Modelling of long-term and conditional volatility dynamics in high-frequency returns at fixed trading time points us-ing a general Semi-GARCH model

by

Xiaojing Hou

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Fakultät für Betriebswirtschaftslehre

Fachgebiet konometrie und quantitative Methoden

Warburger Straße 100

33098 Paderborn

Modelling of long-term and conditional volatility dynamics in high-frequency returns at fixed trading time points us-ing a general Semi-GARCH model

Master's Thesis

by

XIAOJING HOU

Hoehen Str.28

33098 Paderborn

Thesis Supervisor:

Prof. Yuanhua Feng

and

xxx

Paderborn, June 28, 2015

Declaration

(Translation from German)

I hereby declare that I prepared this thesis entirely on my own and have not used outside sources without declaration in the text. Any concepts or quotations applicable to these sources are clearly attributed to them. This thesis has not been submitted in the same or substantially similar version, not even in part, to any other authority for grading and has not been published elsewhere.

Original Declaration Text in German:

Erklärung

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen worden ist. Alle Ausführungen, die wörtlich oder sinngemäß übernommen worden sind, sind als solche gekennzeichnet.

City, Date	Signature

Contents

List of Figures

List of Tables

1 Introduction

Trading off risks against returns appears to be essential and vital for making a financial decision. Hence the econometric analysis of risk (volatility) becomes an important part in forecasting market tendency and supporting making financial decisions, such as portfolio diversification, risk management and derivative pricing. In the last 20 years volatility was a research hotspot in financial industry. Volatility is regarded as a parameter for evaluating the risk of assets return. Generally, the stronger the volatility is, the higher the risk is.

In the traditional financial models, the variance of the time series is always assumed as constant. However, it is found that the volatilities of financial time series have always the features of clustering and fat tails (Mandelbrot, 1963; Eugene F. Fama, 1965). These features obviously are not consistent with the assumption of constant, so the traditional econometrical methods cannot analyze the time series variables efficiently in practice. To overcome this problem, several economists have carried out studies on researching and developing frameworks for evaluating volatility. Since Engle introduced the autoregressive conditional heteroskedasticity (ARCH) model (Engle, 1982), he extensions of ARCH model appeared and spread rapidly. Among the carried out researches, the Generalized ARCH (GARCH) model and its derivatives are most widely used (Bollerslev, 1986).

Instead of parametric models or nonparametric models, recently proposed semi parametric models will be introduced in details in this paper. According to many studies (Gourieroux and Monfort, 1992; Eubank and Härdle, 1993), preselected parametric models might be too restricted or too low-dimensional, which may not fit unexpected features and cause the misspecification. However, in nonparametric models, the parameters of the model cannot be estimated, and the model cannot be explained due to lack of specific functions. The semi parametric models do not need a prespecified function and is less sensitive to model misspecification. At the same time, the model can be also explained (Di and Gangopadhyay, 2011).

In this paper, the definition, estimation, some properties of semi-parametric models and the methods of bandwidth selection are discussed. Furthermore, according to the study of the semi-parametric GARCH model and semi-parametric asymmetric power ARCH model, which are introduced by Feng (Feng, 2004) (Feng and Sun, 2013), the semi-parametric method will trying to enter into exponential GARCH and component GARCH models. Then, the discussed semi-parametric models, i.e. semi-parametric APARCH, EGARCH and CGARCH models, are applied to the stock prices of BMW and Allianz from January 2006 to September 2014. Different from other papers, to get the more exact analyzing results the fixed trading time points will be used in this paper. (Lack of a simple conclusion of the analysis)

The plan of this paper is as follows. In section 2 the parametric models are introduced. The semi-parametric models are described in section 3. Section 4 reports the application of the semi-parametric models to the stock prices of BMW and Allianz and the empirical results on the volatility of the selected prices. Finally, this paper is concluded in section 5.

2 Parametric models

It is known that, volatility clustering exists in financial time series, and the distribution of random variables appears the fat tails. Mandelbrot observed the pattern of thick tails or fat tails of stock prices fluctuation, which was not consistent with the traditional assumption of normal distribution, by study the stock price index Mandelbrot (1963). Fama uggested that the financial market returns had the property of volatility clustering. It means that volatility is not only dynamic but also clustered, which appears strong in one certain period but weak in another period. The reasons for these phenomena could be traced back to speculation, political changes, government monetary and fiscal policy, and etc. Eugene F. Fama (1965). Volatility clustering and fat tails cannot be explained by the traditional economic models, which assume the variance is constant, until the Autoregressive Conditional Heteroscedasticity (ARCH) model was introduced by Engle (1982) Bollerslev et al. (1992).

Different from the traditional models, ARCH model suggests that the conditional variance could change over time as a function of past errors with the unconditional variance remaining constant Engle (1982). ARCH model soon became an important tool of volatility measurement, because it improved the traditional model and fit reality better. However, there are some drawbacks in ARCH model. In practical applications of the ARCH model, a relatively long lag in the conditional variance equation is often required, which might lead to increase in complexity of estimating parameters and decrease the freedom degree. When the lag order is high, estimating a totally free lag distribution will often violate the constraint of non-negativity parameters. But the restrict condition is exactly needed in this model to ensure conditional variance to be non-negative Bollerslev (1986). Therefore, there are many economists tried to improve ARCH models. Among these researches, the generalized ARCH (GARCH) model, which is introduced by Bollerslev (1986), is the most widely well known one with a better framework to study time-varying volatility in financial markets.

2.1 Parametric GARCH model

GARCH model added the lagged conditional variances to ARCH model, so that it has a longer memory and a more flexible lag structure than ARCH model.

