NONPARAMETRIC ESTIMATION OF CONDITIONAL EXPECTED SHORTFALL

SCAILLET O.* [1]

* HEC Genève and FAME, Université de Genève, 102 Bd Carl Vogt, CH-1211 Genève 4, Suisse. scaillet@hec.unige.ch

This version: May 2004

Abstract

We consider a nonparametric method to estimate conditional expected shortfalls, i.e. conditional expected losses knowing that losses are larger than a given loss quantile. We derive the asymptotic properties of kernel estimators of conditional expected shortfalls in the context of a stationary process satisfying strong mixing conditions. An empirical illustration is given for several stock index returns, namely CAC40, DAX30, S&P500, DJI, and Nikkei225.

Key words: Nonparametric, Kernel, Time Series, Conditional VaR, Conditional Expected Shortfall, Risk Management, Loss Severity Distribution.

JEL Classification: C14, D81, G10, G21, G22, G28.

[1]We would like to thank the editor and the referee for very constructive criticism, and Marc Henry for many stimulating discussions and proof checking. We have received fruitful comments from H. Hong, O. Renault, M. Rockinger, as well as participants to AFFI meeting, Conference of the Swiss society for financial market research, CEMAF finance meeting, Geneva Research Collaboration seminar, and to seminars at BNP Paribas, University of Melbourne and University of Amsterdam. The author gratefully acknowledges financial support from the Belgian Program on Interuniversity Poles of Attraction (PAI nb. P4/01) and from the Swiss National Science Foundation through the National Center of Competence in Research: Financial Valuation and Risk Management. Part of this research was done when he was visiting THEMA and IRES.

1 Introduction

Most financial institutions are now routinely using risk management systems to adequately control their risks or to suitably allocate their capital. This has been impulsed by either internal requirements (efficient use of capital invested by shareholders, development of new business lines) or external constraints (Capital Adequacy Requirement of the Basle Committee on Banking Supervision, prudential rules imposed by European or American regulators on financial institutions). In this context Value at Risk (VaR) and expected shortfall (ES) have become essential tools to assess riskiness of trading activities (see e.g. JP Morgan (1996), Wilson (1996), Duffie and Pan (1997), Jorion (1997), Dowd (1998), Stulz (1998) for a detailed analysis and applications in risk management). From a formal point of view, VaR is simply a quantile of the loss distribution over a prescribed holding period, while ES is the expected loss knowing that the loss is above VaR.

Decisions concerning the reserve amount and the way of allocating the capital to cover adverse market movements are directly linked to such risk measures. Appropriate risk measurement tools should therefore be allowed to adapt to varying market conditions, and to reflect the latest available information in a non-i.i.d. framework. Up to now most of the risk management literature has focused on marginal VaR and marginal ES, i.e. risk measures referring to marginal or stationary distributional features. [2] These are especially suited for a long term view, but can be less valuable for short term adjustments.

In this paper our main concern is to propose conditional tools that will ideally complement marginal measures. Hence we develop estimators of conditional VaR and conditional ES in a nonparametric framework when the conditioning information is made of past observed returns. Our approach is similar in spirit to conditional parametric approaches for VaR based on GARCH modelling of financial return series (see e.g. Alexander and Leigh (1997), Boudoukh, Richardson and Whitelaw (1997), McNeil and Frey

[2] For time dependent data marginal VaR and marginal ES are defined with respect to the stationary distribution of the losses. We use the term conditional to refer to risk measures defined with respect to conditional distributions instead of stationary distributions. In an i.i.d. setting we do not need to make a difference between marginal and conditional risk measures.

(2000), or Barone-Adesi, Bourgoin, and Giannopoulos (1998)). It is however more flexible and permits to capture other market features than only changing volatility backgrounds (for another flexible, but parametric, approach, see the CAViaR model of Engle and Manganelli (1999)). We propose to enlarge the setting enough to support a broad class of dependence structures, namely strong mixing. [3] The estimation procedure based on a kernel approach is extremely fast and easy to implement. It basically only requires the standard functionalities of any spreadsheet used in financial statement reporting.

This paper extends the nonparametric analysis proposed in a marginal framework by Gouriéroux, Laurent and Scaillet (2000) for VaR, and by Scaillet (2004) for ES. In fact our estimator for the conditional VaR is based on inverting a Nadaraya-Watson type estimator [4] of the conditional distribution function for time series data. It is akin to the estimator studied by Cai (2002) (see the references therein for other proposals [5]). The latter relies however on a smoothing in the direction of the lagged values only and not in all directions. Hence by construction our estimator provides smoother estimates, which are better suited for graphical purposes. Asymptotically and for interior points, i.e. quantiles not to far in the tails, properties of both estimators coincide. For boundary points, the estimator of Cai (2002) exhibits better asymptotic properties in the sense of avoiding the so-called boundary bias. His estimator involves a heavier computational burden since it requires estimating and optimizing additional weights aimed to correct for this bias. Performance of both estimators should be close in practice because sample sizes

[3] This type of mixing conditions, also called α-mixing, is standard for time series since strong mixing is a very weak condition on the dependence structure. Strong mixing sequences encompass dynamics generated by most commonly used parametric models, see Doukhan (1994) for several examples, such as standard ARMA models. GARCH models and stochastic volatility models can also be shown to be strong mixing under some conditions (Carrasco and Chen (2002)).

[4] As noticed by Cai (2002), this type of estimator always produces conditional distribution function estimates which lie between zero and one, and are monotone increasing. This is particularly advantageous if an inversion of the conditional distribution estimator is used to deliver an estimator of a conditional quantile.

[5] Note also that our paper is the first to consider nonparametric estimation of conditional ES.

are large for financial data [6] and the quantiles used in market risk management cannot be considered as extremes.

Let us also remark that relying on VaR for risk measurement purposes has been recently challenged by Artzner, Delbaen, Eber and Heath (1999) since VaR fails to be subbadditive. This subadditivity property expresses the idea that the total risk on a portfolio should not be greater than the sum of the individual risks, and is part of the necessary requirements to be a coherent measure of risk in the sense of Artzner, Delbaen, Eber and Heath (1999). ES can be shown to be a coherent measure of risk for continuous distributions. [7] We further refer the reader to Acerbi and Tasche (2002) for an illuminating and precise mathematical discussion of several risk measures related to what we call ES. Note also that another disadvantage of VaR is that it tells us nothing about the potential size of the loss that exceeds it, while ES does.

Let us finally stress that when we speak about conditional VaR, we refer to a VaR computed with respect to a conditional distribution. In the risk management literature "conditional VaR" (see e.g. Rockafellar and Uryassev (2002)) sometimes refer to what we call ES. Since it seems that there does not exist a universally agreed terminology we prefer to use ES to avoid the somewhat odd term "conditional conditional VaR" in place of conditional ES. We could also have used TailVaR instead of ES as in Artzner, Delbaen, Eber and Heath (1999). The term ES seems to be more in line with actuarial practice.

The paper is organised as follows. In Section 2 we introduce the basic notations and concepts underlying our framework. In Section 3 we give the form of the kernel estimators for conditional VaR and ES, and provide the asymptotic distributions of these nonparametric estimators. An empirical application on several stock index returns, namely CAC40, DAX30, S&P500, DJI, and Nikkei225, is provided in Section 4. Concluding remarks are presented in Section 5. Mathematical developments are gathered in an appendix.

[6] This allows to use small values for the smoothing parameter and thus reduce the boundary area.

[7] On the contrary Acerbi and Tasche (2002) find that simply taking a conditional expectation of losses beyound VaR can fail to yield a coherent measure of risk when there are discontinuities in the loss distribution. Assumptions underlying our smoothing approach rule out this possibility.

2 Framework

We consider a real-valued strictly stationary process $\{Y_t, t \in \mathbb{Z}\}$ and assume that our data consist in a realization of $\{Y_t; t = 1, ..., T\}$. These data may correspond to observed returns at several dates. They may also correspond to simulated values drawn from a parametric model (VARMA, multivariate GARCH or diffusion processes), possibly fitted on another set of data. Simulations are often required when the structure of financial assets is too complex, as for some derivative products and credit sensitive instruments. This, in turn, implies that the sample length T can sometimes be controlled, and asked to be sufficiently large to get satisfying estimation results.

Let us take a positive integer n, and let $0 < \tau_1 < ... < \tau_n$ be integers, so that we define $Z_t = (Y_{t-\tau_1}, ..., Y_{t-\tau_n})'$. In particular we may take $\tau_1 = 1$, ..., $\tau_n = n$, which gives $Z_t = (Y_{t-1}, ..., Y_{t-n})'$. In the following we define conditional objects w.r.t. Z_t, i.e. lagged returns. We denote by $f(y, z)$, $F(y, z)$, the marginal p.d.f. and c.d.f. of $(Y_t, Z_t')'$, while the conditional p.d.f. and c.d.f. are written $f(y|z)$, and $F(y|z)$, respectively.

The conditional VaR knowing that past returns Z_t are equal to $\zeta \in \mathbb{R}^n$ is formally defined by the equality

$$P[-Y_t > VaR(\zeta, p)|Z_t = \zeta] = p, \tag{1}$$

where p is the loss probability. The conditional VaR is a function of the loss probability, which typically ranges from 1% to 5%, while stock returns are usually measured over a one day period or a ten day period. Expression (1) which can be rewritten

$$P[Y_t < -VaR(\zeta, p)|Z_t = \zeta] = F(-VaR(\zeta, p)|\zeta) = p,$$

gives the relationship

$$Q(\zeta, p) = -VaR(\zeta, p), \tag{2}$$

between the quantile $Q(\zeta, p)$ of the conditional distribution of Y_t and $VaR(\zeta, p)$.

The conditional ES is defined as the conditional expected loss knowing that the loss is above the conditional VaR:

$$E[-Y_t| - Y_t > VaR(\zeta, p), Z_t = \zeta] = E[-Y_t|Y_t < Q(\zeta, p), Z_t = \zeta] \equiv m(Q(\zeta, p), \zeta). \tag{3}$$

3 Kernel estimators of conditional VaR and ES

We start with the definition of the kernel estimators before moving to their asymptotic distribution.

3.1 Definition

From (2) and (3), we see that we are in fact interested in estimating conditional quantiles, and conditional expectations knowing that we are below conditional quantiles.

For a quantile of order $p \in (0,1)$, we assume that the cumulative distribution function $F(.|\zeta_i)$ of Y_t given Z_t at distinct points $\zeta_i \in \mathbb{R}^n$, $i = 1, \ldots, d$, is such that the equation $F(y|\zeta_i) = p$ admits a unique solution for each of the ζ_i denoted $Q(\zeta_i, p)$.

