

Archive ouverte UNIGE

https://archive-ouverte.unige.ch

Article scientifique Article

e 2004

_ _ _ _ _ _ _ _ _ _ _ _ _

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Nonparametric estimation and sensitivity analysis of expected shortfall

Scaillet, Olivier

How to cite

SCAILLET, Olivier. Nonparametric estimation and sensitivity analysis of expected shortfall. In: Mathematical finance, 2004, vol. 14, n° 1, p. 115–129. doi: 10.1111/j.0960-1627.2004.00184.x

This publication URL:https://archive-ouverte.unige.ch/unige:41799Publication DOI:10.1111/j.0960-1627.2004.00184.x

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

NONPARAMETRIC ESTIMATION AND SENSITIVITY ANALYSIS OF EXPECTED SHORTFALL

O. SCAILLET

HEC Genève and FAME, Université de Genève

We consider a nonparametric method to estimate the expected shortfall—that is, the expected loss on a portfolio of financial assets knowing that the loss is larger than a given quantile. We derive the asymptotic properties of the kernel estimators of the expected shortfall and its first-order derivative with respect to portfolio allocation in the context of a stationary process satisfying strong mixing conditions. An empirical illustration is given for a portfolio of stocks. Another empirical illustration deals with data on fire insurance losses.

KEY WORDS: nonparametric, kernel, time series, expected shortfall, incremental expected shortfall, risk management, risk adjusted performance measure, portfolio selection, loss severity distribution

1. INTRODUCTION

The *expected shortfall* on a portfolio of assets is the expected loss knowing that the loss is larger than a given quantile of the loss distribution. Such a quantile is called Value at Risk (VaR) in finance (for an introduction see, e.g., Duffie and Pan 1997; Jorion 1997). The expected shortfall and the VaR are two measures of risk. They are useful in setting benchmarks for capital requirements (CAD of the Basle Committee on Banking Supervision), in quantifying and regulating the risk taken by market participants such as traders or insurance underwriters (prudential rules imposed by European or American regulators on financial institutions), as well as in allocating available capital in internal risk management procedures. Their primary use in banks is to ensure that the risk incurred by the financial institution remains within its stated risk appetite, and that the assessment of that risk occurs in a timely manner to allow for prompt corrective action if warranted. When the risks across the trading books of a bank are to be summed up, what is called for is a quantitatively oriented, fast, and coherent analysis.

The use of expected shortfall rather than VaR in risk measurement has been recently advocated by Artzner et al. (1999) because of its better properties. First, expected shortfall, called TailVar by these authors, is subadditive for continuously distributed risks, whereas

I thank F. Delbaen and a referee for constructive criticism, Marc Henry for many stimulating discussions and proof checking, and Paul Embrechts and Alexander McNeil for kindly providing the Danish data on fire insurance losses, which were collected by Mette Rytgaard at Copenhagen Re. Fruitful comments come from R. Anderson, A. Cosma, P. Doukhan, J.-L. Prigent, and O. Renault, as well as participants at the AFFI meeting and seminars at BNP Paribas, CentER, Eurandom, and ETH Risklab. The author gratefully acknowledges financial support from the Belgian Program on Interuniversity Poles of Attraction (PAI nb. P4/01), as well as from the Swiss National Science Foundation through the National Center of Competence in Research: Financial Valuation and Risk Management. The main part of this research was done when the author was visiting THEMA and IRES.

Manuscript received August 2000; final revision received August 2002.

Address correspondence to the author at HEC Genève and FAME, Université de Genève, 102 Bd Carl Vogt, CH-1211 Genève 4, Suisse; e-mail: scaillet@hec.unige.ch.

© 2004 Blackwell Publishing Inc., 350 Main St., Malden, MA 02148, USA, and 9600 Garsington Road, Oxford OX4 2DQ, UK.

VaR is not. This subadditivity property is part of the necessary requirements to be a coherent measure of risk, and it expresses the idea that the total risk on a portfolio should not be greater than the sum of the individual risks. Second, VaR tells us nothing about the potential size of the loss that exceeds it, but expected shortfall does.