2.1.1 Definitions and basic properties

In GARCH model, F_{t-1} denote the set of the past information. The GARCH process is presented in 2.1,

$$\varepsilon_t = \eta_t h_t^{1/2} \qquad \varepsilon_t | F_{t-1} \sim N(0, h_t), \tag{2.1}$$

in which ε_t is a random variables; η_t are i.i.d. random variables with zero mean and unit variance; h_t means the conditional variance of ε_t and is defined as follow:

$$h_t = \omega + \sum_{i=0}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}, \tag{2.2}$$

where $\omega > 0$ is the variances intercept parameter; $p > 0$ is the order of ARCH parameters; $q \geq 0$ is the order of GARCH parameters; α and β ($\alpha_i \geq 0, i =$

$1, \ldots, p; \beta_j \geq 0, j = 1, \ldots, q)$ are two vectors of unknown parameters respectively Bollerslev (1986).

From the model we can see that, GARCH model contains the lagged conditional variances in addition to ARCH model. When $q = 0$, the process is ARCH model; if $p = q = 0$, this process is a simple white noise; while $\sum_{i=1}^{p} \alpha_i < 1$, this process is an infinite dimensional ARCH (∞). That means that ARCH is a special form of GARCH process. Most of the cases, a high order ARCH model can be replaced by a low order GARCH model without breaking the constraint of non-negativity by estimating too many parameters. For example, In the application of ARCH (p) model, sometimes the value of p is very high; but for GARCH (p, q) model, $p = 1, q = 1$ are enough to achieve the same outcome Engle and Bollerslev (1986)

In the paper from Engle, he introduced the stationary of GARCH model. As it is known, a process $\{X_t\}$ is weakly stationary, if $i) E(X_t^2) < \infty$ for all $t \in T$ $, ii) E(X_t)$ and $cov(X_t, X_{t+s})$ are independent of t for $s \in Z, t \in T$ Bougerol and Picard (1992).

So the GARCH (p, q) process with $E(\varepsilon_t) = 0, Var(\varepsilon_t) = \omega(1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j)^{-1}$ and $cov(Y_t, Y_s) = 0$ for $t \neq s$ is wide-sense stationary if and only if $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ Bollerslev (1986)

2.1.2 Estimation of GARCH model

Usually, the maximum likelihood estimation is most widely used to estimate GARCH model. Besides the condition of stationary, to estimate GARCH model, a very important assumption is finite fourth-order moment, which ensures the existence of the squared returns variance. Assuming GARCH model is a normal process:

$$E(\varepsilon_t^4) = \frac{3\omega^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-\beta_1^2-2\alpha_1\beta_1-3\alpha_1^2)}. \tag{2.3}$$

Under this condition, $E(\varepsilon_t^4) < \infty$ if and only if $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$ for GARCH(1,1) mode Milhj (2012).

Letting $\theta = (\omega, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q)$, the conditional Gaussian log-likelihood function is defined as follows:

$$L(\theta) = \frac{1}{n}\sum_{t-1}^{n} l_t, \tag{2.4}$$

where

$$l_t = -\frac{1}{2}ln(h_t^2(F_{t-1};\theta)) - \frac{\varepsilon_t^2}{2h_t^2(F_{t-1};\theta)}. \tag{2.5}$$

Then maximize $L(\theta)$ by calculating and putting the partial derivatives to zero. After n times iteration, the approximated value (MLE) of θ is obtained, which is denoted as $\hat{\theta}$.

The maximum likelihood method is used by most of cases in GARCH family. Based on the assumption of normal distribution, GARCH model can partially explain the clustering and fat tails properties of the returns unconditional distribution. However, after the standardized by GARCH model, the residuals also have the properties of clustering and fat tails, which is not consistent with the normal assumption. So Bollerslev and Nelson used student-t distribution and generalized error distribution (GED) instead of normal distribution to estimate GARCH models Bollerslev (1986)Nelson (1991). When we estimate other GARCH models, which will be discussed below as examples, the main method is the same, maximum likelihood method. The most significant difference is the assumption of the distribution.

Although a simple GARCH model can solve the problem of the ARCH model with high-order linear declining lag structure, it also has some obvious drawbacks. The two constraint conditions of GARCH model, parametric non-negativity and bounded, restrict its applicability. The assumption of non-negative parameters raises the difficulty of estimation. Also with this assumption, the GARCH model cannot well explain the asymmetry of volatility and leverage effects. In order to extend the application of GARCH model, many GARCH derivatives appeared aiming at different problems. For example, to explain the asymmetry of volatility, the exponential GARCH model (EGARCH) was proposed by Nelson Nelson (1991), and threshold GARCH (TGARCH) was suggested by Zakorian Zakoian (1994). Ding, Granger and Engle introduced the APARCH model (Ding et al., 1993), which are able to explain the leverage effects in financial market better. Engle and Lee defined the component GARCH (CGARCH) model to exactly interpret the shock influences to the long-run and the short-run volatility respectively Engle and Lee (1999).

2.2 Parametric asymmetric power ARCH model

Generally, if a data series follows the normal distribution, we can characterize it by its first two moments. However, if the error of the data follows a non-normal distribution, then higher moments of skewness, kurtosis have to be applied, which describes the data beyond to adequately. In this instance, other power transformations appeared to be more appropriate than squared term McKenzie and Mitchell (1999). Ding, Granger and Engle (1993) further studied the autocorrelation of square returns and absolute returns based on GARCH model and Taylor model. It is found that, long lags between absolute or square returns are more correlated than the returns themselves for the return process Taylor (1986). Furthermore,

$|r_t|^d$ has the largest autocorrelation when d is around 1, but smaller monotonically one when d deviates from 1. Based on this they imposed an asymmetric GARCH model, asymmetric power ARCH model, which enables the power of the heteroskedasticity equation to be estimated from the data.

The parametric asymmetric power ARCH model is defined by

$$h_t^{\delta/2} = \omega + \sum_{i=1}^{p} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j h_{t-j}^{\delta/2}, \tag{2.6}$$

where $\omega > 0, \delta \geq 0, \alpha_i \geq 0, i = 1, \dots, p, -1 < \gamma_i < 1, i = 1, \dots, p, \beta_j \geq 0.j = 1, \dots, q$. γ is the leverage parameter, which takes the asymmetric news impact on the volatility into account, and δ is the parameter for the power term, which is a suitable positive number.