For conditional expectations, we look at the quantities

$$m(Q(\zeta_i, p), \zeta_i) f(\zeta_i) p \equiv -E[Y_t | Y_t < Q(\zeta_i, p), Z_t = \zeta_i] f(\zeta_i) p \tag{4}$$

at distinct points $\zeta_i \in \mathbb{R}^n$, $i = 1, \ldots, d$, where f denotes the p.d.f. of Z_t.

Let $k_{ij}(u)$ be a real bounded and symmetric function on $\mathbb{R}$ such that

$$\int k_{ij}(u) du = 1, \quad i = 1, \ldots, d, \quad j = 1, \ldots, n,$$

and

$$K_i(u; h^{(i)}) = \Pi_{j=1}^n k_{ij}(u_j / h_{ij}), \quad i = 1, \ldots, d,$$

where $h^{(i)}$ is a diagonal matrix with elements $(h_{ij})_{j=1}^n$ and determinant $|h^{(i)}|$ (for a scalar x, $|x|$ will denote its absolute value), while the bandwidths h_{ij} are positive functions of T such that

$$|h^{(i)}| + (T|h^{(i)}|)^{-1} \to 0 \text{ when } T \to \infty.$$

In addition, let l and h or l_i and h_i, $i = 1, \ldots, d$, satisfy the same conditions as any of the k_{ij} and h_{ij}.

The p.d.f. of Z_t at ζ_i, i.e. $f(\zeta_i)$, will be estimated by

$$[1; \zeta_i] = (T|h^{(i)}|)^{-1} \sum_{t=1+\tau_n}^{T} K_i(\zeta_i - Z_t; h^{(i)}),$$

while the p.d.f. of (Y_t, Z_t) at (ξ_j, ζ_i), i.e. $f(\xi_j, \zeta_i)$, will be estimated by

$$[1; \xi_j, \zeta_i] = (Th_j|h^{(i)}|)^{-1} \sum_{t=1+\tau_n}^{T} l_j(h_j^{-1}(\xi_j - Y_t))K_i(\zeta_i - Z_t; h^{(i)}).$$

Hence, estimators of the conditional cumulative distribution of Y_t given $Z_t = \zeta_i$ at distinct points ξ_j, $j = 1, \ldots, d$, are obtained as

$$\hat{F}(\xi_j|\zeta_i) = \int_{-\infty}^{\xi_j} [1; u, \zeta_i]du/[1; \zeta_i] \equiv \hat{\phi}(\xi_j, \zeta_i)/[1; \zeta_i], \tag{5}$$

calling $\phi(\xi_j, \zeta_i) = \int_{-\infty}^{\xi_j} f(u, \zeta_i)du$. The first derivatives with respect to ξ of $\hat{F}(\xi|\zeta_i)$ will be denoted by $\hat{f}(\xi|\zeta_i)$. Based on (5) and taking $h_j = h$, $l_j = l$, the conditional quantile $Q(\zeta_i, p)$, namely $-VaR(\zeta_i, p)$, can be estimated by

$$\hat{Q}(\zeta_i, p) = \inf_{y \in \mathbb{R}} \left\{ y : \hat{F}(y|\zeta_i) \geq p \right\}.$$

Finally, to estimate the conditional expectation of $-Y_t$ given $Y_t < Q(\zeta_i, p)$ and $Z_t = \zeta_i$, we will need the following estimate of (4):

$$-\int_{-\infty}^{\hat{Q}(\zeta_i,p)} [Y_t; u, \zeta_i]du = -(Th|h^{(i)}|)^{-1} \int_{-\infty}^{\hat{Q}(\zeta_i,p)} \sum_{t=1+\tau_n}^{T} Y_t l(h^{-1}(u - Y_t))K_i(\zeta_i - Z_t; h^{(i)})du,$$

so that the conditional ES can be estimated by

$$\hat{m}(\hat{Q}(\zeta_i, p), \zeta_i) = -\int_{-\infty}^{\hat{Q}(\zeta_i,p)} [Y_t; u, \zeta_i]du/([1; \zeta_i]p) \equiv -\bar{\phi}(\hat{Q}(\zeta_i, p), \zeta_i)/([1; \zeta_i]p).$$

If a single Gaussian kernel $k_{ij}(u) = \varphi(u)$ and a single bandwidth h are adopted, we simply get:

$$\hat{m}(\hat{Q}(\zeta_i, p), \zeta_i) = \frac{-\sum_{t=1+\tau_n}^{T} Y_t \Phi\left(\frac{\hat{Q}(\zeta_i, p) - Y_t}{h}\right) \prod_{j=1}^{n} \varphi\left(\frac{\zeta_{ij} - Y_{t-\tau_j}}{h}\right)}{p \sum_{t=1+\tau_n}^{T} \prod_{j=1}^{n} \varphi\left(\frac{\zeta_{ij} - Y_{t-\tau_j}}{h}\right)},$$

where φ and Φ denote the p.d.f. and c.d.f. of a standard Gaussian variable, respectively.

3.2 Asymptotic distribution

The asymptotic normality of kernel estimators for conditional VaR and ES can be established under suitable conditions on the kernel, the asymptotic behavior of the bandwidth, the regularity of the conditional expectations and densities, and some mixing properties of the process. Below we mainly follow the presentation of Robinson (1983) (see e.g. Bierens (1985) or Bosq (1998) for alternative sets of assumptions).

Assumption 1 (kernel and bandwidth)

(a) Bandwidths satisfy $|h^{(i)}|||h^{(i)}||^4 T \to 0$.

(b) Kernels and bandwidths satisfy $|k_{ij}(u)| \leq C(1+|u|)^{-(1+\omega_i/n)}$, *and* $||h^{(i)}||^{n+\omega_i-2} \leq C|h^{(i)}|$, $\omega_i > 2$.

Assumption 2 (process)

(a) The process (Y_t) *is strong mixing with coefficients* α_j *such that* $\sum_{l=N}^{\infty} \alpha_l^{1-2/\theta} = O(N^{-1})$, *as* $N \to \infty$, *while* $E|Y_t|^\theta < \infty$, *for some* $\theta > 2$.

(b) For each ζ_i, *we have* $f(\zeta_i) > 0$ *and* $f(Q(\zeta_i,p)|\zeta_i) > 0$.

(c) Second order partial derivatives for the p.d.f. of (Y_t, Z_t) *are continuous in neighbourhoods of all pairs* (ξ_i, ζ_j) *where estimation is performed.*

(d) The p.d.f. of (Z_t, Z_{t+s}) *exists and is bounded in a neighbourhood of all pairs* (ζ_i, ζ_j), $i, j = 1, \ldots, d$, *uniformely in* $s > 1$.

(e) The conditional expectation $m(\xi, z)$ *is twice continuously differentiable at* $\xi = Q(\zeta_i, p)$ *and* $z = \zeta_i$, $i = 1, ..., d$.

(e) The conditional expectation $H(\xi, z) = E[Y_t^2 | Y_t < \xi, Z_t = z]$, *is continuous at* $\xi = Q(\zeta_i, p)$ *and* $z = \zeta_i$, $i = 1, ..., d$.

(f) The conditional expectation $E[|Y_t|^\gamma | Y_t < \xi, Z_t = z]$, *is bounded for* $\gamma > \theta$ *at* $\xi = Q(\zeta_i, p)$ *and* $z = \zeta_i$, $i = 1, ..., d$.

Let S and $\bar{S}$ be the d dimensional vector with components S_i/V_i^2 and $\bar{S}_i/\bar{V}_i^2$, respectively given by:

$$S_i = (T|h^{(i)}|)^{1/2} \left\{ \hat{Q}(\zeta_i, p) - Q(\zeta_i, p) \right\},$$

$$V_i^2 = \frac{p(1-p)\kappa_i}{f(\zeta_i)f(Q(\zeta_i,p)|\zeta_i)^2},$$

and

$$\bar{S}_i = (T|h^{(i)}|)^{1/2}\left\{\hat{m}(\hat{Q}(\zeta_i,p),\zeta_i) - m(Q(\zeta_i,p),\zeta_i)\right\},$$

$$\begin{aligned}\bar{V}_i^2 &= \frac{\kappa_i}{f(\zeta_i)p}\Big(H(Q(\zeta_i,p),\zeta_i) - m(Q(\zeta_i,p),\zeta_i)^2 p \\ &\quad +(1-p)\left(\frac{\bar{m}(Q(\zeta_i,p),\zeta_i)^2}{f(Q(\zeta_i,p)|\zeta_i)^2} + \frac{2\bar{m}(Q(\zeta_i,p),\zeta_i)m(Q(\zeta_i,p),\zeta_i)}{f(Q(\zeta_i,p)|\zeta_i)}\right)\Big),\end{aligned}$$

with $\bar{m}(Q(\zeta_i,p),\zeta_i) = E[Y_t \mathbb{I}_{Y_t=Q(\zeta_i,p)}|Z_t=\zeta_i]$ and $\kappa_i = \int \prod_{j=1}^{n} k_{ij}^2(u)du_j$.

Proposition 1 *Under Assumptions 1 and 2(a)-(d), S converges in distribution to a vector of independent standard normal random variables.*

Proof: see Henry and Scaillet (2000) and the developments in the proof of Proposition 2.

Proposition 2 *Under Assumptions 1 and 2, $\bar{S}$ converges in distribution to a vector of independent standard normal random variables.*

Proof: see Appendix.

Consistent estimates of asymptotic variances V_i^2 and $\bar{V}_i^2$ may be derived after replacement of the various terms by adequate density and conditional moment kernel estimators. However block or local bootstrap procedures (see the review of Bühlmann (2002)) reveal more appropriate in small samples to build pointwise confidence bands.

Finally, let us remark that all asymptotic results will not be affected if simulations of a parametric model are used as input data, when estimated parameters satisfy the usual rate of convergence. The number of simulations could then be chosen large enough to be on the safe side of asymptotic theory.

4 Empirical illustration

This section illustrates the implementation of the estimation procedure described in Section 3. The empirical illustration concerns data on stock index returns. We analyze five major stock indices: CAC40, DAX30, S&P500, DJI, and NIKKEI225. The data are one day returns recorded daily from 03/01/1994 to 07/07/2000, i.e. 1700 observations. Table 1 gathers the summary statistics for these data including the empirical marginal VaR (loss quantile of level p) and the empirical marginal expected shortfall (empirical mean of losses above VaR divided by the probability of occurrence) for a loss probability level p equal to 5%. We may observe that European and Japanese indices are riskier than US indices from a marginal point of view, even if the kurtosis is larger for the latter.

