

Bayesian inference for CoVaR

M. Bernardi,

Sapienza University of Rome, Italy

and

G. Gayraud,

Université de Technologie de Compiègne and CREST, France

and

L. Petrella,

Sapienza University of Rome, Italy

June 17, 2013

Abstract

Recent financial disasters emphasised the need to investigate the consequence associated with the tail co-movements among institutions; episodes of contagion are frequently observed and increase the probability of large losses affecting market participants' risk capital. Commonly used risk management tools fail to account for potential spillover effects among institutions because they provide individual risk assessment. We contribute to analyse the interdependence effects of extreme events providing an estimation tool for evaluating the Conditional Value-at-Risk (CoVaR) defined as the Value-at-Risk of an institution conditioned on another institution being under distress. In particular, our approach relies on Bayesian quantile regression framework. We propose a Markov chain Monte Carlo algorithm exploiting the Asymmetric Laplace distribution and its representation as a location-scale mixture of Normals. Moreover, since risk measures are usually evaluated on time series data and returns typically change over time, we extend the CoVaR model to account for the dynamics of the tail behaviour. Application on U.S. companies belonging to different sectors of the Standard and Poor's Composite Index (S&P500) is considered to evaluate the marginal contribution to the overall systemic risk of each individual institution.

Index Terms VaR, CoVaR, Bayesian quantile regression, time-varying quantile, State Space models.

1 Introduction

During the last years a particular attention has been devoted to measure and quantify the level of financial risk within a firm or investment portfolio. One of the most diffuse risk measurement has become the Value-at-Risk (VaR) which measures the maximum loss in value of a portfolio over a predetermined time period for a given confidence level. In fact, in the current banking regulation framework, the VaR becomes an important risk capital evaluation tool where different institutions are considered as independent entities. Unfortunately, such risk measure fails to consider the institution as part of a system which

might itself experience instability and spread new sources of systemic risk. For a comprehensive and up to date overview of VaR and related risk measures see for example Jorion [31] and McNeil *et al.* [47]. Recent financial disasters emphasised the need for a deep investigation of the co-movement among institutions in order to evaluate their tail interdependence relations. Especially during financial crises, episodes of contagion among institutions are not rare and thus need to be taken into account in order to analyse the overall health level of a financial system: company specific risk can not be appropriately assessed in isolation, without accounting for potential spillover effects from other firms. For this reason different systemic risk measures have been proposed in literature to analyse the tail-risk interdependence (see Acharya *et al.* [1], Acharya *et al.* [2], Adams *et al.* [3], Brownlees and Engle [9] and Billio *et al.* [8]). Recently, Adrian and Brunnermeier [4] introduced the so called Conditional Value-at-Risk (CoVaR), which is defined as the overall VaR of an institution conditional on another institution being under distress. In this way the CoVaR not only capture the systematic risk embedded in each institution, but also reflects individual contributions to the systemic risk, capturing extreme tail risk interdependence.

There are many possible ways to infer on VaR and CoVaR. The most common approaches to estimate VaR is the variance–covariance methodology, historical and Monte Carlo simulations. For an overview of alternative parametric and nonparametric methodologies and processes to generate VaR estimates see Jorion [31] and Lee and Su [43]. See also Chao *et al.* [13] and Taylor [51] for recent develops. Bernardi [5] and Bernardi *et al.* [6] propose to estimate VaR and related risk measures by fitting asymmetric mixture models to the unconditional distribution of returns. Moreover, Girardi and Ergün [26] propose to estimate the CoVaR using multivariate Generalized ARCH models for the conditional returns, and Bernardi *et al.* [7] consider the class of multivariate hidden Markov models.

In this paper, since both VaR and CoVaR are distribution quantiles we address the problem of their estimation using a quantile regression approach. Quantile regression has been popular as a simple, robust and distribution free modeling tool since the seminal work of Koenker and Basset [33] and Koenker [32]. It provides a way to model the conditional quantiles of a response variable with respect to some covariates in order to have a more complete picture of the entire conditional distribution than ordinary least squares regression. In fact, sometimes problem specific features, like skewness, fat-tailed, outliers, truncation, censored and heteroskedasticity, can shadow the nature of the dependence between the variable of interest and the covariates so that the conditional mean would not be enough to understand the nature of that dependence. In particular, the quantile regression approach is appropriate not only when the underlying model is nonlinear or the innovation terms are non-Gaussian, but also when modeling the tail behaviour of the underlying distribution is the primary interest. There is a number of papers on quantile regression both in frequentist and Bayesian framework dealing with parametric and nonparametric approaches. For a detailed review and references, see for e.g., Lum and Gelfand [46] and Koenker [32].

In quantile regression, the quantile of order τ of a dependent variable Y is expressed as a function of covariates \mathbf{X} , say $q^\tau(\mathbf{X})$. In literature different representations have been proposed to specify the quantile function $q^\tau(\mathbf{x})$; the most common specification is the linear one adopted hereafter:

$$q^\tau(\mathbf{x}) = \mathbf{x}^\top \boldsymbol{\theta}, \quad (1.1)$$

where \mathbf{x}^\top denotes the transpose of \mathbf{x} .

The problem of estimating $q^\tau(\mathbf{x})$ through quantile regression has been considered both

from the frequentist and Bayesian point of view. In the former case, Koenker and Basset [33] show that the quantile estimation problem is solved by the following minimisation problem:

$$\operatorname{argmin}_{q^\tau} \sum_{t=1}^T \rho_\tau(y_t - q^\tau(\mathbf{x}_t)), \quad (1.2)$$

where (y_t, \mathbf{x}_t) for $t = 1, \dots, T$ are observations from (Y, \mathbf{X}) and $\rho_\tau(y) = y(\tau - \mathbb{1}(y < 0))$ is the quantile loss function. The Bayesian quantile regression approach (see Yu and Moyeed [53], Kottas and Gelfand [37] and Kottas and Krnjajic [38]) instead considers the distribution of $Y|\mathbf{x}$ as belonging to the Asymmetric Laplace distribution family, denoted by ALD $(\tau, q^\tau(\mathbf{x}), \sigma)$, with positive σ , whose density function is given by:

$$\operatorname{ald}(y|q^\tau(\mathbf{x}), \sigma) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\frac{\rho_\tau(y - q^\tau(\mathbf{x}))}{\sigma}\right\}.$$

The nice feature of the ALD $(\tau, q^\tau(\mathbf{x}), \sigma)$ distribution is that the regression function $q^\tau(\mathbf{x})$ corresponds exactly to the theoretical τ -th quantile of $Y|\mathbf{x}$.

Quantile regression methods have been extensively considered in literature to evaluate the VaR (see, among others, Huang [29], Schaumburg [50], Chernozhukov and Du [10], Kuester *et al.* [41], Taylor [51] and Gerlach *et al.* [23]); recently Chao *et al.* [13] and Hautsch *et al.* [28] consider the same approach to calculate also the CoVaR. In this paper, we propose a Bayesian approach to cast the CoVaR within a quantile regression framework and we show how to model and evaluate it as a quantile of the conditional distribution of an institution k given a particular quantile of another institution j . Bayesian methods are very useful tools since enable us to learn about the whole distribution of quantities. In this context, since the quantities of interest are risk measures, learning about the whole distribution becomes more relevant due to the interpretation of the VaR and CoVaR as financial losses. We also show how to calculate their posterior credible sets useful to assess estimates' accuracy. Up to our knowledge this is the first attempt to infer on CoVaR from a Bayesian point of view.

As second step, since risk measures are usually evaluated on time series data and returns typically change over time, we extend the Adrian and Brunnermeier [4] CoVaR approach to account for the dynamics of the tail behaviour. The idea is to consider time varying quantiles to link the future tail behaviour of a time series to its past movements which is important in risk management contest. In particular, starting from the idea stated in De Rossi and Harvey [16] we propose a dynamic model to capture the evolution of VaR and CoVaR. In order to provide a flexible solution for the quantile modelisation whilst retaining a parsimonious representation, the time evolution of the process should be carefully chosen. Hence, throughout the paper, we propose to model the dynamics of the quantile functions' parameters as local linear trends which represents a good compromise between degree of smoothness of the resulting quantiles and the ability of the model to capture changes over time. Time varying quantiles represent a valid alternative to conditional quantile autoregression proposed in different contexts by Engle and Manganelli [19], Gerlach *et al.* [23], Gouriéroux and Jasiak [27] and Koenker and Xiao [34].

To implement the dynamic Bayesian inference, we cast VaR and CoVaR models in state space representation and we run a Gibbs sampler algorithm using the Exponential-Gaussian mixture representation of ALD distributions (Kotz *et al.* [39]). This approach allows us to obtain a conditional Gaussian state space representation which permits an

efficient numerical solution to the inferential problem. In order to make posterior inference, we use the maximum a posteriori summarising criteria and we prove that it leads to estimated quantiles having good sample properties according to De Rossi and Harvey results [16].

There are several applications of CoVaR which are interesting both in economics and finance. In this paper we analyse different U.S. companies belonging to several sectors of the Standard and Poor's Composite Index (S&P500) in order to evaluate the marginal contribution to the overall systemic risk of a single institution belonging to it. The empirical results show that the proposed models provide realistic and informative characterisation of extreme tail co-movements. Moreover, our findings suggest that the dynamic model we propose is more appropriate when dealing with financial time series data.

The paper is organised as follows. In Section 2 we give a brief definition of Value-at-Risk and Conditional Value-at-Risk measures. In Section 3 we build the time invariant Bayesian model and we provide details on how to make inference using Markov chain Monte Carlo (MCMC) algorithms. In Section 4 we extend the previous framework to the time varying case, allowing the representation of marginal and conditional quantiles as functions of latent processes. In Section 5 we apply the proposed models to real data, while Section 6 concludes.

2 VaR and CoVaR representations

Let (Y_1, \dots, Y_d) be a d -dimensional ($d > 1$) random vector whose real components are expressed through some covariates $\mathbf{X} = (X_1, X_2, \dots, X_M)$ ($M \geq 1$). We have in mind that for any $j \in \{1, \dots, d\}$, Y_j , the variable of institution j , depends on some covariates \mathbf{X} and that for any $k \in \{1, \dots, d\}$, $k \neq j$, the behaviour of the variable Y_k , related to either institution k or the whole system, depends on covariates \mathbf{X} as well as on the behaviour of the variable of institution j , Y_j . Without loss of generality, thereafter, we fix τ in $(0, 1)$ and suppose that we are interested in institutions j and k for $j \neq k$ and $(j, k) \in \{1, \dots, d\} \times \{1, \dots, d\}$.

Let us recall that the Value-at-Risk, $\text{VaR}_j^{\mathbf{x}, \tau}$ of institution j is the τ -th level conditional quantile of the random variable $Y_j | \mathbf{X} = \mathbf{x}$, i.e.,

$$\mathbb{P}(Y_j \leq \text{VaR}_j^{\mathbf{x}, \tau} | \mathbf{X} = \mathbf{x}) = \tau.$$

The Conditional Value-at-Risk ($\text{CoVaR}_{k|j}^{\mathbf{x}, \tau}$) is the Value-at-Risk of institution k conditional on $Y_j = \text{VaR}_j^{\mathbf{x}, \tau}$ at the level τ , i.e., $\text{CoVaR}_{k|j}^{\mathbf{x}, \tau}$ satisfies the following equation

$$\mathbb{P}(Y_k \leq \text{CoVaR}_{k|j}^{\mathbf{x}, \tau} | \mathbf{X} = \mathbf{x}, Y_j = \text{VaR}_j^{\mathbf{x}, \tau}) = \tau. \quad (2.1)$$

Note that the CoVaR corresponds to the τ -th quantile of the conditional distribution of $Y_k | \{\mathbf{X} = \mathbf{x}, Y_j = \text{VaR}_j^{\mathbf{x}, \tau}\}$.

Assuming the linear representation (1.1) of the quantiles under interest, we can write:

$$\text{VaR}_j^{\mathbf{x}, \tau} = \theta_{j,0}^\tau + \theta_{j,1}^\tau x_1 + \theta_{j,2}^\tau x_2 + \dots + \theta_{j,M}^\tau x_M \quad (2.2)$$

$$\text{CoVaR}_{k|j}^{\mathbf{x}, \tau} = \theta_{k,0}^\tau + \theta_{k,1}^\tau x_1 + \theta_{k,2}^\tau x_2 + \dots + \theta_{k,M}^\tau x_M + \beta^\tau \text{VaR}_j^{\mathbf{x}, \tau}, \quad (2.3)$$

where $\theta_{l,m}^\tau$ are unknown parameters with $l \in \{j, k\}$ and $m = 0, \dots, M$. For the sake of simplicity we consider the same τ for both VaR and CoVaR and for the ease of exposition we drop the τ index from all parameters.

3 Time invariant quantile model

The use of Bayesian inference in quantile regression contest is now quite standard although quite recent. In what follows, we adopt the approach used in Yu and Moyeed [53] where data come from an Asymmetric Laplace Distribution which is a convenient tool to deal with quantile regression problems in a Bayesian framework. Suppose that we observe $(\mathbf{y}, \mathbf{x}) = (\mathbf{y}_t, \mathbf{x}_t)_{t=1}^T = (y_{j,t}, y_{k,t}, \mathbf{x}_t)_{t=1}^T$, T independent realizations of (Y_j, Y_k, \mathbf{X}) .

To estimate $\text{VaR}_j^{\mathbf{x}, \tau}$ and $\text{CoVaR}_{k|j}^{\mathbf{x}, \tau}$ we consider the following equations: $\forall t \in \{1, 2, \dots, T\}$,

$$y_{j,t} = \mathbf{x}_t^\top \boldsymbol{\theta}_j + \epsilon_{j,t} \quad (3.1)$$

$$y_{k,t} = \mathbf{x}_t^\top \boldsymbol{\theta}_k + \beta y_{j,t} + \epsilon_{k,t}, \quad (3.2)$$

where β , $\boldsymbol{\theta}_j$ and $\boldsymbol{\theta}_k$ are unknown parameters of dimension 1, $(M+1)$ and $(M+1)$ respectively, and the first component of \mathbf{x}_t is equal to 1 to include a constant term in the regression function. Here, for any $t \in \{1, \dots, T\}$, $\epsilon_{j,t}$ and $\epsilon_{k,t}$ are independent random variables distributed according to $\text{ALD}(\tau, 0, \sigma_j)$ and $\text{ALD}(\tau, 0, \sigma_k)$ respectively, with positive σ_j and σ_k . Due to the property of Asymmetric Laplace distributions, the functions $\mathbf{x}_t^\top \boldsymbol{\theta}_j$ and $\mathbf{x}_t^\top \boldsymbol{\theta}_k + \beta y_{j,t}$ correspond to the τ -th quantiles of $Y_j | \mathbf{X} = \mathbf{x}$ and $Y_k | \{\mathbf{X} = \mathbf{x}, Y_j = y_{j,t}\}$, respectively.

For a fully Bayesian modelling, we need to specify the prior distribution on the unknown parameters vector $\boldsymbol{\gamma} = (\boldsymbol{\theta}, \beta, \sigma_j, \sigma_k)$. We assume the following priors independent on the value of τ :

$$\pi(\boldsymbol{\gamma}) = \pi(\boldsymbol{\theta})\pi(\beta)\pi(\sigma_j)\pi(\sigma_k), \quad (3.3)$$

with $\boldsymbol{\theta} = (\boldsymbol{\theta}_j, \boldsymbol{\theta}_k)^\top \sim \mathcal{N}_{(2M+2)}(\boldsymbol{\theta}^0, \Sigma^0)$, $\beta \sim \mathcal{N}(\beta^0, \sigma_\beta^2)$, $\sigma_j \sim \mathcal{IG}(a_j^0, b_j^0)$ and $\sigma_k \sim \mathcal{IG}(a_k^0, b_k^0)$, and where $\Sigma^0 = \text{diag}(\Sigma_j^0, \Sigma_k^0)$, $\boldsymbol{\theta}^0 = (\boldsymbol{\theta}_j^0, \boldsymbol{\theta}_k^0)^\top$, $\beta^0, \sigma_\beta^2 > 0$, $a_j^0 > 0$, $b_j^0 > 0$, $a_k^0 > 0$ and $b_k^0 > 0$ are given. Notations \mathcal{N} and \mathcal{IG} refer to a Gaussian and an Inverse Gamma distributions. Typically a vague prior will be used on σ_j and σ_k because they are regarded as nuisance parameters.