This model exhibits a power transformation of the conditional standard deviation process, which can linearize nonlinear models. It also reveals the asymmetric absolute residuals, which can reflect the leverage effect of the stock market returns.

In APARCH model, when γ_t is conditional normal, and

$$\frac{1}{\sqrt{2\pi}} \sum_{i=1}^{p} \alpha_i \{(1+\gamma_i)^{\delta} + (1-\gamma)^{\delta}\} 2^{\frac{\delta-1}{2}} \Gamma(\frac{\delta+1}{2}) + \sum_{j=1}^{q} \beta_j < 1, \tag{2.7}$$

which is the condition to ensure the existence of $Eh_t^{\delta/2}$ and $E|\varepsilon_t|^{\delta}$, then when $\delta \geq 2$, ε_t is covariance stationary (sufficient condition) Ding et al. (1993).

In the cases of the values of δ and γ_i are changed, APARCH model derives into several models, the GJR-GARCH model, the TS-GARCH model(Taylor and Schwert model), the NGARCH model (Nonlinear GARCH model) and the TGARCH model (threshold GARCH model).

When $\delta = 2, \gamma_i = 0$, the APARCH model turns into a GARCH model with the covariance stationary for ε_t as $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ Bollerslev (1986).

When $\delta = 2, \gamma_i \neq 0$, the APARCH model can be named as the GJR (Glosten, Jagannathan and Runkle, 1993) model

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}^2 I(\varepsilon_{t-i} < 0) + \sum_{j=1}^{q} \beta_j h_{t-j}, \tag{2.8}$$

with the covariance stationary for ε_t as $\sum_{i=1}^{p} \alpha_i [1 + \gamma_i^2] + \sum_{j=1}^{q} \beta_j < 1$.

This model uses the indicator function I to simulate the asymmetric influence of the positive and negative shocks on the conditional variance Glosten et al. (1993).

When $\delta = 1, \gamma_i = 0$, the APARCH model transforms into the TS-GARCH model

$$h_t^{1/2} = \omega + \sum_{i=1}^{p} \alpha_i |\varepsilon_{t-i}| + \sum_{j=1}^{q} \beta_j h_{t-j}^{1/2}, \tag{2.9}$$

with the covariance stationary for ε_t, the same TS-GARCH model, as $\sqrt{\frac{2}{\pi}} \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{\alpha} \beta_j < 1$ Zakoian (1994)

2.3 Parametric exponential GARCH model

It is considered that, GARCH model has taken the focus on the magnitude matters of excess return, but ignored the information of directions. Moreover, it is proven by empirical evidence that the direction does impact on volatility. Especially for broad-based equity indices and bond market indices, it is observed that market declines forecast higher volatility than comparable market increases do Engle (2001). In addition, the estimated coefficients often violate the parameter restrictions in GARCH model. It is also found that the left tail of the returns distribution was fatter than the right one. The reason is generally considered as the asymmetric response of investors to good news and bad news Jondeau and Rockinger (2003). To solve this problem Nelson (1991) imposed the exponential GARCH model.

Instead of making h_t a linear combination of positive random variables, EGARCH model ensure that h_t remains nonnegative by making $ln(h_t)$ linear in some function g of time and lagged η_ts.

The exponential GARCH model can be written as follows:

$$ln(h_t) = \omega + \sum_{i=1}^{p} \alpha_i g(\eta_{t-i}) + \sum_{j=1}^{q} \beta_j ln(h_{t-j}), \tag{2.10}$$

where

$$g(\eta_t) \equiv \theta\eta_t + \gamma[\mid \eta_t \mid - E \mid \eta_t \mid]. \tag{2.11}$$

In this model, $\beta \equiv 1$; α and β are real, non-stochastic and scalar sequences; $g(\eta_t)$ is a i.i.d. random sequences with zero mean Malmsten (2004).

EGARCH model provides a good solution for the problems of GARCH. There is no constraint for parameters. The function $g(\eta_t)$ contains two parts, $\gamma[|\eta_t| - E|\eta_t|]$ that define the size effect, and $\theta\eta_t$ that defines the sign effect of the shocks on volatility. The size effect is a typical ARCH effect, but the sign effect is asymmetrical. For example, the leverage effect (Black, 1976), which means that volatility rises in response to negative shocks and falls in response to positive ones.

According to the Theorem 2.1 from Nelson, when γ and θ do not equal to zero at the same time, the EGARCH process is strictly stationary and ergodic if and only if $\sum_{i=1}^{p} \alpha_t^2 < \infty$.

The estimation of EGARCH model is mostly same as the estimation of GARCH model. The difference between them is that, in EGARCH model, the generalized error distribution (GED) is used. Because this distribution includes the normal as special case; and when the process is strictly stationary and the distribution of η_t is not too fat tailed, h_t and ε_t have finite unconditional moments of arbitrary order Nelson (1991).

2.4 Parametric component GARCH model

It is known that the stock prices always fluctuate around an average value. This phenomenon is called mean-revert. It is also found that mean-revert of short-term volatility is more rapid than for the long-term one Xinzhong Xu and Stephen J.Taylor (1994). In order to explain this different response between short-term and long-term, Engle and Lee (1999) imposed component GARCH model. In this model, the conditional variance is decomposed into a permanent and transitory component. So that the model can be used to investigate the long-run and short-run movements of volatility affecting securities.