The expected shortfall has been known for decades among actuaries and is very popular in insurance companies. It is used to assess risk on a portfolio of potential claims, and to design reinsurance treaties (choice and price of excess layers). Its estimation has mainly been developed for i.i.d. data in a semiparametric univariate framework based on the estimation of the mean excess function in extreme value theory (see, e.g., Embrechts, Kluppelberg, and Mikosch 1997; McNeil 1997). As far as we know, it has received no theoretical and empirical attention outside such a setting.

In this paper we propose and study the properties of an estimation method based on a kernel approach. It allows for a broad class of dependencies in the observations, namely strong mixing (for examples of commonly used models satisfying α -mixing conditions, see Doukhan 1994), and is therefore well suited for financial data. The estimation procedure is extremely fast and easy to implement. It basically only requires the standard functionalities of any spreadsheet used in financial statement reporting. We examine how to estimate nonparametrically the expected shortfall together with its sensitivity with respect to changes in portfolio allocation. The expected shortfall and its sensitivity are functions of the portfolio allocation and the loss probability. Our interest in getting a smooth and distribution-free estimation of these functions dictates our choice of a kernel-based method.

Knowledge of the sensitivity is helpful in reducing the amount of computational time needed to process large portfolios. Indeed it avoids the need to recompute the expected shortfall each time the portfolio composition is slightly modified. Furthermore, the first-order derivative is useful in portfolio selection problems (see Markowitz 1952 for portfolio selection in a mean-variance framework). It helps to characterize and evaluate efficient portfolio allocations when expected shortfall is substituted for variance as the measure of risk. Numerical optimization algorithms for computation of optimal allocations require smooth estimates of first derivatives in order to converge properly. Gourieroux, Laurent, and Scaillet (2000) have studied the analogous VaR case, and analyzed the sensitivity of the VaR from a theoretical and empirical point of view.

The paper is organized as follows. In Section 2 we outline the framework and define the expected shortfall formally. We provide the analytical form of the first-order derivative of the expected shortfall with respect to portfolio allocation. The expression is a conditional expectation that is amenable to estimation by a kernel approach. As a corollary, incremental (or component) expected shortfalls-that is, contributions of each asset position to the expected shortfall of the portfolio—are fully described and may be easily estimated. Quantifying risk contributions is especially relevant when allocating available capital in internal risk management procedures and deriving risk-adjusted performance measures. We refer to Tasche (1999) for a detailed treatment of the connection between capital allocation (see also Denault 2001 for an axiomatic approach to this problem), risk-adjusted performance measurement, and the characterization of efficient portfolios via the two-fund separation property of portfolio choice theory. In Section 3 we present the kernel estimators and derive their asymptotic distribution under suitable assumptions on the bandwidth and mixing properties of the data generating process. Section 4 presents two empirical illustrations, one on a portfolio of stocks and the other on fire insurance losses. Section 5 concludes. Proofs and mathematical developments are gathered in the Appendix.

2. SENSITIVITY OF THE EXPECTED SHORTFALL

2.1. Definition of the Expected Shortfall

We consider a strictly stationary process $\{Y_t, t \in \mathbb{Z}\}$ taking values in \mathbb{R}^n and assume that our data consist in a realization of $\{Y_t; t = 1, ..., T\}$. The vector $Y_t = (Y_{1,t}, ..., Y_{n,t})'$ corresponds to *n* risks over a given period of time. The data may 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. This, in turn, implies that the sample length *T* can sometimes be controlled, and asked to be sufficiently large to get satisfying estimation results.

The expected shortfall associated with a portfolio with allocation a is defined by

(2.1)
$$m(a, p) = E[-a'Y| - a'Y > VaR(a, p)],$$

where VaR(a, p) is implicitly defined by

$$(2.2) P[-a'Y > VaR(a, p)] = p.$$

The expected shortfall corresponds to the average loss on a portfolio with composition $a = (a_1, \ldots, a_n)'$, knowing that the loss is larger than a given quantile VaR(a, p), called Value at Risk in finance. Note that *a* is a vector of quantities held of particular financial assets when the vector *Y* is made of price differences, and it is a vector of allocations measured in values (percentage) when *Y* is a vector of returns.

