Paper:
Cai, Y. (2014). A General Quantile Function Model for Economic and Financial Time Series. *Econometric Reviews*, 1-21.

http://dx.doi.org/10.1080/07474938.2014.976528

A general quantile function model for economic and financial time series

Yuzhi Cai*, *Swansea University, UK*

Abstract

This paper proposed a general quantile function model that covers both one and multiple dimensional models and that takes several existing models in the literature as its special cases. This paper also developed a new uniform Bayesian framework for quantile function modelling and illustrated the developed approach through different quantile function models. Many distributions are defined explicitly only via their quanitle functions as the corresponding distribution or density functions do not have an explicit mathematical expression. Such distributions are rarely used in economic and financial modelling in practice. The developed methodology makes it more convenient to use these distributions in analyzing economic and financial data. Empirical applications to economic and financial time series and comparisons with other types of models and methods show that the developed method can be very useful in practice.

Key words: German DAX, currency exchange rates, quantile functions models, Bayesian approach.

*Address for correspondence: Dr Yuzhi Cai, School of Management, Swansea University, Swansea, SA2 8PP, United Kingdom. Email: y.cai@swansea.ac.uk

1 Introduction

Making statistical inferences on the conditional quantiles of a financial or economic variable has become more and more popular recently. See, for example, Jorion (2006). One of the approaches to estimating conditional quantiles is based on the estimated distribution function of the variable. In one dimensional cases, this is relatively easy to achieve. For multivariate cases, Li and Racine (2008) proposed a nonparametric kernel-based method for estimating a joint probability distribution function. They also proposed a method for inverting this estimate to obtain quantile surfaces. Their quantile surface estimating method has been shown to be very useful in practice. For example, Maasoumi and Racine (2013) developed a new technique that used the estimated quantile surfaces obtained from Li and Racine (2008) method and that resolved a classic problem of assigning weights to multiple indicators.

Another one of the approaches to estimating conditional quantiles is to estimate the conditional quantiles directly. This approach includes the semi-parametric quantile regression models (Koenker, 2005) and the parametric quantile function models (Gilchrist, 2000).

Given a set of observations $\{y_i, x_{1i}, \ldots, x_{pi}\}$ ($i = 1, \ldots, n$), a one dimensional quantile regression model for the τth conditional quantile of Y, denoted by $q^{\tau}_{Y|X}$, is defined by

$$q^{\tau}_{Y|X} = h(\boldsymbol{\eta}^{\tau}, \mathbf{x}), \tag{1}$$

where h is a known function of the covariate $\mathbf{x} = (x_1, \ldots, x_p)$ and $\boldsymbol{\eta}^{\tau}$ is the model parameter vector depending on τ $(0 \leq \tau \leq 1)$. Note that the effect of the error term of model (1) is estimated through $\boldsymbol{\eta}^{\tau}$ non-parametrically. Hence, the whole model is semi-parametric.

$\boldsymbol{\eta}^{\tau}$ may be estimated by using various methods including Bayesian methods. See for example, Koenker (1984), Koenker and D'Orey (1987, 1994), Kottas and Gelfand (2001), Tsionas (2003), Yu and Moyeed (2001), Cai and Stander (2008), Lancaster and Jun (2010), Wu and Liu's (2009), Bondell et al. (2010) and references therein. A simple example of (1) is given by $q^{\tau}_{Y|X} = a^{\tau}_0 + a^{\tau}_1 x_1 + \cdots + a^{\tau}_p x_p$ with $\boldsymbol{\eta}^{\tau} = (a^{\tau}_0, \ldots, a^{\tau}_p)$.

In multivariate cases, some work can also be found in the literature. For example, Salibian-Barrera and Wei (2008) used the inverse of Rosenblatt's transformation (Rosenblatt, 1952) to map a central confidence region defined in the unit hypercube back onto the original sample space, so that weighted quantile regression estimator can be obtained. Wei (2008) also developed an approach to estimating multivariate quantile contours.

Compared with the semi-parametric approach, much less work can be found in the literature on the parametric quantile function approach. We will focus on this type of models in this paper. A one dimentional parametric quantile function model is defined by

$$Q_Y(\tau \mid \boldsymbol{\xi}, \mathbf{x}) = h_1(\boldsymbol{\eta}_1, \mathbf{x}) + h_2(\boldsymbol{\eta}_2, \mathbf{x}) Q(\tau, \boldsymbol{\gamma}), \tag{2}$$

where $\boldsymbol{\xi} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\gamma})$ is the model parameter vector, h_i $(i = 1, 2)$ are known functions of $\mathbf{x}$ and $\boldsymbol{\eta}_i$, $h_2(\boldsymbol{\eta}_2, \mathbf{x}) > 0$ and $Q(\tau, \boldsymbol{\gamma})$ is the quantile function of the error term with an explicit mathematical expression. A special case of model (2) is the linear quantile function model given by $Q_Y(\tau \mid \boldsymbol{\xi}, \mathbf{x}) = a_0 + a_1 x_1 + \cdots + a_p x_p + Q(\tau, \boldsymbol{\gamma})$ with $h_1(\boldsymbol{\eta}_1, \mathbf{x}) = a_0 + a_1 x_1 + \cdots + a_p x_p$, $h_2(\boldsymbol{\eta}_2, \mathbf{x}) = 1$ and $\boldsymbol{\eta}_1 = (a_0, \ldots, a_p)$. It is seen that, unlike model (1), model (2) can guarantee the monotonicity of the estimated conditional quantiles of Y due to the fact that $Q_Y(\tau \mid \boldsymbol{\xi}, \mathbf{x})$ is a well defined conditional quantile function.

Gilchrist (2000) discussed some methods for estimating $\boldsymbol{\xi}$ based on distributional least squares or distributional absolutes criteria. Cai (2009, 2010b) and Cai et al. (2013) proposed Bayesian approaches to estimating parameters of different types of model (2), including polynomial and linear/non-linear time series quantile functions models.

In multivariate cases, even less work can be found in the literature on quantile function models. Cai (2010a) introduced a multivariate quantile function model and illustrated the usefulness of the model via a financial data set.

The main contributions of this paper include: (i) a general multivariate quantile function model is proposed which includes the quantile function models mentioned above as its special cases; (ii) a new uniform Bayesian framework for estimating such models is developed. We hope that this paper will help readers to develop their own quantile function models in either one or multiple dimensional cases and to have a means for estimating model parameters in practice.

2 A general quantile function model

To simplify the notation, from now on we use the small letter y to represent both a realization of a random variable Y and a variable of a function. Then a general quantile function model for a set of continuous financial variables $\mathbf{y} = (y_1, \ldots, y_m)$ given $\mathbf{x}$ is defined by

$$h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x}) = Q(\tau, \boldsymbol{\gamma}), \tag{3}$$

where $0 \leq \tau \leq 1$, $Q(\tau, \boldsymbol{\gamma})$ is a one dimensional quantile function with parameter $\boldsymbol{\gamma}$, and $h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x})$ is a known function of its arguments with parameter $\boldsymbol{\eta}$. Let $\boldsymbol{\beta} = (\boldsymbol{\eta}, \boldsymbol{\gamma})$. As the $Q(\tau, \boldsymbol{\gamma})$ in model (3) is a quantile function of a distribution, we have

$$\tau = P(U \leq \tau) = P[Q^{-1}\{h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x})\} \leq \tau] = P\{h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x}) \leq Q(\tau, \boldsymbol{\gamma})\},$$

where U is a uniformly distributed random variable between 0 and 1. Therefore, $A_\tau = \{\mathbf{y} \mid h(\mathbf{y}, \boldsymbol{\eta}) \leq Q(\tau, \boldsymbol{\gamma})\}$ forms a region such that the probability of $\mathbf{y} \in A_\tau$ is τ. For example, if $\tau = 0.25$, then the probability of $\mathbf{y} \in A_{0.25}$ is 0.25. Furthermore, the boundary of A_τ may be determined by $h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x}) = Q(\tau, \boldsymbol{\gamma})$. This boundary defines the τth quantile surface that we want to estimate.

Note that if model (3) holds and if h is a continuous function of $\mathbf{y}$, then the resulting quantile regions are unions of compact sets and the resulting quantile surfaces are nested. In the rest of the paper, we assume that $\mathbf{y}$ follows a joint distribution defined by model (3).

