# Borradores de ECONOMÍA

Comparison of Methods for Estimating the Uncertainty of Value at Risk

Por: Santiago Gamba Santamaría,
Oscar Fernando Jaulín Méndez,
Luis Fernando Melo Velandia,
Carlos Andrés Quicazán Moreno

Núm. 927
2016


-tá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Col-

BANCO DE LA REPÚBLICA
COLOMBIA

# COMPARISON OF METHODS FOR ESTIMATING THE UNCERTAINTY OF VALUE AT RISK*

SANTIAGO GAMBA SANTAMARÍA
(gambas@javeriana.edu.co)
OSCAR FERNANDO JAULÍN MÉNDEZ
(ojaulime@banrep.gov.co)
LUIS FERNANDO MELO VELANDIA
(lmelovel@banrep.gov.co)
CARLOS ANDRÉS QUICAZÁN MORENO
(cquicamo@banrep.gov.co)

ABSTRACT. Value at Risk (VaR) is a market risk measure widely used by risk managers and market regulatory authorities. There is a variety of methodologies proposed in the literature for the estimation of VaR. However, few of them get to say something about its distribution or its confidence intervals. This paper compares different methodologies for computing such intervals. Several methods, based on asymptotic normality, extreme value theory and subsample bootstrap, are used. Using Monte Carlo simulations, it is found that these approaches are only valid for high quantiles. In particular, there is a good performance for VaR(99%), in terms of coverage rates, and bad performance for VaR(95%) and VaR(90%). The results are confirmed by an empirical application for the stock market index returns of G7 countries.

*Keywords*: Value at Risk, confidence intervals, data tilting, subsample bootstrap.

*JEL Codes*: C51, C52, C53, G32.

## 1. INTRODUCTION

Value at Risk (VaR) is a widely used risk measure by financial agents such as risk managers or regulatory authorities. It is defined as the maximum loss of a certain asset for a given probability level ($\alpha$); then, VaR corresponds to the $\alpha-$quantile of the distribution of the asset at a certain time.

There are several methods for obtaining the VaR point estimator of an asset for a forecasting horizon. Nevertheless, few of them get to say something about the uncertainty of these estimators. Knowledge about this uncertainty would allow to account for the accuracy of the VaR estimation, and it can be quantified by estimating confidence intervals. Because VaR represents the expected loss in the worst cases at a certain probability level, this measure is used to set capital requirements; therefore, when its confidence intervals are found to be considerably wide, caution is necessary.

---

*Date*: February 17, 2016.

* The opinions expressed here are those of the authors and do not necessarily represent those of the Banco de la República or those of its Board of Directors. The usual disclaimers apply. The authors thank Daniel Mariño for his helpful comments and suggestions.

Methodologies that allow VaR confidence interval estimation rely on different assumptions, some of them are based on asymptotic normality. For example, Moraux [2011] gives a first approach assuming that the returns are independent and identically normally distributed. He computes the confidence intervals by applying the delta-method to estimate the asymptotic variance of VaR. However, it is well known that financial series exhibit heteroscedastic and heavy-tailed behavior.

For this reason, other methods allow the return series to follow an ARMA-GARCH dynamic. Some of them do not assume any particular distribution, but they use asymptotic theory. Spierdijk [2014] mentions that some of these methods fail when the normality assumption does not hold or when the sample size is not large enough. Gao and Song [2008], Francq and Zakoïan [2015] and Chan, Deng, Peng, and Xia [2007] employ different approaches based on QML in order to estimate the quantile of the residuals without the asymptotic normality assumption. Gao and Song [2008] use filtered historical simulation, Francq and Zakoïan [2015] propose a reparametrization of the GARCH errors while Chan, Deng, Peng, and Xia [2007] rely on extreme value theory.

Another method to estimate VaR confidence intervals is based on bootstrap techniques on the standardized errors. However, conventional bootstrap is not consistent when asymptotic normality does not hold, and its inconsistency arises when there are GARCH dynamics with an unbounded fourth moment [Hall and Yao, 2003]. Therefore, Spierdijk [2014] proposes a subsample bootstrap methodology for ARMA-GARCH models using QML. This methodology performs well when asymptotic normality does not hold; and compared to conventional bootstrap, this method does not requiere the fourth moment to be finite.

The purpose of this paper is to compare some of these methodologies, in order to identify the scenarios under which such confidence interval techniques perform properly. This is done by evaluating coverage rates for each method through Monte Carlo simulations. The scenarios consider different persistence degrees in mean and variance, sample sizes, VaR probability levels, confidence levels of the intervals and distributions of the standardized errors.

The results suggest that the evaluated methods have a good performance for VaR(99%) under the classical features of financial time series, such as low mean persistence and high variance persistence. However, Moraux [2011] method, based on asymptotic normality, presents the worst coverage ratios yielding to extremely wide confidence intervals. Additionally, none of the methods have a good performance for VaR(90%) and VaR(95%). These conclusions are supported by an empirical exercise for the stock market index returns of the G7 countries.

The rest of the paper is organized as follows. Section two explains the methods used to compute the confidence intervals of VaR. A simulation exercise of those methods is given in section three. Section four contains an empirical application for the stock market index returns of G7 countries. Finally, some concluding remarks are presented in section five.

## 2. Methodology

In this section four methods are considered to compute the confidence interval of VaR for a return series $r_t$.

2.1. **Asymptotic normality.** Moraux [2011] considers a simple case where the VaR is obtained under normal *iid* assumption, $r_t \overset{iid}{\sim} N(\mu, \sigma^2)$. Then, $\widehat{VaR}(t+1|t)_\alpha = \widehat{\mu} + \widehat{\sigma} q_\alpha$, where $q_\alpha$ stands for the $\alpha^{th}$ quantile of the standard normal distribution.

Given the following asymptotic results, $\sqrt{T}(\widehat{\mu} - \mu) \xrightarrow{d} N(0, \sigma^2)$ and $\sqrt{T}(\widehat{\sigma}^2 - \sigma^2) \xrightarrow{d} N(0, \sigma^4)$, using the delta method he shows that $\sqrt{T}(\widehat{VaR} - VaR) \xrightarrow{d} N\left(0, \sigma^2(1+0.5q_\alpha^2)\right)$.

Then, the $\beta-$asymptotic confidence interval of VaR is

$$\widehat{VaR}(t+1|t)_\alpha \pm \frac{\widehat{\sigma}}{\sqrt{T}} \sqrt{1+0.5q_\alpha^2} \;\; q_{0.5(1+\beta)} \tag{1}$$

Even though this approach is very simple, does not take into account the stylized facts of financial series. In particular, heavy tail distributions and volatility clustering. However, the methods described below consider both facts.

2.2. **Asymptotic Hill estimator.** This method uses the point estimator of a quantile by Chan, Deng, Peng, and Xia [2007]. By using an Extreme Value Theory approach they showed that this estimator is asymptotically normally distributed.

Suppose the returns follow an $AR(P) - GARCH(p,q)$ model, namely,

$$r_t = \mu + \sum_{i=1}^{P} \phi_i r_{t-i} + \varepsilon_t \tag{2}$$

$$\sigma_t^2 = c + \sum_{i=1}^{p} b_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} a_i \sigma_{t-i}^2 \tag{3}$$

where $\varepsilon_t = \sigma_t z_t$, $c < 0, b_i \geq 0, a_i \geq 0$ for every $i$, $\sum_{i=1}^{p} b_i + \sum_{i=1}^{q} a_i < 1$, and $\{z_t\}$ is a sequence of independent identically distributed random variables with mean zero and variance one.

Therefore, the one-step ahead VaR($\alpha$) is of the form

$$VaR\,(t+1 \mid t)_\alpha = \mu + \sum_{i=1}^{P} \phi_i r_{t-i+1} + \sigma_{t+1|t} q_\alpha \tag{4}$$

where $q_\alpha$ is the $\alpha^{th}$ quantile of $z_{t+1}$.

Under the assumption that the tails of the distribution of $z_t$ decrease at constant rate $\gamma$, Extreme Value Theory estimation of excesses over a given threshold can be analyzed in terms of the sequence $\{\hat{z}_t\}$ of standardized residuals.

Let $\hat{z}_{m,1} \leq \hat{z}_{m,2} \leq ... \leq \hat{z}_{m,m}$ be the order statistics of the $m$ last standardized residuals $\hat{z}_v, ..., \hat{z}_T$, with $v = T - m + 1$. Then, the Hill estimator of $\gamma$ is

$$\hat{\gamma} = \left[ \frac{1}{k} \sum_{i=1}^{k} \log \frac{\hat{z}_{m,m-i+1}}{\hat{z}_{m,m-k}} \right]^{-1} \tag{5}$$

where $k$ is the number of extreme observations, and $\hat{z}_{m,m-k}$ can be considered as the threshold.