The component GARCH model is defined by

$$h_t = q_t + \sum_{i=1}^{p} \alpha_i(\varepsilon_{t-i}^2 - q_{t-i}) + \sum_{j=1}^{q} \beta_j(h_{t-j} - q_{t-j}), \tag{2.12}$$

$$q_t = \omega + \rho q_{t-1} + \varphi(\varepsilon_{t-1}^2 - h_{t-1}) \tag{2.13}$$

where q_t the permanent component of the conditional variance and $(h_{t-j} - q_{t-j})$ is the transitory component of the conditional variance (Ghalanos, 2011). $(\alpha + \beta)$ is the mean-revert of s_t. In terms of economy, the smaller the mean-reverting rate, the less persistent the expected volatility to market shocks in the past. That means, when the market receives the information of shocks, volatility responses quickly but has lower persistence. If $0 < (\alpha + \beta) < 1$, the mean-reverts of short-run volatility component is zero at a geometric rate of $(\alpha + \beta)$; If $0 < \rho < 1$, the long-run volatility component follows an AR process, and will converge to a constant level defined by $\omega/(1 - \rho)$. In this model, the long-run component is more persistent than the short-run one, i.e., $0 < (\alpha + \beta) < p < 1$

In GARCH model, the conditional variance must be non-negative. This condition is also required in component GARCH model. But this is not demanded by s_t, which can be positive or negative over time to exhibit the self-correction feature of the mean-reverting volatility process. The condition of stationary in GARCH model is $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{\alpha} \beta_j < 1$, besides which this stationary condition also includes $(\alpha + \beta)(1 - \rho) + \rho < 1$ and requires $\rho < 1, (\alpha + \beta) < 1$ in component model. When $\rho = 1$, the mean series is not a covariance stationary process, but also is still a strictly stationary and ergodic process.

From the above analysis, the stationary conditions of component GARCH model can be concluded as $0 < (\alpha + \beta) < \rho < 1, 0 < \varphi < \beta, \alpha > 0, \beta > 0, \alpha_0 > 0, \varphi > 0$, But they are just sufficient conditions, not necessary ones Engle and Lee (1999).

3 Semi-parametric models

Parametric GARCH model has many advantages. The function form is accessible, and parameters could be easily estimated. If the model assumptions were correct, the estimation is consistent with reality. However, the drawbacks of parametric models are more disgusting. Firstly, a preselected parametric model may not fit unexpected features, due to too restricted or too low dimensional. Secondly, sometimes the regression function seems to be too complicate or difficult to be defined. Thirdly, because different sequence will be witnessed when different conditional distribution is selected in the process of prediction by using parametric models, there will most possibly exist the problem of misspecificationwhich may result in a excessively high model bias and loss of efficiency, unless the assumed function perfectly matches the true error distribution (Di and Gangopadhyay, 2011).

Gourieroux and Monfort had firstly inserted the conditional mean and conditional vari-ance into a nonparametric model, in which the function form is not given. This model can supply the gaps of parametric models. Correspondingly, it is less sensitive to model misspecification and can offer a flexible tool in analyzing unknown regression relationships (Gourieroux and Monfort, 1992). In nonparametric regression models, both the error distribution and the functional form of the mean function are not pre-specified. It will be more useful when the regression function seems to be very complex or a suitable function form cannot be found (Eubank and Härdle, 1993). However, the nonparametric models lose some functions, which can be completed by parametric models. Due to lack of specific function forms, the parameters of the model cannot be estimated; further-more the model cannot be explained. Excess variable estimates could rise up because of poor consideration of nonparametric models, especially for small sample size. So a new model, semi-parametric model, is proposed.

Semi-parametric model consists of not only the nonparametric part but also parametric part. Semi-parametric GARCH model selects the nonparametric form for the scale func-tion and the parametric part for conditional variance lags. Then it does not need a pre-specified function form and is less sensitive to model misspecification. At the same time, the parametric part can also explain the models (Di and Gangopadhyay, 2011).

3.1 Semi-parametric GARCH model

3.1.1 Definition of semi- parametric GARCH model

Semi-parametric GARCH model combines a smooth scale function with the standard GARCH model:

$$Y_i = \mu + \sigma(t_i)\varepsilon_i \tag{3.1}$$

where μ is an unknown constant; $t_i = t/n; \sigma(t) > 0$ is the nonparametric component, a smooth, bounded scale function; and $\{\varepsilon_t\}$ is the parametric component, which is assumed to follow a GARCH(p, q) process.

$$h_i = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}, \tag{3.2}$$

where $h_i^{1/2}$ is the conditional standard deviations of the standardized process

ε_i; $\omega > 0; \alpha_1, \ldots, \alpha_p \geq 0$ and $\beta_1, \ldots, \beta_q \geq 0$. To estimate $v(t)$, $E(\varepsilon_t^8) < \infty$ is assumed to ensure 3.2 strictly stationary, which implies in particular that $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ (Feng, 2004).

3.1.2 Estimation of semi-parametric GARCH model

The estimation of semi-parametric GARCH model can combine the nonparametric estimation of the Y's local variance $v(t) = \sigma^2(t)$,with parametric estimation of the unknown parameter vectors $\theta = (\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)$.

Firstly, $v(t)$ can be estimated by some nonparametric regression approach consistently without any parametric assumptions. In this paper the kernel estimation will be used to estimate $v(t)$. 3.1 needs to be transformed into a general nonparametric regression problem at first. Let $r_i = Y_i - \mu$, $X_i = r_i^2$ and $\xi_i = \varepsilon_i^2 - 1 \geq -1$, which are zero mean stationary time series errors. Then the model 3.1 can be rewritten as

$$X_i = v(t_i) + v(t_i)\xi_i. \tag{3.3}$$

If the constant mean μ is replaced by a sooth function g, we can get a nonparametric regression with scale change and dependence

$$Y_i = g(t_i) + \sigma(t_i)\varepsilon_i, \tag{3.4}$$

where ε_i is a zero mean stationary process.

From the above discuss we can see that, model 3.3 is a special case of the model 3.4; and model 3.4 can be also written as $r_i = \sigma(t_i)\varepsilon_i$ for $r_i = Y_i - \mu$.