As discussed in Yu and Moyeed [53], due to the complexity of the likelihood function, the resulting posterior density for the regression parameters $\boldsymbol{\theta}$ and β does not admit a closed form representation for the full conditional distributions, and need to be sampled by using MCMC-based algorithms. According to Kozumi and Kobayashi [40] we adopt the following well-known representation (3.4) (see Kotz *et al.* [39] and Park and Casella [48]) of $\epsilon \sim \text{ALD}(\tau, 0, \sigma)$ as a mixture of Exponential and Normal distributions:

$$\epsilon = \lambda\omega + \delta\sqrt{\sigma\omega}z, \quad (3.4)$$

where $\omega \sim \mathbb{Exp}(\sigma^{-1})$ and $z \sim \mathcal{N}(0, 1)$ are independent random variables and notation $\mathbb{Exp}(\cdot)$ denotes the Exponential distribution. This allows us to use a Gibbs sampler algorithm as shown in the next subsection. The parameters λ and δ are set by

$$\lambda = \frac{1-2\tau}{\tau(1-\tau)}, \quad \delta = \frac{2}{\tau(1-\tau)}, \quad (3.5)$$

in order to ensure that the τ -th quantile of ϵ is equal to zero. Using this augmented structure, the model given by equations (3.1) and (3.2) admits conditionally on w , the following Gaussian representation: $\forall t \in \{1, 2, \dots, T\}$,

$$y_{j,t} = \mathbf{x}_t^\top \boldsymbol{\theta}_j + \lambda\omega_{j,t} + \delta\sqrt{\sigma_j\omega_{j,t}}z_{j,t} \quad (3.6)$$

$$y_{k,t} = \mathbf{x}_t^\top \boldsymbol{\theta}_k + \beta y_{j,t} + \lambda\omega_{k,t} + \delta\sqrt{\sigma_k\omega_{k,t}}z_{k,t} \quad (3.7)$$

where $z_{j,t}$, $z_{k,t}$ are independent and $\omega_{j,t}$, $\omega_{k,t}$ are drawn independently from $\text{Exp}(\sigma_j^{-1})$ and $\text{Exp}(\sigma_k^{-1})$, respectively. From (3.6) and (3.7), the distribution of \mathbf{Y} conditional on the parameters vector $\boldsymbol{\gamma}$, the observed exogenous variables \mathbf{x} and the augmented variables $\boldsymbol{\omega} = (\omega_{j,t}, \omega_{k,t})_{t=1}^T$ becomes

$$f(\mathbf{y}|\boldsymbol{\omega}, \mathbf{x}, \boldsymbol{\gamma}) = \prod_{t=1}^T \mathcal{N}(y_{j,t}|\omega_{j,t}, \mathbf{x}_t, \boldsymbol{\theta}_j, \sigma_j) \times \prod_{t=1}^T \mathcal{N}(y_{k,t}|\omega_{k,t}, y_{j,t}, \mathbf{x}_t, \beta, \boldsymbol{\theta}_k, \sigma_k).$$

3.1 Computations

Due to the Gaussian representation shown above we are able to implement a partially collapsed Gibbs sampler algorithm based on data augmentation (see Liu [45] and van Dyk and Park [52]). The key idea of the collapsed Gibbs sampler is to avoid simulations from the full conditional distributions of the all model parameters $\boldsymbol{\theta}_j, \boldsymbol{\theta}_k, \beta, \sigma_j, \sigma_k$ by analytically marginalizing them out. This approach has several advantages with respect to a systematic sampling because it reduces the computational time and increases the convergence rate of the sampler. Unfortunately, in our model this complete collapsed approach is not possible since the predictive distribution of the augmented variables $(\omega_{j,t}, \omega_{k,t})$ has not a closed form expression. Instead, it is possible to integrate out the variables $(\omega_{j,t}, \omega_{k,t})$ given the observations, from the full conditionals of the scale parameters (σ_j, σ_k) . The partially collapsed Gibbs sampler we implement is an iterative simulation procedure from the following full conditional distributions:

1. the full conditional distribution of the scale parameters σ_j and σ_k are sampled by integrating out the augmented latent factors $(\omega_{j,t}, \omega_{k,t})_{t=1}^T$, becoming:

$$\pi(\sigma_l|\mathbf{y}_l, \boldsymbol{\theta}_l) \propto \mathcal{IG}(\tilde{a}_l, \tilde{b}_l), \quad \mathbf{y}_l = (y_{l,t})_{t=1}^T, \forall l \in \{j, k\}$$

where

$$\begin{aligned} \tilde{a}_j &= a_j^0 + T, & \tilde{b}_j &= b_j^0 + \sum_{t=1}^T \rho_\tau(y_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j), \\ \tilde{a}_k &= a_k^0 + T, & \tilde{b}_k &= b_k^0 + \sum_{t=1}^T \rho_\tau(y_{k,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_k - \beta y_{j,t}). \end{aligned}$$

2. $\forall t = 1, \dots, T$, $\pi(\omega_{j,t}^{-1}|y_{j,t}, \mathbf{x}_t, \boldsymbol{\theta}_j, \sigma_j) \propto \mathcal{IN}(\psi_{j,t}, \phi_{j,t})$, i.e., an Inverse Gaussian with parameters

$$\psi_{j,t} = \sqrt{\frac{\lambda^2 + 2\delta^2}{(y_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j)^2}}, \quad \phi_{j,t} = \frac{\lambda^2 + 2\delta^2}{\delta^2 \sigma_j}$$

3. $\forall t = 1, \dots, T$, $\pi(\omega_{k,t}^{-1}|\mathbf{y}_t, \mathbf{x}_t, \boldsymbol{\theta}_k, \beta, \sigma_k) \propto \mathcal{IN}(\psi_{k,t}, \phi_{k,t})$ with parameters

$$\psi_{k,t} = \sqrt{\frac{\lambda^2 + 2\delta^2}{(y_{k,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_k - \beta y_{j,t})^2}}, \quad \phi_{k,t} = \frac{\lambda^2 + 2\delta^2}{\delta^2 \sigma_k}$$

4. $\pi(\boldsymbol{\theta}_j | \mathbf{y}_j, \mathbf{x}, \boldsymbol{\omega}_j, \sigma_j) \propto \mathcal{N}_{M+1}(\tilde{\boldsymbol{\theta}}_j, \tilde{\Sigma}_j)$, where $\boldsymbol{\omega}_j = (\omega_{j,t})_{t=1}^T$

$$\begin{aligned} \text{and with } \tilde{\boldsymbol{\theta}}_j &= \boldsymbol{\theta}_j^0 + K_j \left(\mathbf{y}_j - \mathbf{x}^\top \boldsymbol{\theta}_j^0 - \boldsymbol{\omega}_j \lambda \right) \\ \tilde{\Sigma}_j &= (\mathbb{I}_{M+1} - K_j \mathbf{x}) \Sigma_j^0 \\ K_j &= \Sigma_j^0 \mathbf{x}^\top \left(W_j + \mathbf{x} \Sigma_j^0 \mathbf{x}^\top \right)^{-1} \\ W_j &= \text{diag} \left((\omega_{j,t} \times \delta^2 \times \sigma_j)_{t=1}^T \right) \end{aligned}$$

\mathbb{I}_{M+1} denotes the identity matrix of size $(M+1)$.

5. $\pi((\boldsymbol{\theta}_k, \beta)^\top | \mathbf{y}, \mathbf{x}, \boldsymbol{\omega}_k, \sigma_k) \propto \mathcal{N}_{M+2}(\tilde{\boldsymbol{\theta}}_k, \tilde{\Sigma}_k)$, where $\boldsymbol{\omega}_k = (\omega_{k,t})_{t=1}^T$ with

$$\begin{aligned} (\tilde{\boldsymbol{\theta}}_k, \tilde{\beta})^\top &= (\boldsymbol{\theta}_k^0, \beta^0)^\top + K_k \left(\mathbf{y}_k - (\mathbf{x}, \mathbf{y}_j)^\top (\boldsymbol{\theta}_k^0, \beta^0)^\top - \boldsymbol{\omega}_k \lambda \right) \\ \tilde{\Sigma}_k &= (\mathbb{I}_{M+2} - K_k (\mathbf{x}, \mathbf{y}_j)) \begin{pmatrix} \Sigma_k^0 & 0 \\ 0 & \sigma_\beta^2 \end{pmatrix} \\ K_k &= \begin{pmatrix} \Sigma_k^0 & 0 \\ 0 & \sigma_\beta^2 \end{pmatrix} (\mathbf{x}, \mathbf{y}_j)^\top \left(W_k + (\mathbf{x}, \mathbf{y}_j) \begin{pmatrix} \Sigma_k^0 & 0 \\ 0 & \sigma_\beta^2 \end{pmatrix} (\mathbf{x}, \mathbf{y}_j)^\top \right)^{-1} \\ W_k &= \text{diag} \left((\omega_{k,t} \times \delta^2 \times \sigma_k)_{t=1}^T \right) \end{aligned}$$

To initialise the Gibbs sampling algorithm we simulate a random draw from the joint prior distribution of the parameters defined in equation (3.3), and conditional on that, we simulate the initial values of the augmented variables $(\omega_{j,t}, \omega_{k,t})_{t=1}^T$ from their exponential distributions. Updating the parameters in this order ensures that the posterior distribution is the stationary distribution of the generated Markov chain. This is because combining steps 1 and 2 essentially produces draws from the conditional posterior distribution of $\pi(\sigma_j, \sigma_k, \boldsymbol{\omega}_j, \boldsymbol{\omega}_k | \boldsymbol{\theta}_k, \boldsymbol{\theta}_j, \mathbf{y}, \mathbf{x})$.

3.2 VaR and CoVaR posterior estimation

From a Bayesian point of view once we retrieve simulations from the posterior distribution we can choose several ways to summarise them. Lin and Chang [44] used the maximisation of the posterior density to make inference for the quantile regression parameters, showing that this is equivalent to the minimisation problem (1.2) under the frequentist context. Those considerations bring us to consider the maximum a posteriori (MaP) criteria to estimate all the posterior parameters in equations (3.1) and (3.2) assuming ALD distributions for the error terms and diffuse prior on the regressor parameters. From all the MaP parameters involved in the marginal and conditional quantiles, say $(\boldsymbol{\theta}_j^{map}, \boldsymbol{\theta}_k^{map}, \beta^{map})$, the estimators of $\text{VaR}_j^{\mathbf{x}, \tau}$ and the $\text{CoVaR}_{k|j}^{\mathbf{x}, \tau}$ are then derived from equations (2.2) and (2.3) as follows:

$$\begin{aligned} \left(\text{VaR}_j^{\mathbf{x}, \tau} \right)^{map} &= \mathbf{x}^\top \boldsymbol{\theta}_j^{map} \\ \left(\text{CoVaR}_{k|j}^{\mathbf{x}, \tau} \right)^{map} &= \mathbf{x}^\top \boldsymbol{\theta}_k^{map} + \beta^{map} \left(\text{VaR}_j^{\mathbf{x}, \tau} \right)^{map}. \end{aligned}$$

Credible sets at a given confidence level for both $\text{VaR}_{k|j}^{\mathbf{x}, \tau}$ and $\text{CoVaR}_{k|j}^{\mathbf{x}, \tau}$ estimates can be calculated by marginalising out the scale parameters (σ_j, σ_k) and the latent variables

$(\boldsymbol{\omega}_j, \boldsymbol{\omega}_k)$, using the sample draws of the MCMC algorithm. Monte Carlo estimates of the marginal posterior densities of the quantile functions are given by

$$\begin{aligned}\pi\left(\text{VaR}_j^{\mathbf{x},\tau}|\mathbf{y}_j\right) &= \frac{1}{G}\sum_{g=1}^G\pi\left(\mathbf{x}^\top\boldsymbol{\theta}_j|\sigma_j^{(g)},\boldsymbol{\omega}_j^{(g)},\mathbf{y}_j\right), \\ \pi\left(\text{CoVaR}_{k|j}^{\mathbf{x},\tau}|\mathbf{y}_k\right) &= \frac{1}{G}\sum_{g=1}^G\pi\left(\mathbf{x}^\top\boldsymbol{\theta}_k+\beta\left(\text{VaR}_j^{\mathbf{x},\tau}\right)^{(g)}|\sigma_k^{(g)},\boldsymbol{\omega}_k^{(g)},\mathbf{y}_k\right),\end{aligned}$$

$\forall g = 1, 2, \dots, G$, where G denotes the number of post burn-in iterations. The 95% High Posterior Credible intervals $\text{HPD}_{95\%}$ for the τ -th quantile can be obtained from the samples $\left\{\mathbf{x}^\top\boldsymbol{\theta}_j|\sigma_j^{(g)},\boldsymbol{\omega}_j^{(g)},\mathbf{y}_j\right\}_{g=1}^G$ and $\left\{\mathbf{x}^\top\boldsymbol{\theta}_k+\beta\left(\text{VaR}_j^{\mathbf{x},\tau}\right)^{(g)}|\sigma_k^{(g)},\boldsymbol{\omega}_k^{(g)},\mathbf{y}_k\right\}_{g=1}^G$.

4 Time-varying quantile model

As mentioned before VaR and CoVaR are respectively unconditional and conditional quantiles, given current information, of future portfolio values. It is typically the case that returns changes over time and for this reason it can be interesting to build suitable model for time varying VaR and CoVaR. In particular, when modeling time varying quantiles, it is important to link future tail behaviours of time series to its past movements to take into account for risk management arguments. Recently, time varying quantile topic has been received increased attention and different econometric models have been proposed: the most known are the Conditional Autoregressive Valute-at-Risk (CAViaR) model of Engle and Manganelli [19], the Quantile Autoregressive (QAR) model of Koenker and Xiao [34], and the Dynamic Additive Quantile (DAQ) model of Gouriéroux and Jasak [27]. Most of them introduce an autoregressive structure in their modeling, which is intuitively attractive, as series of financial returns tend to exhibit time varying conditional moments, fat tails and volatility clustering. More recently, Gerlach *et al.* [23] deal with the problem of estimating the conditional dynamic VaR using a Bayesian approach. The resulting conditional quantile for the variable of interest is directly modeled as a smooth function of the observed past returns.

In this paper we propose a different approach to introduce dynamics in the quantiles, modelling both the VaR and CoVaR as a function of latent variables having their own time dependence. In particular we estimate, from a Bayesian point of view, the required quantiles simultaneously and we allow the quantiles depending on exogenous variables. In so doing we have in mind a generalisation of De Rossi and Harvey [16] and Kurose and Omori [42] results who proposed to model the unconditional quantile curve using smoothing spline interpolation. More precisely, we model the observed vector at each point in time $(y_{j,t}, y_{k,t})$, as a function of independent latent processes $(\mu_{j,t}, \mu_{k,t})$ and the regressor terms in the following way: $\forall t \in 1, \dots, T$,

$$y_{j,t} = \mu_{j,t} + \mathbf{x}_t^\top \boldsymbol{\theta}_j + \epsilon_{j,t} \quad (4.1)$$

$$y_{k,t} = \mu_{k,t} + \mathbf{x}_t^\top \boldsymbol{\theta}_k + \beta_t y_{j,t} + \epsilon_{k,t}, \quad (4.2)$$

where $\epsilon_{j,t} \sim \text{ALD}(\tau, 0, \sigma_j)$, $\epsilon_{k,t} \sim \text{ALD}(\tau, 0, \sigma_k)$ are independent random variables. The intercept terms $\mu_{l,t}$ with $l \in \{j, k\}$ are introduced to account for time dependence in the

quantile functions. In fact, for $l \in \{j, k\}$, we propose the following smooth time-varying dynamics for $\mu_{l,t}$:

$$\mu_{l,t+1} = \mu_{l,t} + \mu_{l,t}^* + \eta_{l,t} \quad (4.3)$$

$$\mu_{l,t+1}^* = \mu_{l,t}^* + \eta_{l,t}^*, \quad (4.4)$$

where $(\eta_{l,t}, \eta_{l,t}^*)^\top \sim \mathcal{N}_2(0, S_l)$ and $S_l = s_l^2 V = s_l^2 \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$, with $s_l > 0$ allows a certain degree of smoothness of the quantile process. Since one of our main focuses is to analyze the dynamic co-movement of two institutions, we also allow the parameter β_t to change over time. To reflect different impacts between institutions we consider the following evolution for β_t :

$$\beta_{t+1} = \beta_t + \beta_t^* + \eta_{\beta,t} \quad (4.5)$$

$$\beta_{t+1}^* = \beta_t^* + \eta_{\beta,t}^*, \quad (4.6)$$

where $(\eta_{\beta,t}, \eta_{\beta,t}^*)^\top \sim \mathcal{N}_2(0, s_\beta^2 V)$. Throughout the paper we assume that $\forall l \in \{j, k, \beta\}$, $(\eta_{l,t}, \eta_{l,t}^*)$ is independent of $(\epsilon_{j,t}, \epsilon_{k,t})$ (here we use β as an index since there is no ambiguity).