Note that determining the distribution of the error term of a parametric statistical model can be difficult. However, compared with the models for the means, such as conventional regression models, it is a relatively easier task for the quantile function models because quantile functions have some good properties. For example, under certain conditions quantile functions can be added, multiplied, transformed and/or combined to form new quantile functions, i.e. new models.

For example, the distributions of financial returns are usually asymmetric with long tails on both sides. Note that the Pareto distribution is one of the commonly used distributions for extremes. So we may use a Pareto quantile function $S_1(\tau, \gamma_1) = \tau^{\gamma_1}$ ($\tau \in (0,1)$, $\gamma_1 < 0$) and the reflection of another Pareto quantile function $S_2(\tau, \gamma_2) = -(1-\tau)^{\gamma_2}$ ($\gamma_2 < 0$) as building blocks. Then the sum of a linear transformation of S_1 and a linear transformation of S_2 gives a new quantile function $Q(\tau, \boldsymbol{\gamma}) = \frac{\tau^{\gamma_1}-1}{\gamma_1} - \frac{(1-\tau)^{\gamma_2}-1}{\gamma_2}$, where γ_1 and γ_2 jointly model the tails of this new distribution. When $\gamma_1 \neq \gamma_2$ the distribution is skewed. In fact, this quantile function is the quantile function of the generalized lambda distribution (see Freimer et al. (1988)). Of course, if we use other distributions as building blocks then various new distributions may be constructed in practice.

Note that a quantile function obtained by using quantile function properties usually does not have an explicit inverse function. Hence the corresponding distribution/density function does not have an explicit mathematical expression, as for the above generalized lambda distribution. This is why many distributions defined via their quantile functions have not been used in practice. However, the methodology developed in this paper will enable us to use such distributions in financial and economic modelling easily.

The fitted quantile function models may be checked via conventional methods. For example, the residuals of a fitted model can be checked for independence and randomness. A quantile-quantile plot (QQ-plot) or a probability-probability plot (PP-plot) may be used to check the specifications of the model. Formal statistical tests may also be carried out.

Now let us consider several special cases. If $m = 1$, let $\mathbf{y} = y$, and $h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x}) = (y - h_1(\boldsymbol{\eta}_1, \mathbf{x}))/h_2(\boldsymbol{\eta}_2, \mathbf{x})$, then model (3) becomes $y = h_1(\boldsymbol{\eta}_1, \mathbf{x}) + h_2(\boldsymbol{\eta}_2, \mathbf{x})Q(\tau, \boldsymbol{\gamma})$, which in fact defines the quantile function of y given by model (2).

Furthermore, if $p = 1$ and $h_1(\boldsymbol{\eta}_1, x) = a_0 + \sum_{j=1}^{k_1} a_j x^j, \quad h_2(\boldsymbol{\eta}_2, x) = b_0 + \sum_{j=1}^{k_1} b_j x^j$, then model (3) defines a polynomial quantile function model studied by Cai (2010b):

$$Q_y(\tau \mid \boldsymbol{\beta}, \mathbf{x}) = a_0 + a_1 x + \cdots + a_{k_1} x^{k_1} + (b_0 + b_1 x + \cdots + b_{k_2} x^{k_2}) Q(\tau, \boldsymbol{\gamma}). \tag{4}$$

If y is a time series, let $y = y_t$, $x_i = y_{t-i}$ $(i = 1, \ldots, p)$, $h_1(\boldsymbol{\eta}_1, \mathbf{x}) = a_0 + \sum_{j=1}^{p} a_j y_{t-j}$, and $h_2(\boldsymbol{\eta}_2, \mathbf{x}) = \eta$, then model (3) defines an AR quantile function model, see Cai (2009):

$$Q_{y_t}(\tau \mid \boldsymbol{\beta}, \mathbf{x}) = a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p} + \eta Q(\tau, \boldsymbol{\gamma}). \tag{5}$$

Compared with a conventional AR(p) model $y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$, where $\varepsilon_t \sim N(0, \sigma^2)$, we see that model (5) and the conventional AR(p) model are equivalent if $Q(\tau, \boldsymbol{\gamma})$ is the quantile function of N(0, 1) and $\eta = \sigma$. So the autocorrelation structures can also be dealt with by quantile function models easily.

For financial time series, the conditional variance of y_t may also depend on the lag values $y_{t-1}, \ldots, y_{t-q}$ for some $q > 0$. So if $\eta = \sqrt{b_0 + \sum_{j=1}^{q} b_j y_{t-j}^2}$, then we have

$$Q_{y_t}(\tau \mid \boldsymbol{\beta}, \mathbf{x}) = a_0 + \sum_{j_1=1}^{p} a_{j_1} y_{t-j_1} + \sqrt{b_0 + \sum_{j_2=1}^{q} b_{j_2} y_{t-j_2}^2}\; Q(\tau, \boldsymbol{\gamma}). \tag{6}$$

See Cai et al. (2013) for further details. It is seen that in this case the autocorrelation structure can also be dealt with.

If $m > 1$, $p = 0$ and $h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x}) = \sum_{k=1}^{m} \left(y_k - \sum_{j=0}^{k-1} a_{k\,k-j-1} y_{k-j-1} \right)^2$, then model (3) defines a multivariate quantile function model

$$\sum_{k=1}^{m} \left(y_k - \sum_{j=0}^{k-1} a_{kk-j-1} y_{k-j-1} \right)^2 = Q(\tau, \boldsymbol{\gamma}), \tag{7}$$

which was further studied by Cai (2010a).

Therefore, model (3) covers both one-dimensional and multi-dimensional quantile function models. The function $h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x})$ is able to describe not only dependence structures between variables but also autocorrelation structures of a financial time series.

Many other dependence structures between financial or economic variables can also be studied by using a quantile function model. For illustration purposes, we consider the two-dimensional case, i.e. $m = 2$, via the following four models, where y_1 and y_2 may represent any two financial variables of interest.

Model 1: $a_0 + a_1 \sin(y_1) + a_2 y_2 = Q(\tau, \boldsymbol{\gamma})$.

In this model we let $Q(\tau, \boldsymbol{\gamma})$ be the quantile function of the log-normal distribution with mean 0 and variable 1. This model describes a cyclic dependence structure between two financial variables.

Model 2: $a_0 + a_1 y_1 + a_2 y_2 + a_3 y_1 y_2 = Q(\tau, \boldsymbol{\gamma})$.

In this model we let $Q(\tau, \boldsymbol{\gamma})$ be the quantile function of the t-distribution with 5 degrees of freedom. This model describes a linear relationship between two financial returns and their interactions.

Model 3: $a_0 + a_1 y_1^2 + a_2 y_2 = Q(\tau, \boldsymbol{\gamma})$.

In this model we let $Q(\tau, \boldsymbol{\gamma})$ be the quantile function of the Weibull distribution with shape parameter 1 and scale parameter 2. This model describes the quadratic relationship between two financial variables.

Model 4: $a_0 + a_1 (y_1 - b_1)^2 + a_2 (y_2 - b_2)^2 = Q(\tau, \boldsymbol{\gamma})$.

In this model we let $Q(\tau, \boldsymbol{\gamma})$ be the quantile function of the exponential distribution with rate 0.05. This model describes the nonlinear dependence structure (clustered and with some extremes) between two financial variables.

We generated four data sets, one from each of the above four models with the parameter

Table 1: Parameter values of Models 1-4

Models	a_0	a_1	a_2	a_3	b_1	b_2
Model 1	1	1.2	-0.5	-	-	-
Model 2	1	1.2	7.1	-2.9	-	-
Model 3	-2	1.5	0.5	-	-	-
Model 4	-1.5	0.5	0.1	-	1	-1

values given in Table 1. The sample size is 200 for all four models. Note that the sample size, the parameter values and the $Q(\tau, \boldsymbol{\gamma})$ were all chosen arbitrarily. Figure 1 shows the joint probability density plots of two financial variables, Figure 2 shows the contour plots of the density functions and Figure 3 shows the quantile curves of the joint distribution of the two variables. Note that the grey points in the figures are the simulated data.