Hill [1975] proposes the following point estimator for $q_\alpha$

$$\hat{q}_\alpha = (1-\alpha)^{-\frac{1}{\hat{\gamma}}} \left( \frac{k}{m} \right)^{\frac{1}{\hat{\gamma}}} \hat{z}_{m,m-k} \tag{6}$$

And then, the point estimator for $VaR\,(t+1 \mid t)_\alpha$ is

$$\widehat{VaR}\,(t+1|t)_\alpha = \hat{\mu} + \sum_{i=1}^{P} \hat{\phi}_i r_{t-i+1} + \widehat{\sigma}_{t+1|t} \hat{q}_\alpha \tag{7}$$

Chan, Deng, Peng, and Xia [2007] show that this estimator is consistent and converges in distribution to a standard normal distribution

$$\frac{\hat{\gamma}\sqrt{k}}{\left| \log\left( \frac{k}{m(1-\alpha)} \right) \right|} \left[ \frac{\widehat{VaR}\,(t+1|t)_\alpha}{VaR\,(t+1 \mid t)_\alpha} - 1 \right] \xrightarrow{d} N(0,1) \tag{8}$$

Then, a $\beta$-level confidence interval for $\widehat{VaR}\,(t+1|t)_\alpha$ is

$$\left( \widehat{VaR}\,(t+1|t)_\alpha \exp\left\{ q_{(1-\beta)/2} \left| \log\left( \frac{k}{m(1-\alpha)} \right) \right| \Big/ \left( \hat{\gamma}\sqrt{k} \right) \right\},\ \widehat{VaR}\,(t+1|t)_\alpha \exp\left\{ q_{(1+\beta)/2} \left| \log\left( \frac{k}{m(1-\alpha)} \right) \right| \Big/ \left( \hat{\gamma}\sqrt{k} \right) \right\} \right) \tag{9}$$

2.3. **Data tilting.** Chan, Deng, Peng, and Xia [2007] propose to use data tilting to estimate *VaR* confidence intervals. Data tilting method is a non-parametric approach which can be seen as a generalization of the empirical likelihood methodology,[1] where the observations are weighted in order to minimize a distance function.[2]

Let $z_t$ and $z_{m,m-k}$ be defined as in section 2.2 and $\delta_t = I(\hat{z}_t \geq \hat{z}_{m,m-k})$. Then, this approach involves the following steps.

First, for any fixed vector of weights $w = (w_v, \ldots, w_T)$ such that $w_t \geq 0$ and $\sum_{t=v}^{T} w_t = 1$, the next optimization problem is solved

[1] A complete discussion and analysis of empirical likelihood method is found in Owen [1988], Owen [1990], and Owen [2001], among others.

[2] This method is useful for constructing confidence regions and one of its advantages is that it enables to compute them allowing a certain degree of asymmetry.

$$(\widehat{\gamma}(w), \widehat{c}(w)) = \underset{\gamma,c}{argmax} \sum_{t=v}^{T} w_t \log\left(\left(c\gamma\,\widehat{z}_t^{\,-\gamma-1}\right)^{\delta_t}\left(1 - c\,\widehat{z}_{m,m-k}^{\,-\gamma}\right)^{1-\delta_t}\right) \tag{10}$$

This results in

$$\widehat{\gamma}(w) = \frac{\sum_{t=v}^{T} w_t \delta_t}{\sum_{t=v}^{T} w_t \delta_t \left(\log \widehat{z}_t - \log \widehat{z}_{m,m-k}\right)} \tag{11}$$

$$\widehat{c}(w) = \widehat{z}_{m,m-k}^{\,\widehat{\gamma}(w)} \sum_{t=v}^{T} w_t \delta_t \tag{12}$$

It is important to note that $\widehat{\gamma}(w), \widehat{c}(w)$ will be part of the constraints of data tilting optimization as explained below.

Then, defining the distance function[3] as

$$D_l(w) = \begin{cases} -m^{-1}\sum_{t=v}^{T} \log(m w_t), & \text{if } l = 0 \\ \sum_{t=v}^{T} w_t \log(m w_t), & \text{if } l = 1 \\ (l(1-l))^{-1}\left(1 - m^{-1}\sum_{t=v}^{T}(m w_t)^l\right), & \text{if } l \neq 0, 1 \end{cases} \tag{13}$$

And solving for the weights to minimize this distance,

$$(2m)^{-1} L\left(VaR(t+1 \mid t)_\alpha\right) = \min_{w} D_l(w) \tag{14}$$

subject to

$$\begin{gathered} w_t \geq 0, \\ \sum_{t=v}^{T} w_t = 1, \\ \widehat{\gamma}(w) \log\left(\left(VaR(t+1|t)_\alpha - \widehat{\mu}_{t+1|t}\right) / \left(\widehat{\sigma}_{t+1|t}\,\widehat{z}_{m,m-k}\right)\right) = \log\left(\sum_{t=v}^{T} w_t \delta_t / (1-\alpha)\right) \end{gathered} \tag{15}$$

Gives the following solution after using standard Lagrange multiplier method for $D_1(w)$ [4]

$$w_t = \begin{cases} \frac{1}{m} e^{-1-\lambda_1}, & \text{if } \delta_t = 0 \\ \frac{1}{m}\exp\left\{-1 - \lambda_1 + \lambda_2\left(\frac{\log\left(\left(VaR(t+1|t)_\alpha - \widehat{\mu}_{t+1|t}\right)/\left(\widehat{\sigma}_{t+1|t}\,\widehat{z}_{m,m-k}\right)\right)}{A_2(\lambda_1)}\right) - \frac{1}{A_1(\lambda_1)}\right. & \\ \left. -\frac{A_1(\lambda_1)}{A_2^2(\lambda_1)} \log\left(\widehat{z}_t/\widehat{z}_{m,m-k}\right) \log\left(\left(VaR(t+1|t)_\alpha - \widehat{\mu}_{t+1|t}\right)/\left(\widehat{\sigma}_{t+1|t}\,\widehat{z}_{m,m-k}\right)\right)\right\}, & \text{if } \delta_t = 1 \end{cases} \tag{16}$$

where $\lambda_1$ and $\lambda_2$ satisfy

$$\sum_{t=v}^{T} w_t = 1, \;\; \widehat{\gamma}(w) \log\left(\left(VaR(t+1|t)_\alpha - \widehat{\mu}_{t+1|t}\right)/\left(\widehat{\sigma}_{t+1|t}\,\widehat{z}_{m,m-k}\right)\right) = \log\left(\sum_{t=v}^{T} w_t \delta_t / (1-\alpha)\right) \tag{17}$$

[3]This function measures the distance between the unconstraint weights and the weights given a uniform distribution ($1/m$) and it is based on power divergence measures [Hall and Yao, 2003].

[4]Chan, Deng, Peng, and Xia [2007] claim that the distance function for $l = 1$ gives good robustness properties.

and

$$A_1(\lambda_1) = 1 - \frac{m-k}{m} e^{-1-\lambda_1}, \;\; A_2(\lambda_1) = A_1(\lambda_1) \frac{\log\left( \left( VaR(t+1|t)_\alpha - \widehat{\mu}_{t+1|t} \right) / (\widehat{\sigma}_{t+1|t}\, \widehat{z}_{m,m-k}) \right)}{\log\left( A_1(\lambda_1)/(1-\alpha) \right)} \quad (18)$$

Under some conditions Chan, Deng, Peng, and Xia [2007] show that

$$L\left( VaR(t+1|t)^0_\alpha \right) \xrightarrow{d} \chi^2(1) \quad (19)$$

where $VaR(t+1|t)^0_\alpha$ denotes the true value of $VaR(t+1|t)_\alpha$.

Finally, based on this result, the confidence interval with level $\beta$ for $VaR(t+1|t)^0_\alpha$ is

$$I^t_\beta = \{ VaR(t+1|t)_\alpha : L(VaR(t+1|t)_\alpha) \leqslant u_\beta \} \quad (20)$$

where $u_\beta$ is the $\beta$-level critical value of $\chi^2(1)$.

2.4. **Subsample bootstrap.** Spierdijk [2014] proposes a residual subsample bootstrap methodology for estimating confidence interval for VaR. She assumes that the return series follows an ARMA-GARCH model. The procedure includes the following steps:

1. Use QML method to estimate an ARMA-GARCH model to the return series, $r_1, \ldots, r_T$.
2. Draw a $l-$random subsample without replacement from the standardized residuals of the previous step, $\tilde{z}_1, \ldots, \tilde{z}_l$.
3. Using the estimated parameters of step-1 and $\tilde{z}_1, \ldots, \tilde{z}_l$ generate the bootstrap returns, $\tilde{r}_1, \ldots, \tilde{r}_l$.
4. Use again QML method to estimate an ARMA-GARCH model to $\tilde{r}_1, \ldots, \tilde{r}_l$
5. Compute $\tilde{q}^\alpha_{\tilde{z}}$ as the $\alpha-$ sample quantile from the standardized residuals of the previous step.
6. Using the estimated parameters of step-4 and $r_1, \ldots, r_T$ calculate $\tilde{\mu}_{t+1|t}$, $\tilde{\sigma}_{t+1|t}$ for $t = 1 \ldots T$.
7. Calculate $\widetilde{VaR}(t+1|t)_\alpha = \tilde{\mu}_{t+1|t} + \tilde{\sigma}_{t+1|t}\tilde{q}^\alpha_{\tilde{z}}$.

As a final stage, steps $1-7$ are repeated $B-$times, resulting in $\widetilde{VaR}(t+1|t)_{\alpha,1}, \ldots, \widetilde{VaR}(t+1|t)_{\alpha,B}$.