In order to estimate the model we rewrite equations (4.1)-(4.6) using a state space representation so that $\forall t \in 1, \dots, T$,

$$\mathbf{y}_t = Z_t \boldsymbol{\xi}_t + \mathbf{x}_t^\top \boldsymbol{\theta} + \boldsymbol{\epsilon}_t \quad (4.7)$$

$$\boldsymbol{\xi}_{t+1} = A \boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \quad (4.8)$$

where

- $\boldsymbol{\epsilon}_t = (\epsilon_{j,t}, \epsilon_{k,t})$ is the vector of independent ALDs as defined in equations (3.1)–(3.2),
- $Z_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & y_{j,t} & 0 \end{pmatrix}$ is the time varying matrix of loading factors,
- $\boldsymbol{\xi}_t = (\mu_{j,t}, \mu_{j,t}^*, \mu_{k,t}, \mu_{k,t}^*, \beta_t, \beta_t^*)^\top$ is the vector of latent states whose dynamic is given by the transition matrix A , with $A = \mathbb{I}_3 \otimes B$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$; and \otimes denotes the kronecker product,
- $\boldsymbol{\theta} = (\boldsymbol{\theta}_j, \boldsymbol{\theta}_k)$ is a $(M \times 2)$ matrix of time invariant coefficients,
- $\boldsymbol{\eta}_t = (\eta_{j,t}, \eta_{j,t}^*, \eta_{k,t}, \eta_{k,t}^*, \eta_{\beta,t}, \eta_{\beta,t}^*)^\top$ is the time-varying error vector distributed according to $\mathcal{N}_6(0, \Omega)$, where $\Omega = \text{diag}(s_j^2, s_k^2, s_\beta^2) \otimes V$.

To complete the Bayesian model specification we choose as prior distribution for the non-varying parameters $\boldsymbol{\theta}$ the same as in Section 3; the parameters in the variance-covariance matrices of all latent processes are distributed according to $\mathcal{IG}(r_l^0, v_l^0)$ with positive r_l^0 and $v_l^0 \forall l \in \{j, k, \beta\}$. In addition we assume that the vector of first states $\boldsymbol{\xi}_1$ is distributed according to $\mathcal{N}_6(0, \kappa \mathbb{I}_6)$, with $\kappa > 0$ sufficiently large.

The linear state space model introduced in (4.7)-(4.8) for modeling time-varying conditional quantiles is non-Gaussian because of the assumption made on the innovation terms.

So in those circumstances optimal filtering techniques used to analytically marginalize out the latent states based on the Kalman filter recursions can not be applied (see Durbin and Koopman, [17]). Considering the (3.4) representation of the innovation terms in (4.7) it easy to recognise that the non-Gaussian state space model admits a conditionally Gaussian representation. More specifically equations (4.7) and (4.8) become: $\forall t \in 1, \dots, T$,

$$\mathbf{y}_t = \mathbf{c}_t + Z_t \boldsymbol{\xi}_t + \mathbf{x}_t^\top \boldsymbol{\theta} + G_t \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim \mathcal{N}_2(0, \mathbb{I}_2) \quad (4.9)$$

$$\boldsymbol{\xi}_{t+1} = A \boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}(0, \Omega) \quad (4.10)$$

where the time-varying vector \mathbf{c}_t , and matrix G_t are respectively $\mathbf{c}_t = (\lambda \omega_{j,t}, \lambda \omega_{k,t})^\top$ and $G_t = \begin{pmatrix} \delta \sqrt{\sigma_j \omega_{j,t}} & 0 \\ 0 & \delta \sqrt{\sigma_k \omega_{k,t}} \end{pmatrix}$; $\omega_{j,t}$ and $\omega_{k,t}$ are independent with $\omega_{l,t} \sim \mathbb{E} \text{xp}(\sigma_l^{-1})$ for $l \in (j, k)$ and $\sigma_l > 0$; λ and δ are defined in equation (3.5).

The complete-data likelihood of the unobservable components $(\boldsymbol{\xi}_t)_{t=1}^T$ and $\boldsymbol{\omega} = (\omega_{j,t}, \omega_{k,t})_{t=1}^T$ and all parameters $\boldsymbol{\gamma} = (\boldsymbol{\theta}, s_j^2, s_k^2, s_\beta^2, \sigma_j, \sigma_k)$ can be factorized as follows

$$\begin{aligned} & \mathcal{L}((\boldsymbol{\xi}_t)_{t=1}^T, \boldsymbol{\omega}, \boldsymbol{\gamma} | \mathbf{y}, \mathbf{x}) \\ & \propto \prod_{t=1}^T f(y_{j,t} | \boldsymbol{\xi}_t, \omega_{j,t}, \sigma_j, \mathbf{x}_t) \times \prod_{t=1}^T f(y_{k,t} | y_{j,t}, \boldsymbol{\xi}_t, \omega_{k,t}, \sigma_k, \mathbf{x}_t) \\ & \times \prod_{t=1}^T f(\omega_{j,t} | \sigma_j) \times \prod_{t=1}^T f(\omega_{k,t} | \sigma_k) \times f(\boldsymbol{\xi}_1) \prod_{t=1}^{T-1} f(\boldsymbol{\xi}_{t+1} | \boldsymbol{\xi}_t, s_j^2, s_k^2, s_\beta^2) \\ & \propto \prod_{t=1}^T (\sigma_j \times \sigma_k \times \omega_{j,t} \times \omega_{k,t})^{-\frac{1}{2}} \\ & \times \exp \left\{ -\frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \mathbf{c}_t - Z_t \boldsymbol{\xi}_t - \mathbf{x}_t^\top \boldsymbol{\theta})^\top (G_t G_t^\top)^{-1} (\mathbf{y}_t - \mathbf{c}_t - Z_t \boldsymbol{\xi}_t - \mathbf{x}_t^\top \boldsymbol{\theta}) \right\} \\ & \times (\sigma_j)^{-T} \exp \left\{ -\frac{\sum_{t=1}^T \omega_{j,t}}{\sigma_j} \right\} \times (\sigma_k)^{-T} \exp \left\{ -\frac{\sum_{t=1}^T \omega_{k,t}}{\sigma_k} \right\} \\ & \times \exp \left\{ -\frac{1}{2\kappa} \boldsymbol{\xi}_1^\top \boldsymbol{\xi}_1 \right\} \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T-1} (\boldsymbol{\xi}_{t+1} - A \boldsymbol{\xi}_t)^\top \Omega^{-1} (\boldsymbol{\xi}_{t+1} - A \boldsymbol{\xi}_t) \right\}. \quad (4.11) \end{aligned}$$

4.1 Computations

Using the complete-data likelihood in (4.11) and the prior distributions stated in Sections 3 and 4 we are able to write the joint posterior distribution of the parameters and the unobservable components. The form of the posterior allows us to sample from the complete conditional distributions and to use the Gibbs sampler algorithm, as shown below.

After choosing a set of initial values for the parameter vector $\boldsymbol{\gamma}^{(0)}$, simulations from the posterior distribution at the i -th iteration of $\boldsymbol{\gamma}^{(i)}$, $\{\boldsymbol{\xi}_t, t = 1, 2, \dots, T\}^{(i)}$ and $\{\omega_{l,t}, l \in j, k, t = 1, 2, \dots, T\}^{(i)}$ for $i = 1, 2, \dots$, are obtained by the following Gibbs sampling scheme:

1. For $l \in \{j, k, \beta\}$, generate s_l^2 from $\mathcal{IG}(\tilde{r}_l, \tilde{v}_l)$ with parameters

$$\tilde{r}_l = r_l^0 + (T - 1), \quad \tilde{v}_l = v_l^0 + \sum_{t=1}^{T-1} (\boldsymbol{\xi}_{t+1}^l - B \boldsymbol{\xi}_t^l)^\top V^{-1} (\boldsymbol{\xi}_{t+1}^l - B \boldsymbol{\xi}_t^l)$$

where $\boldsymbol{\xi}_t^l$ denotes the vector consisting of the elements $(\mu_{j,t}, \mu_{j,t}^*)$ for $l = j$, $(\mu_{k,t}, \mu_{k,t}^*)$ for $l = k$, and (β_t, β_t^*) for $l = \beta$.

2. For $l = j, k$, generate σ_l from $\mathcal{IG}(\tilde{a}_l, \tilde{b}_l)$ with parameters

$$\tilde{a}_l = a_l^0 + T, \quad \tilde{b}_l = b_l^0 + \sum_{t=1}^T \rho_\tau(y_{l,t} - z_{l,t} \boldsymbol{\xi}_t)$$

where $z_{l,t}$ denotes the first row of the matrix Z_t for $l = j$ or the second one for $l = k$.

3. For $l = j, k$, denote $\mathbf{y}_l = (y_{l,t})_{t=1}^T$, $\mathbf{z}_l \boldsymbol{\mu}_l = (z_{l,t} \times \mu_{j,t})_{t=1}^T$, $\boldsymbol{\beta} \mathbf{y}_j = (\beta_t \times y_{j,t})_{t=1}^T$. Then, generate $\boldsymbol{\theta}_l \sim \mathcal{N}_M(\tilde{\boldsymbol{\theta}}_l, \tilde{\Sigma}_l)$ with parameters

$$\begin{aligned} \tilde{\boldsymbol{\theta}}_j &= \boldsymbol{\theta}_j^0 + K_j (\mathbf{y}_j - z_j \boldsymbol{\mu}_j - \mathbf{x}^\top \boldsymbol{\theta}_j^0 - \boldsymbol{\omega}_j \lambda) \\ \tilde{\Sigma}_j &= (\mathbb{I}_M - K_j \mathbf{x}) \Sigma_j^0 \\ K_j &= \Sigma_j^0 \mathbf{x}^\top (W_j + \mathbf{x} \Sigma_j^0 \mathbf{x}^\top)^{-1} \\ W_j &= \text{diag}((\omega_{j,t} \times \delta^2 \times \sigma_j)_{t=1}^T) \end{aligned}$$

and

$$\begin{aligned} \tilde{\boldsymbol{\theta}}_k &= \boldsymbol{\theta}_k^0 + K_k (\mathbf{y}_k - z_k \boldsymbol{\mu}_k - \boldsymbol{\beta} \mathbf{y}_j - \mathbf{x}^\top \boldsymbol{\theta}_k^0 - \boldsymbol{\omega}_k \lambda) \\ \tilde{\Sigma}_k &= (\mathbb{I}_M - K_k \mathbf{x}) \Sigma_k^0 \\ K_k &= \Sigma_k^0 \mathbf{x}^\top (W_k + \mathbf{x} \Sigma_k^0 \mathbf{x}^\top)^{-1} \\ W_k &= \text{diag}((\omega_{k,t} \times \delta^2 \times \sigma_k)_{t=1}^T) \end{aligned}$$

where Σ_j^0 and Σ_k^0 are square matrices of dimension M .

4. For $l = j, k$ and for all $t \in \{1, 2, \dots, T\}$, generate $\omega_{l,t}^{-1} \sim \mathcal{LN}(\mu_{l,t}, \gamma_{l,t})$, with parameters

$$\psi_{j,t} = \sqrt{\frac{\lambda^2 + 2\delta^2}{(y_{j,t} - z_{j,t} \mu_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j)^2}}, \quad \phi_{j,t} = \frac{\lambda^2 + 2\delta^2}{\delta^2 \sigma_j}$$

and

$$\psi_{k,t} = \sqrt{\frac{\lambda^2 + 2\delta^2}{(y_{k,t} - \beta_t y_{j,t} - z_{k,t} \mu_{k,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_k)^2}}, \quad \phi_{k,t} = \frac{\lambda^2 + 2\delta^2}{\delta^2 \sigma_k}.$$

5. Since conditionally on the augmented latent states $(\omega_{j,t}, \omega_{k,t})_{t=1}^T$, the state space model defined in equations (4.9)–(4.10), is linear and Gaussian, the latent dynamics can be marginalized out by running the Kalman filter-smoothing algorithm. We draw $(\boldsymbol{\xi}_t, \beta_t)_{t=1}^T$ jointly using the multi-move simulation smoother of Durbin and Koopman [18]. This entails running a Kalman filter forward with the state equation defined as in (4.10). As in Johannes and Polson [30], equation (4.2) for $y_{k,t}$ is a measurement equation with time-varying coefficients, because $y_{j,t}$ is known and represents a time varying factor loading. Once the Kalman filter is run forward, we run the Kalman

smoother backward in order to get the moments of joint full conditional distribution of the latent states (4.11). Finally, we simulate a sample path by drawing from this joint distribution. For a similar simulation algorithm based on forward-filtering backward-smoothing see also Carter and Kohn [11], [12] and Fruhwirth-Schnatter [20].

4.2 Maximum a Posteriori

As in the time-invariant case, once we retrieve simulations from the posterior distribution, in order to make posterior inference, we use the maximum a posteriori summarising criteria. In what follows we prove that using this criteria, the estimated quantiles have good sample properties according to De Rossi and Harvey [16] Proposition 3, i.e. a generalization of the "fundamental property", that is the sample quantile has the appropriate number of observations above and below

Proposition 4.1 *For the state space model defined in equations (4.1)-(4.6) with prior distributions specified in Sections 3 and 4, κ large enough and a diffuse prior on θ , the MaP quantile estimates $\mu_{j,t}^{map} + \mathbf{x}_t^\top \theta_j^{map}$ and $\mu_{k,t}^{map} + \mathbf{x}_t^\top \theta_k^{map} + y_{j,t} \beta_t^{map}$ satisfy*

$$\sum_{t \notin C} (x_{m,t} + 1) \chi_\tau(y_{j,t} - (\mu_{j,t}^{map} + \mathbf{x}_t^\top \theta_j^{map})) = 0, \quad \forall m \in \{1, \dots, M\}$$

$$\sum_{t \notin C} (y_{j,t} + x_{m,t} + 1) \chi_\tau(y_{k,t} - (\mu_{k,t}^{map} + \mathbf{x}_t^\top \theta_k^{map} + y_{j,t} \beta_t^{map})) = 0, \quad \forall m \in \{1, \dots, M\}$$

where $C \subset \{1, \dots, T\}$ is the set of indices through which the Map quantile estimate passes and

$$\chi_\tau : z \rightarrow \begin{cases} \tau - 1 & \text{if } z < 0 \\ \tau & \text{if } z > 0 \end{cases}$$

Corollary 4.1 *The MaP quantiles estimates $(q_t^r(\mathbf{x}_t))^{map} = \mu_{j,t}^{map} + \mathbf{x}_t^\top \theta_j^{map}$ and $(q_t^r(\mathbf{x}_t, y_{j,t}))^{map} = \mu_{k,t}^{map} + \mathbf{x}_t^\top \theta_k^{map} + y_{j,t} \beta_t^{map}$ satisfies a generalization of the fundamental property of sample time-varying quantiles, that is:*

$$\sum_{t \in A} h_t \leq (1 - \tau) \sum_{t=1}^T h_t - \sum_{t \in C^-} h_t \quad \text{and} \quad \sum_{t \in B} h_t \leq \tau \sum_{t=1}^T h_t - \sum_{t \in C^-} h_t, \quad (4.12)$$

where $h_t = \begin{cases} x_{m,t} + 1 & \text{for the VaR} \\ y_{j,t} + x_{m,t} + 1 & \text{for the CoVaR} \end{cases}$, and $A \cup B \cup C = \{1, \dots, T\}$ such that A denotes the set of indices for which the observations are (strictly) above the MaP quantiles estimate, B is the index set for which the observations are (strictly) below the corresponding MaP quantiles estimates, $C = C^+ \cup C^- = \{t \in C : h_t \geq 0\} \cup \{t \in C : h_t < 0\}$ is the index set for which the observations coincide with the quantile estimator.

Remark 4.1 *If $h_t > 0 \quad \forall t$, then inequalities in (4.12) coincide with the one stated in De Rossi and Harvey [16]. Moreover, if the distribution of $(Y_{j,t}, Y_{k,t})$ is continuous, then C is a.s. the empty set and hence, the generalised property stated in (4.12) is the one of De Rossi and Harvey [16].*

When $h_t \equiv h \quad \forall t$, then (4.12) correspond to the fundamental property of sample time-varying quantiles.