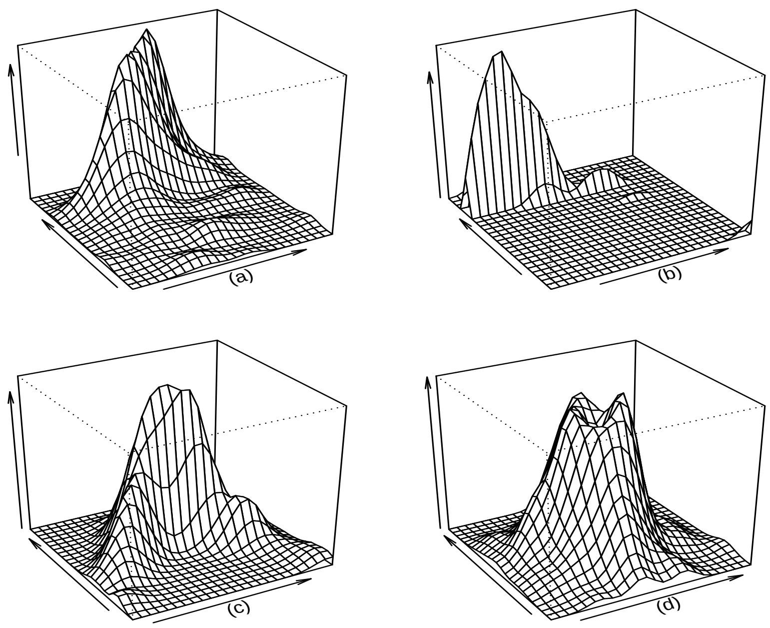

Figure 1: The joint probability density plots of y_1 and y_2 of (a) Model 1, (b) Model 2, (c) Model 3 and (d) Model 4.

The joint density function plots show various dependence structures between the two financial variables, and the contour curves suggest that the joint density functions may have multiple modes. It is clear that the contour curves are very different from the quantile

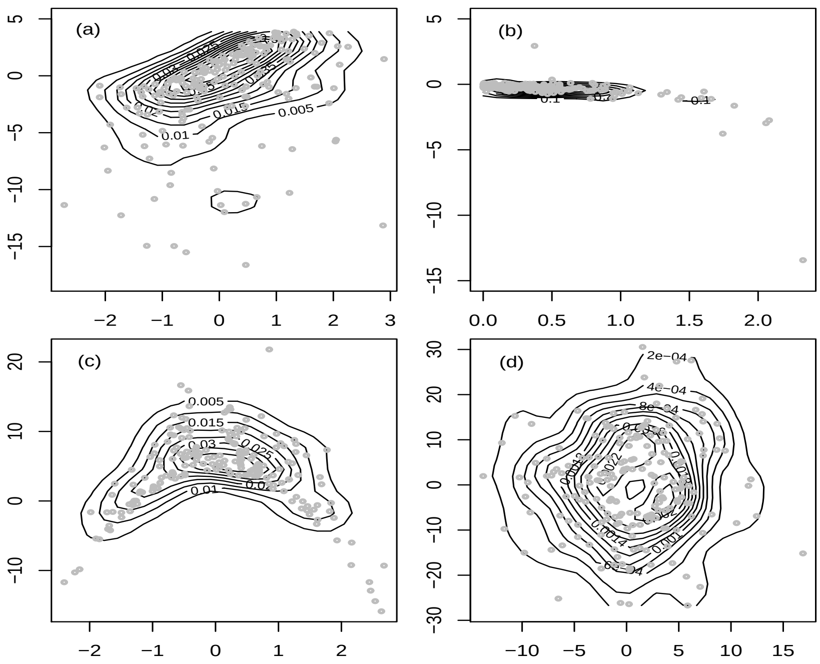

Figure 2: The contour plots of the joint probability density functions of y_1 and y_2 for (a) Model 1, (b) Model 2, (c) Model 3 and (d) Model 4.

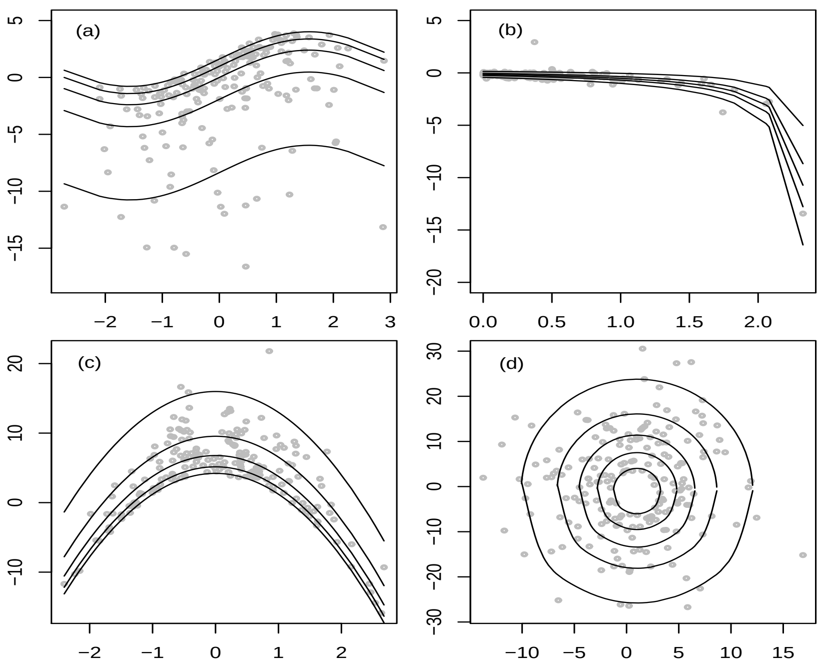

Figure 3: The quantile curves of the joint distribution of y_1 and y_2 for (a) Model 1, (b) Model 2, (c) Model 3 and (d) Model 4.

curves. For each model, we plotted five quantile curves at levels $0.05, 0.25, 0.5, 0.75$ and 0.95. For Models 1, 2, and 3, a quantile curve at a level, say 0.25, tells us that the probability for the two financial variables to take a value within the region below the curve is 0.25. For Model 4, the 0.25th quantile curve indicates that the probability that the two financial variables are in the region enclosed by the quantile curve is 0.25. Note that quantile curves at any other levels can also be obtained similarly.

These examples further show that the general quantile function models can indeed deal with different dependence structures between financial variables easily. In the following section we develop a new uniform Bayesian framework for model (3).

3 The Bayesian MCMC method

3.1 The method

Let $\mathbf{y}_i = (y_{1i}, \ldots, y_{mi})$ and $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})$ $(i = 1, \ldots, n)$ be the observed data, and let $u = h(\mathbf{y} \mid \boldsymbol{\eta}, \mathbf{x})$. Then model (3) says that the quantile function of U is given by $Q(\tau, \boldsymbol{\gamma})$. We assume that $Q(\tau, \boldsymbol{\gamma})$ has an explicit mathematical expression, but its inverse function may not be known explicitly. Hence the probability density function of U may not have an explicit mathematical expression.

Theorem 1 *Let* $\mathbf{u} = (u_1, \ldots, u_n)$*, where* $u_i = h(\mathbf{y}_i \mid \boldsymbol{\eta}, \mathbf{x}_i)$ $(i = 1, \ldots, n)$*. Furthermore, we assume that* u_i *are independent samples of* U*. Then the likelihood of* $\mathbf{u}$ *is given by* $L(\mathbf{u} \mid \boldsymbol{\beta}, \mathbf{x}) = \prod_{i=1}^{n} \{\partial Q(\tau, \boldsymbol{\gamma})/\partial \tau\}^{-1}|_{\tau=\tau_i}$*, where* τ_i *satisfies*

$$u_i = h(\mathbf{y}_i \mid \boldsymbol{\eta}, \mathbf{x}_i) = Q(\tau_i, \boldsymbol{\gamma}). \tag{8}$$

See the Appendix for a proof. Theorem 1 shows that as long as $Q(\tau, \boldsymbol{\gamma})$ is known, the likelihood function of the parameters is an explicit function of τ, but may not be an explicit

Table 2: A general MCMC method, where A is the posterior ratio, and B and C are the ratios of the transition density functions for $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ respectively.

```
Sample
    η' ~ g1(η') and γ' ~ g2(γ') such that η' ∈ Ω1 and γ' ∈ Ω2
Solve for τ'_i (i = 1,...,n) from
        h(y_1i,...,y_mi | η, x_i) = Q(τ'_i, γ')
Sample v ~ U(0,1)
If v ≤ min{ABC,1}, then β = β', τ_i = τ'_i (i = 1,...,n)
else β = β, τ_i = τ_i (i = 1,...,n)
```

function of the parameters themselves. This is because τ_i satisfies (8) which usually can not be solved exactly. Therefore, this poses challenging tasks in estimating such quantile function models. We now develop a Bayesian method for the parameter estimation and we will see that a Bayesian approach is able to resolve the problem.