Table 1: Summary statistics of stock index return data

	CAC40	DAX30	S&P500	DJI	Nikkei225
mean	.001	.001	.001	.001	.000
st. dev.	.012	.013	.010	.010	.014
skew.	-.157	-.498	-.440	-.546	.066
kurt.	1.685	3.267	5.591	5.491	3.243
median	.000	.001	.000	.000	.000
VaR	.020	.022	.016	.015	.023
ES	.028	.031	.023	.023	.031

To conduct the nonparametric analysis we have selected a single bandwidth value according to the usual rule of thumb (empirical standard deviation times $T^{-1/5}$) and a single Gaussian kernel.

We start with the European indices, namely CAC40 and DAX30. Figures 1 and 2 report conditional VaR and conditional ES with a 5% loss probability level. The left columns contain estimates for a single conditioning variable equal to a one period lagged return. Pointwise confidence bands at 90% are also provided. They are built from a block bootstrap procedure (KÜNSCH (1989)) with a block length l of 11 data (this length corresponds to the rule of thumb $l = T^{1/3}$ given in BÜHLMANN (2002)). The right columns give estimates with one period and two period lagged returns as conditioning set. The conditioning values are taken between the first and third quartiles of the return data. Outside this range, the estimation becomes very unstable due to the lack of observations. Conditional VaR and conditional ES of the left column are U-shaped for CAC40. Hence the risk tends to be lower when yesterday return is close to the empirical average and larger otherwise. This means that CAC40 is likely to correct strongly when a large positive return occurs the day before, or to fall deeper when a large loss has already occurred the day before. DAX30 does not share this property. The CAC40 U-shape is inverted in the right column when we consider very negative values of the two period lagged return. A third large loss is thus less likely when two losses have already been successively endured. For DAX30 this is not observed, and the second conditioning variable seems to bring less information (flatter surfaces).

Let us now proceed with US indices, namely S&P500 and DJI. Figures 3 and 4 show a similar decreasing behavior in their left column. Hence we do not see a tendency to strongly correct after a large upmove. Both indices differ in their right column. A third large fall of S&P500, resp. DJI, is more, resp. less, likely after two successive downmoves. Besides the more diversified S&P500 exhibits flatter surfaces than the less diversified DJI.

Finally we find on Figure 5 the same U-shape for the Nikkei225 as for CAC40, but not the inversion when we consider very negative values of the two period lagged return.

5 Concluding remarks

In this paper we have proposed simple nonparametric estimation methods of conditional VaR and conditional ES. The estimation procedure relies on a kernel approach in the context of a general stationary strong mixing process. These estimators have been proved to be empirically relevant in the analysis of stock index returns. We think that they complement ideally the existing battery of risk management tools since they allow revealing the very different types of conditional risk structures present in the data. Eventually we should probably have supplemented our nonparametric methodology with some validation method to be fully convincing about its utility. We leave this interesting, but difficult, task for future research since there is no simple and efficient ways to validate statistically a nonparametric model in the context of conditional VaR and conditional ES as opposed to the case of marginal VaR and conditional means.

APPENDIX

Proof of Proposition 2

Let us adopt for a while the compact notations:

$$Q_i = Q(\zeta_i, p), \qquad \phi_i = \phi(Q_i, \zeta_i), \qquad m_i = m(Q_i, \zeta_i), \qquad f_i = f(\zeta_i).$$

Using the expansions

$$\begin{aligned} \bar{\phi}(\hat{Q}_i, \zeta_i) &= \bar{\phi}(Q_i, \zeta_i) + [Y_t; \bar{Q}_i, \zeta_i] \left\{ \hat{Q}_i - Q_i \right\}, \\ F(Q_i|\zeta_i) = p &= \hat{F}(\hat{Q}_i|\zeta_i) \\ &= \hat{F}(Q_i|\zeta_i) + \hat{f}(\bar{Q}_i|\zeta_i) \left\{ \hat{Q}_i - Q_i \right\}, \end{aligned}$$

and

$$\hat{F}(Q_i|\zeta_i) - F(Q_i|\zeta_i) = B_i^{-1} \left\{ \hat{\phi}_i - \phi_i \right\} - A_i B_i^{-2} \left\{ [1; \zeta_i] - f_i \right\},$$

where

$$|\bar{Q}_i - Q_i| \le |\hat{Q}_i - Q_i|, \qquad |B_i - f_i| \le |[1; \zeta_i] - f_i|, \qquad |A_i - \phi_i| \le |\hat{\phi}_i - \phi_i|,$$

we get

$$\begin{aligned} -\bar{\phi}(\hat{Q}_i, \zeta_i)/[1; \zeta_i] - m_i p &= B_i^{-1} \left\{ -\bar{\phi}(Q_i, \zeta_i) - m_i f_i p \right\} + \frac{[Y_t; \bar{Q}_i, \zeta_i]}{\hat{f}(\bar{Q}_i|\zeta_i)} B_i^{-2} \left\{ \hat{\phi}_i - \phi_i \right\} \\ &\quad - \left(\frac{[Y_t; \bar{Q}_i, \zeta_i]}{\hat{f}(\bar{Q}_i|\zeta_i)} A_i B_i^{-3} + C_i B_i^{-2} \right) \left\{ [1; \zeta_i] - f_i \right\}, \end{aligned}$$

where $|C_i - m_i f_i p| \le | - \bar{\phi}(\hat{Q}_i, \zeta_i) - m_i f_i p|$.

We need to show that

$$(T|h^{(i)}|)^{1/2} \left\{ -E\bar{\phi}(Q_i, \zeta_i) - m_i f_i p \right\} \to 0, \tag{6}$$

$$(T|h^{(i)}|)^{1/2} \left\{ E\hat{\phi}_i - \phi_i \right\} \to 0, \tag{7}$$

$$(T|h^{(i)}|)^{1/2} \left\{ E[1; \zeta_i] - f_i \right\} \to 0. \tag{8}$$

We have:

$$E\bar{\phi}(Q_i,\zeta_i) = \int_{\mathbb{R}^{n+1}} \int_{-\infty}^{Q_i} (Th|h^{(i)}|)^{-1}\omega l(h^{-1}(u-\omega))K_i(\zeta_i-\lambda;h^{(i)})\ f(\omega,\lambda)\,du\,d\lambda\,d\omega$$

$$= \int_{\mathbb{R}^{n+1}} \int_{-\infty}^{Q_i} (u-h\omega)l(\omega)K_i(\lambda;1)f(u-h\omega,\zeta_i-h^{(i)}\lambda))\,du\,d\lambda\,d\omega$$

$$= -m_i f_i p + \int_{-\infty}^{Q_i} u \left\{ \frac{h^2}{2} f^{(\xi\xi)}(u,\zeta_i) \int_{\mathbb{R}} \omega^2 l(\omega)\,d\omega \right.$$

$$\left. + \sum_{j=1}^{n} \frac{h_{ij}^2}{2} f_{jj}^{(\zeta\zeta)}(u,\zeta_i) \int_{\mathbb{R}} \omega^2 k_{ij}(\omega)\,d\omega \right\} du + o(h^2 + \max_j(h_{ij}^2))$$

$$= -m_i f_i p + O(h^2 + \max_j(h_{ij}^2)).$$

In the same way,

$$E\hat{\phi}_i = \phi_i + \int_{-\infty}^{Q_i} \left\{ \frac{h^2}{2} f^{(\xi\xi)}(u,\zeta_i) \int_{\mathbb{R}} \omega^2 l(\omega)\,d\omega \right.$$

$$\left. + \sum_{j=1}^{n} \frac{h_{ij}^2}{2} f_{jj}^{(\zeta\zeta)}(u,\zeta_i) \int_{\mathbb{R}} \omega^2 k_{ij}(\omega)\,d\omega \right\} du + o(h^2 + \max_j(h_{ij}^2)),$$

and

$$E[1;\zeta_i] = f_i + \sum_{j=1}^{n} \frac{h_{ij}^2}{2} f_{jj}^{(\zeta\zeta)}(\zeta_i) \int_{\mathbb{R}} \omega^2 k_{ij}(\omega)\,d\omega + o(\max_j(h_{ij}^2)).$$

This proves the three stated results (6)-(8) using $|h^{(i)}|||h^{(i)}||^4 T \to 0$.

Now following the construction in Robinson (1983) Theorem 5.3 step by step, we take $g_t = -Y_t \mathbb{I}_{|Y_t| \le D}$, $\bar{g}_t = -Y_t - g_t$, for some D, $0 < D < \infty$, and also $W_{it} = g_t L_{it}$, $\bar{W}_{it} = \bar{g}_t L_{it}$, with $L_{it} = K_{it} \int_{-\infty}^{Q_i} l(h^{-1}(u-Y_t))du$, $K_{it} = K_i(\zeta_i - Z_t; h^{(i)})$. We further introduce $V_{it} = c_i(W_{it} - EW_{it})$, $V_{i+d,t} = c_{i+d}(L_{it} - EL_{it})$, $V_{i+2d,t} = c_{i+2d}(K_{it} - EK_{it})$, where c_i, c_{i+d}, c_{i+2d}, $i = 1, ..., d$, are arbitrary constants. Then consider $S_T = (T|h^{(i)}|)^{-1/2} \sum_{t=1}^{T} \sum_{j=1}^{3d} V_{jt}$. Since $V[(T|h^{(i)}|)^{-1/2} \sum_{t=1}^{T} \bar{W}_{it}] \to 0$ as $D \to \infty$, uniformly in large T, we only have to show that S_T has, for fixed D, an asymptotic normal distribution.

First let us put $\sigma_{ij} = \lim_{T\to\infty} E[V_{it}V_{jt}]/(|h^{(i)}||h^{(j)}|)^{1/2}$, $i, j = 1, ..., 3d$. We have

$$\sigma_{ii} = c_i^2 \kappa_i H_i^D f_i p, \qquad \sigma_{i+d,i+d} = c_{i+d}^2 \kappa_i \phi_i, \qquad \sigma_{i+2d,i+2d} = c_{i+2d}^2 \kappa_i f_i,$$

$$\sigma_{i,i+d} = c_i c_{i+d} \kappa_i m_i^D f_i p, \qquad \sigma_{i,i+2d} = c_i c_{i+2d} \kappa_i m_i^D f_i p, \qquad \sigma_{i+d,i+2d} = c_{i+d} c_{i+2d} \kappa_i \phi_i,$$

for $i = 1, ..., d$, where $m_i^D = E[-Y_t \mathbb{1}_{|Y_t|\le D} | Y_t < Q_i, Z_t = \zeta_i]$, $H_i^D = E[Y_t^2 \mathbb{1}_{|Y_t|\le D} | Y_t < Q_i, Z_t = \zeta_i]$, and $\sigma_{ij} = 0$, otherwise.