Let $\hat{\mu} = \bar{y}$ and $\hat{x}_i = \hat{z}_i^2$, in which $\hat{z}_i$ is then defined by $\hat{z}_i = y_i - \bar{y}$. A Nadaraya-

Watson kernel regression which has been proposed by Nadaraya (1964) and Watson (1964), is defined by

$$\hat{v}(t) = \frac{\sum_{i=1}^{n} K(\frac{t_i - t}{b})\hat{x}_i}{\sum_{i=1}^{n} K(\frac{t_i - t}{b})} =: \sum_{i=1}^{n} W_i \hat{x}_i, \tag{3.5}$$

where W_i is the weighting function $W_i = \frac{K(\frac{t_i - t}{b})}{\sum_{i=1}^{n} K(\frac{t_i - t}{b})}$; $K(u)$ is a second order kernel function with compact support [-1,1]; and b is the bandwidth, the size of the weights Fan and Truong (1993).

According to the above assumptions, the estimator ε_i is now replaced by the standardized residuals

$$\hat{\varepsilon}_i = \hat{z}_i / \hat{\sigma}(t_i) = (y_i - \bar{y}) / \hat{\sigma}(t_i) \tag{3.6}$$

Then the estimator of parametric vector θ can be obtained by the standard maximum likelihood method, which has been introduced in section 2.

(Lack of the introduction of boundary problem)

3.2 Asymptotic properties of $\hat{v}$

The estimators of kernel regression function approximate the real value asymptotically. There also exists the bias between the real v and the estimators of kernel regression function $hatv$. To calculate the bias, several assumptions must be satisfied:

1. $E(\varepsilon_t^8) < \infty$ ensures the process strictly stationary, and $\eta \sim N(0,1)$Ling and McAleer (2002)

2. The second derivative of $v(t)$ exists and is continuous on $[0, 1]$.

3. The kernel $K(u)$ is a continuous function, which is symmetric around zero and defined in $[-1, 1]$.

4. When the sample volume $n \to \infty$, the bandwidth $b \to 0$ and $nb \to \infty$. Then both of the bias B and the variance V approaches zero.

Define $R(K) = \int K^2(u)du$ and $I(K) = \int u^2K(u)du$. Under these assumptions, the asymp-totic bias and the asymptotic variance can be obtained. The asymptotic bias B of $\hat{v}(t)$ is:

$$B = E[\hat{v}(t) - v(t)] = \frac{I(K)v''(t)}{2}b^2 + o(b^2)). \tag{3.7}$$

The asymptotic variance V of $\hat{v}(t)$ can be expressed as:

$$V = var[\hat{v}(t)] = R(K)\frac{v(t)}{nb} + o\frac{1}{nb}. \tag{3.8}$$

From 3.7 and 3.8 we can see that, bias is of the order b^2 and the variance is of the order $1/(nb)$. The bias has the positive correlation with the bandwidth, whereas the variance has the negative correlation with the bandwidth. The higher the $v(t)$, the larger the variance; and the more complex the $v(t)$, the larger the bias. Under the assumptions, the asymptotic MISE can be calculated by:

$$MISE = \int (B^2 + V)dt = \frac{I^2(K)R(V''(t))}{4}b^4 + \frac{R(K)}{nb} + o[max(b^4, \frac{1}{nb})]. \tag{3.9}$$

By minimizing the dominant part of the asymptotic MISE, the asymptotically optimal bandwidth of $\hat{v}$ is

$$b = (\frac{R(K)}{I^2(K)R(V''(t))})^{\frac{1}{5}}n^{-\frac{1}{5}}. \tag{3.10}$$

When the bandwidth is of order $n^{-\frac{1}{5}}$, we can get $\hat{v}(t) = v(t)[1 + O_p(n^{-\frac{2}{5}})]$ and $MISE = O(n^{-\frac{4}{5}})$ Gasser and Müller (1984) Fan (1991).

3.2.1 The selection of bandwidth

Applying the estimator $\hat{v}$ requires the specification of kernels and bandwidth. Optimal kernels have been obtained analytically. The selection of bandwidth becomes the most important problem when applying nonparametric regression estimators such as kernel estimators. The regression works well, only if the bandwidth is suitable. Because the kernel estimation uses the points around t_0 to estimate the scale function, a kernel regression is usually biased. The larger the bandwidth, the larger the square bias because further points from t_0 are used, but the smaller the variance because more ob-servations are used for estimation. The bandwidth should be optimized to balance the variance and bias. The optimal bandwidth is the one, which can minimize the mean squared error (MSE) or mean integrated squared error Gasser et al. (1991).

There are many methods available to optimize bandwidth, e.g. Cross-Validation (CV), Generalized CV (GCV), plug-in and etc. In this paper, CV and plug-in will be introduced in detail.

Cross-Validation (CV) The Cross-Validation method is also called leave-one-out. This method estimates the MSE by leaving one observation out at each time point. The main idea of CV is that, the estimation of $v(t_i)$ is without using $(t_i, \hat{x}_i)$, rather assess the quality of the estimation with other observations.

Define the estimation errors by $r_{-i} = (t_i, b) = \hat{x}_i - \hat{v}^{-i}(t_i)$. The function to calculate CV is:

$$CV(b) = \sum_{i=1+i_0}^{n-i_0} r_{-i}^2(t_i, b) = \sum_{i=1+i_0}^{n-i_0} [\hat{x}_i - \hat{v}^{-i}(t_i)]^2, \tag{3.11}$$

where $\hat{v}^{-i}(t_i) = \sum_{j=1, j\neq i}^{n} \frac{K(\frac{t_j - t_i}{b})\hat{x}_j}{\sum_{j=1,j\neq i}^{n} K(\frac{t_j-t_i}{b})}$ is the leave-one-out estimator of $v(t_i)$.