To represent the dependence on $\mathbf{y}$, we write $\mathbf{u} = \mathbf{u}(\mathbf{y})$. Let $\pi(\boldsymbol{\beta})$ be the prior density function of $\boldsymbol{\beta}$, then the posterior density function of $\boldsymbol{\beta}$ is given by $\pi(\boldsymbol{\beta} \mid \mathbf{x}, \mathbf{u}(\mathbf{y})) \propto L(\mathbf{u}(\mathbf{y}) \mid \boldsymbol{\beta}, \mathbf{x})\pi(\boldsymbol{\beta})$. Furthermore, if $\pi(\boldsymbol{\beta} \mid \mathbf{x}, \mathbf{u}(\mathbf{y}))$ is well defined for $\boldsymbol{\beta} = (\boldsymbol{\eta}, \boldsymbol{\gamma}) \in \Omega = \Omega_1 \times \Omega_2$, then a MCMC method can be designed and the model parameters can be estimated. As a Gibbs sampler requires full conditional distributions of $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$, which is not available in this case, the Metropolis-Hasting MCMC method is considered below.

Let $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$ be the current and the proposed parameter values respectively, let τ_i and τ_i' $(i = 1, \ldots, n)$ be the associated probabilities respectively. Then a general MCMC sampler for quantile function models is given in Table 2, where $A = \pi(\boldsymbol{\beta}' \mid \mathbf{x}, \mathbf{u}(\mathbf{y}))/\pi(\boldsymbol{\beta} \mid \mathbf{x}, \mathbf{u}(\mathbf{y}))$, $B = q(\boldsymbol{\eta}' \to \boldsymbol{\eta})/q(\boldsymbol{\eta} \to \boldsymbol{\eta}')$, $C = q(\boldsymbol{\gamma}' \to \boldsymbol{\gamma})/q(\boldsymbol{\gamma} \to \boldsymbol{\gamma}')$, where $q(a' \to a)$ is the transition probability density function of a given a', and g_1 and g_2 are the probability density functions from which $\boldsymbol{\beta}'$ is obtained. For illustration purposes we take g_1 and g_2 as those specified in Theorem 2 below.

Theorem 2 *Let $\boldsymbol{\eta} = (\eta_0, \ldots, \eta_{p_1})$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{p_2})$ and let*

$$g_1(\boldsymbol{\eta}') = \prod_{j=0}^{p_1} \frac{1}{\sqrt{2\pi}\sigma_{\eta_j}} e^{-\frac{(\eta_j'-\eta_j)^2}{2\sigma_{\eta_j}^2}}, \qquad g_2(\boldsymbol{\gamma}') = \prod_{\ell=1}^{p_2} \frac{1}{\sqrt{2\pi}\sigma_{\gamma_\ell}} e^{-\frac{(\gamma_\ell'-\gamma_\ell)^2}{2\sigma_{\gamma_\ell}^2}},$$

where $\boldsymbol{\eta}' \in \Omega_1$ and $\boldsymbol{\gamma}' \in \Omega_2$, while σ_{η_j}, σ_{γ_ℓ} ($j = 0, \ldots, p_1$, $\ell = 1, \ldots, p_2$) are given by user. Then

$$B = \left\{ \int_{\Omega_1} \prod_{j=0}^{p_1} \frac{1}{\sqrt{2\pi}\sigma_{\eta_j}} e^{-\frac{(\eta_j'-\eta_j)^2}{2\sigma_{\eta_j}^2}} d\eta_j' \right\} \left\{ \int_{\Omega_1} \prod_{j=0}^{p_1} \frac{1}{\sqrt{2\pi}\sigma_{\eta_j}} e^{-\frac{(\eta_j-\eta_j')^2}{2\sigma_{\eta_j}^2}} d\eta_j \right\}^{-1}, \tag{9}$$

$$C = \left\{ \int_{\Omega_2} \prod_{\ell=1}^{p_2} \frac{1}{\sqrt{2\pi}\sigma_{\gamma_\ell}} e^{-\frac{({\gamma_\ell}'-\gamma_\ell)^2}{2\sigma_{\gamma_\ell}^2}} d\gamma_\ell' \right\} \left\{ \int_{\Omega_2} \prod_{\ell=1}^{p_2} \frac{1}{\sqrt{2\pi}\sigma_{\gamma_\ell}} e^{-\frac{(\gamma_\ell-{\gamma_\ell}')^2}{2\sigma_{\gamma_\ell}^2}} d\gamma_\ell \right\}^{-1}. \tag{10}$$

See Appendix for a proof. In general cases it can be very difficult to evaluate (9) and (10) because of the complex structure of Ω_1 and Ω_2. In this paper, we use a simulation method to evaluate the integrals involved in (9) and (10) if necessary. For example, to estimate $\int_{\Omega_1} \prod_{j=0}^{p_1} (1/\sqrt{2\pi}\sigma_{\eta_j})\, exp\{-(\eta_j' - \eta_j)^2/2\sigma_{\eta_j}^2\} d\eta_j'$, we simulate $\eta_j^v \sim N(\eta_j, \sigma_{\eta_j}^2)$ $(j = 0, \ldots, p_1,\ v = 1, \ldots, M_1)$. Let N_1 be the number of the simulated samples such that $(\eta_0^v, \ldots, \eta_k^v) \in \Omega_1$. Then the integral is estimated by N_1/M_1.

Therefore, for any quantile function models defined by model (3), if we can specify a proper prior density function for $\boldsymbol{\beta}$ such that $\int_\Omega \pi(\boldsymbol{\beta} \mid \mathbf{x}, \mathbf{u}(\mathbf{y})) d\boldsymbol{\beta} < \infty$, then the above general MCMC method can be used for parameter estimation. Our experience with this sampler shows that several testing runs will enable us to choose proper values of σ_{η_j} and σ_{γ_ℓ} for each application, and that the simulated Markov chain converges to the posterior distribution of the parameters quickly. In the next subsection we illustrate the above general MCMC method using different quantile function models.

3.2 Applications to special quantile function models

First, let us consider model (4), where $Q(\tau, \boldsymbol{\gamma}) = \tau^{\gamma_1}(1-\tau)^{-\gamma_2}, \gamma_1 > 0, \gamma_2 > 0$. So $Q(\tau, \boldsymbol{\gamma})$ is the quantile function of the power-Pareto distribution. It follows from Theorem 1 that

$$L(\mathbf{u}(\mathbf{y}) \mid \boldsymbol{\beta}, \mathbf{x}) = \textstyle\prod_{i=1}^n (\tau_i^{1-\gamma_1}(1-\tau_i)^{1+\gamma_2})/(\gamma_1(1-\tau_i) + \gamma_2\tau_i),$$

where τ_i satisfy

$$u_i = \frac{y_i - (a_0 + a_1 x_i + \cdots + a_{k_1} x_i^{k_1})}{b_0 + b_1 x_i + \cdots + b_{k_2} x_i^{k_2}} = \tau_i^{\gamma_1}(1-\tau_i)^{-\gamma_2}.$$

Furthermore, if the prior distribution of the parameters is given by $\pi(\boldsymbol{\beta}) = \pi(\boldsymbol{\eta}_1)\pi(\boldsymbol{\eta}_2)\pi(\boldsymbol{\gamma})$, where $\boldsymbol{\eta}_1 = (a_0, \ldots, a_{k_1})$, $\boldsymbol{\eta}_2 = (b_0, \ldots, b_{k_2})$, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$ and

$$\pi(\boldsymbol{\eta}_1) = \prod_{j_1=0}^{k_1} \pi(a_{j_1}) = \prod_{j_1=0}^{k_1} \frac{1}{\sqrt{2\pi}\sigma_{j_1}}\, e^{-\frac{a_{j_1}^2}{2\sigma_{j_1}^2}}, \quad \pi(\boldsymbol{\eta}_2) = \prod_{j_2=0}^{k_2} \pi(b_{j_2}) = \prod_{j_2=0}^{k_2} \frac{1}{\sqrt{2\pi}\tilde{\sigma}_{j_2}}\, e^{-\frac{b_{j_2}^2}{2\tilde{\sigma}_{j_2}^2}}, \tag{11}$$