Then, the proposed $\beta$-level confidence interval is

$$\left[ \widetilde{VaR}(t+1|t)_{\alpha,1} + q^{(1-\beta)/2}_{VaR}, \;\; \widetilde{VaR}(t+1|t)_{\alpha,1} + q^{(1+\beta)/2}_{VaR} \right] \quad (21)$$

where $q^x_{VaR}$ is the empirical $x-$quantile of $\left\{ \widetilde{VaR}(t+1|t)_{\alpha,1} - \widehat{VaR}(t+1|t)_\alpha, \ldots, \widetilde{VaR}(t+1|t)_{\alpha,B} - \widehat{VaR}(t+1|t)_\alpha \right\}$ and $\widehat{VaR}(t+1|t)_\alpha = \hat{\mu}_{t+1|t} + \hat{\sigma}_{t+1|t}\hat{q}^\alpha_z$

## 3. Simulation Exercise

For this simulation exercise two processes are considered. The first one corresponds to an *i.i.d.* process with expected value $\mu/(1-\phi_1)$ and variance $c/(1-a_1-b_1)$. The second one follows an $AR(1)-GARCH(1,1)$ process. The parameters, described in equations (2) and (3), are initially set to: $\mu = 1, \phi_1 = 0.1, c = 0.1, b_1 = 0.05, a_1 = 0.92$. The two processes are simulated with three different distributions, normal and Student's t with 3 and 10 degrees of freedom. Both, the number of replications and the sample size are set in 1000 each.

The objective of these simulations is to evaluate the performance of the methodologies by computing the 90% confidence intervals of the 1-step ahead VaR(99%) described in section 2. For this purpose, coverage rates are calculated for four methods: asymptotic normality, asymptotic Hill estimator, data tilting and subsample bootstrap.[5] Coverage rates are defined as the proportion of confidence interval replications that contain the real simulated VaR.

TABLE 1. Simulated Coverage Ratios for different AR Parameters

| AR Parameters | Method | *iid* Normal | *iid* $t_3$ | *iid* $t_{10}$ | AR-GARCH Normal | AR-GARCH $t_3$ | AR-GARCH $t_{10}$ |
|---|---|---|---|---|---|---|---|
| $\mu = 0$, $\phi_1 = 0$ | Asymptotic Normality | 0.90 | 0.49 | 0.83 | 0.90 | 0.51 | 0.87 |
| | Asymptotic Hill | 0.94 | 0.87 | 0.94 | 0.88 | 0.78 | 0.88 |
| | Data Tilting | 0.92 | 0.89 | 0.93 | 0.84 | 0.82 | 0.89 |
| | Subsample Bootstrap | 0.91 | 0.93 | 0.88 | 0.86 | 0.87 | 0.85 |
| $\mu = 1$, $\phi_1 = 0.1$ | Asymptotic Normality | 0.88 | 0.49 | 0.85 | 0.88 | 0.54 | 0.82 |
| | Asymptotic Hill | 0.99 | 0.93 | 0.97 | 0.95 | 0.88 | 0.94 |
| | Data Tilting | 0.90 | 0.90 | 0.92 | 0.87 | 0.86 | 0.88 |
| | Subsample Bootstrap | 0.90 | 0.92 | 0.91 | 0.92 | 0.88 | 0.94 |
| $\mu = 1$, $\phi_1 = 0.4$ | Asymptotic Normality | 0.89 | 0.48 | 0.85 | 0.62 | 0.56 | 0.62 |
| | Asymptotic Hill | 0.99 | 0.95 | 0.98 | 0.97 | 0.90 | 0.95 |
| | Data Tilting | 0.89 | 0.91 | 0.91 | 0.86 | 0.82 | 0.88 |
| | Subsample Bootstrap | 0.90 | 0.92 | 0.91 | 0.10 | 0.16 | 0.05 |
| $\mu = 1$, $\phi_1 = 0.9$ | Asymptotic Normality | 0.89 | 0.46 | 0.84 | 0.26 | 0.28 | 0.25 |
| | Asymptotic Hill | 1.00 | 0.99 | 1.00 | 1.00 | 0.98 | 0.99 |
| | Data Tilting | 0.90 | 0.89 | 0.91 | 0.85 | 0.81 | 0.87 |
| | Subsample Bootstrap | 0.92 | 0.94 | 0.88 | 0.12 | 0.13 | 0.14 |

Simulated coverage rates for 90% confidence interval of one-step ahead VaR(99%). The DGP related to columns 3 through 5 (*iid*) is $r_t \overset{iid}{\sim} F$ with mean $\mu/(1-\phi_1)$ and variance $c/(1-b_1-a_1)$, with $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$, where $F$ is normal, Student's t with 3 degrees of freedom or Student's t with 10 degrees of freedom. Meanwhile, the DGP related to the last three columns (AR-GARCH) is $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ and $\sigma_t^2 = c + b_1\varepsilon_{t-1}^2 + a_1\sigma_{t-1}^2$ with $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$. The standardized errors of the AR-GARCH model are distributed as normal, Student's t with 3 degrees of freedom and Student's t with 10 degrees of freedom.

Tables 1, 2 and 3 show the coverage rates for different simulations. The first column of these Tables specifies the parameter of the simulation, whereas the parameters that are not mentioned

[5]Following Chan, Deng, Peng, and Xia [2007] and Spierdijk [2014], $k$ is set to $1.5(\log T)^2$ for asymptotic Hill estimator and data tilting $B$ is defined as $l_3 = (3.5T^{1/2} + 2T^{2/3})$ for subsample bootstrap, respectively.

take the initial values described earlier. The second column indicates the method evaluated and the rest of the columns contains the coverage rates for the two DGPs (*iid* and AR-GARCH) and the three distributions used to carry out the simulations (normal, $t_3$, $t_{10}$).

Table 1 displays the simulations for a set of AR parameter values, $\{(\mu,\phi_1)\} = \{(0,0),(1,0.1),(1,0.4),(1,0.9)\}$. In Table 2, the exercise is carried out for a set of GARCH parameters $\{(c,b_1,a_1)\} = \{(0.1,0.05,0.92),(0.1,0.005,0.99),(0.1,0.3,0.4)\}$. Finally, Table 3 reports the coverage rates for different VaR probabilities ($\alpha = \{0.90,0.95,0.99\}$) and various confidence interval levels ($\beta = \{0.90,0.95,0.99\}$).

TABLE 2. Simulated Coverage Ratios for different GARCH Parameters

| GARCH Parameters | Method | *iid* Normal | *iid* $t_3$ | *iid* $t_{10}$ | AR-GARCH Normal | AR-GARCH $t_3$ | AR-GARCH $t_{10}$ |
|---|---|---|---|---|---|---|---|
| $c=0.1$, $b_1=0.05$, $a_1=0.92$ | Asymptotic Normality | 0.88 | 0.50 | 0.85 | 0.86 | 0.51 | 0.81 |
| | Asymptotic Hill | 0.98 | 0.93 | 0.97 | 0.95 | 0.87 | 0.94 |
| | Data Tilting | 0.90 | 0.90 | 0.92 | 0.87 | 0.86 | 0.88 |
| | Subsample Bootstrap | 0.90 | 0.92 | 0.91 | 0.92 | 0.88 | 0.94 |
| $c=0.1$, $b_1=0.005$, $a_1=0.99$ | Asymptotic Normality | 0.89 | 0.50 | 0.82 | 0.87 | 0.51 | 0.88 |
| | Asymptotic Hill | 0.96 | 0.89 | 0.94 | 0.95 | 0.86 | 0.93 |
| | Data Tilting | 0.89 | 0.90 | 0.93 | 0.89 | 0.85 | 0.85 |
| | Subsample Bootstrap | 0.91 | 0.92 | 0.89 | 0.94 | 0.91 | 0.92 |
| $c=0.1$, $b_1=0.3$, $a_1=0.5$ | Asymptotic Normality | 0.89 | 0.49 | 0.84 | 0.55 | 0.39 | 0.53 |
| | Asymptotic Hill | 1.00 | 0.98 | 1.00 | 0.99 | 0.94 | 0.97 |
| | Data Tilting | 0.88 | 0.90 | 0.91 | 0.84 | 0.83 | 0.86 |
| | Subsample Bootstrap | 0.86 | 0.93 | 0.89 | 0.48 | 0.66 | 0.63 |

Simulated coverage rates for 90% confidence interval of one-step ahead VaR(99%). The DGP related to columns 3 through 5 (*iid*) is $r_t \overset{iid}{\sim} F$ with mean $\mu/(1-\phi_1)$ and variance $c/(1-b_1-a_1)$, with $\mu = 1$ and $\phi_1 = 0.1$, where $F$ is normal, Student's t with 3 degrees of freedom or Student's t with 10 degrees of freedom. Meanwhile, the DGP related to the last three columns (AR-GARCH) is $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ and $\sigma_t^2 = c + b_1\varepsilon_{t-1}^2 + a_1\sigma_{t-1}^2$ with $\mu = 1$ and $\phi_1 = 0.1$. The standardized errors of the AR-GARCH model are distributed as normal, Student's t with 3 degrees of freedom and Student's t with 10 degrees of freedom.

For the *iid* cases in Table 1, simulations show that the four methodologies perform adequately with the exception of asymptotic normality method at a Student's t distribution with 3 degrees of freedom, as well as the asymptotic Hill estimator which overestimates its coverage ratios getting too close to the unity under high unconditional expected value of the series. For the AR-GARCH simulations, coverage rates are generally close to the expected values. Nevertheless, compared with *iid* cases, most of the coverage rates decrease, which means a deterioration of the confidence interval performance, with the exception of the Hill estimator whose decrease means a better performance due to the overestimation of its intervals. Then, high mean persistence affects negatively the performance of both asymptotic normality and subsample bootstrap confidence intervals, while data tilting intervals performance remains the same under higher persistence.

