{i,j_2})-1=\sigma^2_\lambda m^Z_4+\sigma^2_{Z^2}$, if $j_1=j_2$, and $E(\lambda_{j_1}\lambda_{j_2}Z^2_{i,j_1}Z^2_{i,j_2})-1=\gamma_{Z^2}(0,j_2-j_1)$, otherwise. To discuss the asymptotic variance of the estimator, we define J such that $J\to\infty$ and $J/n_t\to 0$. For any $J<j_1<n_t-J$, we have

$$\sum_{j_2=1}^{n_t}[E(\lambda_{j_1}\lambda_{j_2}Z^2_{i,j_1}Z^2_{i,j_2})-1] \;\to\; \sigma^2_\lambda m^Z_4+\sum_{k=-\infty}^{\infty}\gamma_{Z^2}(k).$$

For $j_1\le J$ or $j_1\ge n_t-J$, the above sum may tend to different constants depending on j_1. But this will not affect the asymptotic analysis, as $n_t\to\infty$. This leads to

$$\begin{aligned}\mathrm{var}\,(\hat\sigma^2_{i,\bullet}) &= \frac{h_i^2\omega_i^2}{n_t^2}n_tV_t81+o(1)]\\ &\approx \frac{h_i^2\omega_i^2}{n_t}V_t,\end{aligned}$$

where

$$V_t=\sigma^2_\lambda m^Z_4+\sum_{k=-\infty}^{\infty}\gamma_{Z^2}(k).$$

iv) Results in this part can be proved similarly. And the constant V_x in (22) is given by

$$V_x=\sigma^2_{\lambda Z^2}+\sum_{k=-\infty}^{\infty}\gamma_h(k).$$

v) Again, we will only focus on the derivation of $\mathrm{var}\,(\hat\sigma_{Y^2})$. Similarly to the proof of Theorem 1 *v*), the variance of $\hat\sigma^2_Y$ can be split into four parts as follows

$$\mathrm{var}\,(\hat\sigma^2_Y)=\frac{1}{n_x^2n_t^2}(T_1+T_2+T_{2a}+T_3), \tag{A.10}$$

where $\sigma^2_{Y^2}$ is divided into T_1, T_2, T_{2a} and T_3 respectively, with the quantities $\gamma^*_{Y^2}(0,0)=\sigma^2_{Y^2}-m^h_2\sigma^2_\omega-\sigma^2_h-\sigma^2_\lambda$, $m^h_2\sigma^2_\omega$, σ^2_h and σ^2_λ as well. For $k_2\neq 0$, $\gamma_{Y^2}(0,k_2)$ is divided into T_2, T_{2a} and T_1 respectively, with the quantities $m^h_2\sigma^2_\omega$, σ^2_h and the remaining part. And for $k_1\neq 0$, $\gamma_{Y^2}(k_1,0)$ is divided into T_3, T_{2a} and T_1 respectively, with the quantities σ^2_λ, $\gamma_h(k_1)$ and the remaining part. For $k_1\neq 0$ and $k_2\neq 0$, we define $\gamma^*_{Y^2}(k_1,k_2)=\gamma_{Y^2}(k_1,k_2)$. Hence $\gamma^*_{Y^2}(k_1,k_2)$ are either $\gamma_{Y^2}(k_1,k_2)$, when they converge to zero, or the differences between $\gamma_{Y^2}(k_1,k_2)$ and their constant limits, such that

$$V=\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty}\gamma^*_{Y^2}(k_1,k_2) \tag{A.11}$$

exists. And T_1 is defined by

$$T_1=\sum_{i_1,i_2=1}^{n_x}\sum_{j_1,j_2=1}^{n_t}\gamma^*_{Y^2}(i_2-i_1,j_2-j_1) \tag{A.12}$$

with $T_1 = n_x n_t V[1 + o(1)]$.