and

$$\pi(\boldsymbol{\gamma}) = \prod_{\ell=1}^{2} \pi(\gamma_\ell) = \prod_{\ell=1}^{2} (\lambda_\ell/\gamma_\ell^2) e^{-\lambda_\ell/\gamma_\ell}, \tag{12}$$

then the posterior distribution of the parameters is given by

$$\pi(\boldsymbol{\beta} \mid \mathbf{x}, \mathbf{y}) \propto \textstyle\prod_{i=1}^n \frac{\tau_i^{1-\gamma_1}(1-\tau_i)^{1+\gamma_2}}{(\sum_{j=0}^{k_2} b_j x_i^j)[\gamma_1(1-\tau_i)+\gamma_2\tau_i]}$$
$$\times \textstyle\prod_{j_1=0}^{k_1}(1/\sqrt{2\pi}\sigma_{j_1})\, e^{-\frac{a_{j_1}^2}{2\sigma_{j_1}^2}} \prod_{j_2=0}^{k_2}(1/\sqrt{2\pi}\tilde{\sigma}_{j_2})\, e^{-\frac{b_{j_2}^2}{2\tilde{\sigma}_{j_2}^2}} \prod_{\ell=1}^{2}(\lambda_\ell/\gamma_\ell^2) e^{-\lambda_\ell/\gamma_\ell}$$

and is well defined on $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\gamma}) \in \Omega_1 \times \Omega_2 \times \Omega_3$, where

$\Omega_1 = \{(a_0, \ldots, a_{k_1}) \mid a_0 + a_1 x_i + \cdots + a_{k_1} x_i^{k_1} < y_i, i = 1, \ldots, n\}$,

$\Omega_2 = \{(b_0, \ldots, b_{k_2}) \mid b_0 + b_1 x_i + \cdots + b_{k_2} x_i^{k_2} > 0, i = 1, \ldots, n\}$,

$\Omega_3 = (0, M] \times (0, \infty)$, where M is a fixed positive real number.

Note that the normal priors (11) can be very useful in practice. These priors say that the values of the model parameters can be positive or negative. Large values of σs represent weak prior information about the model parameters.

Now consider model (5), where $Q(\tau, \boldsymbol{\gamma}) = -\frac{1}{\lambda}\ln(1-\tau), \quad \lambda > 0$. So $Q(\tau, \boldsymbol{\gamma})$ is the quantile function of the exponential distribution with rate λ.

Let $\mathbf{y}_p = (y_1, \ldots, y_{p-1})$ and $\mathbf{y} = (y_{p+1}, \ldots, n)$. Then it follows from Theorem 1 that $L(\mathbf{u}(\mathbf{y}) \mid \mathbf{y}_p, \boldsymbol{\beta}) = \prod_{t=p+1}^{n} \gamma(1-\tau_t)$, where τ_t satisfies

$$y_t - (a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p}) = -(1/\gamma)\ln(1-\tau_t), \tag{13}$$

where $\gamma = \eta/\lambda$. Note that in this special case, τ_t can be found exactly.

If the priors for a_i are given by (11) and $\pi(\gamma) = \alpha e^{-\alpha\gamma}$, then the posterior distribution of the model parameters is well defined on $\Omega = \Omega_1 \times \Omega_2$, where $\Omega_1 = \{(a_0, \ldots, a_p) \mid a_0 + a_1 y_{t-1} + \cdots + a_p y_{t-p} \le y_t,\ t = p, \ldots, n\}$, and $\Omega_2 = (0, \infty)$.

Finally, let us consider model (7), where $Q(\tau, \boldsymbol{\gamma})$ is also the quantile function of the power-Pareto distribution, but in this case, we have $m = 2$ and no covariates. It also follows from Theorem 1 that

$$L(\mathbf{u}(\mathbf{y}) \mid \boldsymbol{\beta}) = \prod_{i=1}^{n} \{\tau_i^{1-\gamma_1}(1-\tau_i)^{1+\gamma_2}\}/\{\gamma_1(1-\tau_i) + \gamma_2\tau_i\} \tag{14}$$

where τ_i satisfy

$$(y_{2i} - a_{21}y_{1i} - a_{20})^2 + (y_{1i} - a_{10})^2 = \tau_i^{\gamma_1}(1-\tau_i)^{-\gamma_2}, \qquad i = 1, \ldots, n. \tag{15}$$

Furthermore, if

$$\pi(\boldsymbol{\eta}) = \prod_{k=1}^{2}\prod_{j=0}^{k-1} \pi(a_{kj}) = \prod_{k=1}^{2}\prod_{j=0}^{k-1} \frac{1}{\sqrt{2\pi}\sigma_{kj}} e^{-\frac{a_{kj}^2}{2\sigma_{kj}^2}} \tag{16}$$

and $\pi(\boldsymbol{\gamma})$ is given by (12), where $\boldsymbol{\eta} = \{a_{kj} : k = 1, 2, j = 0, \ldots, k-1\}$, then the posterior distribution of the parameters is given by

$$\pi(\boldsymbol{\beta} \mid \mathbf{u}(\mathbf{y})) \propto \prod_{i=1}^{n} \frac{\tau_i^{1-\gamma_1}(1-\tau_i)^{1+\gamma_2}}{\gamma_1(1-\tau_i)+\gamma_2\tau_i} \prod_{\ell=1}^{2} \frac{\lambda_\ell}{\gamma_\ell^2} e^{-\lambda_\ell/\gamma_\ell} \prod_{k=1}^{2}\prod_{j=0}^{k-1} \frac{1}{\sqrt{2\pi}\sigma_{kj}} e^{-\frac{a_{kj}^2}{2\sigma_{kj}^2}}, \quad (17)$$

and is well defined on $\Omega_1 \times \Omega_2$, where $\Omega_1 = (-\infty, \infty)^3$, $\Omega_2 = (0, M] \times (0, \infty)$ and M is any fixed positive real number.

From these examples, we see that a MCMC method can also be developed for many other quantile function models similarly. In the following section we apply the quantile function models to some real financial time series.

4 Applications to real data sets

The first example shows a 1-dimensional case and the second example shows a multivariate case. Note that in these applications, we do not have any prior knowledge about the model parameters. Therefore the priors given in (11) and (12) with large standard deviations have been used.

4.1 The German DAX

The German DAX is the most commonly cited benchmark for measuring the returns posted by stocks on the Frankfurt Stock Exchange. The data were collected between 1991 and 1999 and the sample size is 1860. See Figure 4(a) for the time series plot of the data.

Let $\tilde{x}_t$ be the German DAX and $x_t = 100(\log(\tilde{x}_{t+1}) - \log(\tilde{x}_t))$ be the log returns. Figure 4(b) shows the log return time series plot. We will study the log return series in this section.

It is worth mentioning that there is no much autocorrelation structure in the log return series, but indeed there are some autocorrelation structures in the squared return series. The time series plot of the log returns also shows large variations in both directions during this

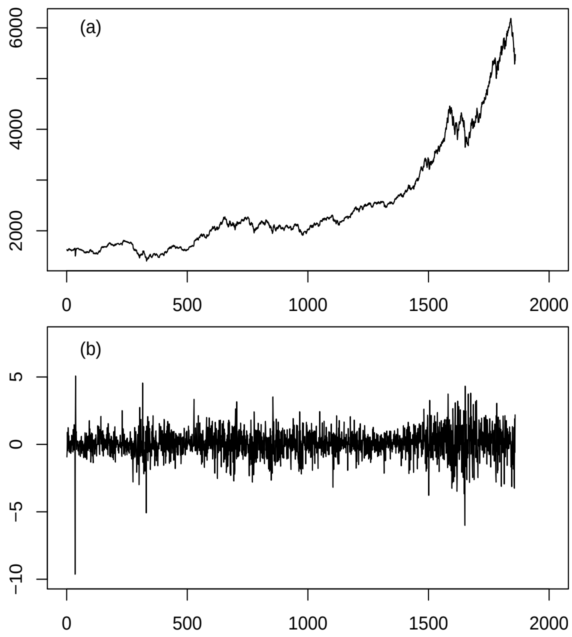

Figure 4: (a) Time series plot of the German DAX between 1991-1998. (b) Time series plot of the log returns.

period of time, which suggests that the distribution of the error term of the model has long tails on both sides, and the distribution may not be symmetric. Due to the flexibility of the generalized lambda distribution, we used this distribution for the error term of model (6). So this model also allows us to investigate the effects of lag values of the returns.