The above-expected loss and quantile are defined with respect to the stationary distribution P of Y, and are thus marginal measures of risk. Conditional expected shortfalls and conditional VaR may be defined by replacing P with a conditional distribution. The expected shortfall and the VaR are functions of the portfolio structure a and the loss probability p. Typical values for the loss probability range from 1% to 5%; stock returns are usually measured over either a 1-day period or a 10-day period.

2.2. Sensitivity of the Expected Shortfall

In the following proposition we provide the first-order derivative $m^{(1)}(a, p) = \frac{\partial m(a, p)}{\partial a'}$ of the expected shortfall with respect to portfolio allocation. This proposition characterizes explicitly the sensitivity of the expected shortfall with respect to local changes in the portfolio structure. For the sake of completeness, we provide in the Appendix a simple proof akin to the derivation of the sensitivity of the VaR in Gourieroux et al. (2000; see also Tasche 1999).

PROPOSITION 2.1. The first-order derivative of the expected shortfall with respect to portfolio allocation is

(2.3)
$$m^{(1)}(a, p) = E[-Y| - a'Y > VaR(a, p)].$$

(1)

Note that the expected shortfall and its first-order derivative are linearly related through $m(a, p) = a'm^{(1)}(a, p)$. This is due to the homogeneity of degree one of the expected shortfall (application of Euler's Theorem). The quantity $a_i \partial m(a, p) / \partial a_i$ is called *incremental* (or component) *expected shortfall* of asset *i* by analogy with the existing VaR terminology (Garman 1997). Indeed, VaR also shares this homogeneity property. Incremental

expected shortfalls and incremental VaR can be used to rank asset positions in terms of their contribution to the total portfolio risk measured by expected shortfall and VaR, respectively.

3. KERNEL ESTIMATOR OF EXPECTED SHORTFALL

In the next lines we describe how to estimate nonparametrically the expected shortfall and its sensitivity with respect to changes in the portfolio structure. We mainly follow the presentation of Robinson (1983) both in terms of notations and assumptions (see, e.g., Bierens 1985 or Bosq 1998 for alternative sets of assumptions). Let us introduce the kernel estimator,

(3.1)
$$[(Y_t, a'Y_t); \xi] = (Th)^{-1} \sum_{t=1}^T Y_t K((\xi - a'Y_t)/h),$$

and define

(3.2)
$$\hat{I}(\xi) = \int_{-\infty}^{\xi} [(Y_t, a'Y_t); u] \, du.$$

Here *h* is a positive number, called the bandwidth, that depends on the sample size *T* and is assumed, at least, to converge to zero as *T* goes to infinity. The real function K(u), the kernel, satisfies, at least, that it integrates to 1. We further define the kernel estimator $\hat{q}(a, p)$ of the quantile of level *p* of the distribution of *a' Y* through the equality

$$\int_{-\infty}^{\hat{q}(a,p)} [(1,a'Y_t);u] \, du = p.$$

This kernel estimator delivers an estimate of q(a, p) = -VaR(a, p). The ratio $\hat{I}(\hat{q}(a, p))/p$ then provides an estimator of the conditional expectation E[Y | a'Y < q(a, p)], and we may consider estimating the expected shortfall and its sensitivity with

(3.3)
$$\hat{m}(a, p) = -a' \hat{I}(\hat{q}(a, p))/p,$$

(3.4)
$$\hat{m}^{(1)}(a, p) = -\hat{I}(\hat{q}(a, p))/p$$

Note that the kernel estimator of the expected shortfall is simply obtained by multiplying the kernel estimator of the sensitivity with the portfolio allocation.