Second the continuity of m_i and H_i implies the continuity of m_i^D and H_i^D. Hence the central limit theorem of lemma 7.1 in Robinson (1983) valid for bounded V_{it} may be applied to S_T. We get the final stated result of the proposition since $m_i^D \to m_i$, $H_i^D \to H_i$, by letting $D \to \infty$, by using the convergence in probability of

$$A_i \to \phi_i, \qquad B_i \to f_i, \qquad C_i \to m_i f_i p,$$

$$\bar{Q}_i \to Q_i, \qquad [Y_t; \bar{Q}_i, \zeta_i] \to \bar{m}_i f_i, \qquad \hat{f}(\bar{Q}_i|\zeta_i) \to f(Q_i|\zeta_i),$$

where $\bar{m}_i = E[Y_t \mathbb{1}_{Y_t = Q_i} | Z_t = \zeta_i]$, and by computing

$$\begin{aligned}
&\begin{pmatrix} \dfrac{1}{f_i} & \dfrac{\bar{m}_i f_i}{f(Q_i|\zeta_i) f_i^2} & -\dfrac{\bar{m}_i f_i \phi_i}{f(Q_i|\zeta_i) f_i^3} - \dfrac{m_i f_i p}{f_i^2} \end{pmatrix} \\
&\qquad\qquad \begin{pmatrix} H_i f_i p & m_i f_i p & m_i f_i p \\ m_i f_i p & \phi_i & \phi_i \\ m_i f_i p & \phi_i & f_i \end{pmatrix} \\
&\qquad\qquad\qquad \begin{pmatrix} \dfrac{1}{f_i} & \dfrac{\bar{m}_i f_i}{f(Q_i|\zeta_i) f_i^2} & -\dfrac{\bar{m}_i f_i \phi_i}{f(Q_i|\zeta_i) f_i^3} - \dfrac{m_i f_i p}{f_i^2} \end{pmatrix}' \\
&= \frac{H_i p}{f_i} - \frac{m_i^2 p^2}{f_i} + \left(1 - \frac{\phi_i}{f_i}\right) \left(\frac{\bar{m}_i^2 \phi_i}{(f(Q_i|\zeta_i) f_i)^2} + \frac{2 \bar{m}_i m_i p}{f(Q_i|\zeta_i) f_i} \right), \\
&= \frac{p}{f_i} \left(H_i - m_i^2 p + (1-p) \left(\frac{\bar{m}_i^2}{f(Q_i|\zeta_i)^2} + \frac{2 \bar{m}_i m_i}{f(Q_i|\zeta_i)} \right) \right),
\end{aligned}$$

from $\phi_i / f_i = p$

REFERENCES :

- Acerbi C. and D. Tasche (2002) : "On the Coherence of Expected Shortfall", *Journal of Banking and Finance*, **26**, 1487-1503.
- Alexander C. and C. Leigh (1997) : "On the Covariance Matrices Used in Value at Risk Models", *Journal of Derivatives*, **4**, 50-62.
- Artzner, P., F. Delbaen, J.M. Eber and D. Heath (1999): "Coherent Measures of Risk", *Mathematical Finance*, **9**, 203-228.
- Barone-Adesi, G. F. Bourgoin, and K. Giannopoulos (1998) : "Don't look back", *RISK*, **11**, 8.
- Bierens H. (1985) : "Kernel Estimators of Regression Functions", *Advances in Econometrics*, Cambridge University Press, Cambridge, 99-144.
- Bosq, D. (1998): "Nonparametric Statistics for Stochastic Processes: Estimation and Prediction", Lecture Notes in Statistics, Springer-Verlag, New-York.
- Boudoukh J. M. Richardson, and R. Whitelaw (1997) : "Investigation of a Class of Volatility Estimators", *Journal of Derivatives*, **4**, 63-62.
- Bühlmann, P. (2002): "Bootstraps for Time Series", *Statistical Science*, 17, 52-72.
- Cai, Z. (2002): "Regression Quantiles for Time Series", *Econometric Theory*, **18**, 169-192.
- Carrasco M., and X. Chen (2002) : "Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models", *Econometric Theory*, **18**, 17-39.
- Dowd, K. (1998) : "Beyond Value at Risk : The New Science of Risk Management", Wiley, Chechester.
- Duffie D., and J. Pan (1997) : "An Overview of Value at Risk", *Journal of Derivatives*, **4**, 7-49.

- Doukhan, P. (1994): "Mixing: Properties and Examples", Lecture Notes in Statistics, Springer-Verlag, New-York.
- Engle R., and S. Manganelli (1999): "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles", DP UCSD.
- Gouriéroux C., J.P. Laurent and O. Scaillet (2000): "Sensitivity Analysis of Values at Risk", *Journal of Empirical Finance*, 7, 225-245.
- Henry M., and O. Scaillet (2000): "Nonparametric Specification Analysis of Dynamic Parametric Models", DP IRES.
- Jorion P. (1997): "Value at Risk : The New Benchmark for Controlling Market Risk", Irwin, Chicago.
- JP Morgan (1996): "Risk Metrics - Technical Document", 4 th edition, New-York, J.P. Morgan.
- Künsch H. (1989): "The Jackknife and the Bootstrap for General Stationary observations", *Annals of Statistics*, **17**, 1217-1241.
- McNeil A. and R. Frey (2000): "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series : an Extreme Value Approach", *Journal of Empirical Finance*, **7**, 271-300.
- Robinson P. (1983) : "Nonparametric Estimators for Time Series", *Journal of Time Series Analysis*, **4**, 185-207. 1
- Rockafellar R. and S. Uryassev (2000): "Optimization of Conditional Value-at-Risk", *Journal of Risk*, **2**, 21-41.
- Scaillet O. (2004): "Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall", *Mathematical Finance*, **14**, 115-129.
- Stulz R. (1998) : "Derivatives, Risk Management, and Financial Engineering", Southwestern Publishing.

- Wilson T.C. (1996) : “Calculating Risk Capital”, in Alexander, Carol (ed), The Handbook of Risk Management and Analysis, Chichester, Wiley, 193-232.

Figure 1 : Conditional measures of risk for the CAC40 index

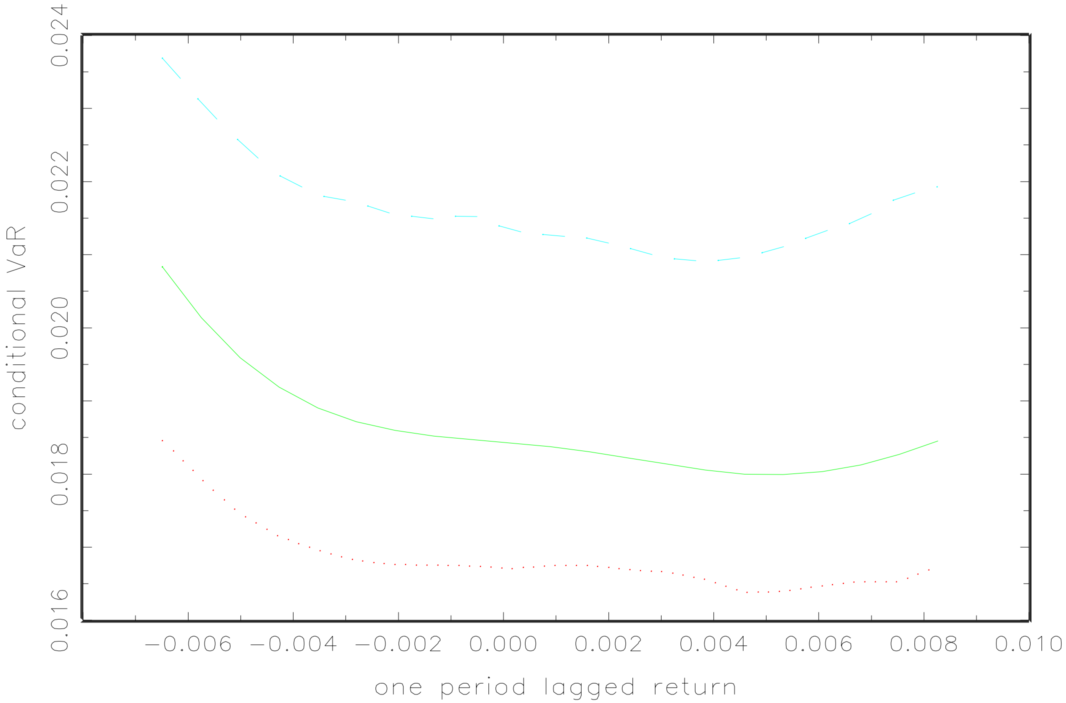
Figure 2 : Conditional measures of risk for the DAX30 index

Figure 3 : Conditional measures of risk for the S&P500 index

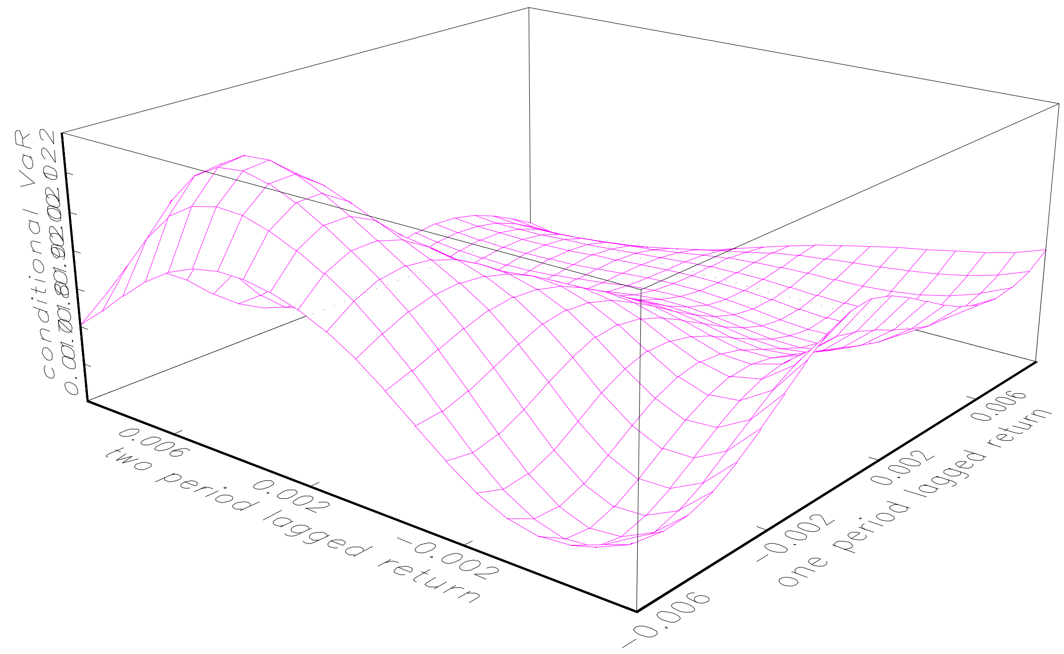
Figure 4 : Conditional measures of risk for the DJI index

