Bayesian Analysis of the Functional-Coefficient Autoregressive Heteroscedastic Model

Cai Jing-Heng

Department of Statistics, Sun Yat-sen University

caijheng@mail.sysu.edu.cn

Joint work with

Song Xin-Yuan, Department of Statistics, The Chinese University of Hong Kong

Feng Xiang-Nan, Department of Statistics, The Chinese University of Hong Kong

Jiang Xue-Jun, Department of Mathematics, South University of Science and Technology

2015.11.14, Zhuhai

ARMA Models and Beyond

Since the pioneering work of the autoregressive moving average (ARMA) models (Box et al., 2008) to analyze time series data, the original ARMA framework has been extended in numerous aspects.

However, most of these analyses were restricted to linear modeling. Given the existence of various nonlinear phenomena, such as asymmetric cycles and bimodality, nonlinear relationships among lagged variables have been observed in real time series data sets (Tong, 1990; Tjøstheim, 1994; Tong, 1995). Moreover, these nonlinear features are beyond the capacity of linear models.

Thus, nonlinear time series analysis has received a great deal of attention over the past thirty years.

Nonlinear Modeling and Limitations

At the early stage of nonlinear time series development, analyses focused on known parametric forms, such as the threshold autoregressive (TAR) model (Tong, 1990) and the exponential autoregressive (EXPAR) model (Haggan and Ozaki, 1981).

An evident limitation of these parametric (both linear and nonlinear) models is that they are too restrictive for many applications since the functional forms for the relationships among variables have to be specified. However, nonlinear functions have many different types, and the functional form of a specific relationship is seldom known in advance. In addition, strict parametric functional forms are likely to miss subtle patterns.

FAR Model

To relax the restrictions in parametric models, functional-coefficient autoregressive (FAR) models have been developed recently.

The advantage of FAR models lies in the flexibility to accommodate most nonlinear features with only minimal prior information assumed.

Further, such models can be employed as an exploratory tool for investigating functional forms. Given their flexibility in dealing with nonlinear relationships, FAR models have been widely used over the past decades (Chen and Tsay, 1993; Cai et al., 2000; Fan and Yao,2003).

Volatility

Volatility is important in asset pricing, monetary policymaking, proprietary trading, portfolio management, and risk analysis. Thus, modeling and predicting volatility is of great importance.

A widely used approach for modeling volatility is the discrete time method, in which the volatility is considered to be a conditional variance of the return.

Since the autoregressive conditional heteroscedasticity (ARCH) model was first introduced by Engle (1982) as an important tool for modeling volatility, numerous variants of the ARCH model have been proposed. Bollerslev (1986) extended the ARCH model into a generalized ARCH (GARCH) model.

Nonlinear Modeling and Limitations

Tong (1990) proposed a self-exciting threshold ARCH (SETAR-ARCH) model with changing conditional variance, which has a piecewise linear conditional mean and an ARCH innovation.

Li and Li (1996) subsequently extended the TAR model to the double-threshold ARCH (DTARCH) model, which can address conditions where both the conditional mean and the conditional variance specifications are piecewise linear given previous information.

Although the aforementioned models are useful for describing the conditional mean and/or conditional variance, they share the limitation of parametric models in that specific parametric forms may be too restrictive to reveal the true conditions.

What We Do?

We consider the functional-coefficient autoregressive heteroscedastic (FARCH) model in our work.

The proposed integrated model framework is more general. Moreover, the FARCH model can accommodate most nonlinear relationships and does not require a priori specified parametric functions for both the conditional mean and conditional variance.

We use the Bayesian P-splines technique to estimate the nonparametric functions in the proposed FARCH model. In addition to the estimation, we also consider model selection (hypothesis testing).

FARCH Model

Let $\{Y_t,\ t = 1, \cdots, T\}$ be a stationary and ergodic time series with $E(Y_t^2) < \infty$, we assume that Y_t's are generated from the following model,

$$Y_t = \alpha_1(Y_{t-d})Y_{t-1} + \alpha_2(Y_{t-d})Y_{t-2} + \cdots + \alpha_p(Y_{t-d})Y_{t-p} + \varepsilon_t, \quad (1)$$

where the delay parameter d is a positive integer, p is the AR order, and $\alpha_j(\cdot)$'s are unknown smoothing functions with second order derivatives.