The failure to handle heavy-tailed distributions of asymptotic normality method is again confirmed by results in Table 2. Additionally, the performance of the methods are negatively affected by low variance persistence. In particular, asymptotic normality, subsample bootstrap and, in a smaller magnitude, data tilting underestimate their coverage ratios, while the Hill estimator intervals tend to overestimate coverage ratios in a higher magnitude.

TABLE 3. Simulated Coverage Ratios for different VaR and Confidence Levels

| $VaR(\alpha)$ $C.I.(\beta)$ | Method | *iid* Normal | *iid* $t_3$ | *iid* $t_{10}$ | AR-GARCH Normal | AR-GARCH $t_3$ | AR-GARCH $t_{10}$ |
|---|---|---|---|---|---|---|---|
| | Asymptotic Normality | 0.90 | 0.47 | 0.83 | 0.90 | 0.54 | 0.85 |
| $\alpha = 0.99$ | Asymptotic Hill | 0.98 | 0.91 | 0.97 | 0.96 | 0.87 | 0.94 |
| $\beta = 0.90$ | Data Tilting | 0.90 | 0.90 | 0.92 | 0.87 | 0.86 | 0.88 |
| | Subsample Bootstrap | 0.90 | 0.92 | 0.91 | 0.92 | 0.88 | 0.94 |
| | Asymptotic Normality | 0.88 | 0.59 | 0.84 | 0.88 | 0.62 | 0.85 |
| $\alpha = 0.95$ | Asymptotic Hill | 0.36 | 0.43 | 0.36 | 0.31 | 0.40 | 0.34 |
| $\beta = 0.90$ | Data Tilting | 0.59 | 0.73 | 0.61 | 0.55 | 0.60 | 0.58 |
| | Subsample Bootstrap | 0.95 | 0.95 | 0.95 | 0.60 | 0.47 | 0.60 |
| | Asymptotic Normality | 0.89 | 0.66 | 0.84 | 0.83 | 0.66 | 0.82 |
| $\alpha = 0.90$ | Asymptotic Hill | 0.26 | 0.41 | 0.27 | 0.28 | 0.34 | 0.28 |
| $\beta = 0.90$ | Data Tilting | 0.64 | 0.73 | 0.74 | 0.65 | 0.64 | 0.66 |
| | Subsample Bootstrap | 0.95 | 0.95 | 0.96 | 0.46 | 0.34 | 0.47 |
| | Asymptotic Normality | 0.94 | 0.57 | 0.89 | 0.93 | 0.61 | 0.89 |
| $\alpha = 0.99$ | Asymptotic Hill | 0.99 | 0.97 | 0.99 | 0.97 | 0.95 | 0.98 |
| $\beta = 0.95$ | Data Tilting | 0.95 | 0.94 | 0.96 | 0.93 | 0.92 | 0.94 |
| | Subsample Bootstrap | 0.96 | 0.96 | 0.95 | 0.96 | 0.93 | 0.97 |
| | Asymptotic Normality | 0.98 | 0.70 | 0.96 | 0.99 | 0.75 | 0.96 |
| $\alpha = 0.99$ | Asymptotic Hill | 1.00 | 0.99 | 1.00 | 1.00 | 0.96 | 0.99 |
| $\beta = 0.99$ | Data Tilting | 0.99 | 0.98 | 1.00 | 0.98 | 0.96 | 0.99 |
| | Subsample Bootstrap | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.99 |

Simulated coverage rates for $\beta$ confidence interval of one-step ahead VaR($\alpha$) . The DGP related to columns 3 through 5 (*iid*) is $r_t \overset{iid}{\sim} F$ with mean $\mu/(1-\phi_1)$ and variance $c/(1-b_1-a_1)$, with $\mu = 1$, $\phi_1 = 0.1$, $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$, where $F$ is normal, Student's t with 3 degrees of freedom or Student's t with 10 degrees of freedom. Meanwhile, the DGP related to the last three columns (AR-GARCH) is $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ and $\sigma_t^2 = c + b_1\varepsilon_{t-1}^2 + a_1\sigma_{t-1}^2$ with $\mu = 1$, $\phi_1 = 0.1$, $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$. The standardized errors of the AR-GARCH model are distributed as normal, Student's t with 3 degrees of freedom and Student's t with 10 degrees of freedom.

In general, the conclusions of Tables 1 and 2 remain unchanged when considering VaR(99%) at 95% and 99% confidence intervals, as shown in Table 3. However, calculating VaR for lower probability levels (VaR(90%) and VaR(95%)) severely decreases the intervals performance. For the $AR-GARCH$ series Hill estimator, data tilting and subsample bootstrap are not capable to reach a coverage rate higher than 0.70. These results lead to conclude that the considered methodologies are only valid for high quantiles (VaR(99%)).

Tables 4, 5 and 6 in Appendix A present the coverage ratios for the same exercises using a sample size of 500 observations instead of 1000. As expected, most of the confidence intervals worsen their performance. Particularly, the data tilting method has a remarkable decrease of the coverage ratio compared to the other three methodologies. Nonetheless, previous conclusions remain substantively unaffected.

To sum up, as presumed, the asymptotic normality method has a bad performance for most of the simulations as well as a poor capacity to capture mean and variance dynamics specially for heavy-tailed distributions. More important, the coverage rates of the asymptotic Hill estimator, data tilting and subsample bootstrap methods are close to the expected values for VaR(99%). Furthermore, these three methods present their best performance when considering the scenarios exhibited by financial time series, being these heavy-tailed distributions, high variance persistence and low mean persistence. On the contrary, none of the studied methods appears to estimate properly the confidence intervals for a VaR with non-extreme probability levels, as simulation results show for VaR(90%) and VaR(95%).

## 4. Empirical Exercise

This section contains an empirical application of the four methods for estimating VaR confidence intervals presented in section 2. For this purpose, the negative returns of the stock market indexes of G7 countries from September 9, 2010 to October 10, 2015 are used, resulting in a sample sizes around 1250 observations. The return series and the normal QQ-plot are presented in Figures 1 and 2, respectively. These graphs exhibit the classic stylized facts for financial time series, i.e. volatility clustering and heavy-tailed behavior.

In order to evaluate the performance of these methods, 90% confidence intervals for one step ahead $VaR(95\%)$ and $VaR(99\%)$ are computed. For this purpose, VaR is initially computed by fitting an AR(p)-GARCH(1,1) model using the available information up to September 25, 2014. Thereafter, the information set is augmented by one observation at a time, for which a new AR(p)-GARCH(1,1) model is fitted. This procedure is implemented recursively, 250 times, until the information set reaches the period September 22, 2015.[6] [7]

[6]The standardized residuals and some specification tests associated to the AR(p)-GARCH(1,1) model for the whole sample are presented in Figure 7 and Tables 7, 8 and 9 of Appendix B, respectively. The results of these tests show no evidence of misspecification

[7]Table 10 in Appendix C show the backtesting results for VaR series computed using the four methodologies for the G7 countries. Following Leccadito, Boffelli, and Urga [2014], the generalized Christoffersen [2011] backtesting tests are implemented jointly for both VaR(95%) and VaR(99%). The null hypotheses of the three tests (independence,

FIGURE 1. Negative returns of G7 stock market indexes

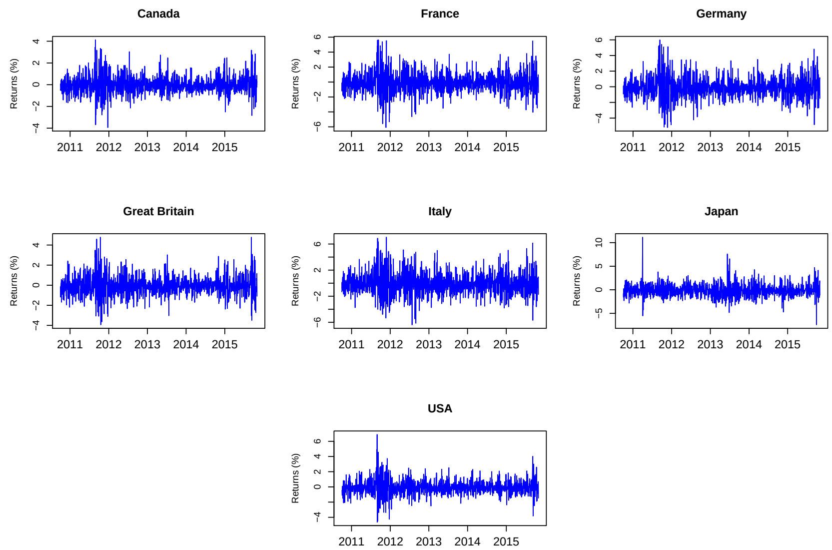
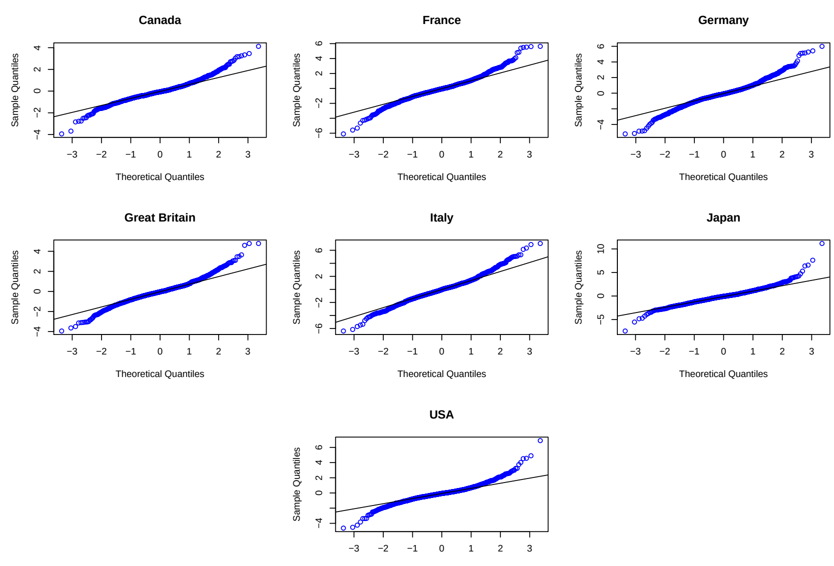