Proof of Corollary 4.1 Following the proof of Proposition 3 in De Rossi and Harvey [16], using Proposition 4.1 and

$$-\tau \sum_{t \in C^+} h_t + (1-\tau) \sum_{t \in C^-} h_t < \sum_{t \notin C} h_t \chi_\tau(y_{j,t} - (q_t^\tau(\mathbf{x}_t))^{map}) < (1-\tau) \sum_{t \in C^+} h_t - \tau \sum_{t \in C^-} h_t, \quad (4.13)$$

where $h_t = x_{m,t} + 1$, the result in (4.1) is obtained by rewriting the central term in the inequality (4.13) as follows:

$$\sum_{t \notin C} h_t \chi_\tau(y_{j,t} - (q_t^\tau(\mathbf{x}_t))^{map}) = \tau \sum_{t \in A} h_t + (\tau - 1) \sum_{t \in B} h_t$$

The same occurs for the CoVaR.

Proof of Proposition 4.1 . For $t \in \{1, \dots, T\}$ recall that $\boldsymbol{\xi}_t^j = (\mu_{j,t}, \mu_{j,t}^*)^\top$ and define $\boldsymbol{\xi}_t^{k,\beta} = (\mu_{k,t}, \mu_{k,t}^*, \beta_t, \beta_t^*)^\top$, $\boldsymbol{\xi}^j = (\boldsymbol{\xi}_t^j)_{t=1}^T$ and $\boldsymbol{\xi}^{k,\beta} = (\boldsymbol{\xi}_t^{k,\beta})_{t=1}^T$.

From equations (4.1)-(4.6), let us write the complete-posterior distribution $p(\boldsymbol{\theta}, \boldsymbol{\xi}^j, \boldsymbol{\xi}^{k,\beta}, s_j^2, s_k^2, s_\beta^2, \sigma_j, \sigma_k | (\mathbf{y}_t)_{t=1}^T)$ as proportional to the product of two parts $Post_{marg}$ and $Post_{cond}$ where

$$\begin{aligned} Post_{marg} &= \prod_{t=1}^T ald(y_{j,t} | \mu_{j,t} + \mathbf{x}_t^\top \boldsymbol{\theta}_j, \sigma_j) \\ &\quad \times \frac{1}{s_j} \exp \left\{ -\frac{1}{2\kappa} (\boldsymbol{\xi}_1^j)^\top \boldsymbol{\xi}_1^j \right\} \exp \left\{ -\frac{12}{2s_j^2} \sum_{t=1}^{T-1} (\boldsymbol{\xi}_{t+1}^j - B\boldsymbol{\xi}_t^j)^\top V^{-1} (\boldsymbol{\xi}_{t+1}^j - B\boldsymbol{\xi}_t^j) \right\} \\ &\quad \times \exp \left(-\frac{1}{2} (\boldsymbol{\theta}_j - \boldsymbol{\theta}_j^0)^\top (\Sigma_j^0)^{-1} (\boldsymbol{\theta}_j - \boldsymbol{\theta}_j^0) \right) \times \mathcal{IG}(r_j^0, v_j^0) \times \mathcal{IG}(a_j^0, b_j^0) \\ Post_{cond} &= \prod_{t=1}^T ald(y_{k,t} | \mu_{k,t} + \mathbf{x}_t^\top \boldsymbol{\theta}_k + \beta_t y_{j,t}, \sigma_k) \\ &\quad \times \exp \left\{ -\frac{1}{2\kappa} (\boldsymbol{\xi}_1^{k,\beta})^\top \boldsymbol{\xi}_1^{k,\beta} \right\} \\ &\quad \times \frac{1}{s_k s_\beta} \exp \left\{ -\frac{12}{2} \sum_{t=1}^{T-1} (\boldsymbol{\xi}_{t+1}^{k,\beta} - (\mathbb{I}_2 \otimes B)\boldsymbol{\xi}_t^{k,\beta})^\top [diag(s_k^2, s_\beta^2) \otimes V]^{-1} (\boldsymbol{\xi}_{t+1}^{k,\beta} - (\mathbb{I}_2 \otimes B)\boldsymbol{\xi}_t^{k,\beta}) \right\} \\ &\quad \times \exp \left(-\frac{1}{2} (\boldsymbol{\theta}_k - \boldsymbol{\theta}_k^0)^\top (\Sigma_k^0)^{-1} (\boldsymbol{\theta}_k - \boldsymbol{\theta}_k^0) \right) \times \mathcal{IG}(r_k^0, v_k^0) \times \mathcal{IG}(a_k^0, b_k^0) \times \mathcal{IG}(r_\beta^0, v_\beta^0). \end{aligned}$$

Note that determining the MaP of $p(\boldsymbol{\theta}, \boldsymbol{\xi}^j, \boldsymbol{\xi}^{k,\beta}, s_j^2, s_k^2, s_\beta^2, \sigma_j, \sigma_k | (\mathbf{y}_t)_{t=1}^T)$ in $(\boldsymbol{\theta}, (\mu_{j,t})_{t=1}^T, (\mu_{k,t})_{t=1}^T, (\beta_t)_{t=1}^T)$ consists in maximizing separately $Post_{marg}$ and $Post_{cond}$ with respect to $(\boldsymbol{\theta}_j, (\mu_{j,t})_{t=1}^T)$, and $(\boldsymbol{\theta}_k, (\mu_{k,t})_{t=1}^T, (\beta_t)_{t=1}^T)$, respectively.

The check function ρ_τ which enters in the *ald* densities, is derivable everywhere except in zero, therefore its derivative corresponds to the function $\chi_\tau : z \rightarrow \begin{cases} \tau - 1 & \text{if } z < 0 \\ \tau & \text{if } z > 0 \end{cases}$.

When it is possible, differentiating $\log(Post_{marg})$ gives:

$$\begin{aligned}
\frac{\partial \log Post_{marg}}{\partial \mu_{j,1}} &= \frac{-\mu_{j,1} - \mu_{j,1}^*}{\kappa} + \frac{1}{\sigma_j} \chi_\tau(y_{j,1} - \mu_{j,1} - \mathbf{x}_1^\top \boldsymbol{\theta}_j) + \frac{6}{s_j^2} \{2(\mu_{j,2} - \mu_{j,1} - \mu_{j,1}^*) - (\mu_{j,2}^* - \mu_{j,1}^*)\} \\
\frac{\partial \log Post_{marg}}{\partial \mu_{j,T}} &= \frac{1}{\sigma_j} \chi_\tau(y_{j,T} - \mu_{j,T} - \mathbf{x}_T^\top \boldsymbol{\theta}_j) + \frac{6}{s_j^2} \{-2(\mu_{j,T} - \mu_{j,T-1} - \mu_{j,T-1}^*) + (\mu_{j,T}^* - \mu_{j,T-1}^*)\} \\
\frac{\partial \log Post_{marg}}{\partial \mu_{j,t}} &= \frac{1}{\sigma_j} \chi_\tau(y_{j,t} - \mu_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j) + \frac{6}{s_j^2} \{-2(\mu_{j,t} - \mu_{j,t-1} - \mu_{j,t-1}^*) + 2(\mu_{j,t+1} - \mu_{j,t} - \mu_{j,t}^*) \\
&\quad - (\mu_{j,t+1}^* - 2\mu_{j,t}^* + \mu_{j,t-1}^*)\}, \quad \forall t \in \{2, \dots, T-1\}, \\
\frac{\partial \log Post_{marg}}{\partial \boldsymbol{\theta}_j} &= \frac{1}{\sigma_j} \sum_{t=1}^T \mathbf{x}_t \chi_\tau(y_{j,t} - \mu_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j) + (\Sigma_j^0)^{-1} (\boldsymbol{\theta}_j - \boldsymbol{\theta}_j^0)
\end{aligned} \tag{4.14}$$

where we use the fact that $S_j^{-1} = \frac{12}{s_j^2} V^{-1} = \frac{12}{s_j^2} \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1/3 \end{pmatrix}$.

It turns out that

$$\sum_t \frac{\partial \log Post_{marg}}{\partial \mu_{j,t}} = \frac{-\mu_{j,1} - \mu_{j,1}^*}{\kappa} + \frac{1}{\sigma_j} \sum_t \chi_\tau(y_{j,t} - \mu_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j),$$

which combined with (4.14) and choosing κ sufficiently large and a diffuse prior on $\boldsymbol{\theta}_j$ implies that the maximizer of $Post_{marg}$ satisfies

$$\sum_{t \notin C} (x_{m,t} + 1) \chi_\tau(y_{j,t} - \mu_{j,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_j) = 0, \quad \forall m \in \{1, 2, \dots, M\},$$

where C denotes the set of observation indices through which the quantile estimate passes.

The derivatives of $\log(Post_{cond})$ with respect to $(\boldsymbol{\theta}_k, (\mu_{k,t})_{t=1}^T)$ can be obtained as the ones for $\log(Post_{marg})$. When it is possible, deriving with respect to $(\beta_t)_{t=1}^T$ leads to:

$$\begin{aligned}
\frac{\partial \log Post_{cond}}{\partial \beta_1} &= \frac{-\beta_1 - \beta_1^*}{\kappa} + \frac{y_{j,1}}{\sigma_k} \chi_\tau(y_{k,1} - \mu_{k,1} - \mathbf{x}_1^\top \boldsymbol{\theta}_k - y_{j,1} \beta_1) + \\
&\quad \frac{6}{s_\beta^2} \{2(\beta_2 - \beta_1 - \beta_1^*) - (\beta_2^* - \beta_1^*)\} \\
\frac{\partial \log Post_{cond}}{\partial \beta_T} &= \frac{y_{j,T}}{\sigma_k} \chi_\tau(y_{k,T} - \mu_{k,T} - \mathbf{x}_T^\top \boldsymbol{\theta}_k - y_{j,T} \beta_T) + \frac{6}{s_\beta^2} \{-2(\beta_T - \beta_{T-1} - \beta_{T-1}^*) + (\beta_T^* - \beta_{T-1}^*)\} \\
\frac{\partial \log Post_{cond}}{\partial \beta_t} &= \frac{y_{j,t}}{\sigma_k} \chi_\tau(y_{k,t} - \mu_{k,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_k - y_{j,t} \beta_t) + \frac{6}{s_\beta^2} \{-2(\beta_t - \beta_{t-1} - \beta_{t-1}^*) \\
&\quad + 2(\beta_{t+1} - \beta_t - \beta_t^*) - (\beta_{t+1}^* - 2\beta_t^* + \beta_{t-1}^*)\}, \quad \forall t \in \{2, \dots, T-1\}.
\end{aligned}$$

It turns out that

$$\sum_t \frac{\partial \log Post_{cond}}{\partial \beta_t} = \frac{-\beta_1 - \beta_1^*}{\kappa} + \frac{1}{\sigma_k} \sum_t y_{j,t} \chi_\tau(y_{k,t} - \mu_{k,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_k - y_{j,t} \beta_t).$$

Choosing a sufficiently large κ and a diffuse prior on $\boldsymbol{\theta}_k$ implies that the maximizer of $Post_{cond}$ satisfies

$$\sum_{t \notin C} (y_{j,t} + x_{m,t} + 1) \chi_\tau(y_{k,t} - \mu_{k,t} - \mathbf{x}_t^\top \boldsymbol{\theta}_k - y_{j,t} \beta_t) = 0, \quad \forall m \in \{1, 2, \dots, M\}.$$

5 Empirical application

Throughout this section we illustrate the methodology previously discussed to real data. In particular we analyse separately the time-invariant specification of CoVaR proposed in Section 3 and the time-varying version considered in Section 4. The idea is to study the tail co-movements between an individual institution j and the whole system k it belongs to. The financial data we rely on refer to the Standard and Poor's Composite Index (k) for the U.S market where different sectors (j) are included. For those institutions and for the whole system we consider both micro and macro variables in order to take into account for individual information and for global economic conditions respectively. The analysis is based on weekly observations; whereas the data are not observable at the same frequency we build a smoothing state space model to fill the missing values. The focus of the empirical application is to show how CoVaR provides interesting insights about the tail risk interdependence. Moreover, we show the relevance of introducing dynamics in the extreme quantiles in order to effectively capture the contribution of individual institutions to the systemic risk evolution. Approaching VaR and CoVaR estimation in a Bayesian framework allows us to calculate their credible sets which are necessary to assess estimates' accuracy.

5.1 The data

The empirical analysis is based on the publicly traded U.S. companies listed in Table 1 belonging to different sectors of the Standard and Poor's Composite Index (S&P500). The sectors considered are: financials, consumer goods, energy, industrials, technologies and utilities. Financials consist of banks, diversified financial services and consumer financial services. Consumer Goods consist of the food and beverage industry, primary food industry and producers of personal and household goods. The energy sector consists of companies producing or supplying energy and it includes companies involved in the exploration and development of oil or gas reserves, oil and gas drilling, or integrated power firms. Industrials consist of industries such as construction and heavy equipment, as well as industrial goods and services that include containers, packing and industrial transport, while technologies are related to the research, development and/or distribution of technologically based goods and services. Utilities consist of the provision of gas and electricity.

Daily equity price data are converted to weekly log-returns (in percentage points) for the sample period between January 2, 2004 to December 28, 2012, covering the recent global financial crisis. Table 2 provides summary statistics for the weekly returns. Except for few companies, the mean return during the estimation period is positive. Consumer goods and Energy have the highest average return, while Banks and Financial services have the lowest one. Focusing on the sample correlation with the market index return, the correlation involving the financials and industrials is the largest on average. The correlation with the market index return varies substantially across sectors, ranging from 0.553 (PEG) to 0.772 (AXP). Bellwether sectors like consumer and utilities show a surprisingly high correlation level. A possible explanation for this empirical evidence is that the correlation among financial stock increases dramatically during times of turbulence and in particular in late 2008 as the global financial crisis intensified. Finally, the last column of Table 2 provides the sample 1% stress level of each institution's return evaluated over the entire time period. By comparing these values with the number of standard deviations away from the their mean, we can see that asset return distributions do not appear highly

skewed. The characteristic of the data summarised in Table 2 induce to study the risk interdependence through the CoVaR tool.

To control for the general economic conditions we use observations of the following macroeconomic regressors as suggested by Adrian and Brunnermeier [4] and Chao, Härdle and Wang [13]:

- (i) the VIX index (VIX), measuring the model-free implied stock market volatility as evaluated by the Chicago Board Options Exchange (CBOE).
- (ii) a short term liquidity spread (LIQSPR), computed as the difference between the 3-months collateral repo rate and the 3-months Treasury Bill rate.
- (iii) the weekly change in the three-month Treasury Bill rate (3MTB).
- (iv) the change in the slope of the yield curve (TERMSPR), measured by the difference of the 10-Years Treasury rate and the 3-months Treasury Bill rate.
- (v) the change in the credit spread (CREDSPR) between 10-Years BAA rated bonds and the 10-Years Treasury rate.
- (vi) the weekly return of the Dow Jones US Real Estate Index (DJUSRE).

Historical data for the volatility index (VIX) can be downloaded from the Chicago Board Options Exchange’s website, while the remaining variables are from the Federal Reserve Board H.15 database. Data are available on a daily frequency and subsequently converted to a weekly frequency.

To capture the individual firms’ characteristics, we include observations from the following microeconomic regressors:

- (i) leverage (LEV), calculated as the value of total assets divided by total equity (both measured in book values).
- (ii) the market to book value (MK2BK), defined as the ratio of the market value to the book value of total equity.
- (iii) the size (SIZE), defined by the logarithmic transformation of the market value of total assets.
- (iv) the maturity mismatch (MM), calculated as short term debt net of cash divided by the total liabilities.

Microeconomic variables are downloaded from the Bloomberg database and are available only on a quarterly basis. Since our analysis builds on weekly frequencies we choose to impute missing observations by smoothing spline interpolation. Details on the procedure are given in the Appendix A.