We fitted a set of models with different orders (p, q). To save space, we only reported the results for the best fitted model in this paper, which corresponds to $p = q = 1$ and which was chosen according to the Bayesian factor. The prior density functions for a_{j_1} and b_{j_2} were given by (11) and (12) with $\sigma_0 = \sigma_1 = \tilde{\sigma}_0 = \tilde{\sigma}_2 = \lambda_1 = \lambda_2 = 2$. Furthermore, we took 0.05 as the variance of all the proposal density functions. The initial value of the parameters was taken as $(\bar{x}, 0, 0, \bar{s}, 0, \gamma_1^{(0)}, \gamma_2^{(0)})$ because it is a point in the support of the posterior distribution, where $a_0^{(0)} = \bar{x} = 0.0652$ is the sample mean, $b_0^{(0)} = \bar{s} = 1.030$ is the sample standard deviation, $\gamma_1^{(0)} = -0.259$ and $\gamma_2^{(0)} = -0.333$ are two random samples from a negative exponential distribution of rate 3 and 4 respectively. Our experience with this data set also shows that the initial values do not have significant effects on the convergence of the method. The sampler was ran 200,000 steps. After a burn-in period of the first 10,000 steps, the posterior samples of the model parameters were collected once every 100 steps. Time series plots (not shown to save space) of the simulated parameters show that the convergence was achieved after the burn-in period. Figure 5 shows the histograms of the posterior samples where the vertical lines give the locations of the estimated parameter values. So the fitted model is given by

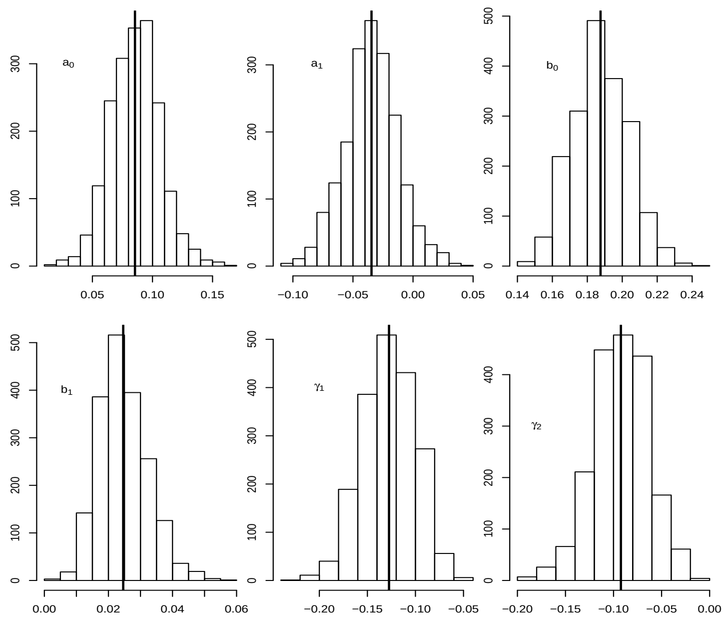

Figure 5: Histograms of the posterior parameter samples for German DAX returns.

$$
\begin{aligned}
Q_{x_t}(\tau \mid \mathbf{x}_{t-1}) &= 0.085 - 0.035x_{t-1} \\
&+ \sqrt{0.188 + 0.025x_{t-1}^2}\left(\frac{\tau^{-0.127}-1}{-0.127} - \frac{(1-\tau)^{-0.092}-1}{-0.092}\right)
\end{aligned} \tag{18}
$$

This fitted model shows that the conditional distribution of x_t is a shifted and scaled generalized lambda distribution $\hat{Q}(\tau, \hat{\gamma}) = \frac{\tau^{-0.127}-1}{-0.127} - \frac{(1-\tau)^{-0.092}-1}{-0.092}$, and both the location $0.085 - 0.035x_{t-1}$ and the scale $\sqrt{0.188 + 0.025x_{t-1}^2}$ of the distribution of x_t depend on x_{t-1}, suggesting that the first order autocorrelation structure has been taken into account.

Figure 6(a) shows the plot of the sample quantiles of the standardized residuals

$$\hat{u}_t = \{x_t - (0.085 - 0.035x_{t-1})\}/\sqrt{0.188 + 0.025x_{t-1}^2}$$

against the quantiles of the generalized lambda distribution $\hat{Q}(\tau, \hat{\gamma})$.

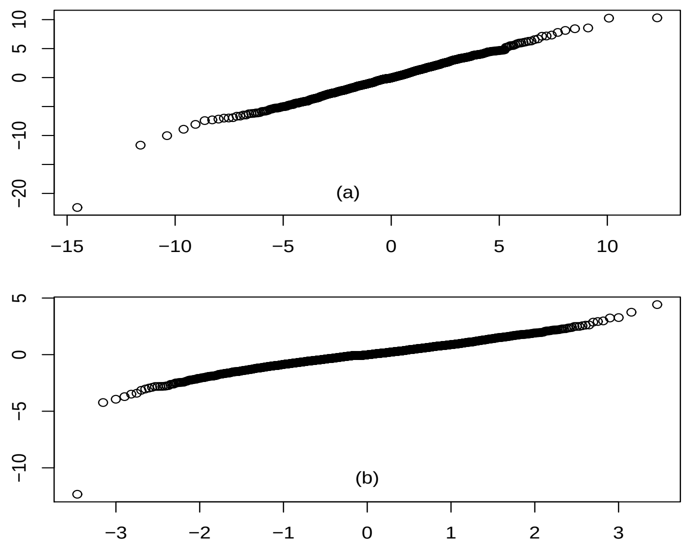

Figure 6: (a) The QQ-plot of the standardized residuals of the quantile function model. (b) The QQ-plot of the standardized residuals of the AR-GARCH model.

We also fitted a set of AR-GARCH models to the return series by using the statistical software R. According to AIC the best fitted model was the AR(1)-GARCH(1,1) model:

$$x_t = 0.0652 - 0.0004x_{t-1} + \sqrt{h_t}\,\epsilon_t, \quad \epsilon_t \sim N(0, 1)$$

where $h_t = 0.0475 + 0.0684v_{t-1}^2 + 0.8877h_{t-1}$ and $v_t = x_t - 0.0652 + 0.0004x_{t-1}$.

Figure 6(a)(b) shows that model (18) has an improved fitting.

Model (18) may be used to obtain predictive quantiles. For illustration purposes Figure 7 shows one-step ahead predictive quanitles at seven different levels, from the top to the bottom, for $\tau = 0.995, 0.95, 0.75, 0.5, 0.25, 0.05$ and 0.005 respectively. A quantile curve at a level, say $\tau = 0.005$, tells us that the conditional probability of the next day's return being lower than this curve is 0.005, thus it provides a measure of value at risk at this level.

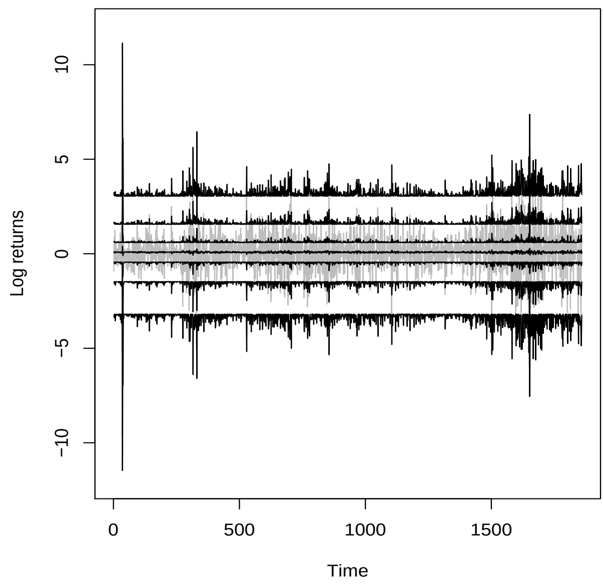

Figure 7: 1-step ahead predictive quantiles for DAX returns. The grey curve corresponds to the actual observed returns.

Table 3: Empirical probability coverage of the predictive quantile curves for DAX returns

τ	0.005	0.05	0.25	0.5	0.75	0.95	0.995
No. returns	5	96	455	952	1380	1762	1849
Proportion	0.003	0.052	0.245	0.512	0.743	0.948	0.995

Table 3 further shows the empirical coverage of these estimated predictive quantile curves, where the second and the third rows give the number of the returns and the proportion of the returns that are below the respective quantile curves. Clearly, all estimated predictive quantile curves provide a good probability coverage, which also implies that the fitted quantile function model is reasonably good for this data set.