Figure 5 : Conditional measures of risk for the Nikkei225 index

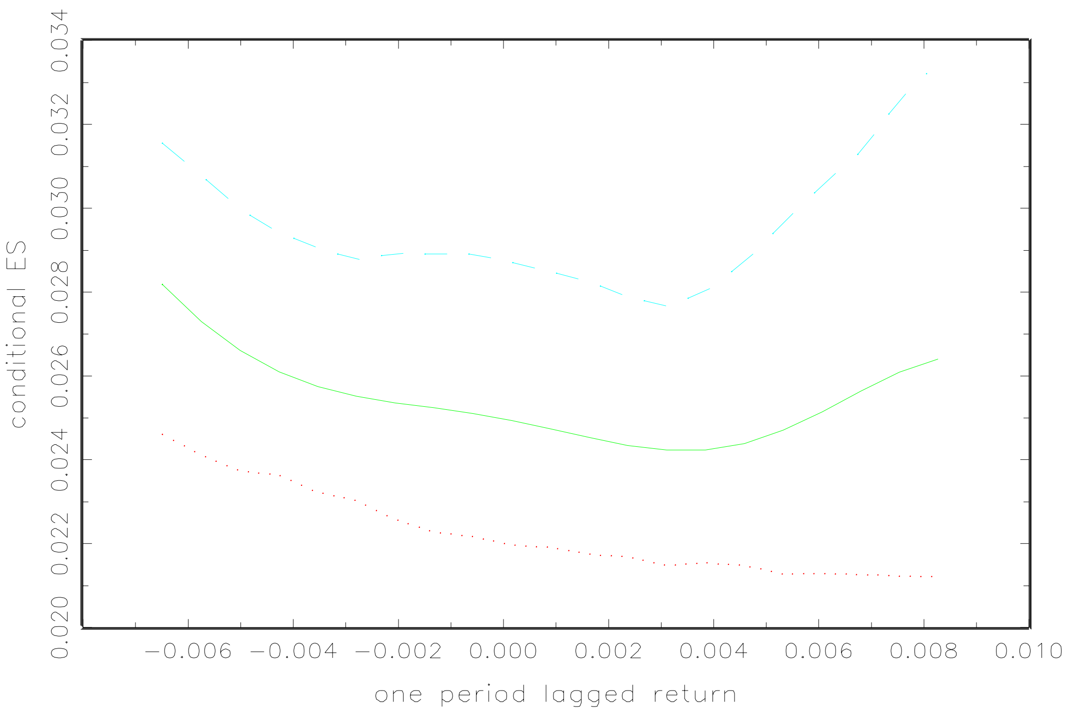
CONDITIONAL VaR



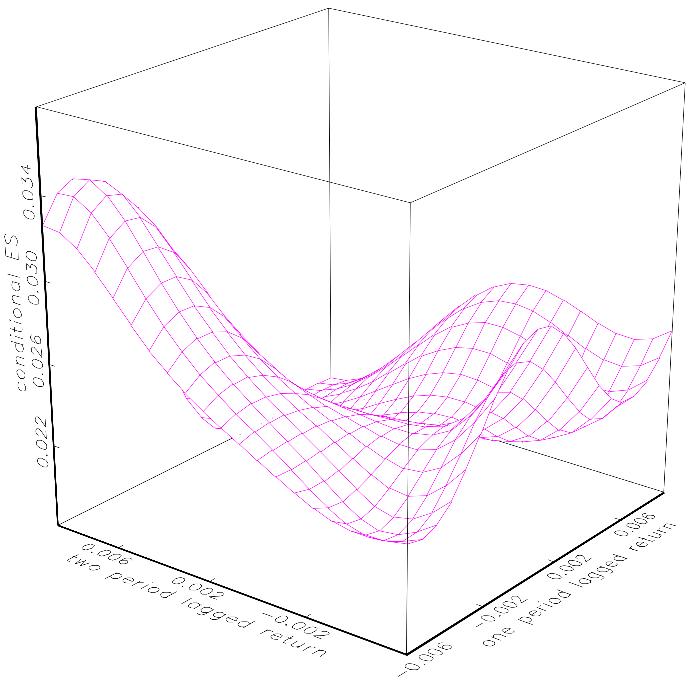
CONDITIONAL VaR

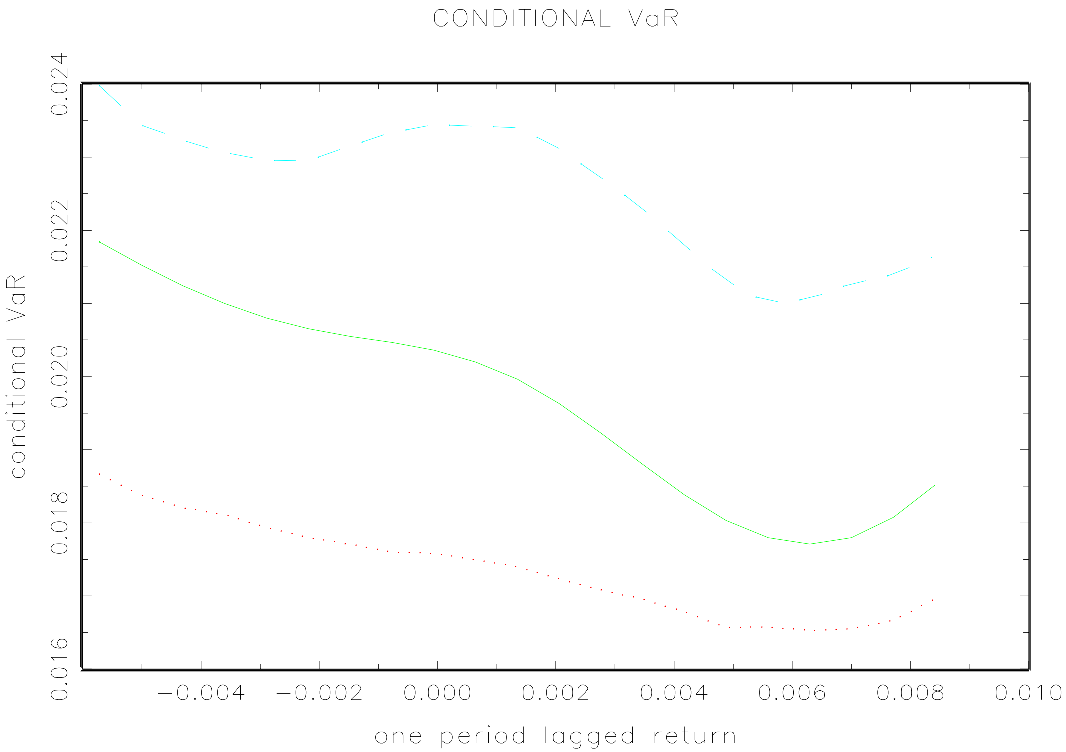


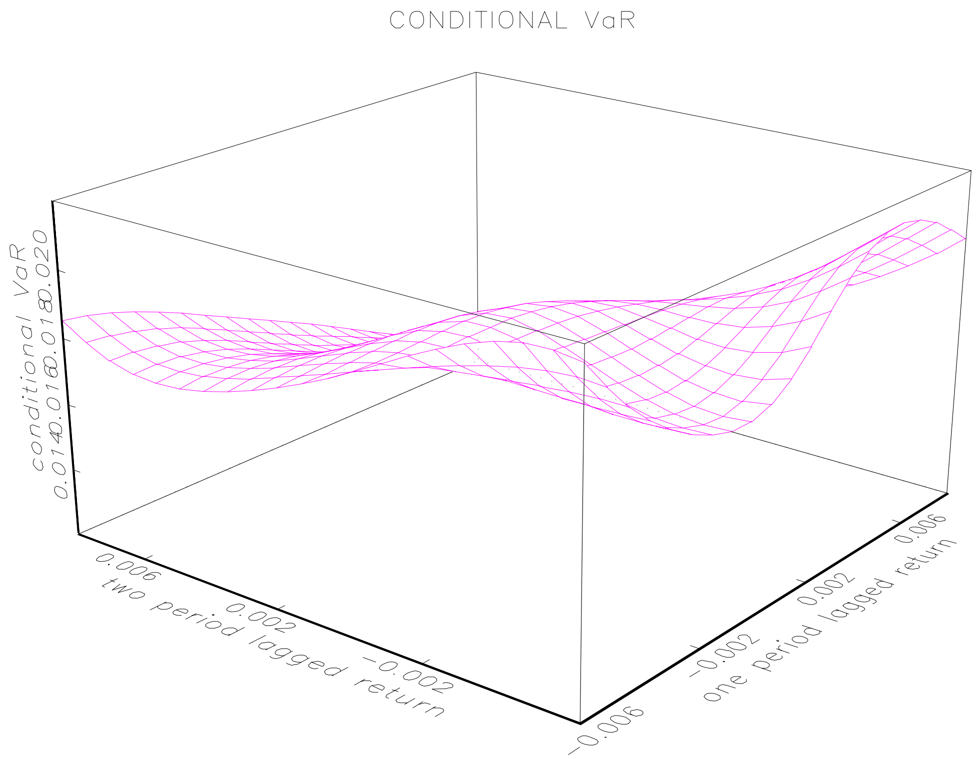
CONDITIONAL EXPECTED SHORTFALL

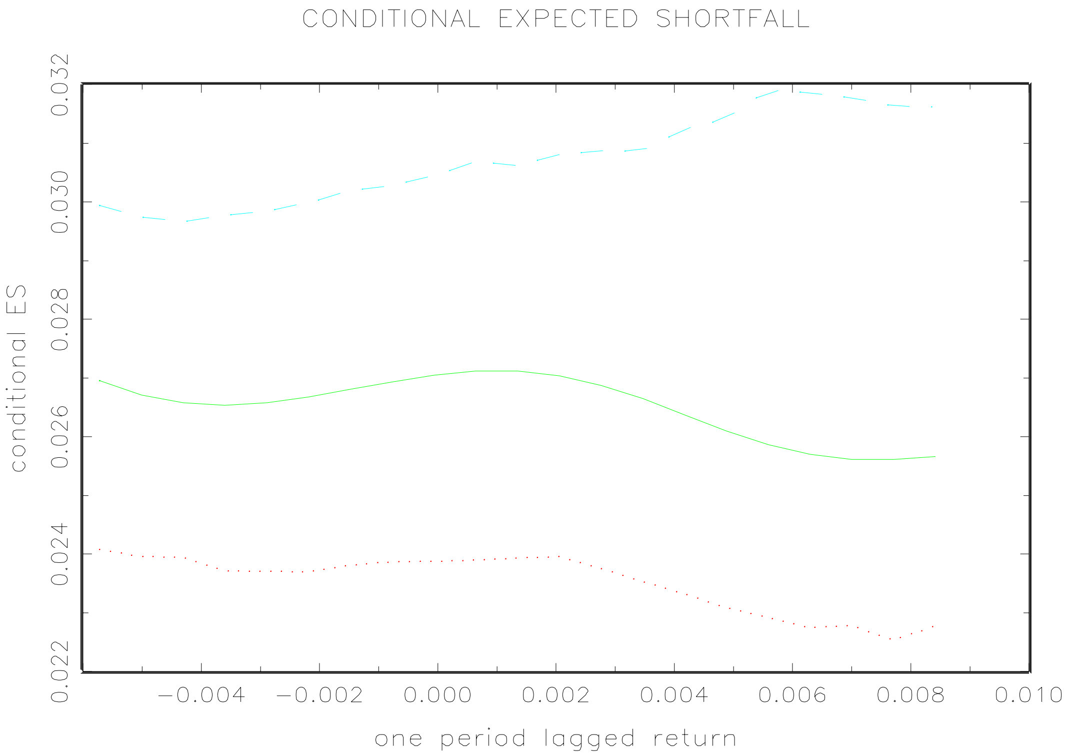


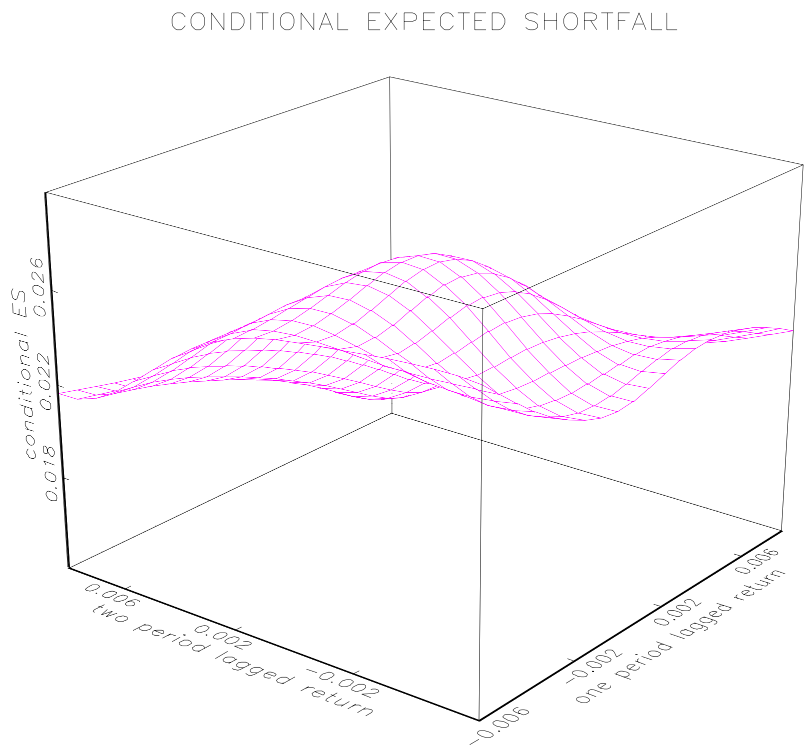
CONDITIONAL EXPECTED SHORTFALL