5.2 Time-invariant risk beta

In what follows we provide the Bayesian empirical analysis for the time-invariant CoVaR model stated in Section 3. In order to implement the inference we specify the hyperparameters values for each prior distribution defined therein. We set the first moment of the exogenous regressor’s parameters θ equal to zero and let their variance-covariance matrices to be diagonal, i.e. $\Sigma_l^0 = 100 \times \mathbb{1}_{11}$, for $l = j, k$. The β parameter is assumed to

Name	Ticker Symbol	Sector
CITIGROUP INC.	C	Financial
BANK OF AMERICA CORP.	BAC	Financial
COMERICA INC.	CMA	Financial
JPMORGAN CHASE & CO.	JPM	Financial
KEYCORP	KEY	Financial
GOLDMAN SACHS GROUP INC.	GS	Financial
MORGAN STANLEY	MS	Financial
MOODY'S CORP.	MCO	Financial
AMERICAN EXPRESS CO.	AXP	Financial
MCDONALD'S CORP.	MCD	Consumer
NIKE INC.	NKE	Consumer
CHEVRON CORP.	CVX	Energy
EXXON MOBIL CORP.	XOM	Energy
BOEING CO.	BA	Industrial
GENERAL ELECTRIC CO.	GE	Industrial
INTEL CORP.	INTC	Technology
ORACLE CORP.	ORCL	Technology
AMEREN CORPORATION.	AEE	Utilities
PUBLIC SERVICE ENTERPRISE INC.	PEG	Utilities

TABLE 1: List of included companies. All the listed companies belong to the Standard and Poor's Composite Index (S&P500) at the start of the trading day of February 15, 2013. The last column reports the sector each company belongs to.

have a Gaussian prior centered on $\beta^0 = 0$ with large variance, $\sigma_\beta^2 = 100$. Vague prior for the nuisance parameters σ_l , for $l = j, k$ are imposed by setting $a_l^0 = b_l^0 = 0.0001$ which correspond to an Inverse Gamma distribution with infinite variance.

The MCMC algorithm illustrated in Section 3 run for 200,000 times with a burn-in phase of 100,000 iterations. For all the considered institutions, Tables 4–5 report the estimated systemic risk β and the exogenous parameters as well as the HPD_{95%}, for $\tau = (0.025, 0.05)$ credible sets. To check the MCMC convergence we also calculate the Geweke's convergence diagnostics (see Geweke [24], [25]) which are not reported to save space but confirm the convergence of the chain.

For all the reported institutions the β 's parameters are positive and significantly different from zero. Note that a positive β indicates that a decrease in $\text{VaR}_j^{\mathbf{x}, \tau}$ (expressed as a larger negative value) yields a greater negative $\text{CoVaR}_{k|j}^{\mathbf{x}, \tau}$, i.e. a higher risk of system losses. Moreover, by comparing HPD_{95%} for the β parameters, it is also evident that the extent of the systemic risk contribution is significantly different across institutions belonging to different sectors. Hence the result highlights the empirical evidence of a sector specific effect of individual losses to the overall systemic risk. We also observe that on average the systemic β has lower value for institutions belonging to the financial sector and is higher for institutions belonging to consumer and energy sectors. This evidence gives the idea of existence of sectors having different sensitivity to the risk exposure. In addition, it is worth noting that sometimes the β parameter displays a huge variation also within the same sector. It is the case for example for the industrial one, where the estimated β coefficient for GE is significantly different from the one of BA whose credible sets are not overlapping. Finally, comparing the β s for the two different values of τ con-

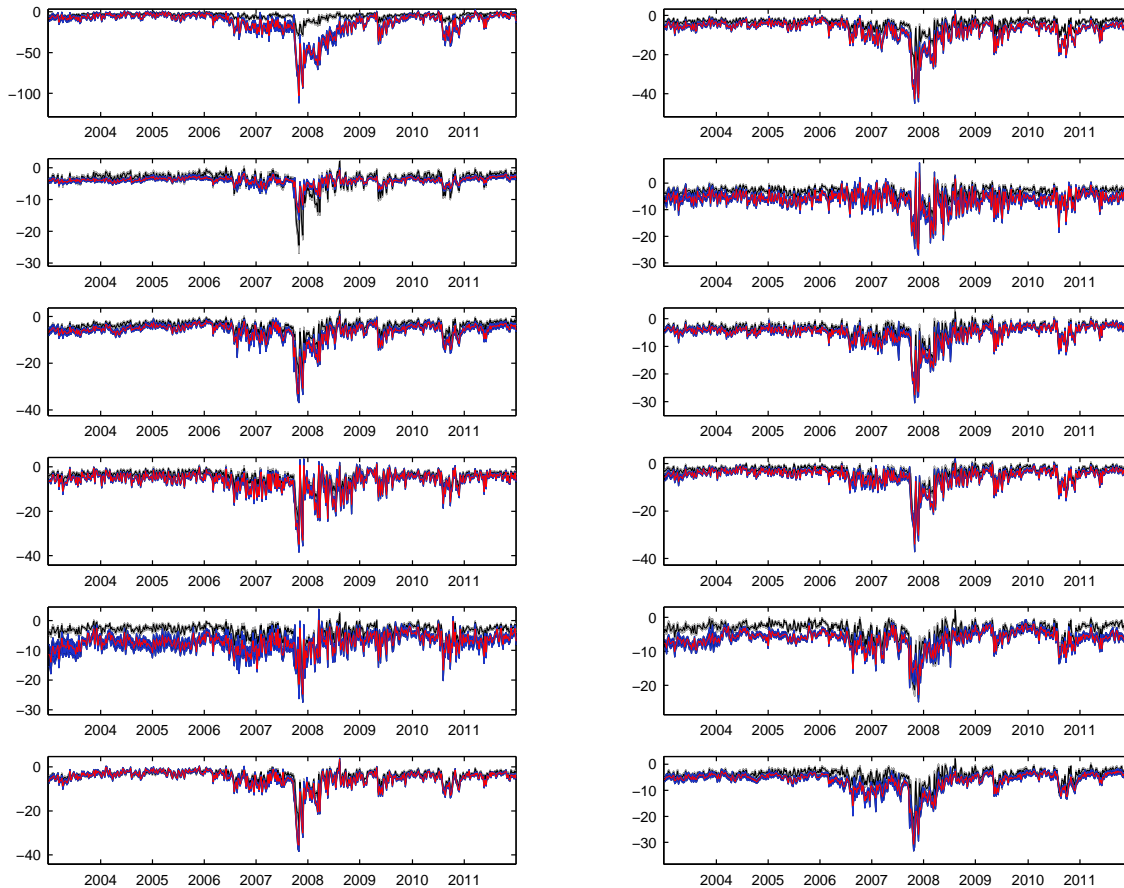


FIGURE 1: Time series plot of the $\text{VaR}_j^{\mathbf{x},\tau}$ (red line) and $\text{CoVaR}_{k|j}^{\mathbf{x},\tau}$ (gray line) at the confidence level $\tau = 0.025$ along with $\text{HPD}_{95\%}$ credible sets, obtained by fitting the model defined in equations (3.1)–(3.2) for the following assets: first panel (financial): C (left), GS (right); second panel (consumer): MCD (left) and NKE (right); third panel (energy): CVX (left), XOM (right); fourth panel (industrial): BA (left), GE (right); fifth panel (technology): INTC (left), ORCL (right); last panel (utilities): AEE (left), PEG (right).

sidered we observe that on average higher values of the parameter tend to be associated with smaller values of the confidence level τ , meaning that the co-movement between asset and market is stronger for extreme returns.

We now concentrate our attention to the influence of macroeconomic variables. From Tables 4–5 we detect some remarkable differences among assets and in particular we observe that:

- except for the VIX and the US Real Estate indices, the impact of the remaining variables change in magnitude and significance as we move from one asset to another. This heterogeneous behaviour seems to be transversal with respect to sectors at least in some cases such as liquidity spread (LIQSPRD). As expected, the VIX index estimated parameters are always significantly negative while the opposite happens for the US Real Estate Index. This is true for both the VaR and CoVaR regressions and for all the considered τ .
- some macroeconomic variables, such as the changes in three-months Treasury Bill

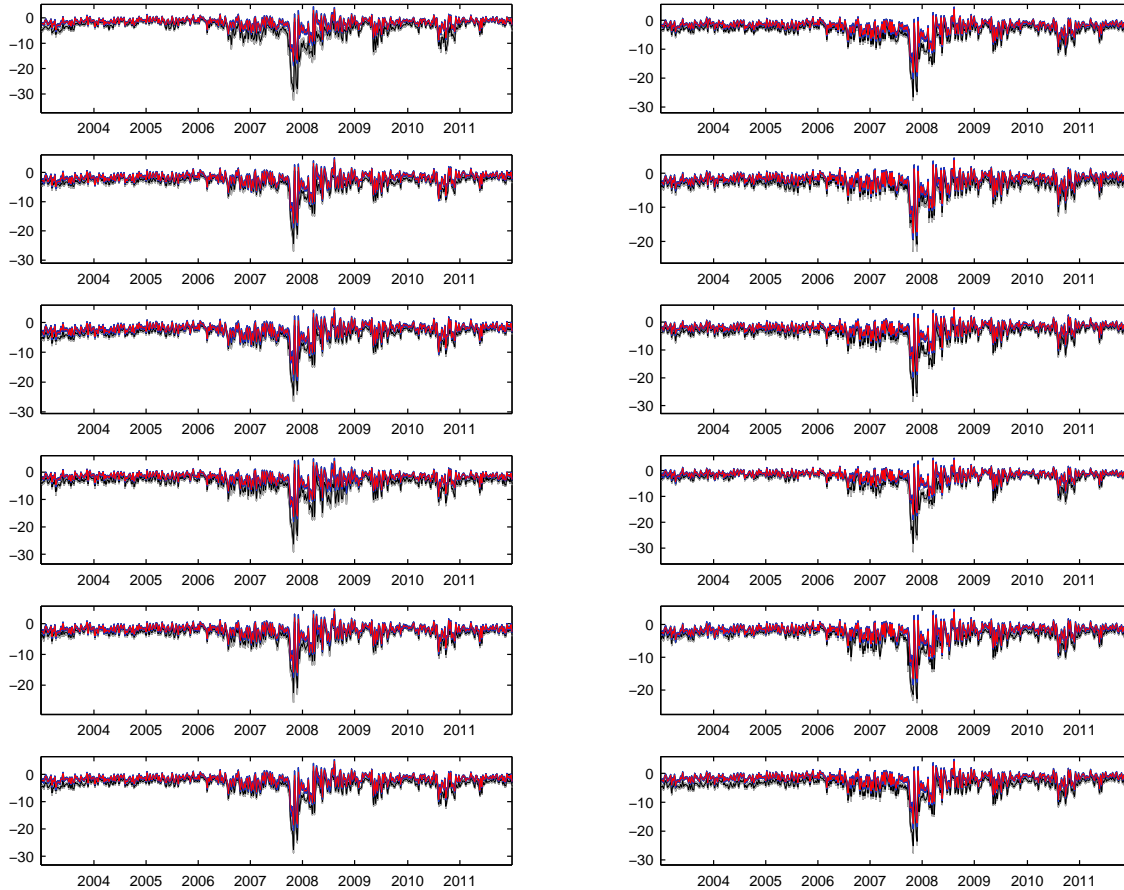


FIGURE 2: Time series plot of the $\text{CoVaR}_{k|j}^{x,\tau}$ at the confidence levels $\tau = 0.025$ (gray line) and $\tau = 0.1$ (red line) along with $\text{HPD}_{95\%}$ credible sets, obtained by fitting the model defined in equations (3.1)–(3.2) for the following assets: first panel (financial): C (left), GS (right); second panel (consumer): MCD (left) and NKE (right); third panel (energy): CVX (left), XOM (right); fourth panel (industrial): BA (left), GE (right); fifth panel (technology): INTC (left), ORCL (right); last panel (utilities): AEE (left), PEG (right).

rate (3MTB) or the term spread (TERMSPR) display a different impact on the CoVaR and VaR being always positive or not significant in the first case and also negative for some sectors (financial and utilities) in the latter case. This means that an increase in the spread between ten years Treasury Bond rates and three-months Treasury Bill rates (CREDSPR) produces a decrease of the CoVaR while reducing individual risks in the case of financial and utilities sectors. This means that in principle large traded firms benefit from an increased interest rate because it increases the opportunity cost of different financing strategies. As expected, the change in credit spread has a negative impact on the firms' riskiness.

- for different values of the confidence level τ the macroeconomic variables show a different impact on the marginal and conditional quantiles becoming in general less significant as the τ -level increases.

For the micro exogenous regressors, we note that:

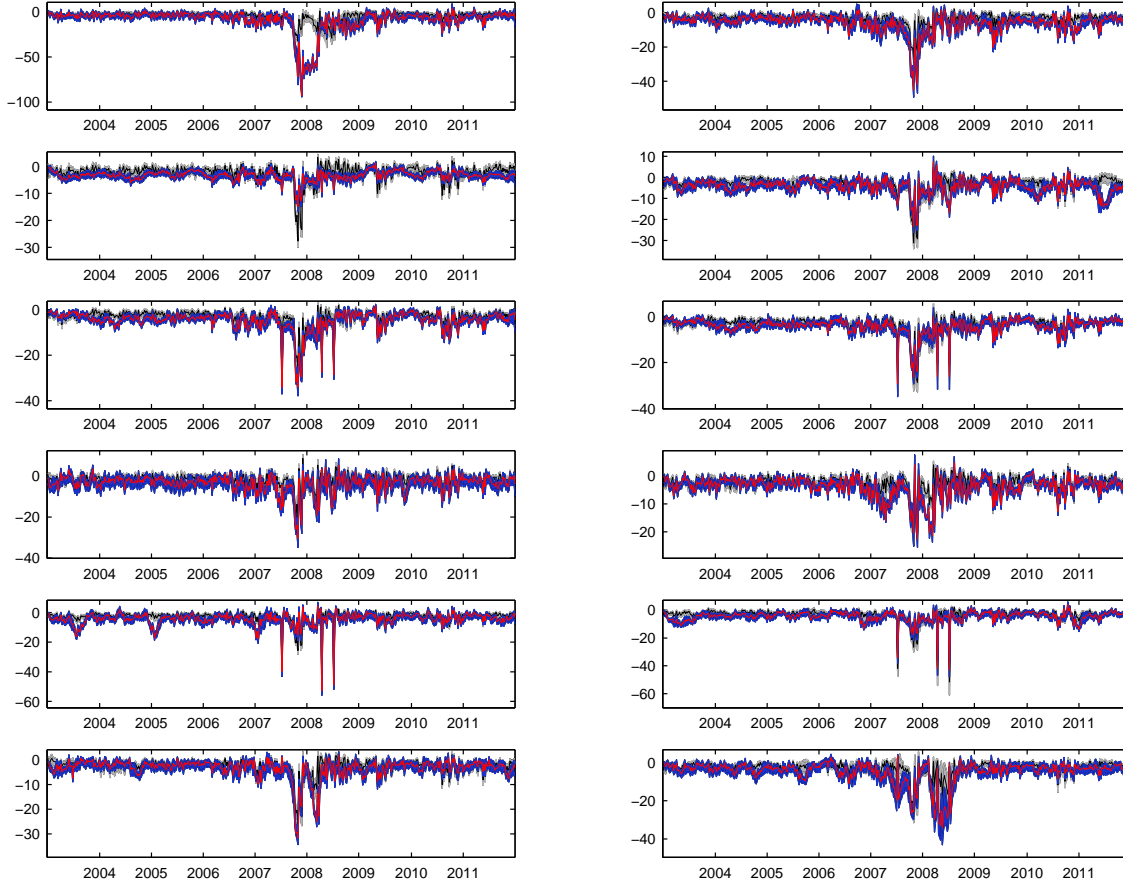


FIGURE 3: Time series plot of the $\text{VaR}_j^{\mathbf{x},\tau}$ (red line) and $\text{CoVaR}_{k|j}^{\mathbf{x},\tau}$ (gray line) at the confidence level $\tau = 0.025$ along with $\text{HPD}_{95\%}$ credible sets, obtained by fitting the model defined in equations (4.1)–(4.6) for the following assets: first panel (financial): C (left), GS (right); second panel (consumer): MCD (left) and NKE (right); third panel (energy): CVX (left), XOM (right); fourth panel (industrial): BA (left), GE (right); fifth panel (technology): INTC (left), ORCL (right); last panel (utilities): AEE (left), PEG (right).