In addition, let $\mathcal{F}_{t-1}$ be the σ-field generated by the random variables $\{\varepsilon_{t-j},\ j = 1, 2, \cdots\}$.

For each t, when information $\mathcal{F}_{t-1}$ is given, we assume that the stochastic error ε_t satisfies $\varepsilon_t = h_t(Y_{t-d})u_t$ with the conditional scale $h_t(Y_{t-d})$ defined as follows,

$$h_t(Y_{t-d}) = \beta_0(Y_{t-d}) + \beta_1(Y_{t-d})|\varepsilon_{t-1}| + \cdots + \beta_q(Y_{t-d})|\varepsilon_{t-q}|, \qquad (2)$$

where q denotes the ARCH order, $\beta_j(\cdot)$'s are unknown smoothing functions satisfying $\beta_0(\cdot) > 0$ and $\beta_j(\cdot) \geq 0$ for $j = 1, \cdots, q$.

The innovations u_t's are independently and identically distributed as $N(0, 1)$.

Inspired by Li and Li (1996) and Fan and Yao (2003), we denote the model defined in (1) and (2) as FARCH(p, d, q).

Notably, $h_t(Y_{t-d})$ is defined as a conditional scale rather than a conditional variance.

Bickel and Lehmann (1976) emphasized that such a scale provides a more natural dispersion concept than variance as well as offers substantial advantages in terms of robustness.

Power of the FARCH Model

The proposed FARCH model naturally extends the FAR model in Chen and Tsay (1993) and the DTARCH model in Li and Li (1996).

First, the conditional scale changes over time in the FARCH model but remains constant in the FAR model. This feature makes the FARCH model more appealing in capturing the dynamic change of volatility.

Second, the unknown functions are piecewise linear in the DTARCH model but nonlinear and unspecified in the FARCH model.

The nonparametric modeling framework provides more flexibility for reflecting the true condition in reality.

The proposed model is able to reveal how the dynamic effects of historical time series values on the future mean value vary according to the lagged variable Y_{t-d} while also investigating how historical volatilities influence the future volatility dynamically according to the lagged variable Y_{t-d}.

Moreover, unlike the DTARCH model, it is unnecessary to carefully choose or estimate the thresholds in a FARCH model.

Thus, the FARCH model is proposed to reveal nonlinear phenomena, such as asymmetric cycles, jump resonance, and amplitude-frequency dependence, especially in financial time series data analysis.

Nonparametric Modeling

The basic idea of B-splines smoothing is to approximate the smooth functions $\alpha_j(\cdot)$ and $\beta_j(\cdot)$ in (1) and (2) by using a sum of B-splines (DeBoor, 2001) with a large number of knots in the domains of the Y_{t-d}'s. Specifically, $\alpha_j(Y_{t-d})$ in (1) can be approximated by

$$\alpha_j(Y_{t-d}) = \sum_{k=1}^{K_\lambda} \lambda_{jk} B_k^\lambda(Y_{t-d}) = \boldsymbol{\lambda}_j^T \mathbf{B}^\lambda(Y_{t-d}), \tag{3}$$

where K_λ is the number of splines determined by the number of knots, $\boldsymbol{\lambda}_j = (\lambda_{j1}, \cdots, \lambda_{jK_\lambda})^T$ is a vector of unknown parameters, $\mathbf{B}^\lambda(Y_{t-d}) = \left(B_1^\lambda(Y_{t-d}), \cdots, B_{K_\lambda}^\lambda(Y_{t-d})\right)^T$, and the functions $B_k^\lambda(\cdot)$ are B-splines basis functions with appropriate order.

Choices of $B_k^\lambda(\cdot)$ and K_λ

A natural choice of $B_k^\lambda(\cdot)$ is the cubic B-splines. Consequently, $\alpha_j(Y_{t-d})$ is a nonlinear function of Y_{t-d}.

In practice, K_λ ranging from 10 to 30 provides sufficient flexibility for modeling $\alpha(\cdot)$.