Further statistical inferences may also be made by using model (18). For example, one-step ahead out-of-sample forecast can be obtained straightaway. Note that the last observed log return is 2.192% which was obtained on day 1860. So the conditional quantile function of the log return on the next day is

$$
\begin{aligned}
Q_{x_{1861}}(\tau \mid \mathbf{x}_{1860}) &= 0.085 - 0.035(2.192) \\
&+ \sqrt{0.188 + 0.025(2.192)^2} \left(\tfrac{\tau^{-0.127}-1}{-0.127} - \tfrac{(1-\tau)^{-0.092}-1}{-0.092} \right) \\
&= 0.0083 + 0.5551 \left(\tfrac{\tau^{-0.127}-1}{-0.127} - \tfrac{(1-\tau)^{-0.092}-1}{-0.092} \right).
\end{aligned}
$$

This predictive quantile function in fact defines the whole conditional distribution of x_{1861} and its density function plot is shown in Figure 8. It is seen that the predictive condi-

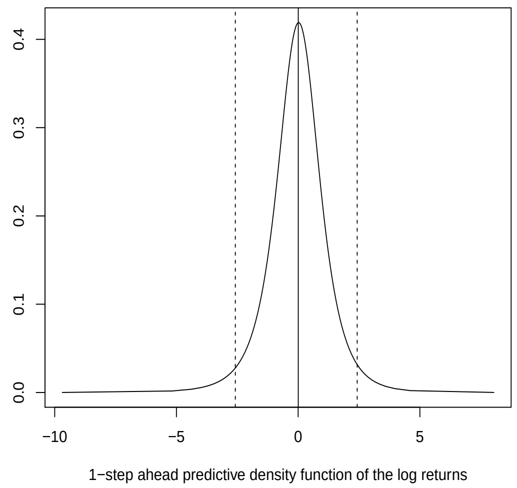

Figure 8: The one-step ahead predictive conditional density function of the DAX return at time 1861. The middle vertical line indicates the location of the predictive median, and the dashed vertical lines give a 95% credible interval for the DAX returns at time 1861.

tional distribution is slightly skewed to the left. The middle vertical line in Figure 8 gives the location of the predictive conditional median, which is about 0.005%; the two dashed lines are at $x = -2.581$ and $x = 2.426$, which show that 95% of the returns will fall between -2.581% and 2.426%. Note that because we have estimated the whole conditional quantile function of the returns, we can also make inferences on any other quantities of interest, such as predictive mean, variance, value at risk at any given levels and expected shortfalls etc.

4.2 The currency exchange rates

Consider the currency exchange rates USD/GBP and CAD/GBP from 2 January 1997 to 21 November 2000. Each time series is of length 975 and the plots of the exchange rates are given in Figure 9 (a) and (b), which suggest that the time series are not stationary. In this study we focus on their returns. Let x_{1t} and x_{2t} be the returns on USD/GBP and

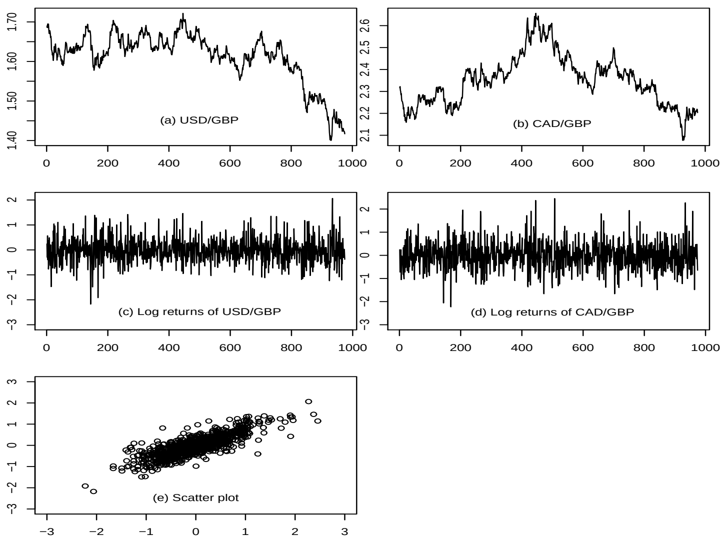

Figure 9: Plots of the exchange rates USD/GBP and CAD/GBP.

CAD/GBP respectively. Figure 9 (c) and (d) show the plots of these return series. Clearly they have similar features that many economic and financial time series have. The auto- and partial correlation function plots of x_{1t} and x_{2t} suggest that there is no significant autocorrelation structure in each of the return series. So we ignore the autocorrelation structure but concentrate on the dependence structure of the two return series.

The scatter plot of the returns in Figure 9(e) shows an obvious positive correlation between the two series. Note that it is not appropriate to use a usual linear regression model in this case as the data are clustered. So we fitted model (7) to the data, where $Q(\tau, \gamma)$ is the quantile function of the power-Pareto distribution.

The prior distributions of the parameters are given by (16) and (12) with $\sigma_{kj} = 13$,

$k = 1, 2$, $j = 0, \ldots, k-1$, and $\lambda_j = 1$, $j = 1, 2$. The chain was run for 50,000 steps. After a burn-in period of the first 3,000 steps, the posterior samples were collected once every 30 steps. The fitted model is

$$(x_2 - 0.6580x_1 + 0.0218)^2 + (x_1 + 0.0227)^2 = \tau^{1.8450}/(1-\tau)^{0.2316}.$$

The QQ-plot in Figure 10 (a) confirms that no major concerns about the fitted model. The fitted τth-quantile curves for $\tau = 0.995, 0.95$, 0.90, $0.75, 0.5, 0.25$ and 0.05 together with the observed data (grey points) are shown in Figure 10(b).

Note that one of the common approaches to bivariate data is to fit a bivariate normal distribution to the data. Figure 10(c) shows the contour curves obtained from the fitted bivariate normal distribution by using the statistical software R. The fitted model has mean $(-0.005, -0.018)^\top$ and variance-covariance matrix $\begin{pmatrix} 0.3522 & 0.2350 \\ 0.2350 & 0.2332 \end{pmatrix}$. Although the plots of the contour curves of the bivariate normal density function look similar to those of the quantile curves, they have completely different interpretations. With a τth-quantile curve we mean that the probability for the two returns fall inside the region defined by the τth-quantile curve is τ, while with a level c contour curve we mean that the value of the joint probability density function at any points in the region defined by the level c contour curve is c. Figure 10(d) further shows a 95% credible interval (dashed curves) for each fitted quantile curve (solid curves).

In studying value at risk, we are interested in extreme quantiles. Figure 10(b) suggests that the joint probability of the two returns on the currency exchange rates being outside the 0.995th quantile curve is 0.005. Indeed, we have checked and found that only 5 points are outside the 0.995th quantile curve (see Figure 10(b)). The empirical probability for this to happen is 5/974 = 0.00513, which is in a good agreement with the true probability.

As our approach is different from the quantile regression approach, we also empirically compared our method with that of Wei (2008). Again, statistical software R was used for obtaining fitted reference quantile contours (Wei, 2008) for this data set. Figure 11 shows

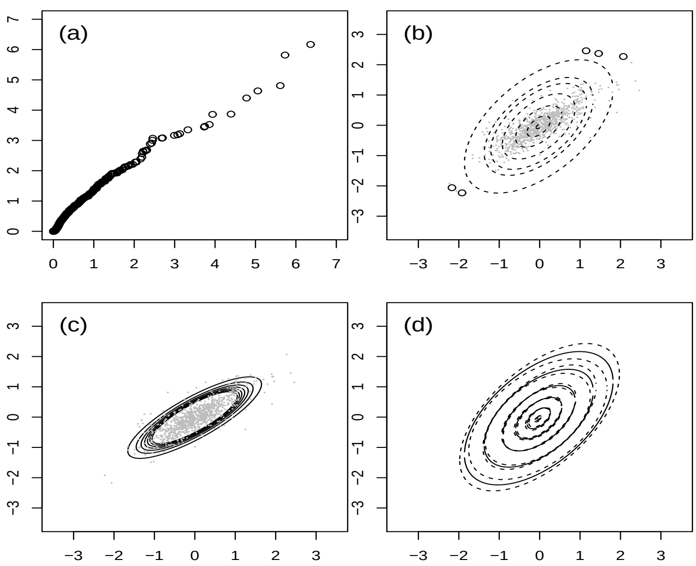

Figure 10: QQ-plot, fitted quantile curves and the corresponding 95% credible intervals obtained from our model and contour curves of a fitted bivariate normal model.

the fitted τth quantile curves by using different methods, where $\tau = 0.995, 0.95, 0.5, 0.05$. It is seen that there is no restriction on the shape of the fitted reference quantile contours if Wei's method is used. This is due to the non-parametric nature of the method. For this data set, when $\tau > 0.05$, both methods provide similar results, while when $\tau = 0.05$ the fitted quantile curves are quite different. We suggest to use both methods if possible in practice for the best results.