FIGURE 2. QQ-plot of G7 stock market indexes




unconditional coverage and conditional coverage) are not rejected for almost all of the countries and methods. Nevertheless, it is important to note that asymptotic normality method is the only one that rejects conditional coverage hypothesis for some countries, while the other three methodologies perform properly since the Christoffersen's tests are not rejected.

Figures 3, 4, 5 and 6 display the (negative) returns of G7 stock market indexes from October, 2014 to October, 2015; the 250 estimates of 1-step ahead VaR (95%) and VaR(99%), and the 90% confidence intervals computed for each method. As a result, it can be appreciated that for normally distributed scenarios, the confidence intervals are wider than those computed using the other techniques. This method also fails to enhance the volatility dynamics of the return series, mainly because this method assumes that time series is independent and identically normal distributed. Meanwhile, the other three methodologies do model this feature.

For the case of VaR(95%) in Figures 8, 9, 10 and 11 in Appendix D, the characteristics of the asymptotic normality intervals remain, but for the other three methodologies, there are important differences. The confidence intervals for both, asymptotic Hill estimation and data tilting, are extremely narrow. This implies a more efficient estimation of the confidence intervals; however, they might be inconsistent since they show a low coverage rate in the simulation exercise. On the other hand, subsample bootstrap intervals are not as narrow as the other two confidence interval methods.

In the case of VaR(99%), the Hill estimator and data tilting confidence intervals are quite similar. This result is expected since both methods are based on QML estimation with extreme value theory. For subsample bootstrap techniques, the confidence intervals obtained are wider (less efficient) than the former two. However, Spierdijk [2014] states that the subsample bootstrap confidence interval width is driven by the method robustness to the lack of asymptotic normality.

FIGURE 3. 90% confidence intervals for $VaR(99\%)$ of G7 stock market indexes negative returns using Asymptotic Normality

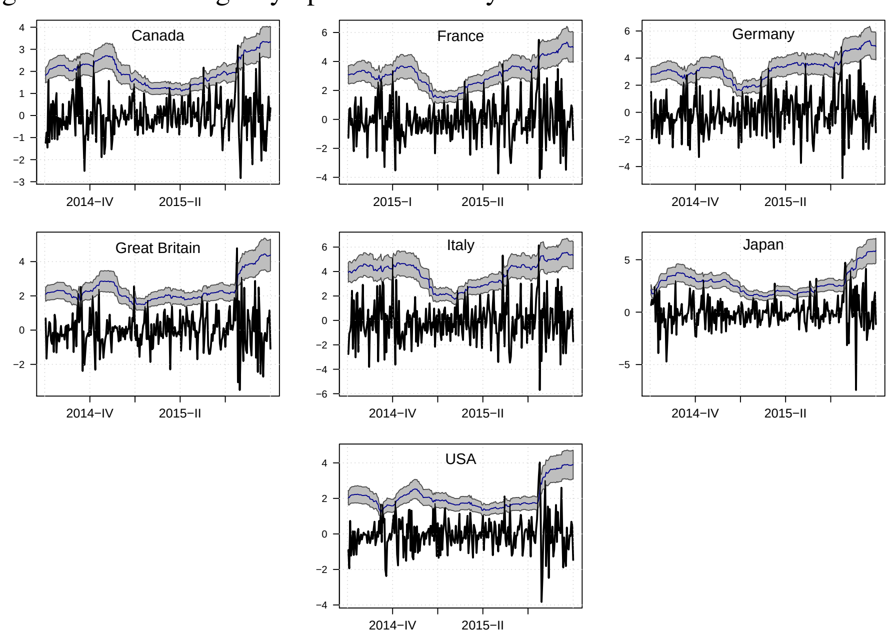


FIGURE 4. 90% confidence intervals for $VaR(99\%)$ of G7 stock market indexes negative returns using Asymptotic Hill Estimator

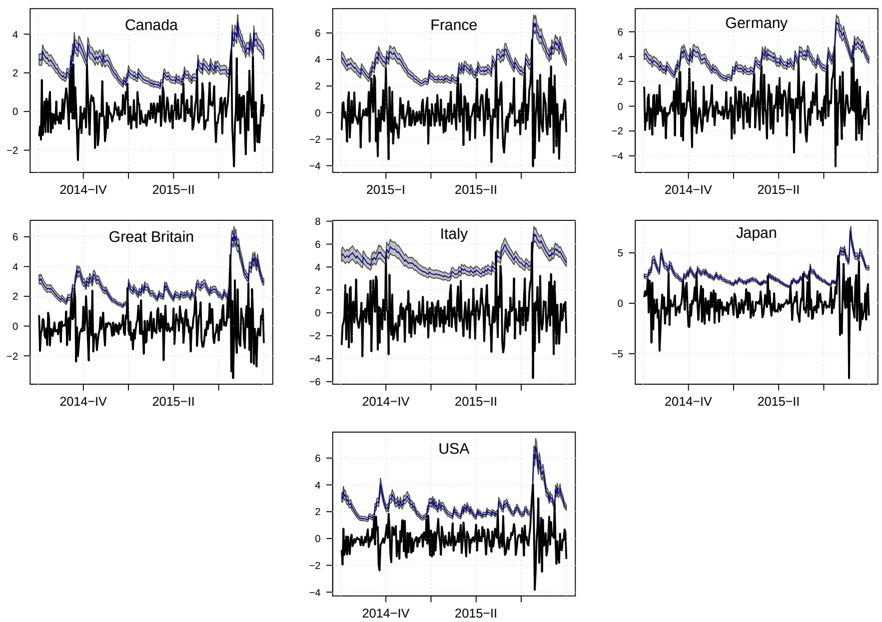



FIGURE 5. 90% confidence intervals for $VaR(99\%)$ of G7 stock market indexes negative returns using Data Tilting

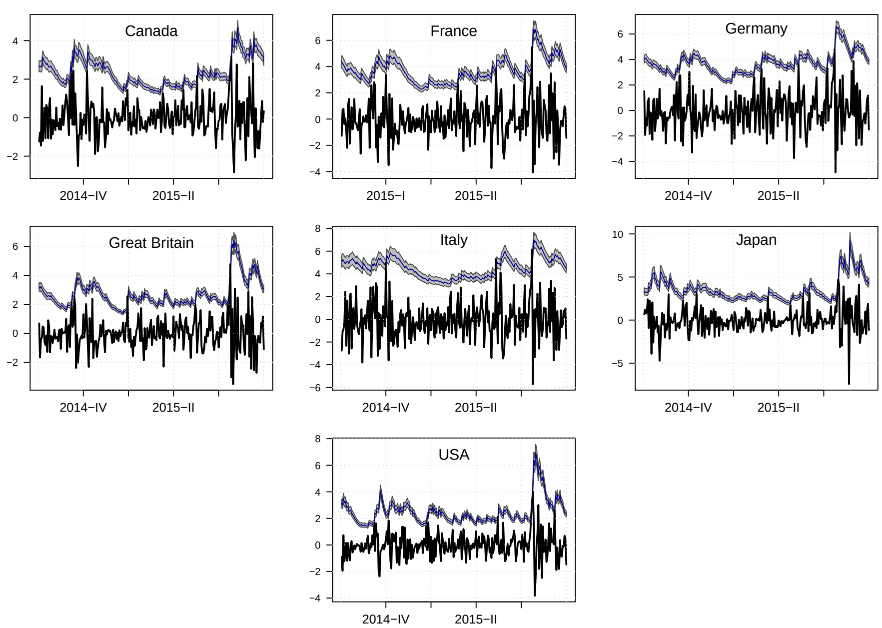
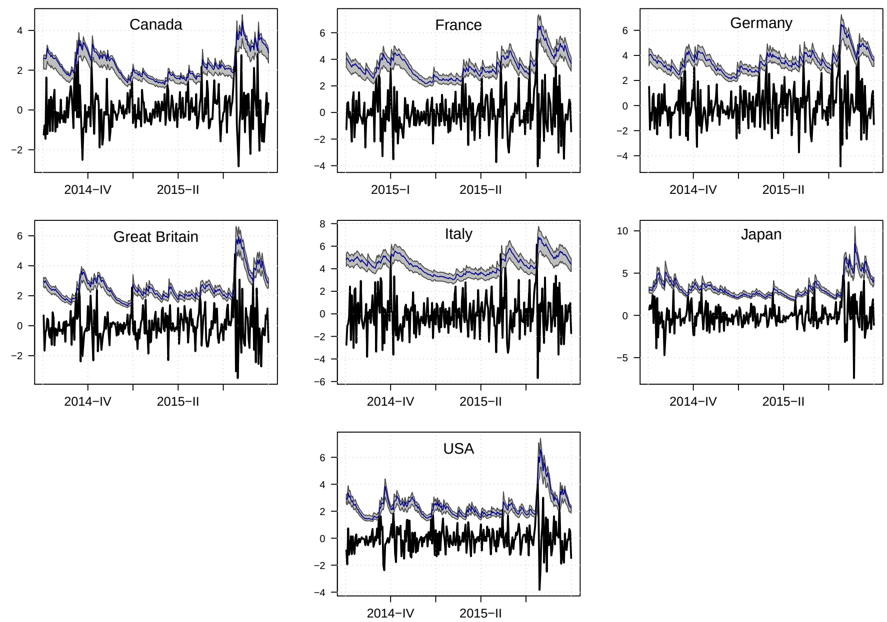