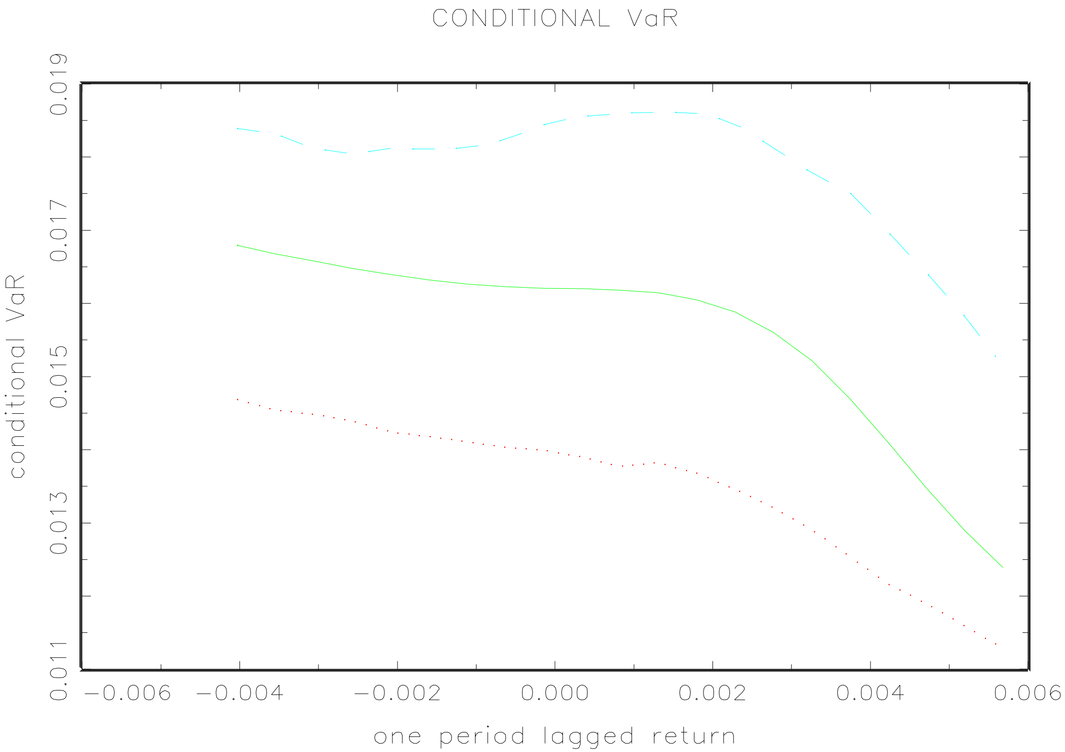


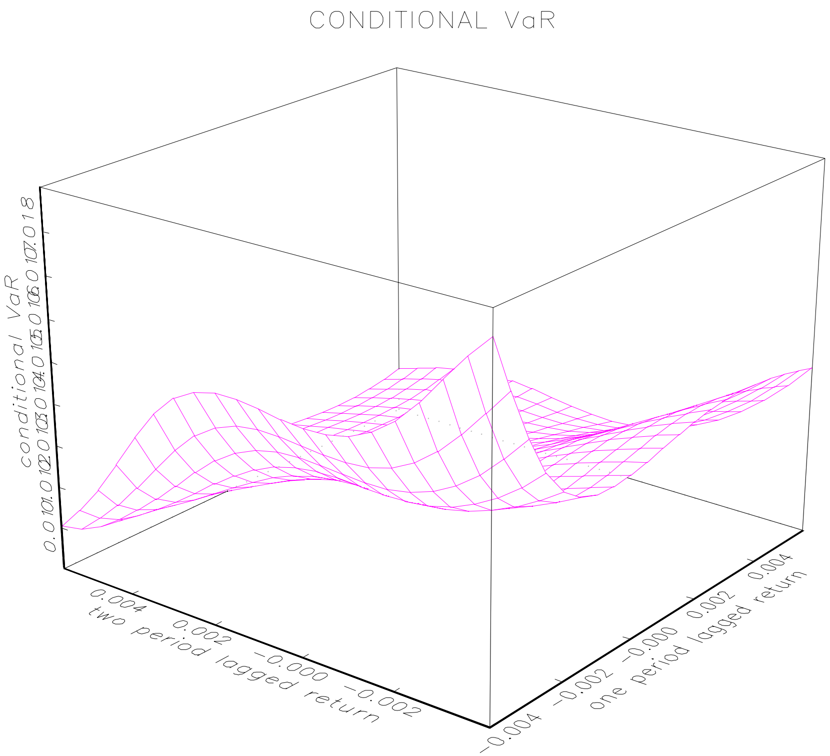

CONDITIONAL VaR
conditional VaR
0.016
0.018
0.020
0.022
0.024
−0.004
−0.002
0.000
0.002
0.004
0.006
0.008
0.010
one period lagged return

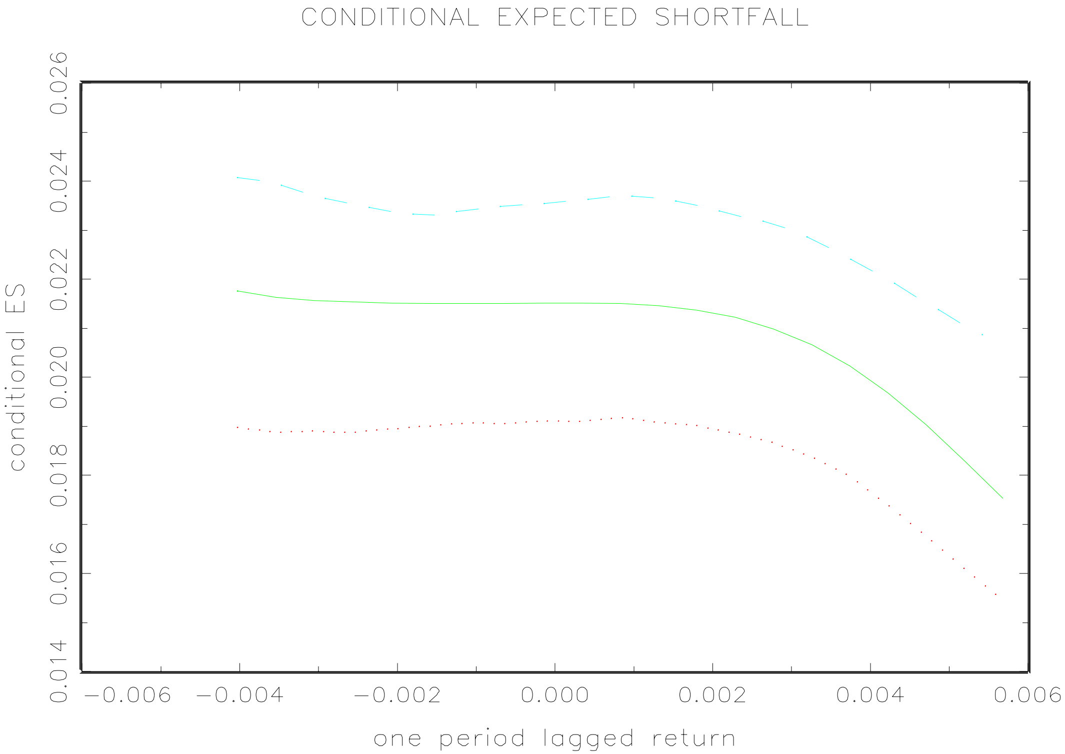

CONDITIONAL VaR
conditional VaR
0.006
0.002
−0.002
two period lagged return
−0.006
−0.002
0.002
0.006
one period lagged return