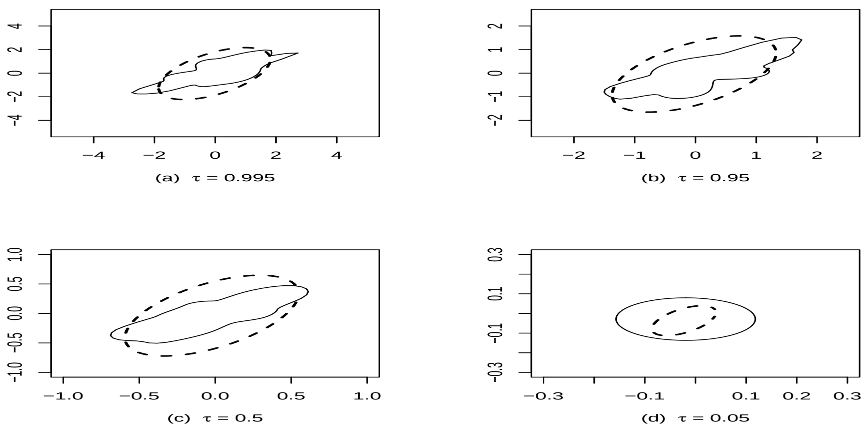

Figure 11: Fitted quantile curves by using our method (dashed curves) and fitted reference quantile contours by using Wei's (2008) method (continuous curves).

5 Comments and conclusions

In this paper we proposed a general quantile function model which covers both one dimensional and multiple dimensional models and which takes the models studied by Cai (2009, 2010a, 2010b) and Cai et al. (2013) as its special cases. We also developed a new uniform Bayesian approach to model parameter estimation.

It is noticed that many distributions have not been used in economic and financial modelling. One such example is the class of distributions that can only be defined by using their quantile functions. The developed methodology makes it possible to use such distributions in economic and financial modelling. Our results show that the developed method can be very useful in practice.

It is worth mentioning that for tail probabilities, a parametric model would always give finer results, but only under the condition that it is a proper model. Our results show that it is possible to build up a proper model for a data set by using the properties of quantile functions. We have found that the robustness of the quantile function models can be significantly improved by using a properly constructed $Q(\tau, \gamma)$. A good example is the use of the generalized lambda distribution. This distribution can provide very good approximations to many standard distributions include normal, lognormal, weibull, t- and F-distributions and many others. We have shown that how this distribution was constructed by using simple quantile functions. Many other distributions can also be constructed similarly. We will investigate this important issue further in the future.

For multivariate quantile surface estimation, we believe thorough and systematic comparisons between our approach, Wei's (2008) approach and Li and Racine's (2008) approach are certainly required in the future.

Acknowledgement. We would like to express our sincere thanks to the Editor, the Associate Editor and an anonymous referee for their very thoughtful and constructive comments which have greatly enhanced the quality and the presentation of the paper.

References

[1] Bondell, H.D., Reich, B.J. and Wang, H. (2010). Non-crossing quantile regression curve estimation. *Biometrika*, 97, 825-838.

[2] Cai, Y. (2010a). Multivariate quantile function models. *Statistica Sinica*, 20, 481-496.

[3] Cai, Y. (2010b). Polynomial power-Pareto quantile function models. *Extremes*, 13, 291-314.

[4] Cai, Y. (2009). Autoregression with non-Gaussian Innovations. *Journal of Time Series Econometrics*, Vol.1, Iss.2, Article 2. DOI: 10.2202/1941-1928.1016.

[5] Cai, Y. and Stander, J. (2008). Quantile self-exciting threshold autoregressive time series models. *Journal of Time Series Analysis*, **29**, 186-202.

[6] Cai, Y, Montes-Rojas, G. and Olmo, J. (2013). Quantile double AR time series models for financial returns. *Journal of Forecasting* 32, 551-560.

[7] Freimer, M., Mudholkar, G.S., Kollia, G., and Lin, C.T. (1988). A study of the generalised Tukey lambda distribution. Communications of Statistics, *Theory Math.*, 17, 3547-3567.

[8] Gilchrist, W.G. (2000). *Statistical Modelling with Quantile Functions*. Chapman & Hall/CRC.

[9] Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd ed. McGraw-Hill.

[10] Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

[11] Koenker, R. (1984). A note on L-estimators for linear models. *Statistics and Probability Letters*, 2, 323-325.

[12] Koenker, R. and D'Orey, V. (1987). Computing regression quantiles. *Applied Statistics*, **36**, 383-393.

[13] Koenker, R. and D'Orey, V. (1994). A remark on Algorithm AS229: Computing dual regression quantiles and regression rank scores. *Applied Statistics*, **43**, 410-414.

[14] Kottas, A. and Gelfand, A.E. (2001). Bayesian semiparametric median regression modelling. *Journal of American Statistical Association* 96, 1458-1468.

[15] Lancaster, T. and Jun, S.J (2010). Bayesian quantile regression methods. *Journal of Applied Econometrics*, 25, 287-307.

[16] Li, Q. and Racine, J. S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business and Economic Statistics*, 26, 423-434.

[17] Maasoumi E. and Racine, J.S. (2013). A solution to aggregation and an application to multidimensional 'well-being' frontiers. http://econ.snu.kr/eng/intro/download.php?filename=137403495268540.pdf&realname =md_ineq.pdf, 12 September 2013.

[18] Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annuals of Mathematical Statistics*, **23**, 470-472.

[19] Salibian-Barrera, M. and Wei, Y. (2008). Weighted quantile regression with non-elliptically structured covariates. *The Canadian Journal of Statistics*, *Canadian Journal of Statistics* 36, 595-611.

[20] Wei, Y. (2008). An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts. *Journal of the American Statistical Association*, **103**, 397-409.

[21] Wu, Y. and Liu, Y. (2009). Stepwise multiple quantile regression estimation using non-crossing constraints. Statistics and Its Interface 2, 299-310.

[22] Yu, K. and Moyeed, R.A. (2001). Bayesian quantile regression. *Statistics and Probability Letters*, **54**, 437-447.

Appendix

Proof of Theorem 1

The likelihood of $\mathbf{u}$ can be written as $L(\mathbf{u} \mid \boldsymbol{\beta}) = f(u_1, \boldsymbol{\gamma}) f(u_2, \boldsymbol{\gamma}) \cdots f(u_n, \boldsymbol{\gamma})$, where u_i depends on $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$. Now, it follows from (8) that for each u_i there exists τ_i such that τ_i is the actual probability for having the observed value u_i. Hence it follows from $f(u, \gamma) = 1/(\partial Q(\tau, \boldsymbol{\gamma})/\partial \tau)$ that the result holds.

Proof of Theorem 2

First note that the proposals η_j' $(j = 0, \ldots, p_1)$ are obtained from $N\left(\eta_j, \sigma_{\eta_j}^2\right)$ independently such that $\boldsymbol{\eta}' \in \Omega_1$, and the proposals γ_ℓ' $(\ell = 1, \ldots, p_2)$ are obtained from $N\left(\gamma_\ell, \sigma_{\gamma_\ell}^2\right)$ independently such that $\boldsymbol{\gamma}' \in \Omega_2$. Therefore, it follows from

$$q(\boldsymbol{\eta}' \to \boldsymbol{\eta}) = \frac{\prod_{j=0}^{p_1} (1/\sqrt{2\pi}\sigma_{\eta_j}) exp\{-(\eta_j - \eta_j')^2/2\sigma_{\eta_j}^2\}}{\int_{\Omega_1} \prod_{j=0}^{p_1} (1/\sqrt{2\pi}\sigma_{\eta_j}) exp\{-(\eta_j - \eta_j')^2/2\sigma_{\eta_j}^2\} d\eta_j}$$

that (9) holds. Similarly, we can show that (10) also holds.