Similarly, $\beta_j(Y_{t-d})$ in (2) is approximated by

$$\beta_j(Y_{t-d}) = \sum_{k=1}^{K_\gamma} \gamma_{jk} B_k^\gamma(Y_{t-d}) = \boldsymbol{\gamma}_j^T \mathbf{B}^\gamma(Y_{t-d}), \qquad (4)$$

where K_γ, $\boldsymbol{\gamma}_j = (\gamma_{j1}, \cdots, \gamma_{jK_\gamma})^T$, and

$$\mathbf{B}^\gamma(Y_{t-d}) = \left(B_1^\gamma(Y_{t-d}), \cdots, B_{K_\gamma}^\gamma(Y_{t-d})\right)^T$$

are defined in a similar manner as those in (3).

To satisfy the model assumption of $\beta_0(\cdot) > 0$, $\beta_j(\cdot) \geq 0$, $j = 1, \cdots, q$, we impose the constraints: $\boldsymbol{\gamma}_0^T \mathbf{B}^\gamma(Y_{t-d}) > 0$, $\boldsymbol{\gamma}_j^T \mathbf{B}^\gamma(Y_{t-d}) \geq 0, j = 1, \cdots, q$ for $t = d+1, \cdots, T$.

The Likelihood

Let $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1^T, \cdots, \boldsymbol{\lambda}_p^T)^T$, $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_0^T, \cdots, \boldsymbol{\gamma}_q^T)^T$, $\mathbf{B}_{Yj}^{\lambda}(Y_{t-d}) = \mathbf{B}^{\lambda}(Y_{t-d})Y_{t-j}$ for $j = 1, \cdots, p$, $\mathbf{B}_{\varepsilon j}^{\gamma}(Y_{t-d}) = \mathbf{B}^{\gamma}(Y_{t-d})|\varepsilon_{t-j}|$ for $j = 1, \cdots, q$, and $\mathbf{B}_{\varepsilon 0}^{\gamma}(Y_{t-d}) = \mathbf{B}^{\gamma}(Y_{t-d})$. With the constant terms being disregarded, the conditional log-likelihood function can be written as

$$
\begin{aligned}
L(\boldsymbol{\Lambda}, \boldsymbol{\Gamma}) = & -\frac{1}{2} \sum_{t=s+1}^{T} \left[\log(h_t^2(Y_{t-d})) + \varepsilon_t^2 h_t(Y_{t-d})^{-2} \right] \\
= & -\frac{1}{2} \sum_{t=s+1}^{T} \Bigg[2 \log \Big(\sum_{j=0}^{q} \boldsymbol{\gamma}_j^T \mathbf{B}_{\varepsilon j}^{\gamma}(Y_{t-d}) \Big) + \\
& \frac{\big(Y_t - \sum_{j=1}^{p} \boldsymbol{\lambda}_j^T \mathbf{B}_{Yj}^{\lambda}(Y_{t-d}) \big)^2}{\big[\sum_{j=0}^{q} \boldsymbol{\gamma}_j^T \mathbf{B}_{\varepsilon j}^{\gamma}(Y_{t-d}) \big]^2} \Bigg],
\end{aligned} \tag{5}
$$

where $s = \max\{p, d, q\}$.

Over-fitting

The problem of over-fitting may occur if an excessive number of knots are used in (3) and (4). Eilers and Marx (1996) proposed the P-spline by penalizing the coefficients of adjacent B-splines to prevent over-fitting of the B-splines approximation and to guarantee sufficient smoothness of the fitted curves. This process facilitates the penalized likelihood estimation, of which the penalized likelihood

$$L_p = L(\mathbf{\Lambda}, \mathbf{\Gamma}) - \sum_{j=1}^{p} \rho_{\lambda j} \sum_{l=k+1}^{K} (\Delta^k \lambda_{jl})^2 - \sum_{j=0}^{q} \rho_{\gamma j} \sum_{l=k+1}^{K} (\Delta^k \gamma_{jl})^2 \qquad (6)$$

is maximized with respect to the unknown parameters λ and γ.

Problems with ML Methods

In the context of ML estimation, these smoothing parameters are chosen via a cross-validation procedure.

However, the computational burden for determining the optimal values of $\rho_{\lambda j}$ and $\rho_{\gamma j}$ is heavy when the number of smooth functions in the model is large.

More importantly, identifying the explicit form of cross-validation for the conditional variance model is difficult.

Therefore, for the proposed FARCH model, the optimal values of $\rho_{\lambda j}$ and $\rho_{\gamma j}$ are difficult to obtain using the ML-based methods.