Let $0 < h_{min} < h_{max} < 1, \triangle h = (h_{max} - h_{min})/m$, where m is an integer. Then the optimal bandwidth can be obtained by calculating the value of CV at $h = h_{min}, h = h_{min} + \triangle h, h = h_{min} + 2 \triangle h, \ldots, h = h_{max}$. The optimal bandwidth is the one, at which CV(b) is minimized Sarda (1993).

Cross- Validation is the most simple method for bandwidth selection and easy to understand. However, this method has large variability, which leads often a suboptimal non-parametric fit of the regression function. For example, the selected bandwidth by CV is often very large. Sometimes the resulted band width can be very small and wrong Altman and Léger (1995).

Plug-in bandwidth selected method In this method, some adapted assumptions are required.

1. The function $v(t)$ is strictly positive on [0,1] and exist at least continuous fourth moment.
2. v'' is estimated with a symmetric fourth order kernel for estimating the second derivative with compacted sup-port [-1,1].
3. The bandwidth b satisfied $b \to 0$ and $nb^5 \to \infty$ as $n \to \infty$.

Denote $c_f = f(0)$, where $f(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} exp(ik\lambda)\gamma_\varepsilon(k)$ is the spectral density of ε_i with $\gamma_\varepsilon(k)$ is the autocovariance function of ε_i . Then the asymptotic optimal bandwidth can be rewritten as

$$b_A = (2\pi c_f \frac{R(K)}{I^2(K)} \frac{I(v^2)}{I((v'')^2)})^{\frac{1}{5}} n^{-\frac{1}{5}} = Cn^{-\frac{1}{5}}, \tag{3.12}$$

where

$$c_f = \frac{E(\varepsilon_i^4)}{3\pi} \frac{(1 - \sum_{j=1}^{q} \beta_j)^2}{(1 - \sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j)^2} \tag{3.13}$$

with the assumption, the functions $\phi(z) = 1 - \sum_{i=1}^{max(p,q)} \alpha_i z^i$ and $\varphi(z) = 1 - \sum_{j=1}^{max(p,q)} \beta_j z^j$ have no common roots. $\hat{E}(\varepsilon^4 = \sum_{i=1}^{n} \hat{\varepsilon})_i^4/n$ is a nonparametric estimator of ε_i^4 Feng (2004).

This method starts from the asymptotic optimal bandwidth $b_0 = Cn^{-1/5}$ with C=0.5.

Then follow j-th iteration process. In the j-th iteration, $\hat{v}$ and $\hat{\theta}$ should be calculated with the bandwidth b_{j-1} at first. Step 2, calculate $\hat{E}(\varepsilon^4)$, where $\hat{E}(\varepsilon^4) = \sum_{i=1}^{n} \hat{\varepsilon}_i^4/n$ is a nonparametric estimator of $E(\varepsilon^4)$, and $\hat{I}(v)^2$ with $\hat{I}(v)^2 = \frac{1}{n}\sum_{i=n_1}^{n_2} \hat{v}(t_i)^2$, where n_1 and n_2 denote the integer part of na and n(1-a) respecyively, in which [a, 1-a] is the compacted support of the mean integrated squared error (MISE). In this calculation, the selected bandwidth is $b_{\varepsilon,j} = b_{v,j} = b_{j-1}^{5/4}$. Step 3, calculate $\hat{c}_f$ from $\hat{\theta}$ and $\hat{E}(\varepsilon^4)$. Step 4, calculate $\hat{I}((v'')^2)$, where $\hat{v''}$ is a kernel estimator of v'' with $\hat{I}((v'')^2) = \frac{1}{n}\sum_{i=n_1}^{n_2} \hat{v}''(t_i)^2$, in which $\hat{v}''$ is obtained by using the bandwidth $b_{d,j} = b_{j-1}^{5/7}$. Step 5, improve b_{j-1} by

$$b_A = (2\pi\hat{c}_f \frac{R(K)}{I^2(K)} \frac{\hat{I}(v^2)}{\hat{I}((v'')^2)})^{\frac{1}{5}} n^{-\frac{1}{5}}$$

.

Finally, increase j by one and repeat the second part until convergence is reached, where the condition $b_j - b_{j-1} < 1/n$ is used as a criterion for the convergence of $\hat{b}$; or a given maximal number of iterations has been done, where the maximal number of iterations is put to be twenty. Put $\hat{b} = b_j$. The asymptotic performance of $\hat{b}$ is quantified by $(\hat{b} - b_A)/b_A = O_p(n^{-2/7} + O_p(n^{-1/2})$.

The plug-in estimator of the bandwidth has much lower variability than Cross-Validation for a broad variety of situations, including nonsmooth functions Gasser et al. (1991) Feng (2004).

3.3 Semi-parametric APARCH model

3.3.1 Definition of semi-parametric model

Denoting $r_i = Y_i - \mu, i = 1, \dots, n$ is the logarithmic returns from an asset. According to the above discussion 3 we can get the semi-parametric APARCH model, which is defined as follows:

$$r_i = \sigma(t_i)\varepsilon_i, \tag{3.14}$$

where $\sigma(t_i)$ is a smooth scale function and $\sigma(t_i) > 0$; $t_i = t/n$ is the rescaled time; ε_i is the rescaled process, which follows the parametric APARCH model with i.i.d. random variables $\eta$and are the conditional variance of the rescaled process h_t.

$$h_t^{\delta/2} = \omega + \sum_{i=1}^{p} \alpha_i(|\varepsilon_{t-i}| - \gamma_i\varepsilon_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j h_{t-j}^{\delta/2}, \tag{3.15}$$

where $\omega > 0, \delta \geq 0, \alpha_i \geq 0, i = 1, \dots, p, -1 < \gamma_i < 1, i = 1, \dots, p, \beta_j \geq 0.j = 1, \dots, q$. The same as introduced in Section 2, the parametric component should satisfy the assumptions of the parametric APARCH model, such as $\sum \alpha_i + \sum \beta_j < 1$, etc. γ is the leverage parameter and δ is the parameter for the power term Ding et al. (1993)

The semi-parametric model provides us a tool to decompose financial risk into an unconditional component $s(\tau)$, a conditional component $h_t^{1/2}$ and the i.i.d.

innovations η_t.