The term T_3 is the same as in the proof of Theorem 1 v). According to the above decomposition of the ACV we have $T_2 = n_x n_t^2 m_2^h \sigma_\omega^2$. To obtain an approximation of T_{2a}, choose I such that $I \to \infty$ and $I/n_x \to 0$, as $n_x \to \infty$. For $I \leq i_1 \leq n_x - I$ and any $k_2 \neq 0$ we have

$$\sum_{i_2 \neq i_1} \gamma_{Y^2}(k_1, k_2) + \sigma_h^2 \approx V_h, \tag{A.13}$$

where V_h is the sum of all ACV of the daily GARCH component h_i. This leads to $V_{2a} = n_x n_t^2 V_h[1 + o(1)]$. Inserting these results into (A.10) we have

$$\text{var}\,(\hat{\sigma}_Y^2) \approx \frac{V}{n_x n_t} + \frac{m_2^h \sigma_\omega^2 + V_h}{n_x} + \frac{\sigma_\lambda^2}{n_t}. \tag{A.14}$$

This finishes the proof of Theorem 2. $\diamond$

A sketched proof of Theorem 3.

In the following only a sketched proof of the results in Theorem 3 will be given. Note that the formulas of the bias in the conditional and unconditional kernel variance estimators can be proved easily by adapting known results in the literature, because they are not affected by the dependence structure. In the following we will hence mainly focus on investigating the formulas of the asymptotic variances of the proposed estimators.

i) Under the assumptions of Theorem 3 and conditioning on i, Eq. (5) can be written as

$$\tilde{r}_{i,j}^2 | i = h_i \omega_i \sigma^2(x_i, t_j) + h_i \omega_i \sigma^2(x_i, t_j) \eta_{i,j}^*, \tag{A.15}$$

where $\eta_{i,j}^* = \lambda_j Z_{i,j}^2 - 1$ with $E(\eta_{i,j}^*) = 0$ and $\text{var}\,(\eta_{i,j}^*) = m_2^\lambda m_4^Z - 1$. This is a heteroskedastic nonparametric regression with conditional variance of $Y_{i,j}$ given i, i.e. $h_i \omega_i \sigma^2(x_i, t)$, to be its (conditional) mean and volatility functions at the same time. Following related study in the literature (Feng, 2004), it can be shown that the asymptotic variance of $\hat{\sigma}^2(t|x_i)$ is

$$\text{var}\,[\hat{\sigma}^2(t|x_i)] \approx h_i^2 \omega_i^2 \frac{\sigma^4(x_i, t) V_t}{n_x b_x} R(K), \tag{A.16}$$

where V_t is as defined before. Although it is also not difficult to prove (A.16) by means of the ACV response function $\Gamma_K(u)$, this will be omitted to save space.

ii) Conditioning on j, Eq. (5) can be represented as

$$\tilde{r}_{i,j}^2 | j = \lambda_j \sigma^2(x_i, t_j) + \lambda_j \sigma^2(x_i, t_j) \eta_{i,j}^{**}, \tag{A.17}$$

where $\eta_{i,j}^{**} = h_i\omega_i Z_{i,j}^2 - 1$. Similar analysis as above will lead to the formula of the asymptotic variance of $\hat{\sigma}^2(x|t_j)$, where V_x is as defined in the proof of Theorem 2 *iv*).

iii) Now, consider the kernel estimator of the variance of the stationary part $Y_{i,j}$

$$\hat{\sigma}_Y^2(x,t) = \sum_{j=1}^{n_t}\sum_{i=1}^{n_x} w_{i,j} Y_{i,j}^2, \tag{A.18}$$

where $w_{i,j}$ are the same as in $\hat{\sigma}^2(x,t)$. Then the following lemma holds.

Lemma 1. *Under the conditions of Theorem 3 we have*

$$var\,[\hat{\sigma}^2(x,t)] \approx \sigma^4(x,t) var\,[\hat{\sigma}_Y^2(x,t)]. \tag{A.19}$$