If a Gaussian kernel $K(x) = \varphi(x)$ is adopted, the estimator $\hat{I}(\hat{q}(a, p))$ simplifies to

$$\hat{I}(\hat{q}(a, p)) = T^{-1} \sum_{t=1}^{T} Y_t \Phi((\hat{q}(a, p) - a'Y_t)/h),$$

where φ and Φ denote the p.d.f. and c.d.f. of a standard Gaussian variable, respectively. For the Epanechnikov kernel: $K(x) = \frac{3}{4}(1 - x^2)$, for $|x| \le 1$, and K(x) = 0, otherwise, it simplifies to:

$$\begin{split} \hat{I}(\hat{q}(a, p)) &= T^{-1} \sum_{t=1}^{T} Y_t \left(\mathbb{I}_{\frac{\hat{q}(a, p) - a'Y_t}{h} > 1} + \left(\frac{1}{2} + \frac{3}{4} \left(\frac{\hat{q}(a, p) - a'Y_t}{h} \right) \right. \\ &\left. - \frac{1}{4} \left(\frac{\hat{q}(a, p) - a'Y_t}{h} \right)^3 \right) \mathbb{I}_{|\frac{\hat{q}(a, p) - a'Y_t}{h}| \le 1} \right), \end{split}$$

where II denotes the indicator function. Explicit forms may also be derived for the biweight, triweight, and triangular kernels.

In the aforementioned kernel estimators, we plug a kernel-based quantile estimator in (3.2) to get (3.3) and (3.4). Such a quantile estimate is a smooth function of the portfolio allocation *a* and the loss probability *p*. We could also have thought of plugging in an empirical quantile but that choice would deteriorate the smoothness because empirical quantiles are not smooth functions of *a* and *p*.

The asymptotic normality of the kernel estimators (3.3) and (3.4) can be established under suitable conditions on the kernel, the asymptotic behavior of the bandwidth, the regularity of the conditional expectations, and some mixing properties of the process.

ASSUMPTION 3.1 (kernel and bandwidth).

- (a) The function K is a real bounded and symmetric function on \mathbb{R} such that $\int K(u) du = 1$.
- (b) The bandwidth h is a positive function of T such that Th → ∞ and Th⁵ → 0 when T → ∞.
- (c) The kernel and the bandwidth satisfy $|K(u)| \le C(1+|u|)^{-1-\omega}$, and $h^{\omega-2} \le C$ for $\omega > 2, C > 0$.

Assumption 3.2 (process).

- (a) The process Y 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 $E|Y_i|^{\theta} < \infty$, $i = 1, ..., n, \theta > 2$.
- (b) The p.d.f. g(y) of Y exists and is twice continuously differentiable.
- (c) The equation $P[a'Y \le z] = p$ admits a unique solution denoted by q(a, p), and $f(z) = \partial P[a'Y \le z]/\partial z > 0$ at z = q(a, p).
- (d) The conditional expectation $G_i(u) = E[Y_i | a'Y < u], i = 1, ..., n$, has continuous second order partial derivatives at u = q(a, p).
- (e) The conditional expectation $H_{ij}(u) = E[Y_i Y_j | a' Y < u], i, j = 1, ..., n, is contin$ uous at <math>u = q(a, p).
- (f) The conditional expectation $E[|Y_i|^{\gamma} | a'Y < u]$, i = 1, ..., n, is bounded for $\gamma > \theta$ at u = q(a, p).

PROPOSITION 3.1. Under Assumptions 3.1 and 3.2, the kernel estimators $\hat{m}(a, p)$ and $\hat{m}^{(1)}(a, p)$ are asymptotically normally distributed as

(3.5)
$$\sqrt{Th}(\hat{m}(a, p) - m(a, p)) \to N\left(0, a' Va \int K^2(z) dz\right),$$

(3.6)
$$\sqrt{Th}(\hat{m}^{(1)}(a, p) - m^{(1)}(a, p)) \to N\left(0, V \int K^2(z) dz\right),$$

where

$$V = \left(\frac{E[YY' \mid a'Y < q(a, p)]}{p} + \frac{E[Y\mathbb{I}_{a'Y=q(a, p)}]E[Y'\mathbb{I}_{a'Y=q(a, p)}]}{pf^2(q(a, p))} - \frac{E[Y \mid a'Y < q(a, p)]E[Y'\mathbb{I}_{a'Y=q(a, p)}]}{f(q(a, p))} - \frac{E[Y\mathbb{I}_{a'Y=q(a, p)}]E[Y' \mid a'Y < q(a, p)]}{f(q(a, p))}\right).$$

Consistent estimates of the matrix V may be derived after replacement of the various terms by adequate density and conditional moment kernel estimators.