The reason to use semi-parametric APARCH model rather than the para-metric APARCH is that, if the scale function changes over time, the parametric component cannot be estimate consistently from the data, when the non-stationary scale function is not estimated. However, after estimating the nonstationary scale function an approximate stationary process for further analysis can be obtained. When the process follows a parametric model, the semi-parametric framework still works but with some loss of the efficiency (Feng and Sun, 2013).

3.3.2 Estimation of the semi-parametric APARCH model

3.4 Semi-parametric EGARCH model

According to the above discussions of semi-parametric GARCH and APARCH models, the semi-parametric EGARCH model can be defined as follows:

$$Y_t = \sigma(t_i)\varepsilon_i, \tag{3.16}$$

where Y is the return process; $\sigma(t_i) > 0$ is the scale function; t_i are the rescaled times; and ε_t are the standard residuals, which follows the parametric exponential GARCH model.

$$\varepsilon_t = \eta_t h_t^{\frac{1}{2}}, \tag{3.17}$$

$$ln(h_t) = \omega + \sum_{i=1}^{p} \alpha_i g(\eta_{t-i}) + \sum_{j=1}^{q} \beta_j ln(h_{t-j}), \tag{3.18}$$

where α and β are real, non-stochastic and scalar sequences; $g(\eta_t)$ is a i.i.d. random sequences with zero mean, which is defined as follows:

$$g(\eta_t) \equiv \theta\eta_t + \gamma[|\eta_t| - E|\eta_t|]. \tag{3.19}$$

The assumptions of this parametric component approximate to the parametric exponential GARCH model Nelson (1991).

3.5 Semi-parametric CGARCH model

where Y is the return process; $\sigma(t_i) > 0$ is the scale function; t_i are the rescaled times; and ε_t are the standard residuals, The semi-parametric CGARCH model can written as

$$Y_t = \sigma(t_i)\varepsilon_i, \tag{3.20}$$

where ε_t is assumed to follow the parametric component GARCH model,

$$h_t = q_t + \sum_{i=1}^{p} \alpha_i(\varepsilon_{t-i}^2 - q_{t-i}) + \sum_{j=1}^{q} \beta_j(h_{t-j} - q_{t-j}), \tag{3.21}$$

$$q_t = \omega + \rho q_{t-1} + \varphi(\varepsilon_{t-1}^2 - h_{t-1}) \tag{3.22}$$

In which q_t is the permanent component of the conditional variance and $(h_{t-j} - q_{t-j})$ is the transitory component of the conditional variance; and the parametric specifications assumption of stationary $0 < (\alpha + \beta) < \rho < 1, 0 < \varphi < \beta, \alpha > 0, \beta > 0, \omega > 0$ should be satisfied Engle and Lee (1999) (Ghalanos, 2014).

3.6 Conclusion

4 Application

4.1 Tables and Figures Sample

Table 4.1: Estimated coefficients of the APARCH - norm models for Sap index at open

	APARCH(1,1)		APARCH(1,2)		APARCH(2,1)		APARCH(2,2)	
	mCoeff	s.e.	Coeff	s.e.	Coeff	s.e.	Coeff	s.e.
μ	0.012459	0.013555	0.0124602	0.01356	0.011818	0.013940	0.012812	0.013729
ω	0.050010	0.006252	0.0500216	0.00687	0.054857	0.007283	0.097006	0.012668
α_1	0.076004	0.006336	0.0760131	0.00838	0.058272	0.011220	0.052233	0.009390
α_2					0.021706	0.011395	0.086890	0.008718
γ_1	1	0.005733	0.99999999	0.00572	1	0.007523	1	0.009469
γ_2					1	0.023297	1	0.011564
β_1	0.889283	0.009774	0.88929646	0.09400	0.879062	0.01205	0.098934	0.079439
β_1			0.00000001	0.08676			0.687549	0.073186
σ_1	1.017673	0.12041	1.01704876	0.1207	1.058141	0.127632	1.069076	0.128453

Table 4.2: Selected bandwidths in all times (T) of ALV

	T=09:30	T=11:00	T=12:30	T=14:00	T=15:30
d=1	0.09845639	0.0932583	0.09118069	0.09206592	0.1017511
d=2	0.1069041	0.1046261	0.09312175	0.1070296	0.08780973

Table 4.3: BIC of all selected models for ALV

	T=09:30	T=11:00	T=12:30	T=14:00	T=15:30
APARCH-t(1,1)	2.7117	2.7542	2.7543	2.7424	2.7558
APARCH-t(1,2)	2.7151	2.7577	2.7577	2.7458	2.7593
APARCH-t(2,1)	2.7183	2.7611	2.7606	2.7493	2.7622
APARCH-t(2,2)	2.7218	2.7645	2.7641	2.7524	2.7656
EGARCH-t(1,1)	2.7108	2.7528	2.7524	2.7396	2.7545
EGARCH-t(1,2)	2.7142	2.7562	2.7558	2.7430	2.7580
EGARCH-t(2,1)	2.7176	2.7597	2.7579	2.7458	2.7609
EGARCH-t(2,2)	2.7204	2.7631	2.7614	2.7492	2.7646
CGARCH-t(1,1)	2.7283	2.7671	2.7690	2.7592	2.7715
CGARCH-t(1,2)	2.7309	2.7706	2.7725	2.7627	2.7750
CGARCH-t(2,1)	2.7317	2.7706	2.7721	2.7625	2.7742
CGARCH-t(2,2)	2.7342	2.7740	2.7756	2.7660	2.7777