Proof. Note that $w_{i,j}$ is only positive, if $|x_i - x| < b_x$ and $|t_j - t| < b_t$. This implies that $\sigma^2(x_i,t_j) = \sigma^2(x,t)[1+o(1)]$ for $w_{i,j} > 0$. We have

$$\begin{aligned}
\text{var}\,[\hat{\sigma}^2(x,t)] &= \text{var}\left[\sum_{j=1}^{n_t}\sum_{i=1}^{n_x} w_{i,j}\tilde{r}_{i,j}^2\right] \\
&= \text{var}\left[\sum_{j=1}^{n_t}\sum_{i=1}^{n_x} w_{i,j}\sigma^2(x,t)[1+o(1)]Y_{i,j}^2\right] \\
&\approx \sigma^4(x,t)\text{var}\left[\sum_{j=1}^{n_t}\sum_{i=1}^{n_x} w_{i,j}Y_{i,j}^2\right] \\
&= \sigma^4(x,t)\text{var}\,[\hat{\sigma}_Y^2(x,t)].
\end{aligned} \tag{A.20}$$

Lemma 1 is proved.

Now, define $k_x = [n_x b_x]$ and $k_t = [n_t b_t]$, where $[\cdot]$ denote the integer part. By means of the ACV response function of $\mathcal{K}(u_1,u_2)$ we have

Lemma 2. *Under the conditions of Theorem 3 the variance of $\hat{\sigma}_Y^2(x,t)$ is dominated by*

$$var\,[\hat{\sigma}_Y^2(x,t)] \approx \frac{1}{n_x b_x}\frac{1}{n_t b_t}\sum_{k_1=-2k_x}^{2k_x}\sum_{k_2=-2k_t}^{2k_t}\Gamma_{\mathcal{K}}(u_{1k_1},u_{2k_2})\gamma_{Y^2}(k_1,k_2), \tag{A.21}$$

where $u_{1k_1} = k_1/n_x$, $u_{2k_2} = k_2/n_t$ and $\Gamma_{\mathcal{K}}(\cdot,\cdot)$ is the ACV response function of $\mathcal{K}(\cdot,\cdot)$.

Proof. The non-zero weights at an interior observation point (x_0,t_0) with $x_0 = (i_0-0.5)/n_x$ and $t_0 = (j_0-0.5)/n_t$ are the same and will be denoted by $\tilde{w}_{r,s}$, $r = -k_x$, ..., 0, ..., k_x and

$s = -k_t$, ..., 0, ..., k_t, for observations $Y_{i,j}$, $i = i_0 - k_x$, ..., i_0, ..., $i_0 + k_x$ and $j = j_0 - k_t$, ..., j_0, ..., $j_0 + k_t$. It can be shown that $\tilde{w}_{r,s} = \tilde{w}_r^x \tilde{w}_s^t$ with

$$\tilde{w}_r^x \approx \frac{1}{n_x b_x} K(v_r) \text{ and } \tilde{w}_s^t \approx \frac{1}{n_t b_t} K(v_s), \tag{A.22}$$

where $v_r = r/(n_x b_x)$ and $v_s = s/(n_t b_t)$. Under these notations we have

$$\text{var}\left[\hat{\sigma}_Y^2(x,t)\right] = \sum_{r_1,r_2=-k_x}^{k_x} \sum_{s_1,s_2=-k_t}^{k_t} \frac{K(v_{r_1})K(v_{r_2})}{n_x^2 b_x^2} \frac{K(v_{s_1})K(v_{s_2})}{n_t^2 b_t^2} \gamma_{Y^2}(r_1 - r_2, s_1 - s_2). \tag{A.23}$$

Note that the range of $r_1 - r_2$ is from $-2k_x$ to $2k_x$ and that of $s_1 - s_2$ is from $-2k_t$ to $2k_t$. The above formula can hence be rewritten as

$$\text{var}\left[\hat{\sigma}_Y^2(x,t)\right] = \frac{1}{n_x b_x} \frac{1}{n_t b_t} \sum_{k_1=-2k_x}^{2k_x} \sum_{k_2=-2k_t}^{2k_t} \beta(k_1,k_2) \gamma_{Y^2}(k_1,k_2), \tag{A.24}$$

where

$$\beta(k_1,k_2) = \sum_{r_1-r_2=k_1} \sum_{s_1-s_2=k_2} \frac{K(v_{r_1})K(v_{r_2})}{n_x b_x} \frac{K(v_{s_1})K(v_{s_2})}{n_t b_t}. \tag{A.25}$$