Note that we do not face the curse of dimensionality often encountered in kernel methods since the kernel dimension is small (one) and independent of the number of assets in the portfolio. Our approach, which basically only requires computations of empirical averages, may be applied even in the presence of very large portfolios, possibly involving simulated returns. Indeed, the asymptotic results of Proposition 3.1 will not be affected if simulations of a parametric model are used as input data, when estimated parameters satisfy the usual parametric rate of convergence. The number of simulations should then be chosen large enough to be on the safe side of asymptotic theory. Let us further remark that Monte Carlo experiments (available on request from the author) show that the proposed estimators perform well in terms of bias and mean squared error for sample lengths possibly encountered in practical situations (two trading years of daily observations).

Finally, Proposition 3.1 can be easily adapted if we wish to consider higher conditional moments such as $E[(a'Y)^2 | -a'Y > VaR(a, p)]$, or, more generally, E[g(-a'Y) | -a'Y > VaR(a, p)] for a given differentiable function g, instead of m(a, p).

4. EMPIRICAL ILLUSTRATIONS

This section illustrates the implementation of the estimation procedure described in section 3. We provide two empirical illustrations, the first one in finance and the second one in insurance. They concern French data on stock returns and Danish data on fire insurance losses.

4.1. Stock Return Data

We analyze two companies listed on the Paris Bourse: Thomson-CSF (electronic devices) and L'Oréal (cosmetics). Both stocks are highly traded and belong to the index CAC 40 of the Paris Bourse. The data are one-day returns recorded daily from 04/01/1997 to 05/04/1999 (546 observations). The mean and standard deviation are 0.0049% and 1.262% for Thomson-CSF, 0.0586% and 1.330% for L'Oréal. Minimum returns are -4.524% and -4.341%; maximum returns are 3.985% and 4.013%, respectively. We have -0.2387 and 0.0610 for skewness, and 1.099 and 1.295 for kurtosis. This indicates that the data cannot be considered as normally distributed. The correlation is 38.54%. The autocorrelograms, partial autocorrelograms, and Ljung-Box statistics have not detected the presence of linear dynamics in the one-day returns. In the following, kernel estimates are compared to empirical estimates and estimates under a Gaussian assumption for the stationary distribution.

Empirical estimates of the expected shortfall and its sensitivity are given by $T^{-1} \sum_{t=1}^{T} a' Y_t \mathbb{I}_{a'Yt < \tilde{q}(a,p)}/p$ and $T^{-1} \sum_{t=1}^{T} Y_t \mathbb{I}_{a'Yt < \tilde{q}(a,p)}/p$, where $\tilde{q}(a, p)$ is the quantile of level *p* computed from the empirical distribution of the portfolio returns. Note that these empirical estimators are not smooth functions of portfolio allocation and loss probability, and are neither differentiable in *a* nor in *p*.

A simple expression for the VaR is available under a Gaussian assumption for the stationary distribution, namely $VaR(a, p) = -a'\mu + (a'\Omega a)^{1/2}z_{1-p}$, where z_{1-p} is the quantile of level 1 - p of a standard Gaussian distribution. Its estimation only requires estimates of the mean μ and covariance matrix Ω . This explains why VaR estimates under a Gaussian assumption serve as benchmark in many banks. An explicit form for the expected shortfall may also be computed in a Gaussian setting:

$$m(a, p) = -a'\mu + (a'\Omega a)^{1/2} \frac{\varphi(z_{1-p})}{p}.$$

A straightforward differentiation gives

$$m^{(1)}(a, p) = -\mu + \frac{\Omega a}{(a'\Omega a)^{1/2}} \frac{\varphi(z_{1-p})}{p},$$

which may be shown to coincide with -E[Y | a' Y < -VaR(a, p)] using

$$E[Y \mid a' Y < k] = \mu - \frac{\Omega a}{(a'\Omega a)^{1/2}} \frac{\varphi\left((k - a'\mu)/(a'\Omega a)^{1/2}\right)}{\Phi\left((k - a'\mu)/(a'\Omega a)^{1/2}\right)}$$

for a given threshold k.