- the leverage regressor (LEV) has always a negative (when significant) coefficient for the VaR, indicating that individual risk is greatly enhanced in highly leveraged companies, while for the CoVaR regression it is positive for AXP (which belong to consumer finance sub-sector) but negative for GS, MS and MCO (which belongs to the diversified financial services sub-sector);
- the market capitalisation (SIZE) has a significantly positive impact on the VaR while its sign varies also across institutions belonging to the same sector for the CoVaR regression. This evidence suggests that large institutions are more risky if considered in isolation. Moreover, the extent to which large companies contribute to the overall risk is not clear, depending on the “degree of connection” among institutions and on diversification of their portfolios;
- the maturity mismatch coefficient (MM) is always negative for the VaR regression while it is positive for the CoVaR regression. This denotes the existence of positive correlation between financial imbalances and individual riskiness (measured by the

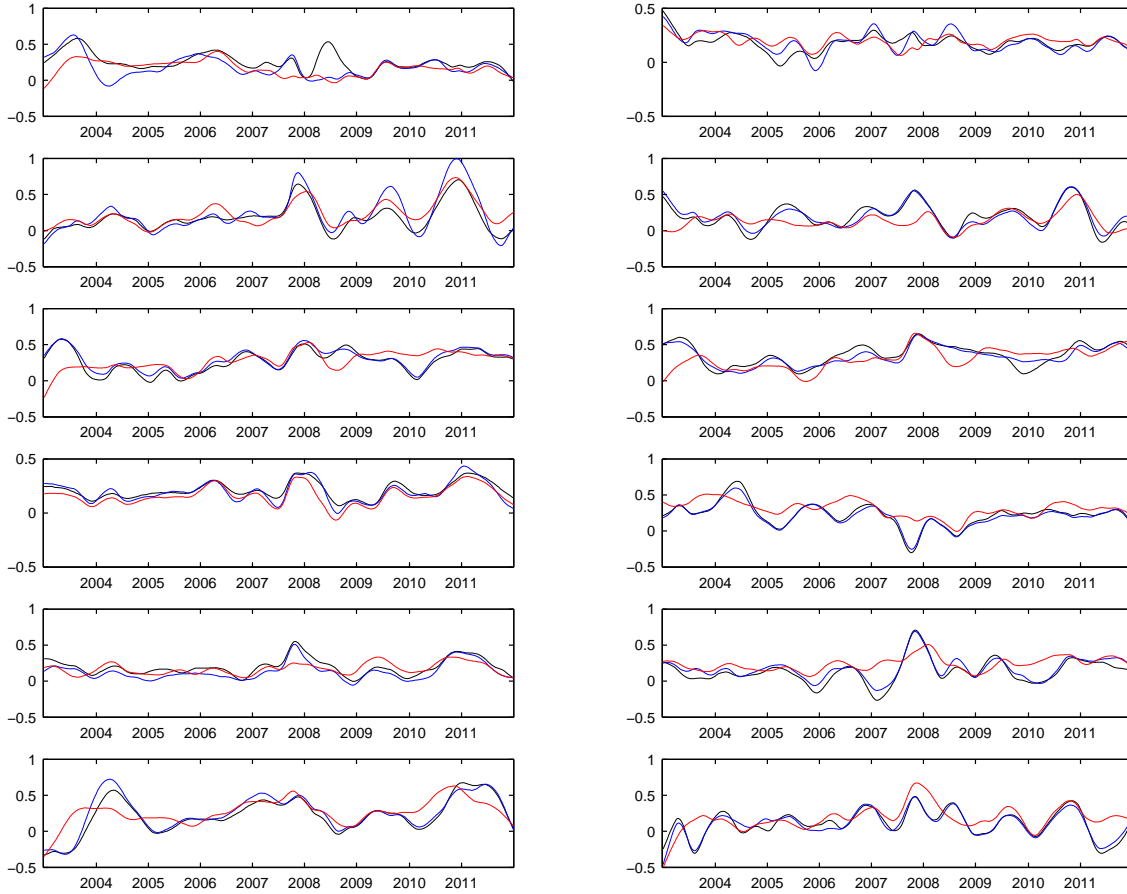


FIGURE 4: Time series plot of the dynamic β at three confidence level $\tau = 0.025$, (dark line), $\tau = 0.05$, (blue line) and $\tau = 0.5$ (red line), obtained by fitting the model defined in equations (4.1)–(4.6) for the following assets, (denoted by j): first panel (financial): C (left), GS (right); second panel (consumer): MCD (left) and NKE (right); third panel (energy): CVX (left), XOM (right); fourth panel (industrial): BA (left), GE (right); fifth panel (technology): INTC (left), ORCL (right); last panel (utilities): AEE (left), PEG (right).

VaR), as we would expect;

- the market-to-book ratio (MK2BK) is significantly different from zero for the VaR only in two cases: CMA and KEY which belong to the bank sector, and it is always positive for the CoVaR regression.

To have a complete picture of the individual and systemic risk contributions we plot in Figure 1 together the estimated VaR and CoVaR along with their HPD_{95%} for some of the assets listed in Table 1. Starting with individual risk assessment, we clearly find that the VaR profiles are relatively similar across institutions, displaying strong negative downside effects upon the occurrence of the recent financial crisis of 2008, and 2010 and the sovereign debt crisis of 2012. However, the analysis of the time series evolution of the marginal contribution to the systemic risk, measured by CoVaR, reveals different behaviours for the considered assets. In particular, Citygroup (C), which belongs to the bank sector, seems to contribute more to the overall risk than other assets do. Inspecting Figure 1 and Tables 4–5, we note that institutions having low β coefficients provide major

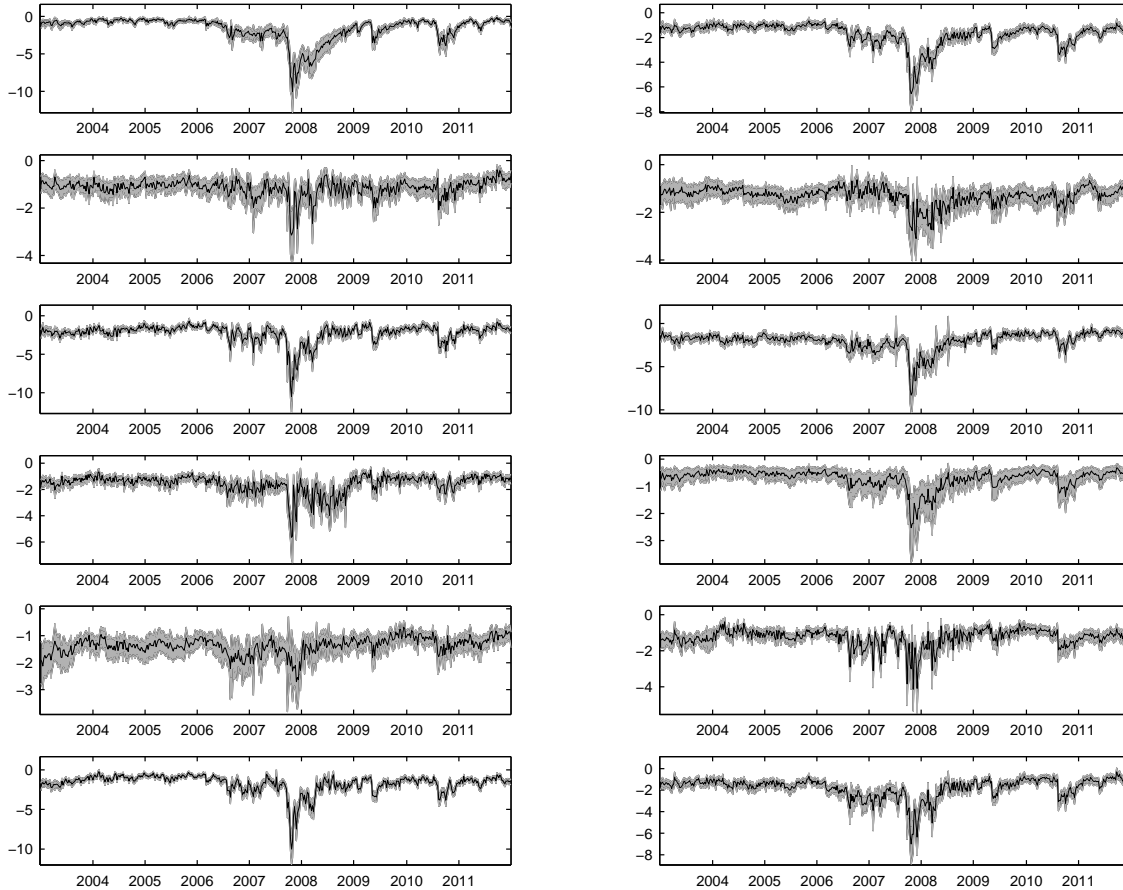


FIGURE 5: Time series plot of the static $\Delta\text{CoVaR}_{k|j}^{x,\tau}$ at the confidence level $\tau = 0.025$ along with $\text{HPD}_{95\%}$ credible sets, obtained by fitting the model defined in equations (3.2) for the following assets, (denoted by j): first panel (financial): C (left), GS (right); second panel (consumer): MCD (left) and NKE (right); third panel (energy): CVX (left), XOM (right); fourth panel (industrial): BA (left), GE (right); fifth panel (technology): INTC (left), ORCL (right); last panel (utilities): AEE (left), PEG (right).

contribution to the 2008 financial crisis. On the contrary, McDonalds Corp. (MCD), which belongs to the consumer sector, has a large estimated β , and inspecting the CoVaR plot in Figure 1 its contribution to the overall systemic risk is much lower than that of financial institutions.

For the selected companies, Figure 2 plots the CoVaR for two different confidence levels $\tau = 0.025$ and $\tau = 0.1$. From this figure we can highlight the different impact of the crisis among sectors. For the financial one, for example, we note that the difference between $\text{CoVaR}_j^{x,0.025}$ and $\text{CoVaR}_j^{x,0.1}$ is much larger than for assets belonging to other sectors, meaning that the financial sector probably had a huge impact on the extreme systemic risk during the 2008 crisis, as witnessed by the extremely large losses.

5.3 Time-varying risk beta

We estimate time-varying systemic risk betas according to the dynamic model defined in equations (4.1)–(4.6) using the same exogenous variables described in Section 5.1. Tables

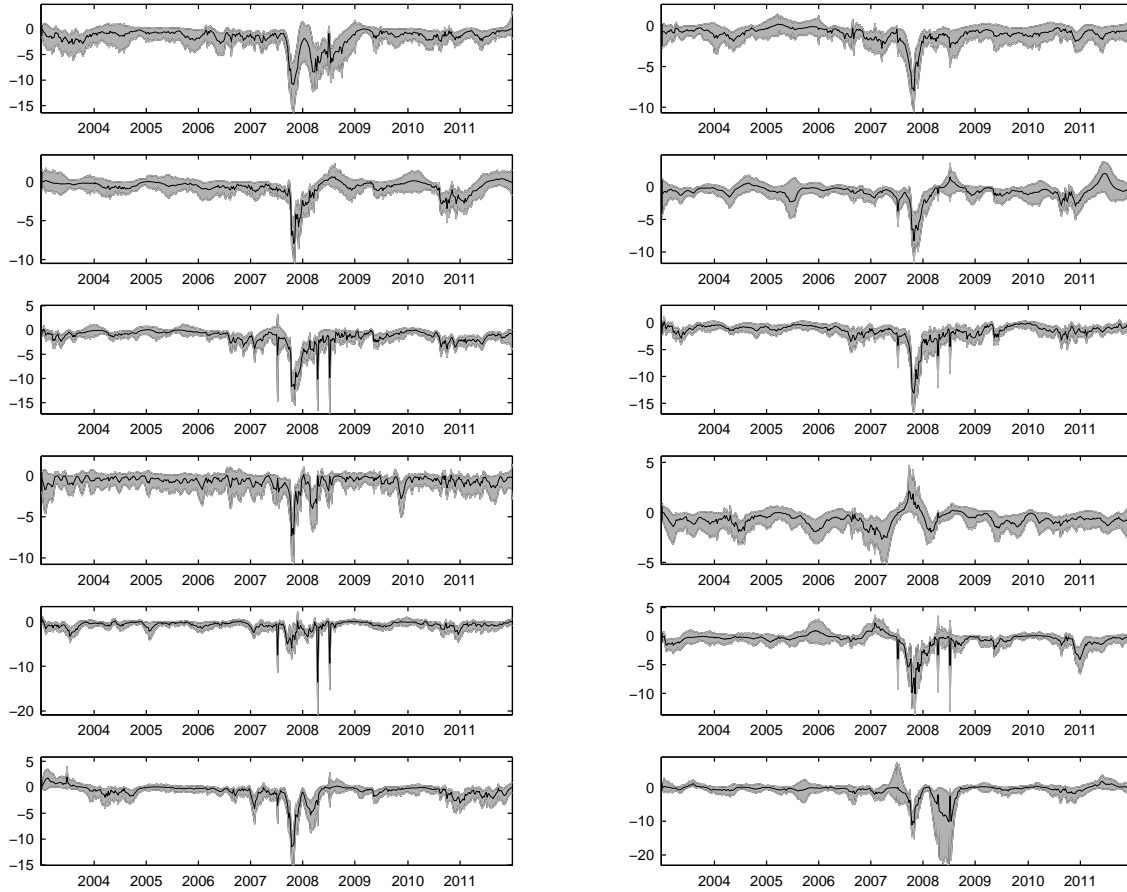


FIGURE 6: Time series plot of the dynamic $\Delta\text{CoVaR}_{k|j}^{\mathbf{x},\tau}$ at the confidence level $\tau = 0.025$ along with $\text{HPD}_{95\%}$ credible sets, obtained by fitting the model defined in equations (4.1)–(4.6) for the following assets, (denoted by j): first panel (financial): C (left), GS (right); second panel (consumer): MCD (left) and NKE (right); third panel (energy): CVX (left), XOM (right); fourth panel (industrial): BA (left), GE (right); fifth panel (technology): INTC (left), ORCL (right); last panel (utilities): AEE (left), PEG (right).

6 and 7 list the posterior estimates of the exogenous regressor parameters θ for both the $\text{VaR}_j^{\tau,\mathbf{x}}$ and $\text{CoVaR}_{k|j}^{\tau,\mathbf{x}}$ regressions, at the confidence level $\tau = 0.025$ and $\tau = 0.05$, respectively. We observe that all the macroeconomic variables, except for the volatility index (VIX) and the liquidity spread (LIQSPR), either have a positive impact on both VaR and CoVaR or are non significant. This means that an upward shift in the term structure of interest rate provides a marginal positive contribution to individual and systemic risks. Interestingly, American Express Co. (AXP) is the only financial institution displaying a negative coefficient for the variable credit spread (CREDSPR) in the individual VaR regression. This essentially means that an increase in the credit spread increases the individual riskiness of that institution. This result seems to be coherent with the American Express' institutional activity, since it is a global financial services institution whose main offerings are charge and credit cards. As expected, the volatility index (VIX) and the rate of change of the Dow Jones US real estate index (DJUSRE) have different effects on the individual and overall risks. Concerning the individual variables a clear positive effect for all the considered institutions is evident only for the SIZE and just for the VaR

regression. The maturity mismatch (MM) is non significant almost for all the reported institutions for both VaR and CoVaR regression, while the market-to-book ratio (MK2BK) and the leverage (LEV) variables reveal an heterogeneous sector-specific impact on the risk measures. In this respect the analysis is not exhaustive and the identification of sector specific risk factors deserves further investigation using a larger number of institutions.

Figure 3, which is the dynamic counterpart of Figure 1, shows that the dynamic CoVaR risk measure suddenly adapts to capture extreme negative losses especially during the 2008 financial crisis. Comparing this evidence with the one shown in Figure 1 it is clear that the dynamic model provides a better characterisation of extreme tail co-movements when dealing with time series data.

Turning our attention to the time varying β 's, Figure 4 plots the evolution of the MaP estimates for three different confidence levels $\tau = 0.025$, $\tau = 0.05$ and $\tau = 0.5$. As expected the evolution of the β 's for $\tau = 0.025$ (dark line) and $\tau = 0.05$ (blue line), almost coincides for all the reported institutions. Moreover, the systemic risk betas behaviours display a huge cross-sectional heterogeneity. For example, the systemic risk beta of the consumer sectors (second panel in Figure 4) increase over time, showing their highest values during the financial crisis at the end of 2008 and 2011. The time series behaviour of the systemic risk betas reveals a different impact of the crisis periods on the overall marginal risk contribution of each institution.