Furthermore, note that the two restrictions $r_1 - r_2 = k_1$ and $s_1 - s_2 = k_2$ are independent of each other. We have

$$\beta(k_1,k_2) = \sum_{r_1-r_2=k_1} \frac{K(v_{r_1})K(v_{r_2})}{n_x b_x} \sum_{s_1-s_2=k_2} \frac{K(v_{s_1})K(v_{s_2})}{n_t b_t}. \tag{A.26}$$

Consider the case with $-2k_x \le k_1 \le 0$ in detail. Now, we have $r_2 = r_1 - k_1$ and the range of r_1 is from $-k_x$ to $k_x + k_1$. Define $u_1 = k_1/k_x$. We have $u_1 \in [-2, 0]$, $v_{r_2} = v_{r_1} - u_1$ and the range of v_{r_1} is from -1 to $u_1 + 1$. This leads to, for $k_1 \le 0$,

$$\begin{aligned} \sum_{r_1-r_2=k_1} \frac{K(v_{r_1})K(v_{r_2})}{n_x b_x} &= \sum_{-1 \le v_{r_1} \le u_1+1} \frac{K(v_{r_1})K(v_{r_1} - u_1)}{n_x b_x} \\ &\approx \int_{-1}^{u_1+1} K(v_r)K(v_r - u_1)dv_r. \end{aligned} \tag{A.27}$$

Similarly, define $u_2 = k_2/k_t$. We have, for $k_2 \le 0$,

$$\sum_{s_1-s_2=k_2} \frac{K(v_{s_1})K(v_{s_2})}{n_t b_t} \approx \int_{-1}^{u_2+1} K(v_s)K(v_s - u_1)dv_s. \tag{A.28}$$

Results for $k_1, k_2 \ge 0$ can be proved analogously. This indicates that $\beta(k_1,k_2) \approx \Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2})$. Lemma 2 is proved by inserting this into (A.24).

Further calculation leads to the following decomposition

$$\text{var}\left[\hat{\sigma}_Y^2(x,t)\right] \approx \frac{1}{n_x b_x} \frac{1}{n_t b_t} (S_1 + S_2 + S_{2a} + S_3), \tag{A.29}$$

where S_1, S_2, S_{2a} and S_3 correspond to the four ACV-components defined in the proof of Theorem 2 v) and are given by

$$S_1 = \sum_{k_1=-2k_x}^{2k_x} \sum_{k_2=-2k_t}^{2k_t} \Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2}) \gamma^*_{Y^2}(k_1, k_2), \tag{A.30}$$

$$S_2 = m_2^h \sigma_\omega^2 \sum_{k_2=-2k_t}^{2k_t} \Gamma_{\mathcal{K}}(0, u_{2k_2}), \tag{A.31}$$

$$S_{2a} = \sum_{k_1=-2k_x}^{2k_x} \sum_{k_2=-2k_t}^{2k_t} \Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2}) \gamma_h(k_1) \tag{A.32}$$

and

$$S_3 = \sigma_\lambda^2 \sum_{k_1=-2k_x}^{2k_x} \Gamma_{\mathcal{K}}(u_{1k_1}, 0). \tag{A.33}$$

Firstly, choose K_x and K_t so that $K_x, K_t \to \infty$, $K_x/n_x \to 0$ and $K_t/n_t \to 0$, as $n_x, n_t \to \infty$. It can be shown that $\Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2}) \approx \Gamma_{\mathcal{K}}(0,0) = R^2(K)$ for $|k_1| \le K_1$ and $|k_2| \le K_2$, and the sum of $\gamma^*_{Y^2}(k_1, k_2)$ in this case tends to the sum of all $\gamma^*_{Y^2}$, because $\gamma^*_{Y^2}(k_1, k_2)$ are absolutely summable. Consequently, the remaining part of S_1 converges to zero, because $\Gamma_{\mathcal{K}}(u_1, u_2)$ is bounded. We have