Note that

$$m^{(1)}(a, p) = -\mu + \frac{\Omega a}{a'\Omega a}(m(a, p) + a'\mu),$$

which says that the first-order derivative under a Gaussian assumption is an affine function of the expected shortfall with coefficients depending on the portfolio allocation, but independent of the loss probability p. It is a nonlinear function of the VaR:

$$m^{(1)}(a, p) = -\mu + \frac{\Omega a}{(a'\Omega a)^{1/2}} \frac{\varphi\left(\frac{Va\,R(a, p) + a'\mu}{(a'\Omega a)^{1/2}}\right)}{p}.$$

Figure 4.1 conveys estimates of the expected shortfall for various loss probabilities and for various portfolios including these two stocks. The loss probability levels range from 0.01 to 0.1 and the allocations from 0% (100%) to 100% (0%) in Thomson-CSF (L'Oréal) stock. The left plot gives the kernel estimates. We have selected bandwidth values such that $h = 0.5(a'\hat{\Omega}a)^{1/2}T^{-1/5}$, where $\hat{\Omega}$ is the empirical covariance matrix computed from observed returns. Oversmoothing (undersmoothing) has been obtained with twice (half) this bandwidth—that is, with (a quarter of) the rule of thumb. The central and right plots correspond to empirical estimates and estimates under a Gaussian assumption for the stationary distribution, respectively. We may observe that the kernel estimates are smoother than the empirical estimates. Estimates based on a normality assumption are below the kernel estimates and empirical estimates. This could have been expected from the skewness and kurtosis exhibited by the individual stock returns. Estimates under the Gaussian assumption lead thus to underestimation of the reserve amount aimed to cover potential losses.

Let us now examine the sensitivities. Estimated first-order partial derivatives of the expected shortfall are provided in Figure 4.2 for the first asset, Thomson-CSF, and in Figure 4.3 for the second asset, L'Oréal. Comparison of the plots of kernel and empirical estimates indicates clearly the advantage of using a smooth estimator. As already mentioned, numerical algorithms that would rely on empirical estimates of the first-order partial derivatives would encounter difficulties to converge properly. For a given probability level estimates under a Gaussian assumption are always either increasing for the first asset or decreasing for the second asset. Kernel estimates show that this is an artifact of the normality assumption.













FIGURE 4.4. Estimated expected shortfall for insurance loss data.

4.2. Insurance Loss Data

The Danish data consist of 1794 daily losses over one million Danish Krone (DKK) for the years 1980 to 1990. The loss figure is a total loss figure with a positive sign and includes damage to buildings and damage to furniture and personal property, as well as loss of profits. The mean and standard deviation are equal to 4,235,299.9 and 9,256,401.5, respectively. The minimum loss is 325,000 and the maximum loss is 200,700,000. The skewness is 13.629 and the kurtosis is 251.336. The autocorrelogram, partial autocorrelogram, and Ljung-Box statistics have not revealed any linear dynamics in the loss data. Figure 4.4 shows estimates of the expected shortfall for loss probabilities ranging from 0.005 to 0.05. The jagged line corresponds to empirical estimates, the upper dashed line corresponds to kernel-based estimates. They are very close to each other, but the shape is rather erratic for empirical estimates. The bandwidth is taken here equal to $\hat{\sigma} T^{-1/5}$, where $\hat{\sigma}$ is the standard deviation of the observed losses (other choices have not led to results that can be distinguished by the naked eye). The lower dashed line carries the estimates based on a normality assumption, which are again clearly below the kernel estimates. Our kernel estimates agree with estimates based on extreme value theory. For example, we get an estimated expected shortfall of DKK 66.96 million when the loss probability is equal to 1%. From the estimation results given by McNeil (1997) for a peaks-over-threshold (POT) method, expected shortfall estimates are 58.69 and 69.59 million for a threshold *u* set equal to 10 and 20 million, respectively. POT estimates are carried by the two dotted lines, which enclose our kernel estimates for small probabilities (less than 0.012). The upper dotted line is interrupted around p = 0.016 where quantile values become in that case inferior to u = 20.