FIGURE 6. 90% confidence intervals for $VaR(99\%)$ of G7 stock market indexes negative returns using Subsample Bootstrap




## 5. CONCLUDING REMARKS

This paper evaluates the performance of different methodologies to estimate the confidence intervals for Value at Risk by using Monte Carlo exercises. The methods evaluated are asymptotic normality, asymptotic Hill estimator, data tilting and subsample bootstrap. The simulation study relies on coverage rates as a measure of performance to find the robustness of the methodologies under some scenarios, such as, mean and variance persistence, VaR probability levels, confidence interval levels and probability distributions.

In general, the subsample bootstrap method presents the best performance of the four evaluated methodologies. As expected, the asymptotic normality approach yields the worst coverage rates, since this assumes *iid* dynamics. It is also important to note that the coverage rates of the four methods studied present a slight decrease when dealing with heavy-tail distributions. Finally, the simulation results show that the four methods that were considered are only valid for high quantiles. In particular, there is a good performance at VaR(99%), in terms of coverage rates, and bad performance for VaR(95%) and VaR(90%).

The empirical exercise confirms the main results found in the simulations. VaR(95%), VaR(99%) and their confidence intervals obtained by the asymptotic normality method do not capture the volatility dynamics of the analyzed series. Additionally, these intervals are considerably wider than the rest of the studied methodologies. On the other hand, confidence intervals for VaR(95%)

computed by data tilting, asymptotic Hill estimator and subsample bootstrap are extremely narrow, which could indicate inconsistency in the estimations, as shown in the simulation exercises. Finally, the VaR(99%) confidence intervals for these three methods present similar dynamics. This fact suggests that any of them can be used for measuring the uncertainty of VaR for the returns of the G7 stock market indexes.

## REFERENCES


CHAN, N. H., S.-J. DENG, L. PENG, AND Z. XIA (2007): "Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations," *Journal of Econometrics*, 137(2), 556–576.

CHRISTOFFERSEN, P. F. (2011): *Elements of financial risk management*. Academic Press.

FRANCQ, C., AND J.-M. ZAKOÏAN (2015): "Risk-parameter estimation in volatility models," *Journal of Econometrics*, 184(1), 158–173.

GAO, F., AND F. SONG (2008): "Estimation risk in GARCH VaR and ES estimates," *Econometric Theory*, 24(5), 1404–1424.

HALL, A., AND Q. YAO (2003): "Data tilting for time series," *Journal of the Royal Statistical Society: Series B*, 65(2), 425–442.

HILL, B. M. (1975): "A simple general approach to inference about the tail of a distribution," *The annals of statistics*, 3(5), 1163–1174.

LECCADITO, A., S. BOFFELLI, AND G. URGA (2014): "Evaluating the accuracy of value-at-risk forecasts: New multilevel tests," *International Journal of Forecasting*, 30(2), 206–216.

MORAUX, F. (2011): "How valuable is your VaR? Large sample confidence intervals for normal VaR," *Journal of Risk Management in Finnacial Institutions*, 4(2), 189–200.

OWEN, A. (1988): "Empirical likelihood ratio confidence intervals for single function," *Biometrika*, 75(2), 967–975.

——— (1990): "Empirical likelihood regions," *Annals of Statistics*, 18(1), 90–120.

——— (2001): *Empirical likelihood*. Chapman and Hall, London.

SPIERDIJK, L. (2014): "Confidence intervals for ARMA-GARCH Value-at-Risk: The case of heavy tails and skewness," *Computational Statistics and Data Analysis, http://dx.doi.org/10.1016/j.csda.2014.08.011*.

## APPENDIX A. SIMULATED COVERAGE RATIOS FOR A SMALL SAMPLE SIZE

TABLE 4. Simulated Coverage Ratios for different AR Parameters and a Sample Size of 500 Observations

| AR Parameters | Method | *iid* Normal | *iid* $t_3$ | *iid* $t_{10}$ | AR-GARCH Normal | AR-GARCH $t_3$ | AR-GARCH $t_{10}$ |
|---|---|---|---|---|---|---|---|
| $\mu = 0$, $\phi_1 = 0$ | Asymptotic Normality | 0.88 | 0.48 | 0.84 | 0.91 | 0.53 | 0.85 |
| | Asymptotic Hill | 0.85 | 0.90 | 0.88 | 0.79 | 0.83 | 0.85 |
| | Data Tilting | 0.67 | 0.80 | 0.69 | 0.67 | 0.78 | 0.59 |
| | Subsample Bootstrap | 0.80 | 0.84 | 0.83 | 0.77 | 0.77 | 0.75 |
| $\mu = 1$, $\phi_1 = 0.1$ | Asymptotic Normality | 0.89 | 0.49 | 0.84 | 0.87 | 0.52 | 0.83 |
| | Asymptotic Hill | 0.96 | 0.94 | 0.96 | 0.90 | 0.89 | 0.92 |
| | Data Tilting | 0.66 | 0.85 | 0.78 | 0.65 | 0.85 | 0.67 |
| | Subsample Bootstrap | 0.83 | 0.84 | 0.82 | 0.88 | 0.79 | 0.88 |
| $\mu = 1$, $\phi_1 = 0.4$ | Asymptotic Normality | 0.90 | 0.47 | 0.84 | 0.64 | 0.57 | 0.64 |
| | Asymptotic Hill | 0.98 | 0.98 | 0.98 | 0.94 | 0.94 | 0.95 |
| | Data Tilting | 0.66 | 0.90 | 0.73 | 0.67 | 0.79 | 0.64 |
| | Subsample Bootstrap | 0.82 | 0.84 | 0.82 | 0.07 | 0.30 | 0.16 |
| $\mu = 1$, $\phi_1 = 0.9$ | Asymptotic Normality | 0.90 | 0.49 | 0.84 | 0.24 | 0.27 | 0.27 |
| | Asymptotic Hill | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 1.00 |
| | Data Tilting | 0.69 | 0.79 | 0.73 | 0.63 | 0.78 | 0.67 |
| | Subsample Bootstrap | 0.83 | 0.84 | 0.82 | 0.04 | 0.04 | 0.05 |

Simulated coverage rates for 90% confidence interval of one-step ahead VaR(99%). The DGP related to columns 3 through 5 (*iid*) is $r_t \overset{iid}{\sim} F$ with mean $\mu/(1-\phi_1)$ and variance $c/(1-b_1-a_1)$, with $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$, where $F$ is normal, Student's t with 3 degrees of freedom or Student's t with 10 degrees of freedom. Meanwhile, the DGP related to the last three columns (AR-GARCH) is $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ and $\sigma_t^2 = c + b_1\varepsilon_{t-1}^2 + a_1\sigma_{t-1}^2$ with $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$. The standardized errors of the AR-GARCH model are distributed as normal, Student's t with 3 degrees of freedom and Student's t with 10 degrees of freedom.