Bayesian Solutions

In the Bayesian framework, the coefficients λ and γ are regarded as random, and the difference penalties are replaced by their stochastic analogues, $\Delta^k \lambda_{jl} = \Delta^k \lambda_{j,l-1} + e_{\lambda,jl}$ and $\Delta^k \gamma_{jl} = \Delta^k \gamma_{j,l-1} + e_{\gamma,jl}$, where $e_{\lambda,jl}$ and $e_{\gamma,jl}$ are independently distributed as $N[0, \tau_{\lambda j}]$ and $N[0, \tau_{\gamma j}]$, respectively.

The amount of smoothness is then controlled by the additional variance parameters $\tau_{\lambda j}$ and $\tau_{\gamma j}$, which correspond to the inverse of the smoothing parameters.

In this work, $\tau_{\lambda j}$ and $\tau_{\gamma j}$ can be considered as new smoothing parameters. In the Bayesian model framework, these smoothing parameters, along with the regression coefficients in $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$, can be obtained via the MCMC algorithm without much difficulty.

Posterior Inference

Let $\mathbf{Y} = \{Y_1, \cdots, Y_T\}$ be the set of observed time series. Let $\boldsymbol{\tau}_\lambda = \{\tau_{\lambda 1}, \cdots, \tau_{\lambda p}\}$, $\boldsymbol{\tau}_\gamma = \{\tau_{\gamma 0}, \cdots, \tau_{\gamma_q}\}$, and $\boldsymbol{\theta} = \{\boldsymbol{\Lambda}, \boldsymbol{\Gamma}, \boldsymbol{\tau}_\lambda, \boldsymbol{\tau}_\gamma\}$ include all unknown parameters in the model.

The Bayesian estimate of $\boldsymbol{\theta}$ can be obtained via the sample mean of a sufficiently large number of observations drawn from the posterior distribution $p(\boldsymbol{\theta}|\mathbf{Y})$. However, this posterior distribution is intractable because of model complexity.

MCMC Methods

To solve this problem, we use the Gibbs sampler (Geman and Geman, 1984) algorithm to simulate each component of $\{\boldsymbol{\Lambda}, \boldsymbol{\Gamma}, \boldsymbol{\tau}_{\lambda}, \boldsymbol{\tau}_{\gamma}\}$ given others from its full conditional distribution iteratively.

Considering the nonlinearity and complexity of the model, several full conditional distributions are nonstandard, and sampling observations from such distributions is not straightforward. Hence, the Metropolis-Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) is adopted to simulate observations from nonstandard distributions.

Model comparison

A specific question regarding model selection is whether an advanced (more complex) model is actually “better” than an elementary (simpler) one. This issue is particularly relevant here:

Question

If a parametric model with simple constant coefficients can provide a better fit to the observed data, then the nonparametric model with functional coefficients is unnecessary; or if a model with constant variance is adequate, then modeling both the conditional mean and conditional variance is unnecessary.

In this work, we use the Bayes factor (Kass and Raftery, 1995), which is defined by the ratio of marginal likelihoods, as the model comparison statistic.

Bayesian Forecasting

Following Tsay (2010), we apply parametric bootstraps to compute nonlinear forecasts. Let T be the forecast origin and $l > 0$ be the forecast horizon. That is, we are at time index T and interested in forecasting Y_{T+l}.

Given the observed data $Y_1, \cdots, Y_T$, the parametric bootstrap computes realizations $Y_{T+1}, \cdots, Y_{T+l}$ as follows.

(1) Draw a new stochastic error ε_{T+1} from $N(0, h_{T+1}(Y_{T+1-d}))$, where $h_{T+1}(Y_{T+1-d})$ is calculated based on (2).

(2) Compute Y_{T+1} based on (1).

(3) Repeat steps (1) and (2) to obtain $Y_{T+2}, \cdots, Y_{T+l}$.

The above procedure produces a realization for Y_{T+l}. The procedure is repeated M times to obtain M realizations for Y_{T+l}, denoted by $Y_{T+l}^{(m)}$, $m = 1, \cdots, M$.