$$\begin{aligned} S_1 &= \sum_{|k_1| \le K_1} \sum_{|k_2| \le K_2} \Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2}) \gamma^*_{Y^2}(k_1, k_2) \\ &\quad + \sum_{|k_1| > K_1} \sum_{|k_2| > K_2} \Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2}) \gamma^*_{Y^2}(k_1, k_2) \\ &\approx R^2(K) V, \end{aligned} \tag{A.34}$$

where V is the same as in Theorem 2. Note that $\Gamma_{\mathcal{K}}(0, u_{2k_2}) = R(K)\Gamma_K(u_{2k_2})$. We have

$$S_2 \approx n_t m_2^h \sigma_\omega^2 R(K) I(\Gamma_K). \tag{A.35}$$

Note that $\Gamma_{\mathcal{K}}(u_{1k_1}, u_{2k_2}) = \Gamma_K(u_{1k_1})\Gamma_K(u_{2k_2})$ and $\gamma_h(k_1)$ sum up to V_h. We have

$$\begin{aligned} S_{2a} &= \sum_{k_2=-2k_t}^{2k_t} \sum_{k_1=-2k_x}^{2k_x} \Gamma_K(u_{1k_1}) \Gamma_K(u_{2k_2}) \gamma_h(k_1) \\ &\approx \sum_{k_2=-2k_t}^{2k_t} \Gamma_K(u_{2k_2}) R(K) V_h \\ &\approx R(K) I(\Gamma_K) V_h. \end{aligned} \tag{A.36}$$

Finally, note that $\Gamma_{\mathcal{K}}(u_{1k_1}, 0) = R(K)\Gamma_K(u_{1k_1})$, we have $S_3 \approx n_x b_x \sigma_\lambda^2 R(K) I(\Gamma_K)$. Theorem 3 is proved by inserting these results into (A.29) and then inserting (A.29) into (A.19). $\diamond$

References

Andersen, T.G., Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *J. Empir. Finance*, 4, 115-158.

Andersen, T.G., Bollerslev, T. (1998). Deutschemark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer-run dependencies. *J. Finance*, 53, 219-265.

Andersen, T.G., Bollerslev, T. and Cai, J. (2000). Intraday and interday volatility in the Japanese stock market. *J. Intern. Finan. Markets, Instit. Money*, 10, 107-130.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31, 307-327.

Borovkova, S.A. and Lopuhaä, H.P. (2009). A spatial approach to multivariate volatility modelling. Preprint, Vrije University of Amsterdam.

Brockwell, P., Chadraa, E. and Lindner, A. (2006). Continuous-time GARCH processes. *Ann. Appl. Prob.*, 16, 790-826.

Caporin, M. and Paruolo, P. (2006). GARCH models with spatial structure. SIS Statistica, 2006, 447-450.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimation of U.K. inflation. *Econometrica*, 50, 987-1008.

Engle, R.F. and Sokalska, M.E. (2012). High Frequency Multiplicative Component GARCH. *J. Fin. Econometrics*, 10, 54-83.

Feng, Y. (2004). Simultaneously modelling conditional heteroskedasticity and scale change. *Econometric Theory*, 20, 563-596.

Feng, Y. and Heiler, S. (1998). Locally weighted autoregression. In: R. Galata and H. Küchenhoff (eds.) *Econometrics in Theory and Practice*, pp. 101-117, Physica-Verlag, Heidelberg.

Feng, Y. and Heiler, S. (2009). A simple bootstrap bandwidth selector for local polynomial fitting. *J. Statist. Comput. Simul.*, 79, 1425-1439.

Feng, Y. and McNeil, A. J. (2008). Modelling of scale change, periodicity and conditional heteroskedasticity in return volatility. *Economic Modelling*, 25, 850-867.

Härdle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press, New York.

Härdle, W., Tsybakov, A.B. and Yang, L. (1998). Nonparametric vector autoregression. *J. Statist. Pl. Inf.*, 68, 221-245.

He, C. and Teräsvirta, T. (1999a). Properties of Moments of a Family of GARCH Processes. *J. Econometrics*, 92, 173-192.

He, C. and Teräsvirta, T. (1999b). Fourth Moment Structures of the GARCH (p,q) Process. *Econometric Theory*, 15, 824-846.