5. CONCLUDING REMARKS

Risk measures answer the need for quantifying the risk of potential losses on a portfolio of assets. This need may arise due to internal concerns (risk-reward trade-off) or external

constraints (prudential rules imposed by regulators). In this paper we have proposed an estimation procedure for one of these measures, namely the expected shortfall, and for its sensivity with respect to changes in portfolio allocation. The procedure relies on a kernel approach, and provides smooth estimates of the expected shortfall and its sensitivity viewed as functions of the portfolio allocation and the loss probability. It is fast, efficient, and valid even in the dependent case for general stationary strong mixing multivariate processes. The empirical examples in finance and insurance have shown its relevance in practical situations. Possible use of other nonparametric approaches (local polynomials) may be contemplated, but this awaits future research.

APPENDIX

Proof of Proposition 2.1. Let us set $X = -\sum_{j=1, j\neq i}^{n} a_j Y_j$, and $Z_i = -Y_i$, i = 1, ..., n. Equation (2.1) defining the expected shortfall can be rewritten as $E[X + a_i Z_i | X + a_i Z_i > VaR(a, p)]$. The expression for the first-order derivative is then a direct consequence of the following lemma.

LEMMA A.1. Let us consider a bivariate vector (X, Z), and the conditional expectation

$$m(\epsilon, p) = E[X + \epsilon Z | X + \epsilon Z > Q(\epsilon, p)],$$

where the quantile $Q(\epsilon, p)$ is defined by

$$P[X + \epsilon Z > Q(\epsilon, p)] = p.$$

Then

$$\frac{\partial m(\epsilon, p)}{\partial \epsilon} = E\left[Z \mid X + \epsilon Z > Q(\epsilon, p)\right].$$

Proof of Lemma A.1. Let us denote f(x, z) the joint p.d.f. of the pair (X, Z). We get

$$E[X + \epsilon Z | X + \epsilon Z > Q(\epsilon, p)] = \frac{E[(X + \epsilon Z) \mathbb{I}_{X + \epsilon Z > Q(\epsilon, p)}]}{p}$$
$$= \frac{1}{p} \int \left(\int_{Q(\epsilon, p) - \epsilon z} (x + \epsilon z) f(x, z) \, dx \right) \, dz.$$

The differentiation with respect to ϵ gives

$$\frac{1}{p} \int \left(\int_{Q(\epsilon,p)-\epsilon z} zf(x,z) \, dx \right) \, dz - \frac{1}{p} \int \left(\frac{\partial Q(\epsilon,p)}{\partial \epsilon} - z \right) Q(\epsilon,p) f(Q(\epsilon,p)-\epsilon z,z) \, dz.$$

Since the differentiation with respect to ϵ of the condition defining the quantile, $P[X + \epsilon Z > Q(\epsilon, p)] = p$, implies the restriction

$$\int \left(\frac{\partial Q(\epsilon, p)}{\partial \epsilon} - z\right) f(Q(\epsilon, p) - \epsilon z, z) dz = 0,$$

the second term vanishes, and we conclude that

$$\frac{\partial m(\epsilon, p)}{\partial \epsilon} = \frac{1}{p} \int \left(\int_{Q(\epsilon, p) - \epsilon z} zf(x, z) \, dx \right) \, dz,$$
$$= E[Z \, | \, X + \epsilon \, Z > Q(\epsilon, p)].$$