The point forecast of Y_{T+l} and the estimated variance of the forecast are calculated as follows,

$$\hat{Y}_{T+l} = \frac{1}{M}\sum_{m=1}^{M} Y_{T+l}^{(m)}, \quad \widehat{\mathrm{Var}}(\hat{Y}_{T+l}) = \frac{1}{M-1}\sum_{m=1}^{M}(Y_{T+l}^{(m)} - \hat{Y}_{T+l})^2.$$

Tsay (2010) pointed out that $M = 3000$ could provide satisfactory results. In this paper, we employ mean square error (MSE) to measure the performance of point forecasts. For an l-step-forecast, the MSE is defined as follows,

$$\mathrm{MSE}(l) = \frac{1}{n}\sum_{j=0}^{n-1}(\hat{Y}_{T+j+l} - Y_{T+j+l})^2, \tag{7}$$

where n is the number of l-step-ahead forecasts available in the forecasting subsample. In application, the model with the smallest magnitude on that measure is regarded as the best l-step-ahead forecasting model.

However, it is possible that different l may result in the selection of different models. In addition, we employ mean variance values (MV) to assess the robustness of the forecasting. For l-step-forecast, the MV is defined as follows,

$$\mathrm{MV}(l) = \frac{1}{n} \sum_{j=0}^{n-1} \widehat{\mathrm{Var}}(\hat{Y}_{T+j+l}). \tag{8}$$

The model with the smallest MV value provides the most reliable forecasts.

Simulation study

The main objective of this subsection is to demonstrate the empirical performance of the proposed methodology. We considered the following FARCH(2, 2, 1) model:

$$\begin{aligned} Y_t &= \alpha_1(Y_{t-2})Y_{t-1} + \alpha_2(Y_{t-2})Y_{t-2} + \varepsilon_t, \\ h_t &= \beta_0(Y_{t-2}) + \beta_1(Y_{t-2})|\varepsilon_{t-1}|, \end{aligned} \tag{9}$$

where $\alpha_1(u) = u/20 - 0.15$, $\alpha_2(u) = 0.5\cos(u - 0.5)$, $\beta_0(u) = 0.5 + (u - 0.5)^2/20$, and $\beta_1(u) = 0.2 + (u - 1)^2/30$.

We conducted 100 replications with two different sample sizes $T = 1,000$ and 2,000.