5.4 Measuring marginal contribution to systemic risk

In their paper Adrian and Brunnermeier [4] introduced as a measure of the marginal contribution to system risk, the $\Delta\text{CoVaR}_{k|j}^{\mathbf{x},\tau}$ defined as

$$\Delta\text{CoVaR}_{k|j}^{\mathbf{x},\tau} = \text{CoVaR}_{k|Y_j=\text{VaR}_j^{\mathbf{x},\tau}}^{\mathbf{x},\tau} - \text{CoVaR}_{k|Y_j=\text{VaR}_j^{\mathbf{x},0.5}}^{\mathbf{x},\tau}$$

where $\text{CoVaR}_{k|Y_j=\text{VaR}_j^{\mathbf{x},\tau}}^{\mathbf{x},\tau} = \text{CoVaR}_{k|j}^{\mathbf{x},\tau}$ and $\text{CoVaR}_{k|Y_j=\text{VaR}_j^{\mathbf{x},0.5}}^{\mathbf{x},\tau}$ satisfies equation (2.1) with $\text{VaR}_j^{\mathbf{x},0.5}$ instead of $\text{VaR}_j^{\mathbf{x},\tau}$. To illustrate the behaviour of $\Delta\text{CoVaR}_{k|j}^{\mathbf{x},\tau}$ in our application we plot in Figure 5 the time-invariant one, while in Figure 6 its dynamic version. As expected, for all the considered companies, the marginal contribution to systemic risk increases during market turbulences, showing their lowest values during the financial crisis of 2008. Moreover, some relevant differences among companies belonging to different sectors are evident: in particular, the financial sector, (first panel of Figures 5–6), the energy sector (third panel of Figures 5-6) and the utilities sector (bottom panel of Figures 5-6), display the largest drop. Vice versa, the technology sector displays the lowest variations of the ΔCoVaR measure during the 2008 recession. The grey areas correspond to the $\text{HPD}_{95\%}$ associated to the ΔCoVaR contributions providing information about the size of the risk contribution for the given confidence level. We note a huge cross-sectional heterogeneity on the credible sets behaviour with some sectors such as the consumer, industrial and technology being characterised by large uncertainty of the ΔCoVaR estimates.

Comparing the two figures, it is evident that the dynamic ΔCoVaR estimates are smoother than the corresponding time-invariant one, leading to a less noise-corrupted signal useful for policy maker purposes. Moreover, Figure 6 provides a clear indication of the high flexibility of the dynamic model implying a promptly reaction of the risk measure to the economic and financial downturns. These considerations argue in favour of the dynamic model when dealing with time series data.

Finally, we consider the time series relationship between VaR and CoVaR or ΔCoVaR .

In what is perhaps the key result of Adrian and Brunnermeier [4], they find that the CoVaR (ΔCoVaR) of two institutions may be significantly different even if the VaR of the two institutions are similar. On this basis, they suggest the policy maker to employ the CoVaR (ΔCoVaR) risk measure, as a valid alternative to the VaR, while forming policy regarding institution's risk. Results obtained with our modelling support their thesis. In fact, building the following simple regression model for each institution j

$$\text{CoVaR}_{k|j}^{\tau,\mathbf{x}} = \delta_0 + \delta_1 \text{VaR}_k^{\tau,\mathbf{x}} + \nu, \quad (5.1)$$

or

$$\Delta\text{CoVaR}_{k|j}^{\tau,\mathbf{x}} = \delta_0 + \delta_1 \text{VaR}_k^{\tau,\mathbf{x}} + \nu, \quad (5.2)$$

with $\mathbb{E}(\nu|\text{VaR}_k^{\tau,\mathbf{x}}) = 0$, where k is the Standard and Poor's Index we test $H_0 : \delta_1 = 1$ to show that $\text{CoVaR}_{k|j}^{\tau,\mathbf{x}}$ ($\Delta\text{CoVaR}_{k|j}^{\tau,\mathbf{x}}$) is significantly different from the $\text{VaR}_k^{\tau,\mathbf{x}}$. Bold numbers in Table 3 indicate cases where the corresponding coefficient δ_1 is not significantly different from one. Except for few cases the null hypothesis is rejected and this is more evident for the dynamic model than for the time invariant one. In particular, for the ΔCoVaR regression there is enough evidence that the estimated δ_1 parameter is related to the sector the institutions belong to, being higher for financials, energies and utilities. Interestingly, the regression coefficients is almost zero for the technology sector.

6 Conclusion

One of the major issue policy makers deal with during financial crisis is the evaluation of the extent to which riskness tail events spread across financial institutions. In fact, during financial turmoils, the correlations among asset returns tend to rise, a phenomenon known in the economic and financial literature as contagion. From a statistical point of view the risk of contagion essentially implies that the joint probability of observing large losses increases during recessions. The common risk measures recently imposed by the public regulators, (the Basel Committee, for the bank sector) such as the VaR, fail to account for such risk spillover among institutions. The CoVaR risk measure recently introduced by Adrian and Brunnermeier [4] overcomes this problem being able to account for the dependence among institutions' extreme events.

In this paper we address the problem of estimating the CoVaR in a Bayesian framework using quantile regression. We first consider a time-invariant model allowing for interactions only among contemporaneous variables. The model is subsequently extended in a time-varying framework where the constant part and the CoVaR parameter β are modelled as functions of unobserved processes having their own dynamics. In order to make posterior inference, we use the maximum a posteriori summarising criteria and we prove that it leads to estimated quantiles having good sample properties according to De Rossi and Harvey results [16] and we efficient Gibbs sampler algorithms based on data augmentation for the two models considered.

The Bayesian approach used throughout the paper allows to infer on the whole posterior distribution of the quantities of interest and their credible sets which are important to assess the accuracy of the point estimates. Since the quantities of interest in this contest are risk measures, learning about the whole distribution becomes more relevant due to the interpretation of the VaR and CoVaR as financial losses. In addition, credible sets provide upper and lower limits for the capital requirements for banks and financial institutions.

To verify the reliability of the built models we analyse weekly time series for nineteen institutions belonging to six different sectors of the Standard and Poor 500 composite index spanning the period from 2nd January, 2004 to 31st December, 2012. We use micro and macro exogenous variable to characterise the quantile functions. From the empirical results it is clear that the model and the proposed approach are able to sharply estimate marginal and conditional quantiles providing a more realistic and informative characterisation of extreme tail co-movements. In particular, the dynamic version of the model we propose outperforms the time invariant specification when the analysis is based on time series data. Up to our knowledge this is the first attempt to implement a Bayesian inference for the CoVaR.

A Missing values treatment

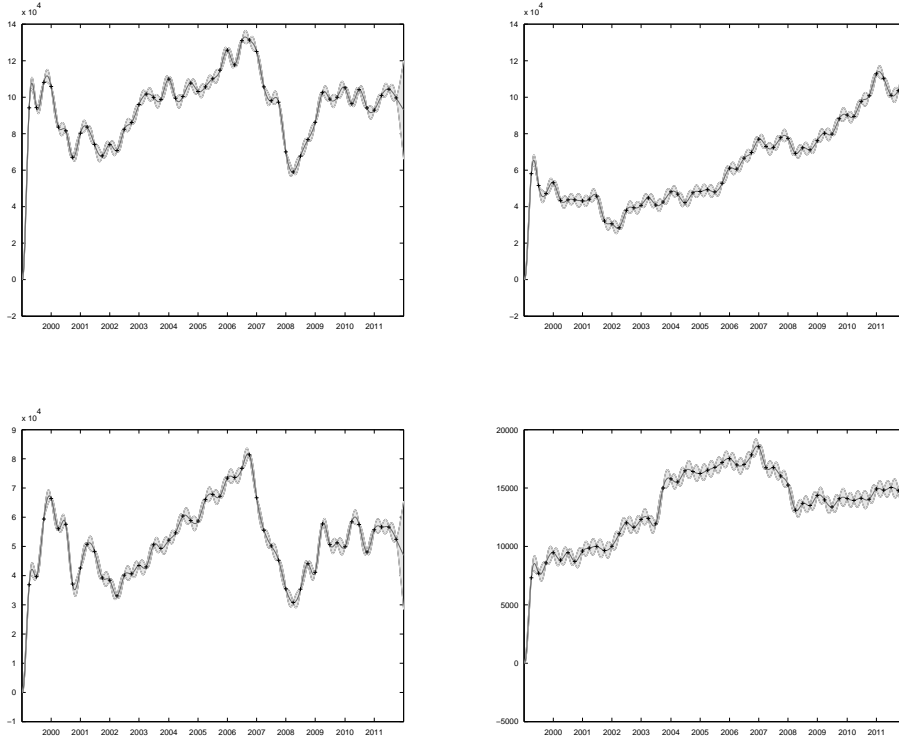


FIGURE 7: Missing values imputation using smoothing spline of the SIZE variable for AXP (top, left), MCD (top, right), BA (bottom, left) and AEE (bottom, right). In each plot predicted values (gray line) as well as 95% HPD credible sets (gray area) are displayed. Dark cross denotes observed values.

In this appendix we give details on the procedure used to impute missing observations. In particular for each variable, starting from balance sheets data available only on a quarterly basis, we implement a nonparametric smoothing cubic spline, see for example Koopman [35] and Koopman *et al.* [36]. Formally, suppose we have an univariate time series $y_{\tau_1}, y_{\tau_2}, \dots, y_{\tau_n}$ not necessarily equispaced in time and define $\delta_t = \tau_t - \tau_{t-1}$, for $t = 1, 2, \dots, n$ as the difference between two consecutive observations, and assume the time series is not entirely observed, we approximate the series by a sufficiently smooth function $s(\tau_t)$. Standard approach chooses $s(\tau_t)$ by minimizing the following penalized-least squares criterion:

$$\mathcal{L}(\lambda) = \sum_{i=1}^n [y_{\tau_i} - s(\tau_i)]^2 + \lambda \sum_{i=1}^n [\Delta^m s(\tau_i)]^2, \quad (\text{A.1})$$

with respect to $s(\tau_t)$, for a given penalization term λ . The function $s(\tau_t), \forall t = 1, 2, \dots, n$ is a polynomial spline of order $m + 1$ and when $m = 2$, we have a smoothing cubic spline model.

To estimate the smoothing parameter λ and to forecast missing observations model

(A.1) can be cast in State Space form that is, for $m = 2$:

$$\begin{aligned} y(\tau_t) &= s(\tau_t) + \epsilon(\tau_t) \\ s(\tau_{t+1}) &= s(\tau_t) + \delta_t \ell(\tau_t) + \zeta_1(\tau_t) \\ \ell(\tau_{t+1}) &= \ell(\tau_t) + \zeta_2(\tau_t) \end{aligned}$$

where the transition equation innovations vector is $[\zeta_1(\tau_t), \zeta_2(\tau_t)]^\top \sim \mathcal{N}(\mathbf{0}_2, \mathbf{S}_\zeta)$ with $\mathbf{S}_\zeta = \sigma_\zeta^2 \begin{bmatrix} \frac{1}{3}\delta_t^3 & \frac{1}{2}\delta_t^2 \\ \frac{1}{2}\delta_t^2 & \delta_t \end{bmatrix}$, the measurement innovation $\epsilon(\tau_t) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ and the penalty parameter λ coincides with the signal-to-noise ratio $\lambda = \sigma_\zeta^2 / \sigma_\epsilon^2$. We fix the parameter σ_ϵ to one and estimate $\lambda = \sigma_\zeta^2$ in a Bayesian framework imposing a diffuse Inverse Gamma prior distribution, i.e. $\lambda = \sigma_\zeta^2 \sim \mathcal{IG}(\alpha_\lambda^0, \beta_\lambda^0)$. We fit the model using MCMC techniques, in particular, the Gibbs sampler with data augmentation (see Geman and Geman [22], Tanner and Wong [49] and Gelfand and Smith [21]). The Gibbs sampler consist in the following two steps:

- S1. simulate the latent process $\boldsymbol{\xi}_t^{(i+1)} = [s(\tau_t), \ell(\tau_t)]^\top$ using the disturbance simulation smoothing algorithm of de Jong and Shephard [15] appropriately adjusted to handle missing observations and the diffuse initialization of the state vector (see also the augmented Kalman filter and smoother of de Jong, [14]).
- S2. simulate the $\lambda^{(i+1)}$ parameter from the complete full conditional distribution which is an Inverse Gamma distribution $\lambda^{(i+1)} \sim \mathcal{IG}(\tilde{\alpha}_\lambda^{(i+1)}, \tilde{\beta}_\lambda^{(i+1)})$ with parameters:

$$\tilde{\alpha}_\lambda^{(i+1)} = \alpha_\lambda^0 + \frac{n-1}{2}, \quad \tilde{\beta}_\lambda^{(i+1)} = \beta_\lambda^0 + \frac{1}{2} \sum_{t=1}^{n-1} \left(\boldsymbol{\xi}_{t+1}^{(i+1)} - A \boldsymbol{\xi}_t^{(i+1)} \right)^\top \mathbf{S}_\zeta^{-1} \left(\boldsymbol{\xi}_{t+1}^{(i+1)} - A \boldsymbol{\xi}_t^{(i+1)} \right).$$

for $i = 1, 2, \dots, G$, with G equal to the number of draws. The algorithm is initialized at $i = 0$ by simulating the λ parameter from its prior distribution. To estimate the missing observations we average simulated points of $s(\tau_t)$ across Gibbs sampler draws.

B Tables

Name	Mean	Min	Max	Std. Dev.	Corr.	1% Str. Lev.
C	-0.490	-92.632	78.797	9.756	0.673	-26.267
BAC	- 0.212	-59.287	60.672	7.990	0.705	-27.231
CMA	-0.072	-31.744	33.104	5.945	0.718	-19.794
JPM	0.086	-41.684	39.938	5.826	0.711	-12.785
KEY	-0.210	-61.427	40.976	7.428	0.694	-24.231
GS	0.071	-36.564	39.320	5.641	0.719	-16.850
MS	-0.170	-90.465	69.931	8.342	0.699	-19.706
MCO	0.125	-27.561	28.300	5.615	0.688	-20.719
AXP	0.091	-28.779	24.360	5.104	0.772	-16.108
MCD	0.320	-12.130	11.878	2.606	0.554	-4.978
NKE	0.260	-18.462	18.723	3.746	0.655	-11.596
CVX	0.254	-31.674	15.467	3.585	0.745	-8.275
XOM	0.197	-22.301	8.717	3.081	0.696	-7.123
BA	0.164	-25.294	16.034	4.258	0.727	-12.468
GE	-0.023	-18.680	30.940	4.311	0.715	-15.578
INTC	-0.052	-17.038	16.935	4.117	0.671	-12.533
ORCL	0.203	-15.518	12.135	3.762	0.632	-9.986
AEE	0.015	-29.528	9.485	3.118	0.682	-8.172
PEG	0.143	-26.492	10.568	3.318	0.553	-8.191
S&P500	0.052	-20.083	11.355	2.637	1.000	-7.258

TABLE 2: Summary statistics of the company’s returns and market index (S&P500) returns (in percentage). The sixth column, denoted by “Corr”, is the correlation coefficient with the market returns while the last column, denoted by “1% Str. Lev.” is the 1% empirical quantile of the returns distribution.

Name	CoVaR		Δ CoVaR	
	Time Invariant	Time Varying	Time Invariant	Time Varying
C	1.325	1.035	0.401	0.325
BAC	1.084	0.983	0.230	0.298
CMA	1.109	1.264	0.109	0.177
JPM	1.259	0.573	0.136	0.111
KEY	1.078	-0.481	0.229	0.165
GS	1.085	1.276	0.234	0.188
MS	1.003	0.604	0.275	0.132
MCO	1.085	-4.139	0.130	0.052
AXP	1.110	18.352	0.171	0.205
MCD	0.933	3.500	0.062	0.192
NKE	0.842	1.251	0.095	0.222
CVX	0.986	2.376	0.291	0.354
XOM	1.099	1.280	0.276	0.365
BA	1.017	2.666	0.169	0.156
GE	1.179	21.245	0.078	-0.039
INTC	0.874	5.928	0.060	0.097
ORCL	0.998	3.985	0.066	0.229
AEE	1.054	1.441	0.346	0.296
PEG	1.001	5.123	0.205	0.271

TABLE 3: δ_1 estimates for the regressions in equations (5.1)–(5.2) for both the time invariant and time varying models. Bold numbers indicate that the corresponding coefficients are not significantly different from 1.