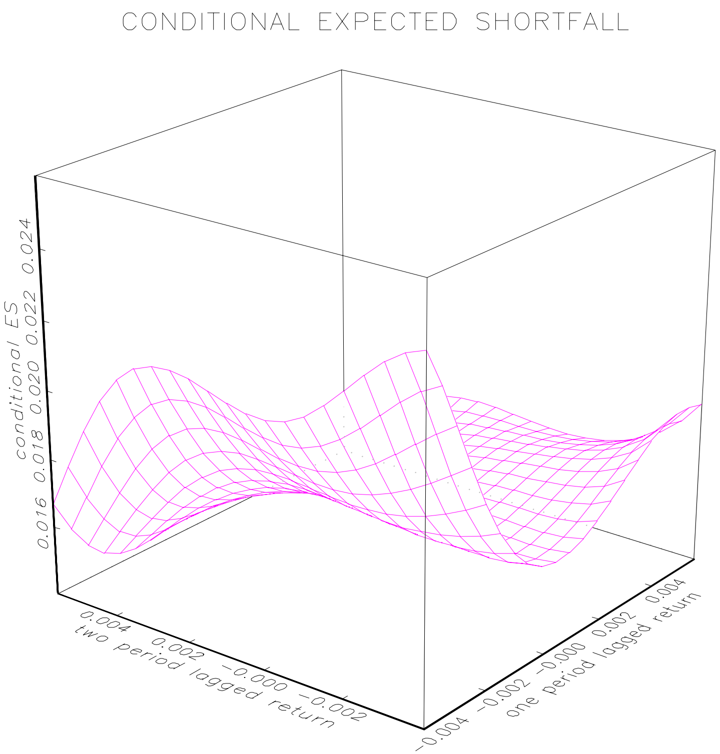

CONDITIONAL EXPECTED SHORTFALL
conditional ES
0.022
0.024
0.026
0.028
0.030
0.032
−0.004
−0.002
0.000
0.002
0.004
0.006
0.008
0.010
one period lagged return

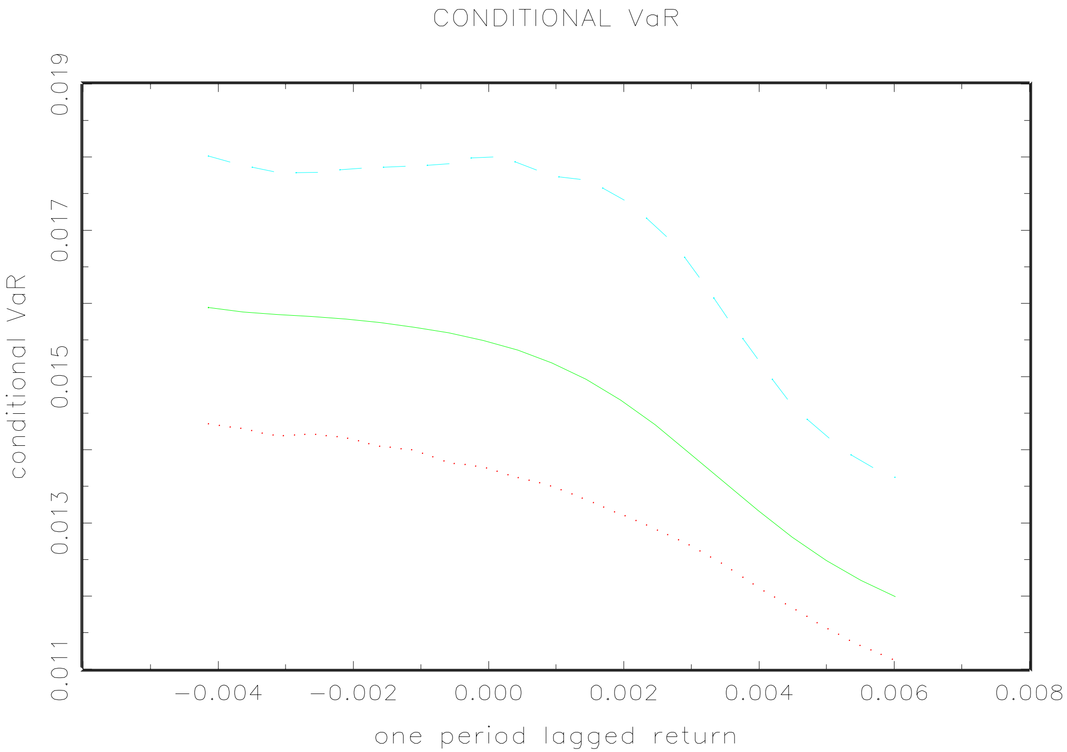

CONDITIONAL EXPECTED SHORTFALL
conditional ES
0.018
0.022
0.026
0.006
0.002
−0.002
two period lagged return
−0.006
−0.002
0.002
0.006
one period lagged return

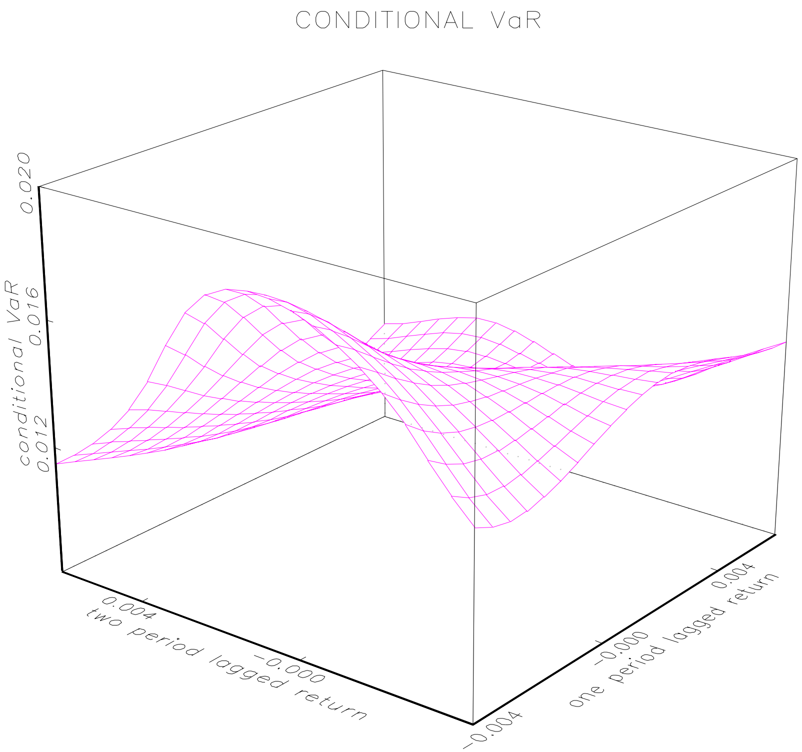

CONDITIONAL VaR
conditional VaR
0.011
0.013
0.015
0.017
0.019
−0.006
−0.004
−0.002
0.000
0.002
0.004
0.006
one period lagged return

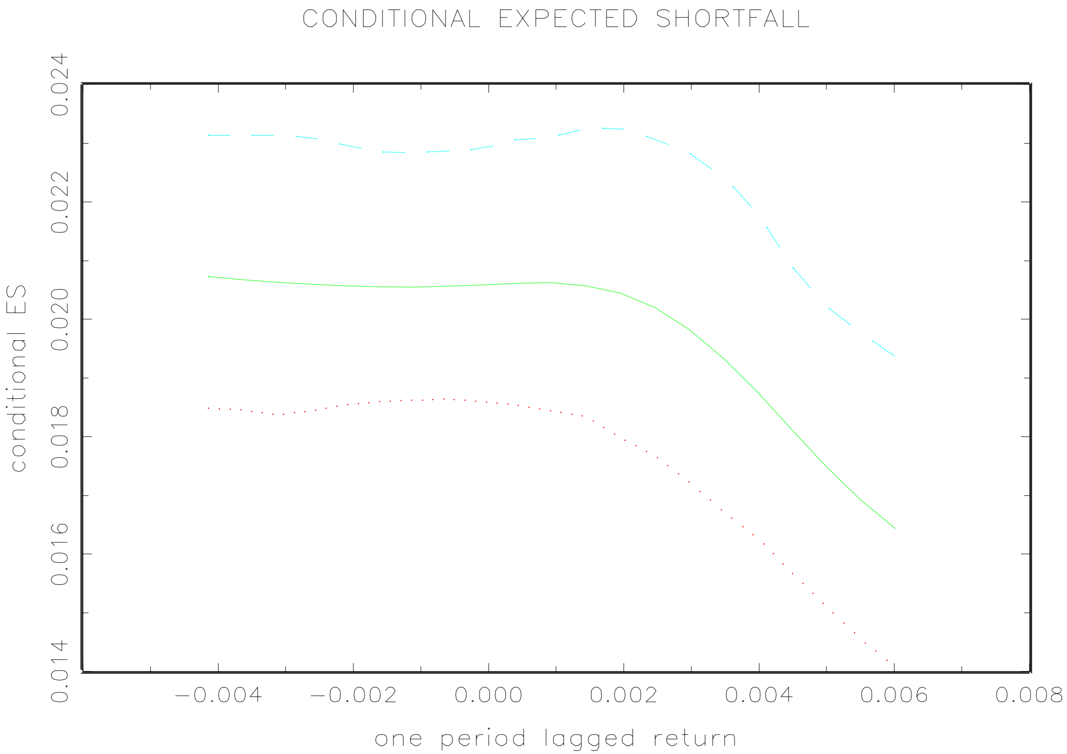

CONDITIONAL VaR
conditional VaR
two period lagged return
one period lagged return
0.004
0.002
−0.000
−0.002
−0.004