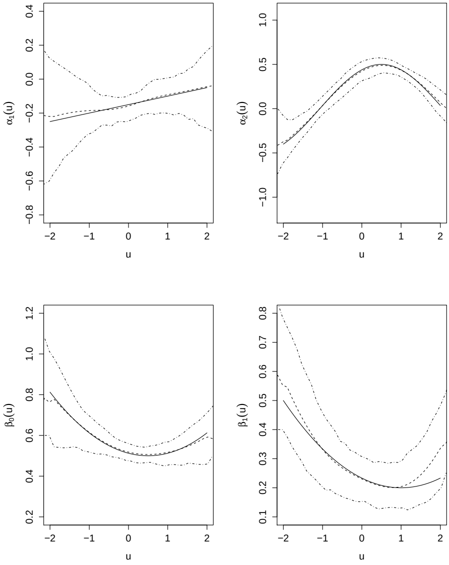

FIGURE: $T = 1,000$

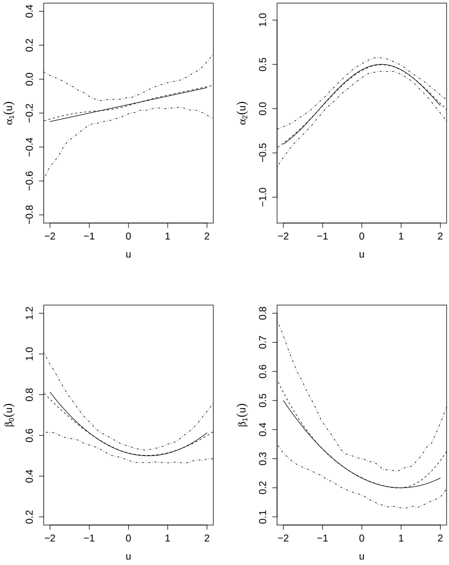

FIGURE: $T = 2,000$

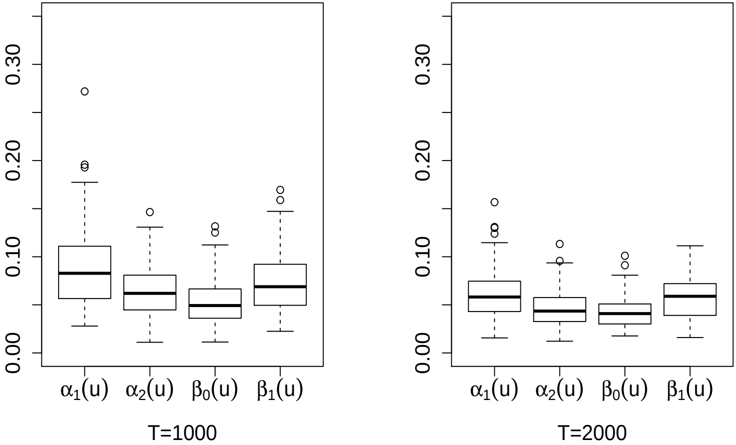

FIGURE: RASE

Model Comparison

To evaluate the performance of the Bayes factor, we conducted model selection based on the above simulated data sets. The following competing models are considered:

M_1: FARCH model defined in (9).

M_2: Semiparametric model with functional coefficient mean model but constant coefficient variance model:

$$Y_t = \alpha_1(Y_{t-2})Y_{t-1} + \alpha_2(Y_{t-2})Y_{t-2} + \varepsilon_t,$$
$$h_t = \beta_0 + \beta_1|\varepsilon_{t-1}|,$$

where $\alpha(\cdot)$'s are unknown functions, and β's are unknown parameters.

M_3: DTARCH(2,2,1) model

$$Y_t = [\alpha_{11} Y_{t-1} + \alpha_{21} Y_{t-2}] \cdot \mathrm{I}(Y_{t-2} \leq 0) + [\alpha_{12} Y_{t-1} + \alpha_{22} Y_{t-2}] \cdot \mathrm{I}(Y_{t-2} > 0) + \varepsilon_t,$$
$$h_t = \left[\beta_{01} + \beta_{11}|\varepsilon_{t-1}|\right] \cdot \mathrm{I}(Y_{t-2} \leq 0) + \left[\beta_{02} + \beta_{12}|\varepsilon_{t-1}|\right] \cdot \mathrm{I}(Y_{t-2} > 0),$$

where α's and β's are unknown parameters.

M_4: Parametric ARCH(2,1) model

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t,$$
$$h_t = \beta_0 + \beta_1|\varepsilon_{t-1}|,$$

where α's and β's are unknown parameters.

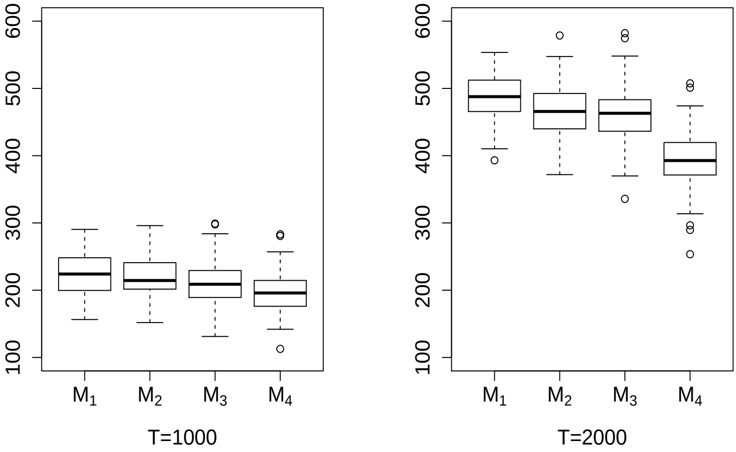

FIGURE: Boxplots of the log Bayes factor values.

TABLE: Out-of-sample forecast comparison among M_1 to M_4 in the simulation study.