Heiler, S. and Feng, Y. (1998). A simple root n bandwidth selector for nonparametric regression. *J. Nonpar. Statist.*, 9, 1-21.

Herrmann, E., Engel, J., Wand, M. P. and Gasser, T. (1995). A Bandwidth Selector for Bivariate Kernel Regression. *J. Roy. Statist. Soc., Ser. B*, 171-180.

Klüppelberg, C., Lindner, A. and Maller, R. (2004). A continuous time GARCH process driven by a Levy process: Stationarity and second order behaviour. *J. Appl. Probab.*, 41, 601-622.

Müller, H.-G. (1985). Empirical bandwidth choice for nonparametric kernel regression by means of pilot estimators. *Statist. Decisions*, 2, Supp. Issue, 193-206.

Ruppert, D. and Wand, M.P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.*, 22, 1346-1370.

Scott, D.W. (1992). *Multivariate Density Estimation - Theory, Practice, and Visualization.* John Wiley & Sons, New York.

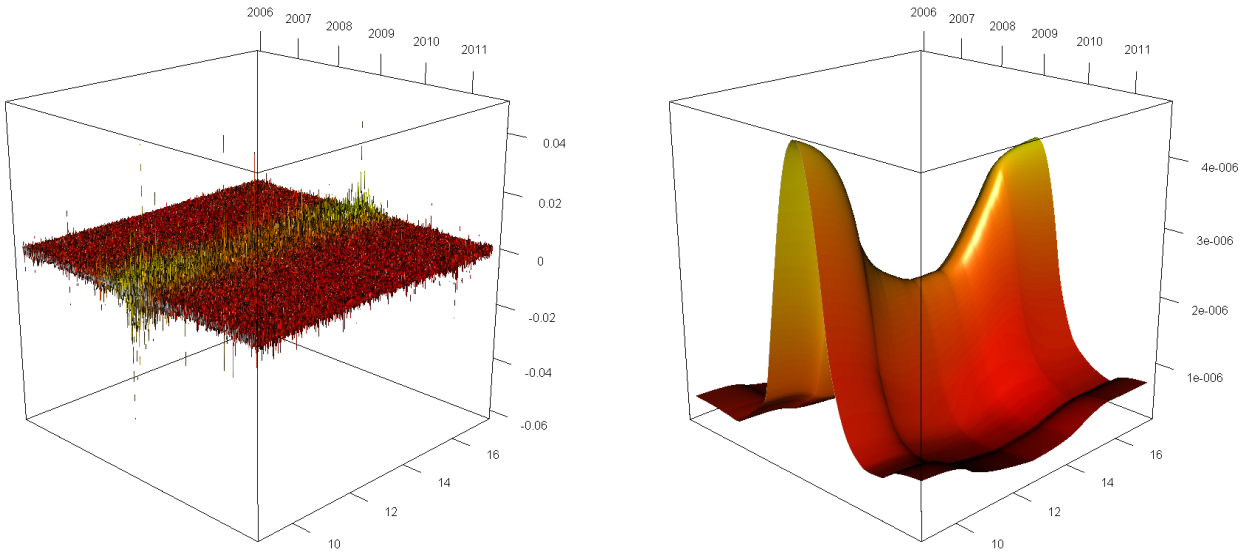

Figure 1: Spatial representation of the 1-minute Allianz returns (left) and the finally estimated volatility surface (right).