Table 4.4: Estimated coefficients for ALV at 09:30

	APARCH-t(1,1,)	EGARCH-t(1,1,)	CGARCH-t(1,1,)
η	0.030811	0.023525	0.046245
ω	0.072068	-0.008490	0.004690
α_1	0.065588	-0.121483	0.101784
β_1	0.841074	0.928114	0.835055
γ_1	0.720705	0.172949	_
δ	1.815581	_	_
η_{11}	_	_	0.995571
η_{21}	_	_	0.000000
shape	6.506437	6.391486	6.060666

Table 4.5: Tables and figures for ALV at 11:00

	APARCH-t(1,1,)	EGARCH-t(1,1,)	CGARCH-t(1,1,)
η	0.027935	0.028150	0.044178
ω	0.086802	-0.006788	0.005350
α_1	0.071707	-0.120221	0.091516
β_1	0.833555	0.916025	0.841185
γ_1	0.655442	0.142740	_
δ	1.660996	_	_
η_{11}	_	_	0.994889
η_{21}	_	_	0.000000
shape	7.303654	7.136397	6.944785

Table 4.6: Tables and figures for ALV at 12:30

	APARCH-t(1,1,)	EGARCH-t(1,1,)	CGARCH-t(1,1,)
η	0.024217	0.020136	0.043784
ω	0.085778	- 0.006793	0.003444
α_1	0.072438	- 0.119206	0.092086
β_1	0.839806	0.917423	0.833037
γ_1	0.743891	0.143880	-
δ	1.468959	-	-
η_{11}	-	-	0.996659
η_{21}	-	-	0.000000
shape	7.915656	7.797071	7.239844

Table 4.7: Tables and figures for ALV at 14:00

	APARCH-t(1,1,)	EGARCH-t(1,1,)	CGARCH-t(1,1,)
η	0.023732	0.023156	0.042847
ω	0.095697	-0.009432	0.003821
α_1	0.088357	-0.133916	0.099963
β_1	0.829213	0.903927	0.820708
γ_1	0.820682	0.164739	-
δ	1.147269	-	-
η_{11}	-	-	0.996322
η_{21}	-	-	0.000000
shape	7.431702	7.412819	6.838452

Table 4.8: Tables and figures for ALV at 15:40

	APARCH-t(1,1,)	EGARCH-t(1,1,)	CGARCH-t(1,1,)
η	0.011890	0.008475	0.029493
ω	0.070333	-0.005184	0.004089
α_1	0.063602	-0.110778	0.083003
β_1	0.857014	0.928123	0.846245
γ_1	0.664906	0.137833	-
δ	1.736078	-	-
η_{11}	-	-	0.995975
η_{21}	-	-	0.000000
shape	8.127734	7.980692	7.692529

Bibliography

Altman, N. and Léger, C. (1995). Bandwidth selection for kernel distribution function estimation. *Journal of Statistical Planning and Inference*, 46(2):195–214.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). ARCH modeling in finance.

Bougerol, P. and Picard, N. (1992). Stationarity of Garch processes and of some nonnegative time series. *Journal of Econometrics*, 52(1-2):115–127.

Di, J. and Gangopadhyay, A. (2011). On the efficiency of a semi-parametric GARCH model. *The Econometrics Journal*, 14(2):257–277.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

Engle, R. (2001). GARCH 101: The Use of ARCH / GARCH Models in Applied Econometrics. 15(4):157–168.

Engle, R. and Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. and Lee, G. G. (1999). *A long-run and short-run component model of stock return volaility.*

Eubank, R. L. and Härdle, W. (1993). Applied Nonparametric Regression. *Technometrics*, 35(2):225.

Eugene F. Fama (1965). The Behavior of Stock-Market Prices. *chicago journals*, 38(1):34–105.

Fan, J. (1991). Asymptotic normality for deconvolution kernel density estimators. *Journal of statistical*, 53(1):97–110.

Fan, J. and Truong, Y. K. (1993). Nonparametric Regression with Errors in Variables.

Feng, Y. (2004). SIMULTANEOUSLY MODELING CONDITIONAL HETEROSKEDASTICITY AND SCALE CHANGE.

Feng, Y. and Sun, L. (2013). A Semi-APARCH approach for comparing long-term and short-term risk in Chinese financial market and in mature financial markets § Yuanhua Feng 1 and Lixin Sun 2 1. pages 1–25.

Gasser, T., Kneip, a., and Kohler, W. (1991). A Flexible and Fast Method for Automatic Smooting. *Journal of the American Statistical Association*, 86(415):643–652.

Gasser, T. and Müller, H.-G. (1984). Estimating Regression Functions and Their Derivatives by the Kernel Method. *Scandinavian Journal of Statistics*, 11(3):171–185.

Ghalanos, A. (2011). Introduction to the rugarch package . *October*, pages 1–36.

Ghalanos, A. (2014). Introduction to the rugarch package . *October*, pages 1–36.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48(5):1779–1801.

Gourieroux, C. and Monfort, A. (1992). Qualitative threshold ARCH models. *Journal of Econometrics*, 52(1-2):159–199.

Jondeau, E. and Rockinger, M. (2003). Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, 10(5):559–581.

Ling, S. and McAleer, M. (2002). Necessary and sufficient moment conditions for the GARCH (r, s) and asymmetric power GARCH (r, s) models. *Econometric Theory*, 18(3):722–729.

Malmsten, H. (2004). Evaluating exponential GARCH models. *Finance*, pages 1–25.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices.

McKenzie, M. and Mitchell, H. (1999). Generalised asymmetric power ARCH modeling of exchange rate volatility. *Applied Financial Economics*.

Milhj, A. (2012). The Moment Structure of ARCH processes. *Statistics*, 12(4):281–292.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–370.

Sarda, P. (1993). Smoothing parameter selection for smooth distribution functions. *Journal of Statistical Planning and Inference*, 35(1):65–75.

Taylor, S. (1986). Modelling Financial Time Series.

Xinzhong Xu and Stephen J.Taylor (1994). The term structure of volatility implied by foreign exchange options. 29(1):57–74.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of economic dynamics and control*, 18:931–955.