		1-step		2-step	
Sample Sizes	Model	MSE	MV	MSE	MV
$T = 1000$	M_1	0.463	0.475	0.472	0.488
	M_2	0.533	0.488	0.490	0.499
	M_3	0.549	0.493	0.494	0.508
	M_4	0.563	0.507	0.509	0.519
$T = 2000$	M_1	0.464	0.472	0.469	0.484
	M_2	0.484	0.480	0.486	0.492
	M_3	0.492	0.493	0.499	0.508
	M_4	0.500	0.506	0.505	0.518

More on Model Comparison

In order to check whether the FARCH model can provide a good approximation when the true model is relatively simple, we generated time series from a parametric DTARCH model and analyzed the data sets via the FARCH model. Then, the relevant questions might be (i) whether the FARCH model produces reasonable estimation results, and (ii) whether the Bayes factor selects the true (simpler) model.

To address these issues, we generated 100 data sets with the sample size of $T = 2,000$ from the following DTARCH(1,2,1) model:

$$\begin{aligned} Y_t &= -0.3 Y_{t-1} \cdot \mathrm{I}(Y_{t-2} \leq 0) + 0.35 Y_{t-1} \cdot \mathrm{I}(Y_{t-2} > 0) + \varepsilon_t, \\ h_t &= \left[0.2 + 0.42|\varepsilon_{t-1}|\right] \cdot \mathrm{I}(Y_{t-2} \leq 0) + \left[0.4 + 0.24|\varepsilon_{t-1}|\right] \cdot \mathrm{I}(Y_{t-2} > 0). \end{aligned} \tag{10}$$

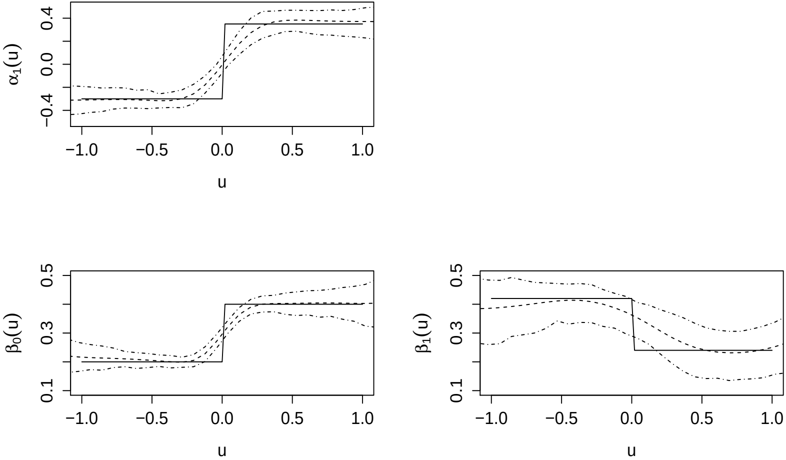

FIGURE: Estimates of the unknown smooth functions in the simulation study.

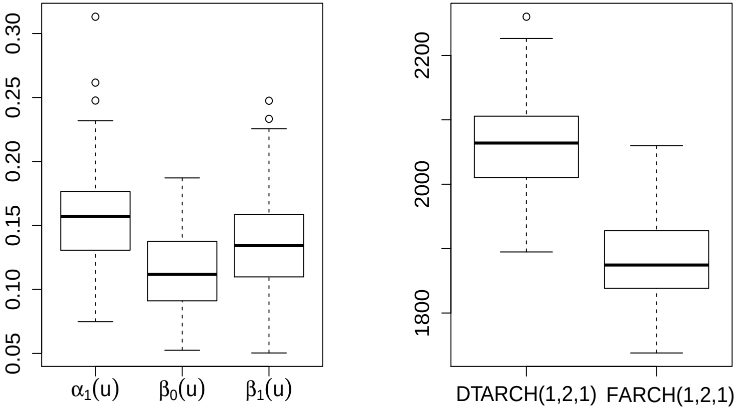

FIGURE: Boxplots of the RASE values (left panel) and the log Bayes factor values (right panel) in the simulation study.

Real Data Analysis

As an illustration, we apply the proposed methodology to the study of the daily S&P 500 Composite Index from Jan 3, 2000 to July 27, 2011.

Our main goal is to investigate the nonlinear features of return and volatility in terms of the S&P 500 Composite Index. Let X_t be the closing price at time t, the return series Y_t was defined as $Y_t = 100\log(X_t/X_{t-1})$. The total sample size was $T = 2,909$. We used the first 2,700 observations to develop the model and used the remaining 209 observations as the out-of-sample set to assess the performance of the Bayesian forecasting.