VaR	Financial									Consumer			Energy		Industrial		Technology		Utilities	
	C	BAC	CMA	JPM	KEY	GS	MS	MCO	AXP	MCD	MKE	CVX	XOM	BA	GE	INTC	ORCL	AEE	PEG	
VIX	-0.372	-0.402	-0.366	-0.143	-0.606	-0.275	-0.715	-0.034	-0.211	-0.178	-0.257	-0.401	-0.308	-0.411	-0.082	-0.074	-0.226	-0.131	-0.164	
LIQSPR	(-0.602,-0.120)	(-0.464,-0.084)	(-0.469,-0.314)	(-0.256,-0.025)	(-0.824,-0.460)	(-0.400,-0.220)	(-1.019,-0.668)	(-0.018,0.900)	(-0.283,-0.130)	(-0.206,-0.114)	(-0.298,-0.156)	(-0.514,-0.350)	(-0.418,-0.282)	(-0.511,-0.262)	(-0.164,-0.010)	(-0.086,0.069)	(-0.317,-0.167)	(-0.331,-0.081)	(-0.179,0.016)	
3MTB	-0.010	0.051	-0.008	0.027	0.094	0.004	0.065	-0.012	-0.013	0.045	0.027	0.038	0.017	0.005	-0.052	0.047	0.024	0.024	0.015	
TERMSPR	(-0.046,0.051)	(-0.021,0.062)	(-0.031,0.018)	(0.006,0.069)	(0.045,0.136)	(-0.048,0.039)	(0.037,0.140)	(-0.026,0.124)	(-0.036,0.024)	(0.030,0.060)	(-0.017,0.049)	(-0.009,0.063)	(-0.019,0.067)	(-0.048,0.026)	(-0.068,-0.005)	(0.002,0.082)	(-0.024,0.059)	(0.008,0.078)	(-0.043,0.077)	
CREDSR	0.126	0.017	-0.038	0.088	-0.054	0.038	-0.002	0.104	-0.039	-0.020	0.090	0.062	0.021	0.072	0.028	0.076	0.100	-0.030	0.043	
DJUSRE	(0.036,0.186)	(-0.054,0.074)	(-0.097,0.006)	(-0.015,0.088)	(-0.118,-0.012)	(0.031,0.136)	(-0.087,0.060)	(0.024,0.215)	(-0.070,0.017)	(-0.030,0.023)	(0.010,0.095)	(-0.007,0.108)	(-0.019,0.109)	(0.016,0.102)	(0.008,0.081)	(0.015,0.105)	(0.028,0.119)	(-0.043,0.035)	(-0.006,0.091)	
LEV	0.143	0.066	0.113	0.107	0.004	0.050	0.018	0.084	0.029	-0.015	0.063	0.026	0.028	0.062	0.062	0.065	0.065	-0.026	0.021	
MK2BK	(0.069,0.194)	(-0.011,0.099)	(0.062,0.124)	(0.044,0.116)	(-0.048,0.027)	(0.039,0.134)	(-0.046,0.056)	(0.054,0.179)	(-0.001,0.059)	(-0.029,0.018)	(0.009,0.084)	(-0.004,0.060)	(-0.009,0.070)	(0.028,0.090)	(0.034,0.079)	(0.032,0.090)	(0.013,0.082)	(-0.043,0.015)	(-0.021,0.043)	
SIZE	0.042	0.135	-0.006	0.121	0.067	0.014	-0.068	0.073	-0.107	-0.009	0.079	0.047	0.051	0.020	-0.011	0.129	0.107	-0.164	-0.067	
MM	(-0.029,0.169)	(-0.078,0.126)	(-0.074,0.058)	(0.024,0.150)	(-0.036,0.114)	(-0.087,0.063)	(-0.130,0.050)	(0.045,0.230)	(-0.168,-0.079)	(-0.027,0.068)	(0.019,0.122)	(0.000,0.108)	(0.006,0.144)	(-0.073,0.068)	(-0.050,0.035)	(0.063,0.163)	(0.024,0.146)	(-0.154,-0.033)	(-0.117,-0.002)	
σ_j	0.821	0.443	0.326	0.621	0.574	0.543	0.632	0.581	0.585	0.130	0.468	0.221	0.326	0.431	0.512	0.317	0.216	0.295	0.337	
σ_s	(0.588,0.980)	(0.387,0.690)	(0.309,0.539)	(0.541,0.830)	(0.396,0.671)	(0.380,0.637)	(0.524,0.797)	(0.497,0.863)	(0.485,0.672)	(0.061,0.205)	(0.337,0.518)	(0.122,0.276)	(0.171,0.344)	(0.354,0.527)	(0.406,0.571)	(0.210,0.374)	(0.136,0.329)	(0.229,0.399)	(0.278,0.476)	
σ_e	4.010	-2.095	-2.126	-5.057	-9.054	3.715	7.84	13.564	4.993	-3.784	-3.996	-8.426	1.008	8.872	1.752	10.882	3.213	29.995	7.278	
σ_τ	(3.060,5.498)	(-2.776,0.234)	(-2.110,-0.036)	(-8.468,-5.457)	(-11.707,-5.887)	(3.426,4.314)	(4.867,1.740)	(13.316,13.575)	(3.357,5.100)	(-5.398,6.451)	(-9.360,1.539)	(-8.127,3.678)	(-4.034,8.182)	(0.875,1.003)	(-2.053,2.712)	(-0.518,13.613)	(-2.475,9.482)	(27.499,33.212)	(0.963,8.038)	
σ_ω	-10.536	-3.206	-2.164	-6.315	-2.946	-9.493	-8.836	0.000	0.190	4.503	2.413	-5.572	1.460	0.008	-1.991	-35.730	-11.746	7.257	21.233	
σ_ν	(6.514,9.782)	(6.104,8.709)	(0.795,2.780)	(26.485,30.494)	(17.537,24.030)	(4.867,1.740)	(2.716,4.767)	(29.507,30.644)	(10.611,12.669)	(7.376,9.753)	(2.399,4.800)	(12.581,14.745)	(3.936,6.351)	(7.432,7.850)	(5.967,8.552)	(34.792,38.321)	(12.874,16.147)	(1.877,3.235)	(13.295,17.800)	
σ_ξ	2.854	3.491	2.974	-2.903	-3.318	-3.859	1.419	2.211	-2.196	-0.940	0.266	-4.235	3.713	-1.985	1.846	8.831	-5.146	-1.110	2.157	
σ_ζ	(-6.838,5.443)	(-5.925,6.244)	(-7.669,4.549)	(-6.271,5.793)	(-6.506,5.477)	(-7.390,4.397)	(-6.704,5.186)	(-2.740,12.412)	(-8.834,2.961)	(-9.186,0.908)	(-7.130,4.815)	(-7.558,3.962)	(-7.566,4.130)	(-5.881,6.317)	(-4.292,6.978)	(2.176,12.974)	(-5.219,3.953)	(-6.208,5.558)	(-5.796,6.345)	
σ_η	0.241	0.214	0.186	0.145	0.162	0.176	0.188	0.095	0.126	0.092	0.128	0.135	0.115	0.121	0.115	0.114	0.121	0.094	0.117	
σ_θ	(0.203,0.249)	(0.197,0.239)	(0.171,0.211)	(0.133,0.165)	(0.144,0.190)	(0.143,0.179)	(0.181,0.222)	(0.085,0.115)	(0.114,0.139)	(0.090,0.112)	(0.108,0.135)	(0.122,0.149)	(0.104,0.133)	(0.093,0.115)	(0.093,0.115)	(0.107,0.133)	(0.114,0.141)	(0.086,0.107)	(0.121,0.153)	

TABLE 6: Parameter estimates obtained by fitting the Dynamic CoVaR model to each of the 19 assets vs SP500 and all the exogenous variables, for the confidence levels $\tau = 0.025$. For each regressor the first row reports parameter estimates by Maximum a Posteriori, while the second row reports the 95% High Posterior Density (HPD) credible sets.

References

- [1] ACHARYA, V.V., ENGLE, R.F. AND RICHARDSON, M., (2012). Capital shortfall: a new approach to ranking and regulating systemic risks. *American Economic Review*, **102**, pp. 59–64.
- [2] ACHARYA, V.V., E., PHILIPPON, T. AND RICHARDSON, M., (2010). Measuring systemic Risk. *New York University Working Paper*.
- [3] ADAMS, Z., FÜSS, R. AND GROPP, R., (2010). Modeling spillover effects among financial institutions: a state-dependent sensitivity Value-at-Risk (SDSVaR) approach. *Working paper, European Business School*.
- [4] ADRIAN, T. AND BRUNNERMEIER, M.K., (2011). CoVaR. *Unpublished manuscript*.
- [5] BERNARDI, M., (2013). Risk measures for skew normal mixtures. *Statistics & Probability Letters*, **83**, pp. 1819–1824.
- [6] BERNARDI, M., MARUOTTI, A. AND PETRELLA, L., (2012). Skew mixture models for loss distributions: a Bayesian approach. *Insurance: Mathematics and Economics*, **51**, pp. 617–623.
- [7] BERNARDI, M., MARUOTTI, A. AND PETRELLA, L., (2013). Multivariate Markov-switching models for tail risk contagion and codependence. *Unpublished manuscript*.
- [8] BILLIO, M., GETMANSKY, M., LO, A.W. AND PELLIZON, L., (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, **104**, pp. 535–559.
- [9] BROWNLEES, C.T. AND ENGLE, R. (2012). Volatility, correlation and tails for systemic risk measurement. *Working paper*.
- [10] CHERNOZHUKOV, V. AND DU, S., (2008). Extremal Quantiles and Value-at-Risk. *The New Palgrave Dictionary of Economics. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume*. Palgrave Macmillan.
- [11] CARTER, C.K. AND KOHN, R., (1994). On Gibbs sampling for state space models. *Biometrika*, **81**, pp. 541–553.
- [12] CARTER, C.K. AND KOHN, R., (1996). Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika*, **83**, pp. 589–601.
- [13] CHAO, S-K., HÄRDLE, W.K. AND WANG, W. (2012). Quantile regression in risk calibration. *SFB 649 Discussion Paper*, 2012–006.
- [14] DE JONG P., (1991). The diffuse Kalman filter. *Annals of Statistics*, **19**, pp. 1073–1083.
- [15] DE JONG P. AND SHEPHARD, N., (1995). The simulation smoother for time series models. *Biometrika*, **82**, pp. 339–350.
- [16] DE ROSSI, G. AND HARVEY, A., (2009). Quantiles, expectiles and splines. *Journal of Econometrics*, **152**, pp. 179–185.

- [17] DURBIN, J. AND KOOPMAN, S.J., (2012). *Time Series Analysis by State Space Methods, 2nd Eds*, Oxford: Oxford University Press.
- [18] DURBIN, J. AND KOOPMAN, S.J., (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, **89**, pp. 1603–616.
- [19] ENGLE, R.F. AND MANGANELLI, S., (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, **22**, pp. 367–381.
- [20] FRÜHWIRTH-SCHNATTER, S., (1994). Data augmentation and dynamic linear models, *Journal of time series analysis*, **15**, pp. 183–202.
- [21] GELFAND, A.E. AND SMITH, F.M., (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, pp. 398–409.
- [22] GEMAN, S. AND GEMAN, D., (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, **6**, pp. 721–741.
- [23] GERLACH, R.H., CHEN, C.W.S. AND CHAN, N.Y.C., (2011). Bayesian time-varying quantile forecasting for Value-at-Risk in financial markets. *Journal of Business & Economic Statistics*, **29**, pp. 481–492.
- [24] GEWEKE, J., (1992). Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In J. M. Bernardo, J. Berger, A. P. Dawid, and A. F. M. Smith, eds., *Bayesian Statistics 4*, Oxford University Press, pp. 169–193.
- [25] GEWEKE, J., (2005). *Contemporary Bayesian Econometrics and Statistics*. Wiley Series in Probability and Statistics. Wiley, Hoboken.
- [26] GIRARDI, G., ERGÜN, A.T., (2013). *Systemic risk measurement: Multivariate GARCH estimation of CoVaR*. *Journal of Banking & Finance*, <http://dx.doi.org/10.1016/j.jbankfin.2013.02.027>.
- [27] GOURIEROUX, C. AND JASIAK, J., (2008). Dynamic quantile models. *Journal of Econometrics*, **147**, pp. 198–205.
- [28] HAUTSCH, N., SCHAUMBURG, J. AND SCHIENLE, M., (2011). Quantifying time-varying marginal systemic risk contributions. *Unpublished manuscript*.
- [29] HUANG, A.Y., (2012). Value at risk estimation by quantile regression and kernel estimator. *Review of Quantitative Finance and Accounting*, doi10.1007/s11156-012-0308-x
- [30] JOHANNES, M. AND POLSON, N., (2009). MCMC methods for Financial Econometrics, in *Handbook of Financial Econometrics*, Volume 2, pp. 1–72, edited by Y. Ait-Sahalia and L.P. Hansen, North Holland.
- [31] JORION, P., (2007). *Value-at-Risk: The new benchmark for managing financial risk*. McGraw-Hill, Chicago.
- [32] KOENKER, P., (2005). *Quantile Regression*. Cambridge University Press, Cambridge.

- [33] KOENKER, B. AND BASSET, G., (1978). Regression Quantiles. *Econometrica*, **46**, pp. 33–50.
- [34] KOENKER, R. AND XIAO, Z., (2006). Quantile autoregression. *Journal of the American Statistical Association*, **101**, pp. 980–990.
- [35] KOOPMAN, S.J., (1991). Efficient smoothing algorithms for time series models. *Doctoral Thesis*.
- [36] KOOPMAN, S.J., SHEPHARD, N. AND DOORNIK, J.A., (1998). Statistical algorithms for models in state space using Ssf Pack 2.2. *Econometrics Journal*, **1**, pp. 1–55.
- [37] KOTTAS, A. AND GELFAND, A.E., (2001). Bayesian semiparametric median regression modeling. *Journal of the American Statistical Association*, **96**, pp. 1458–1468.
- [38] KOTTAS, A., KRNJAJIC, M., (2009). Bayesian semiparametric modelling in quantile regression. *Scandinavian Journal of Statistics*, **36**, pp.297-319.
- [39] KOTZ, S., KOZUBOWSKI, T.J. AND PODGÓRSKI, K. (2001). *The Laplace distribution and generalizations. A revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhäuser.
- [40] KOZUMI, H. AND KOBAYASHI, G., (2011). Gibbs Sampling Methods for Bayesian Quantile Regression, *Journal of Statistical Computation and Simulation*, **81**, pp. 1565–1578.
- [41] KUESTER, K., MITTNIK, S. AND PAOLELLA, M.S., (2011). Value-at-Risk Prediction: A Comparison of Alternative Strategies, *Journal of Financial Econometrics*, **4**, pp. 53–89.
- [42] KUROSE, Y. AND OMORI, Y., (2012). Bayesian analysis of time-varying quantiles using a smoothing spline, *Journal of the Japan Statistical Society*, **42**, pp. 23–46.
- [43] LEE, C.F. AND SU, J.B., (2012). Alternative statistical distributions for estimating value-at-risk: theory and evidence, *Review of Quantitative Finance and Accounting*, **39**, pp. 309–331.
- [44] LIN, N. AND CHANG, C., (2012). Comment on Article by K. Lum and A.E. Gelfand, “Spatial Quantile Multiple Regression Using the Asymmetric Laplace Process”, *Bayesian Analysis*, **7**, pp. 263–270.
- [45] LIU, J.S. (1994). The Collapsed Gibbs Sampler in Bayesian Computations with Applications to a Gene Regulation Problem, *Journal of the American Statistical Association*, **89**, pp. 958–966.
- [46] LUM, K. AND GELFAND, A.E., (2012). Spatial Quantile Multiple Regression Using the Asymmetric Laplace Process, *Bayesian Analysis*, **7**, pp. 235–258.
- [47] MCNEIL, A., FREY, R. AND EMBRECHTS, P., (2005). *Quantitative risk management: Concepts, Techniques, Tools*. Princeton Series in Finance, Princeton.
- [48] PARK, T. AND CASELLA, G., (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, **103**, pp. 681–686.

- [49] TANNER, M.A. AND WONG, W.H., (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Society*, **82**, pp. 528–550.
- [50] SCHAUMBURG, J., (2010). Predicting extreme VaR: nonparametric quantile regression with refinements from extreme value theory, *SFB 649 Discussion Paper*, 2010-009.
- [51] TAYLOR, J.W., (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, **6**, pp. 382–406.
- [52] VAN DYK, D.A. AND PARK, T., (2008). Partially collapsed Gibbs samplers: Theory and Methods. *Journal of the American Statistical Association*, **110**, pp. 790–796.
- [53] YU, K. AND MOYEED, R., (2001). Bayesian quantile regression. *Statistics and Probability Letters*, **54**, pp. 437–447.