TABLE 5. Simulated Coverage Ratios for different GARCH Parameters and a Sample Size of 500 Observations

| GARCH Parameters | Method | *iid* | | | AR-GARCH | | |
|---|---|---|---|---|---|---|---|
| | | Normal | $t_3$ | $t_{10}$ | Normal | $t_3$ | $t_{10}$ |
| $c = 0.1$, $b_1 = 0.05$, $a_1 = 0.92$ | Asymptotic Normality | 0.89 | 0.49 | 0.84 | 0.87 | 0.52 | 0.83 |
| | Asymptotic Hill | 0.96 | 0.94 | 0.96 | 0.90 | 0.89 | 0.92 |
| | Data Tilting | 0.66 | 0.85 | 0.78 | 0.65 | 0.85 | 0.67 |
| | Subsample Bootstrap | 0.83 | 0.84 | 0.82 | 0.88 | 0.79 | 0.88 |
| $c = 0.1$, $b_1 = 0.005$, $a_1 = 0.99$ | Asymptotic Normality | 0.88 | 0.49 | 0.82 | 0.87 | 0.53 | 0.84 |
| | Asymptotic Hill | 0.88 | 0.93 | 0.91 | 0.91 | 0.93 | 0.93 |
| | Data Tilting | 0.73 | 0.86 | 0.70 | 0.62 | 0.83 | 0.63 |
| | Subsample Bootstrap | 0.84 | 0.83 | 0.82 | 0.88 | 0.82 | 0.83 |
| $c = 0.1$, $b_1 = 0.3$, $a_1 = 0.5$ | Asymptotic Normality | 0.89 | 0.50 | 0.85 | 0.57 | 0.40 | 0.52 |
| | Asymptotic Hill | 1.00 | 0.99 | 1.00 | 1.00 | 0.99 | 1.00 |
| | Data Tilting | 0.62 | 0.78 | 0.65 | 0.67 | 0.74 | 0.60 |
| | Subsample Bootstrap | 0.78 | 0.81 | 0.78 | 0.62 | 0.66 | 0.68 |

Simulated coverage rates for 90% confidence interval of one-step ahead VaR(99%) . The DGP related to columns 3 through 5 (*iid*) is $r_t \overset{iid}{\sim} F$ with mean $\mu/(1-\phi_1)$ and variance $c/(1-b_1-a_1)$, with $\mu = 1$ and $\phi_1 = 0.1$, where $F$ is normal, Student's t with 3 degrees of freedom or Student's t with 10 degrees of freedom. Meanwhile, the DGP related to the last three columns (AR-GARCH) is $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ and $\sigma_t^2 = c + b_1 \varepsilon_{t-1}^2 + a_1 \sigma_{t-1}^2$ with $\mu = 1$ and $\phi_1 = 0.1$. The standardized errors of the AR-GARCH model are distributed as normal, Student's t with 3 degrees of freedom and Student's t with 10 degrees of freedom.

TABLE 6. Simulated Coverage Ratios for different VaR and Confidence Levels, and a Sample Size of 500 Observations

| $VaR(\alpha)$ $C.I.(\beta)$ | Method | *iid* Normal | $t_3$ | $t_{10}$ | AR-GARCH Normal | $t_3$ | $t_{10}$ |
|---|---|---|---|---|---|---|---|
| | Asymptotic Normality | 0.89 | 0.49 | 0.84 | 0.87 | 0.52 | 0.83 |
| $\alpha = 0.99$ | Asymptotic Hill | 0.96 | 0.94 | 0.96 | 0.90 | 0.89 | 0.92 |
| $\beta = 0.90$ | Data Tilting | 0.66 | 0.85 | 0.78 | 0.65 | 0.85 | 0.67 |
| | Subsample Bootstrap | 0.83 | 0.84 | 0.82 | 0.88 | 0.79 | 0.88 |
| | Asymptotic Normality | 0.88 | 0.60 | 0.84 | 0.85 | 0.61 | 0.83 |
| $\alpha = 0.95$ | Asymptotic Hill | 0.63 | 0.72 | 0.66 | 0.55 | 0.67 | 0.63 |
| $\beta = 0.90$ | Data Tilting | 0.67 | 0.79 | 0.71 | 0.65 | 0.57 | 0.55 |
| | Subsample Bootstrap | 0.91 | 0.91 | 0.92 | 0.68 | 0.53 | 0.69 |
| | Asymptotic Normality | 0.89 | 0.63 | 0.86 | 0.87 | 0.67 | 0.83 |
| $\alpha = 0.90$ | Asymptotic Hill | 0.14 | 0.21 | 0.19 | 0.14 | 0.16 | 0.14 |
| $\beta = 0.90$ | Data Tilting | 0.69 | 0.78 | 0.63 | 0.57 | 0.62 | 0.71 |
| | Subsample Bootstrap | 0.91 | 0.92 | 0.92 | 0.58 | 0.43 | 0.56 |
| | Asymptotic Normality | 0.94 | 0.56 | 0.90 | 0.93 | 0.58 | 0.91 |
| $\alpha = 0.99$ | Asymptotic Hill | 0.99 | 0.98 | 0.99 | 0.97 | 0.95 | 0.97 |
| $\beta = 0.95$ | Data Tilting | 0.76 | 0.89 | 0.83 | 0.65 | 0.89 | 0.76 |
| | Subsample Bootstrap | 0.89 | 0.91 | 0.89 | 0.94 | 0.85 | 0.93 |
| | Asymptotic Normality | 0.98 | 0.69 | 0.96 | 0.98 | 0.76 | 0.96 |
| $\alpha = 0.99$ | Asymptotic Hill | 1.00 | 0.98 | 1.00 | 0.99 | 0.99 | 1.00 |
| $\beta = 0.99$ | Data Tilting | 0.92 | 0.97 | 0.94 | 0.90 | 0.94 | 0.84 |
| | Subsample Bootstrap | 0.96 | 0.96 | 0.96 | 0.97 | 0.93 | 0.98 |

Simulated coverage rates for $\beta$ confidence interval of one-step ahead VaR($\alpha$) . The DGP related to columns 3 through 5 (*iid*) is $r_t \overset{iid}{\sim} F$ with mean $\mu/(1-\phi_1)$ and variance $c/(1-b_1-a_1)$, with $\mu = 1$, $\phi_1 = 0.1$, $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$, where $F$ is normal, Student's t with 3 degrees of freedom or Student's t with 10 degrees of freedom. Meanwhile, the DGP related to the last three columns (AR-GARCH) is $r_t = \mu + \phi_1 r_{t-1} + \varepsilon_t$ and $\sigma_t^2 = c + b_1\varepsilon_{t-1}^2 + a_1\sigma_{t-1}^2$ with $\mu = 1$, $\phi_1 = 0.1$, $c = 0.1$, $b_1 = 0.05$ and $a_1 = 0.92$. The standardized errors of the AR-GARCH model are distributed as normal, Student's t with 3 degrees of freedom and Student's t with 10 degrees of freedom.

## APPENDIX B. RESIDUALS AND DIAGNOSTIC TESTS OF THE AR-GARCH MODEL FOR G7 STOCK MARKET INDEXES RETURNS

FIGURE 7. Standardized residuals of the AR-GARCH model for G7 stock market indexes negative returns

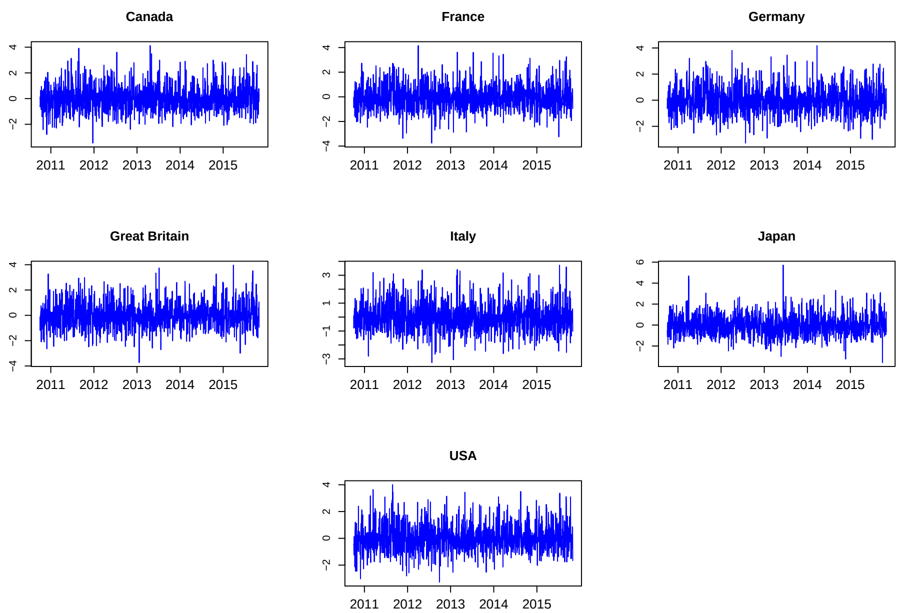



TABLE 7. Sign Bias Test

| | Canada | | France | | Germany | | Great Britain | | Italy | | Japan | | USA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | t-stat | p-value | t-stat | p-value | t-stat | p-value | t-stat | p-value | t-stat | p-value | t-stat | p-value | t-stat | p-value |
| Sign Bias | 1.12 | (0.26) | 0.11 | (0.91) | 0.57 | (0.57) | 0.60 | (0.55) | 0.01 | (0.99) | 1.27 | (0.20) | 2.14 | (0.03) |
| Negative Sign Bias | 1.39 | (0.17) | 2.35 | (0.02) | 2.25 | (0.02) | 2.08 | (0.04) | 2.25 | (0.02) | 1.84 | (0.07) | 1.44 | (0.15) |
| Positive Sign Bias | 0.46 | (0.64) | 1.16 | (0.25) | 1.35 | (0.18) | 1.70 | (0.09) | 1.08 | (0.28) | 2.05 | (0.04) | 0.96 | (0.34) |

TABLE 8. Ljung-Box Test for the Standardized Residuals

| | Canada | | France | | Germany | | Great Britain | | Italy | | Japan | | USA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value |
| 20 | 11.60 | (0.93) | 14.79 | (0.79) | 14.02 | (0.83) | 11.53 | (0.93) | 16.57 | (0.68) | 21.76 | (0.35) | 18.97 | (0.52) |
| 50 | 48.87 | (0.52) | 42.01 | (0.78) | 36.46 | (0.92) | 41.10 | (0.81) | 49.78 | (0.48) | 42.97 | (0.75) | 51.12 | (0.43) |
| 100 | 79.95 | (0.93) | 99.26 | (0.50) | 89.21 | (0.77) | 106.82 | (0.30) | 88.21 | (0.79) | 74.24 | (0.97) | 103.56 | (0.38) |