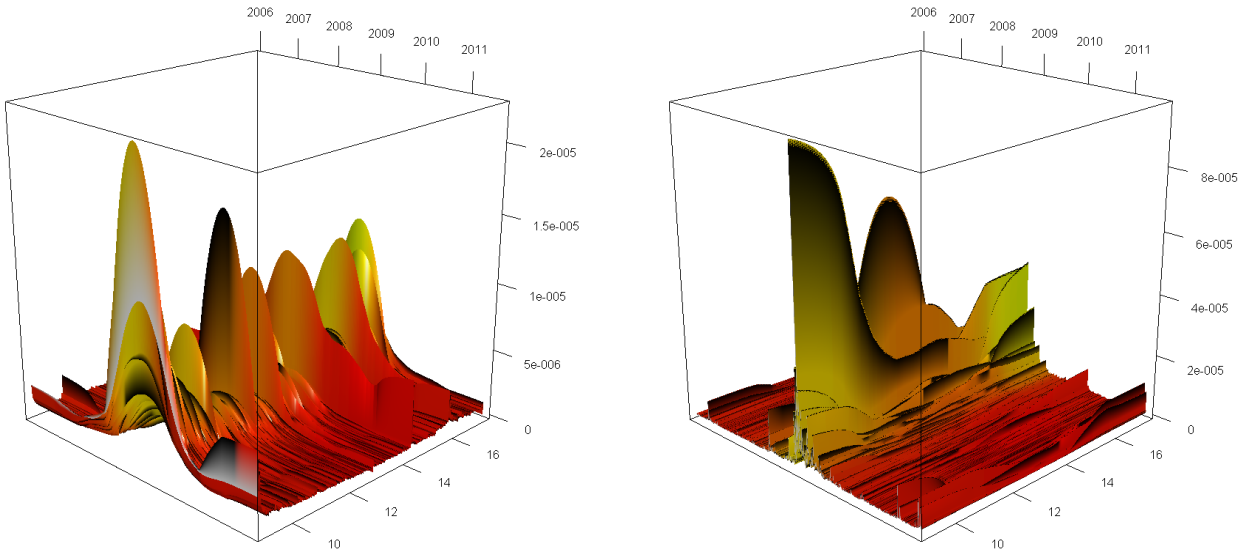

Figure 2: Intermediate smoothing results conditioning on trading time (left) and those conditioning on trading day (right).

Recent discussion papers

2013-08	Yuanhua Feng	Double-conditional smoothing of high-frequency volatility surface in a spatial multiplicative component GARCH with random effects
2013-07	Sonja Brangewitz Claus-Jochen Haake	Cooperative Transfer Price Negotiations under Incomplete Information
2013-06	Mark Schopf	Preserving Eastern or Offshore Oil for Preventing Green Paradoxes?
2013-05	Mark Schopf Hendrik Ritter	Unilateral Climate Policy: Harmful or even Disastrous? *[forthcoming in: Environmental and Resource Economics]*
2013-04	Volker Seiler	Comment on Ameriks, Caplin, Leahy & Tyler (2007): Measuring Self-Control Problems
2013-03	B. Michael Gilroy Anastasia Heiman Mark Schopf	Gibt es eine optimale Frauenquote?
2013-02	Yuanhua Feng Chen Zhou	Forecasting financial market activity using a semiparametric fractionally integrated Log-ACD
2013-01	Berno Buechel Nils Röhl	Robust Equilibria in Location Games
2012-11	B. Michael Gilroy Mark Schopf Anastasia Semenova	Grundeinkommen und Arbeitsangebot: Die Perspektive Deutschlands
2012-10	N. Bilkic B. Carreras Painter T. Gries	Unsustainable Sovereign Debt - is the Euro Crisis only the Tip of the Iceberg?
2012-09	B. Michael Gilroy Heike Schreckenberg Volker Seiler	Water as an Alternative Asset
2012-08	B. Michael Gilroy Mark Schopf Anastasia Semenova	Basic Income and Labor Supply: The German Case
2012-07	Tim Krieger	Es geht nicht ohne internationalen Ordnungsrahmen: Zum Verhältnis von Markt und Staat im Angesicht der Finanzkrise *[forthcoming in: Amos international: Internationale Zeitschrift für christliche Sozialethik (2012), 3]*
2012-06	Thomas Gries Margarete Redlin	Trade Openness and Economic Growth: A Panel Causality Analysis
2012-05	Mark Schopf Hendrik Ritter	Reassessing the Green Paradox *[republished as CIE Working Paper 2013-05 "Unilateral Climate Policy: Harmful or even Disastrous?"]*
2012-04	Yuanhua Feng David Hand Keming Yu	A Multivariate Random Walk Model with Slowly Changing Drift and Cross-correlation Applied to Finance