First, we determined the delay parameter d, the AR order p, and the ARCH order q by comparing FARCH models with different values of p, d, and q. By rotating their values from 1 to 3, a total of 27 candidate models were considered.

The Bayes factors chose FARCH(1,2,1), which is defined as

$$\begin{aligned} Y_t =&\alpha_1(Y_{t-2})Y_{t-1} + \varepsilon_t, \\ h_t =&\beta_0(Y_{t-2}) + \beta_1(Y_{t-2})|\varepsilon_{t-1}|. \end{aligned} \tag{11}$$

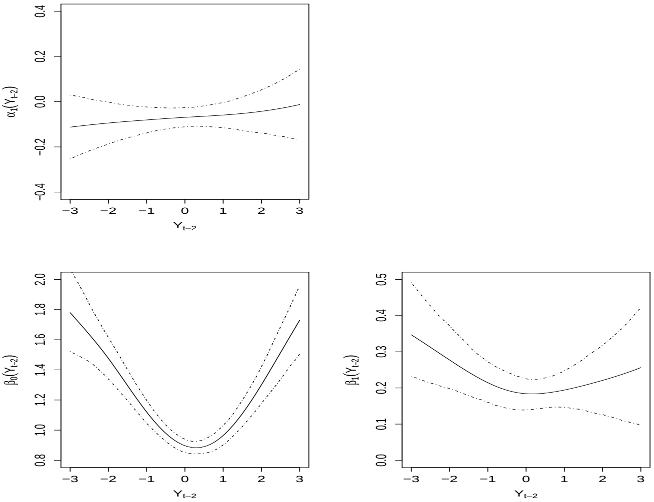

FIGURE: Estimated curves in the analysis of the S&P 500 data set under FARCH(1, 2, 1).

Findings

First, the estimated values of $\hat{\alpha}_1(\cdot)$ are slightly below zero, implying mild negative associations between the future mean return and the historical returns. Thus, the future mean return tends to be positive (negative) if the historical returns are negative (positive). This patten may explain why the index values tend to be stable as well as why the overall market is robust to minor financial events.

Second, the estimated curve $\hat{\beta}_1(\cdot)$ exhibits a U-shaped pattern, which means the persistence in volatility tends to be high when the historical returns are high or low, reaching the minimum when Y_{t-d} is approximately zero. In addition, the estimated curve presents obvious asymmetry; the values of $\hat{\beta}_1(\cdot)$ at the negative range of Y_{t-d} are clearly larger than those at the positive range of Y_{t-d}. That is, the volatility persistence tends to be higher in bear markets than in bull markets.

Third, the Bayes factors selected the model with $d = 2$. Fan and Yao (2003) highlighted that Y_{t-d} should be regarded as the "model-dependent variable," which provides useful information on the modeling and the dependence structure of the observed data. In this study, the delay parameter $d = 2$ indicates that the historical returns will be internally reflected in the model with one day lag.

Finally, the dynamic coefficients, $\hat{\beta}_0(\cdot)$ and $\hat{\beta}_1(\cdot)$ in the volatility model are clearly neither linear nor piecewise linear. These nonlinear patterns cannot be well captured by the existing parametric or semiparametric time series models.

Summary

We proposed a novel FARCH model to analyze time series data. The FARCH model generalizes both FAR and DTARCH models, thus providing a flexible model framework to capture various nonlinear phenomena for both the conditional mean and conditional variance.

We developed a Bayesian approach coupled with Bayesian P-splines and MCMC algorithms to obtain the estimation of the functional coefficients.

We employed a Bayesian model selection statistic, the Bayes factor, to address the hypothesis testing problem. The work expands the scope of Bayesian nonparametric time series modeling in the statistics and economics literature.

Limitations

First, we did not provide a sufficient and/or a necessary condition to theoretically ensure the stationarity of the FARCH model because of the model complexity.

Second, our proposed method cannot satisfactorily estimate a very wiggly function like $\beta(u) = 0.2 + 0.1\sin(2\pi u)$, in the conditional variance/scale model. The unsatisfactory performance may be due to the nature of volatility.

Third, the innovations were assumed to follow the standard normal distribution in our model. This assumption may not be valid in practice.

Finally, the current research only handles univariate time series.


中山大學
SUN YAT-SEN UNIVERSITY
Thank You!
谢谢！
國立中山大學