TABLE 9. Ljung-Box Test for the Squared Standardized Residuals

| | Canada | | France | | Germany | | Great Britain | | Italy | | Japan | | USA | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value | $\chi^2$-stat | p-value |
| 20 | 29.59 | (0.08) | 18.71 | (0.54) | 12.80 | (0.89) | 17.46 | (0.62) | 22.69 | (0.30) | 32.60 | (0.04) | 28.72 | (0.09) |
| 50 | 63.29 | (0.10) | 32.54 | (0.97) | 36.92 | (0.92) | 54.79 | (0.30) | 51.53 | (0.41) | 61.5 | (0.13) | 49.77 | (0.48) |
| 100 | 116.11 | (0.13) | 88.36 | (0.79) | 81.41 | (0.91) | 130.35 | (0.02) | 92.38 | (0.69) | 98.22 | (0.53) | 93.27 | (0.67) |

## APPENDIX C. BACKTESTING TESTS

TABLE 10. Leccadito Test for VaR(95%) and VaR(99%)

| | Independence | | Unconditional Coverage | | Conditional Coverage | |
|---|---|---|---|---|---|---|
| | $\chi^2_1$-stat | p-value | $\chi^2_1$-stat | p-value | $\chi^2_2$-stat | p-value |
| **Asymptotic Normality** | | | | | | |
| Canada | 5.93 | (0.43) | 10.39 | (0.01) | 16.32 | (0.01) |
| France | 7.74 | (0.26) | 3.69 | (0.16) | 11.43 | (0.08) |
| Germany | 15.33 | (0.02) | 5.50 | (0.06) | 20.83 | (0.00) |
| Great Britain | 8.70 | (0.19) | 5.64 | (0.06) | 14.34 | (0.03) |
| Italy | 4.08 | (0.67) | 3.64 | (0.16) | 7.71 | (0.26) |
| Japan | 4.69 | (0.58) | 16.73 | (0.00) | 21.42 | (0.00) |
| USA | 6.48 | (0.37) | 9.49 | (0.01) | 15.97 | (0.01) |
| **Asymptotic Hill** | | | | | | |
| Canada | 4.92 | (0.55) | 5.11 | (0.08) | 10.03 | (0.12) |
| France | 5.33 | (0.50) | 0.50 | (0.78) | 5.83 | (0.44) |
| Germany | 6.01 | (0.42) | 3.28 | (0.19) | 9.29 | (0.16) |
| Great Britain | 4.37 | (0.63) | 2.04 | (0.36) | 6.41 | (0.38) |
| Italy | 3.69 | (0.72) | 0.20 | (0.90) | 3.89 | (0.69) |
| Japan | 4.00 | (0.68) | 7.74 | (0.02) | 11.73 | (0.07) |
| USA | 6.40 | (0.38) | 1.96 | (0.38) | 8.36 | (0.21) |
| **Data Tiling** | | | | | | |
| Canada | 4.82 | (0.57) | 4.04 | (0.13) | 8.86 | (0.18) |
| France | 5.33 | (0.50) | 0.50 | (0.78) | 5.83 | (0.44) |
| Germany | 7.67 | (0.26) | 1.58 | (0.45) | 9.25 | (0.16) |
| Great Britain | 4.88 | (0.56) | 2.08 | (0.35) | 6.96 | (0.32) |
| Italy | 3.75 | (0.71) | 0.50 | (0.78) | 4.25 | (0.64) |
| Japan | 8.29 | (0.22) | 3.10 | (0.21) | 11.40 | (0.08) |
| USA | 6.40 | (0.38) | 1.96 | (0.38) | 8.36 | (0.21) |
| **Subsample Bootstrap** | | | | | | |
| Canada | 4.10 | (0.66) | 2.88 | (0.24) | 6.98 | (0.32) |
| France | 5.33 | (0.50) | 0.50 | (0.78) | 5.83 | (0.44) |
| Germany | 7.67 | (0.26) | 1.58 | (0.45) | 9.25 | (0.16) |
| Great Britain | 2.96 | (0.81) | 6.43 | (0.04) | 9.39 | (0.15) |
| Italy | 3.69 | (0.72) | 0.20 | (0.90) | 3.89 | (0.69) |
| Japan | 6.65 | (0.35) | 2.88 | (0.24) | 9.53 | (0.15) |
| USA | 6.40 | (0.38) | 1.96 | (0.38) | 8.36 | (0.21) |

## APPENDIX D. 90% CONFIDENCE INTERVALS FOR $VaR(95\%)$

FIGURE 8. 90% confidence intervals for $VaR(95\%)$ of G7 stock market indexes negative returns using Asymptotic Normality

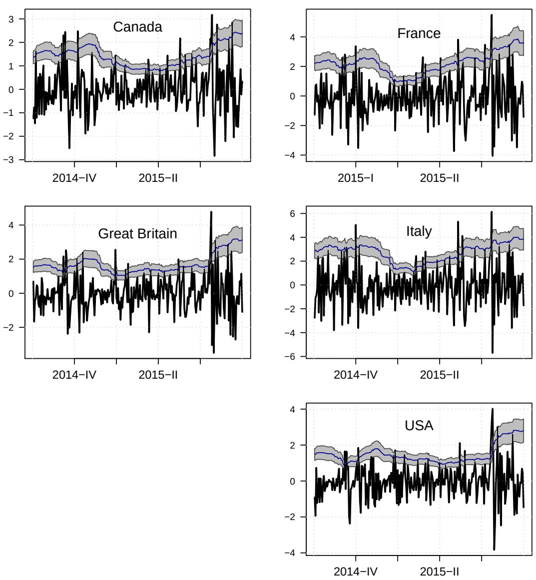



FIGURE 9. 90% confidence intervals for $VaR(95\%)$ of G7 stock market indexes negative returns using Asymptotic Hill Estimator

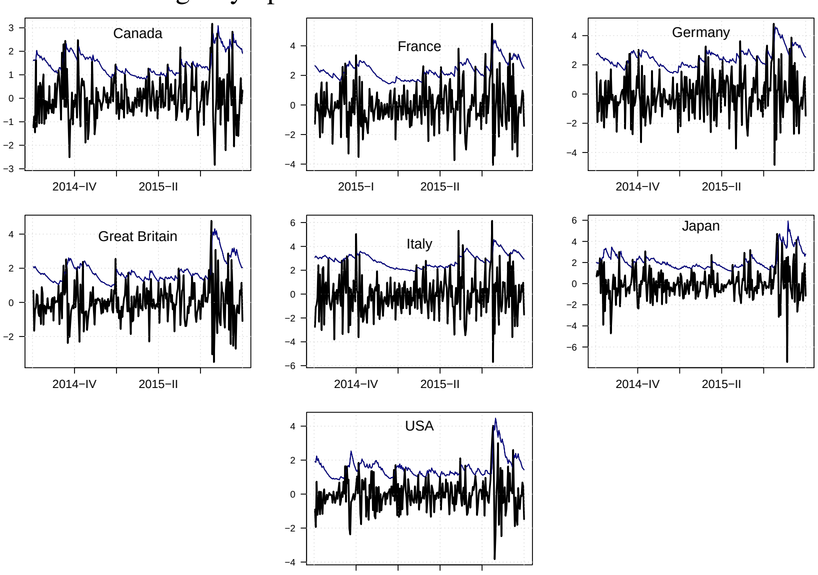


FIGURE 10. 90% confidence intervals for $VaR(95\%)$ of G7 stock market indexes negative returns using Data Tilting

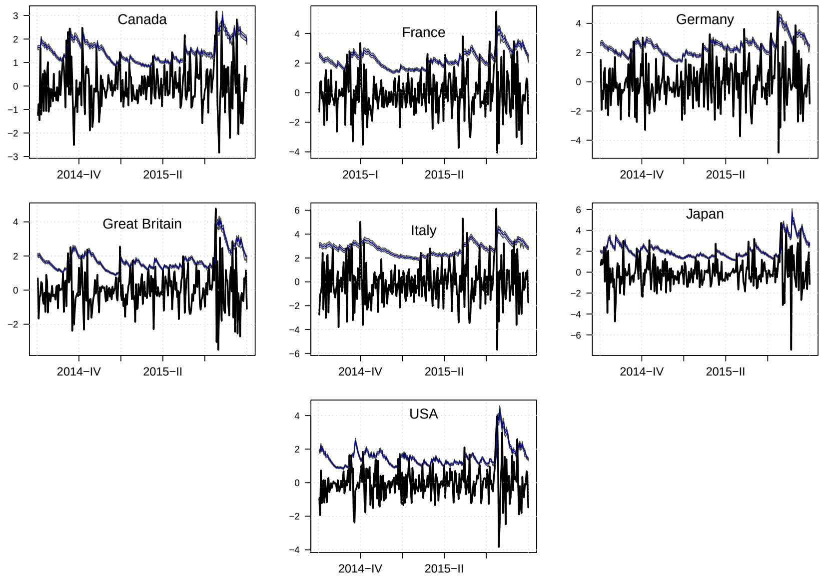



FIGURE 11. 90% confidence intervals for $VaR(95\%)$ of G7 stock market indexes negative returns using Subsample Bootstrap