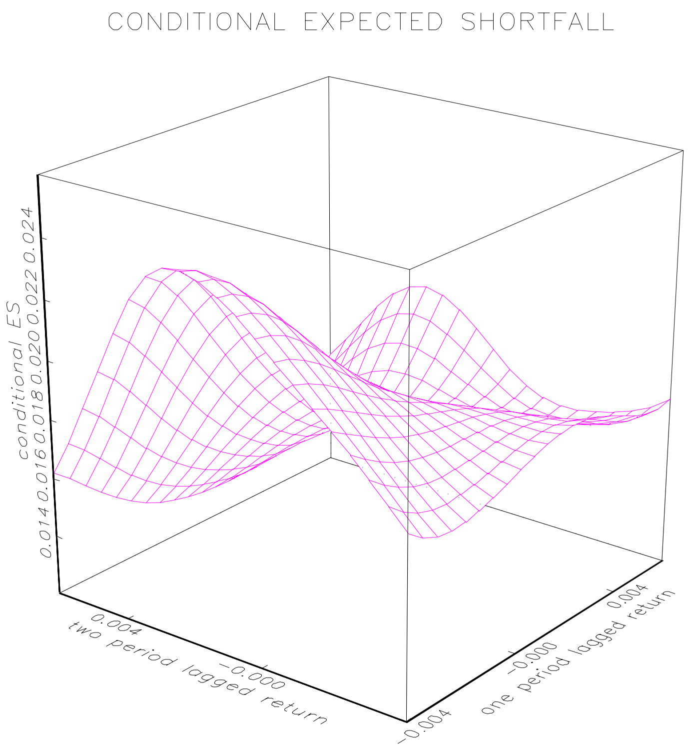

CONDITIONAL EXPECTED SHORTFALL
conditional ES
0.014
0.016
0.018
0.020
0.022
0.024
0.026
−0.006
−0.004
−0.002
0.000
0.002
0.004
0.006
one period lagged return


CONDITIONAL EXPECTED SHORTFALL
conditional ES
0.016
0.018
0.020
0.022
0.024
two period lagged return
one period lagged return
0.004
0.002
−0.000
−0.002
−0.004

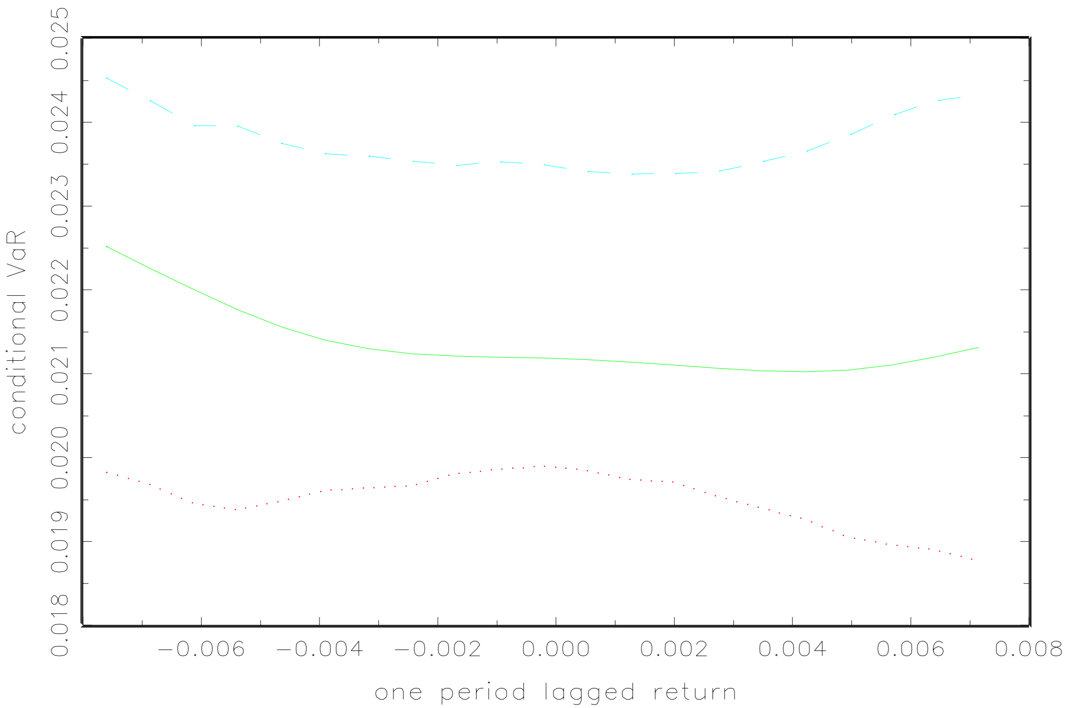

CONDITIONAL VaR
conditional VaR
0.011
0.013
0.015
0.017
0.019
−0.004
−0.002
0.000
0.002
0.004
0.006
0.008
one period lagged return


CONDITIONAL VaR
conditional VaR
0.012
0.016
0.020
0.004
two period lagged return
−0.000
−0.004
one period lagged return

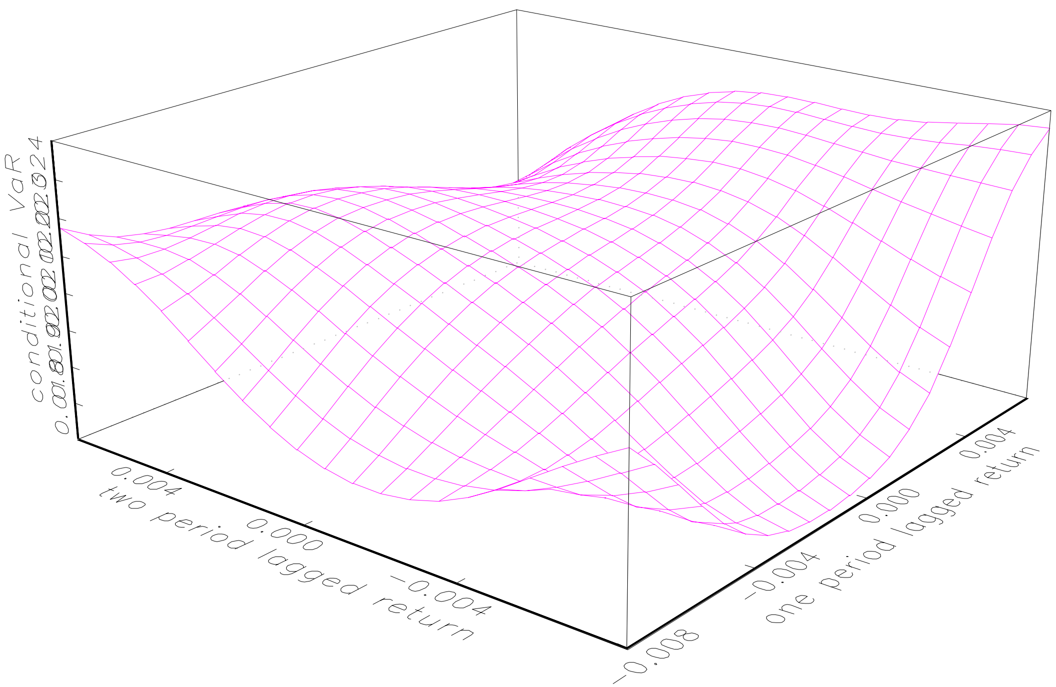

CONDITIONAL EXPECTED SHORTFALL
conditional ES
0.014
0.016
0.018
0.020
0.022
0.024
−0.004
−0.002
0.000
0.002
0.004
0.006
0.008
one period lagged return


CONDITIONAL EXPECTED SHORTFALL
conditional ES
0.014 0.016 0.018 0.020 0.022 0.024
0.004
two period lagged return
−0.000
−0.004
one period lagged return

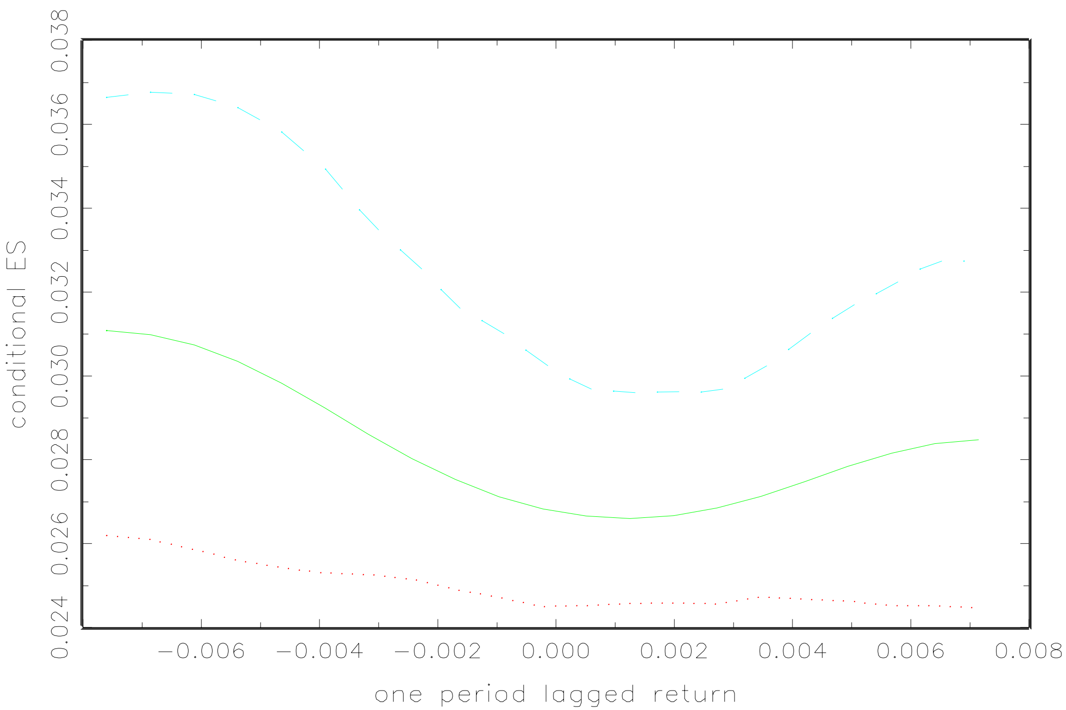
CONDITIONAL VaR



CONDITIONAL VaR



CONDITIONAL EXPECTED SHORTFALL



CONDITIONAL EXPECTED SHORTFALL