Proof of Proposition 3.1. The asymptotic normality of $\hat{m}(a, p) = -a'\hat{I}(\hat{q}(a, p))/p$ and $\hat{m}^{(1)}(a, p) = -\hat{I}(\hat{q}(a, p))/p$ is deduced from the expansion of $\hat{I}(\hat{q}(a, p))$ around q(a, p):

$$\hat{I}(\hat{q}(a, p)) = \hat{I}(q(a, p)) + \frac{\partial \hat{I}(\bar{q}(a, p))}{\partial u}(\hat{q}(a, p) - q(a, p)),$$

where $\bar{q}(a, p)$ is a mean value located between $\hat{q}(a, p)$ and q(a, p), and the asymptotic normality of $\hat{I}(q(a, p))/p$ and $\hat{q}(a, p)$:

$$\begin{split} \sqrt{Th} \bigg(\frac{\hat{I}[q(a, p)]}{p} - E[Y|a'Y < q(a, p)] \bigg) &\to N \bigg(0, \frac{E[YY'|a'Y < q(a, p)]}{p} \int K^2(z) dz \bigg), \\ \sqrt{Th}(\hat{q}(a, p) - q(a, p)) &\to N \bigg(0, \frac{p}{f^2(q(a, p))} \int K^2(z) dz \bigg). \end{split}$$

(a) We deduce the first of the two aforementioned normality results along the same lines as Robinson (1983, Thm. 5.3), which establishes asymptotic normality of conditional expectation estimators in the context of strong mixing processes.

We need to show that

$$\sqrt{Th}(E[\hat{I}(q(a, p))] - pE[Y \mid a'Y < q(a, p)]) \to 0.$$

Let us consider arbitrary coordinates *i* and *j*, *i*, *j* = 1, ..., *n*, $i \neq j$. We have, successively,

$$h^{-1}E\left[\int_{-\infty}^{q(a,p)} Y_{i}K((a'Y-u)/h) du\right]$$

$$= h^{-1}\int_{-\infty}^{q(a,p)} \int_{\mathbb{R}^{n}} y_{i}K\left(\left(a_{j}y_{j}+\sum_{l\neq j}a_{l}y_{l}-u\right)/h\right)g(y_{1},\ldots,y_{j},\ldots,y_{n})dy_{1}\cdots dy_{n}du,$$

$$= \int_{-\infty}^{q(a,p)} \int_{\mathbb{R}^{n}} \frac{y_{i}}{a_{j}}K(z)g\left(y_{1},\ldots,\left(u-hz-\sum_{l\neq j}a_{l}y_{l}\right)/a_{j},\ldots,y_{n}\right)\prod_{s\neq j}dy_{s}du dz,$$

$$= \int_{-\infty}^{\frac{q(a,p)}{a_{j}}} \int_{\mathbb{R}^{n}} y_{i}g\left(y_{1},\ldots,v-\sum_{l\neq j}\frac{a_{l}}{a_{j}}y_{l},\ldots,y_{n}\right)dy_{1}\cdots dy_{n}dv$$

$$+ \frac{h^{2}}{2}\int_{-\infty}^{\frac{q(a,p)}{a_{j}}} \int_{\mathbb{R}^{n}}\frac{y_{i}}{a_{j}^{2}}\left[\frac{\partial^{2}}{\partial y_{j}^{2}}g\left(y_{1},\ldots,v-\sum_{l\neq j}\frac{a_{l}}{a_{j}}y_{l},\ldots,y_{n}\right)\right]dy_{1}\cdots dy_{n}dv\int_{\mathbb{R}}z^{2}K(z) dz$$

$$+ o(h^{2}), \qquad \text{(by symmetry of the kernel)}$$

$$= p E[Y_i | a' Y < q(a, p)] + O(h^2)$$

from which we deduce the needed convergence result using $Th^5 \rightarrow 0$.

Now, following Robinson (1983, Thm. 5.3) step by step, we build $\bar{g}_{it} = Y_{it} \mathbb{I}_{|Y_{it}| \le D}$, $\bar{\bar{g}}_{it} = Y_{it} - \bar{g}_{it}$, for some $D, 0 < D < \infty$, and also $\bar{W}_{it} = \bar{g}_{it}K_t$, $\bar{\bar{W}}_{it} = \bar{\bar{g}}_{it}K_t$, with $K_t = \int_{-\infty}^{q(a,p)} K((u - a'Y_t)/h) du$. We further introduce $\bar{S}_{iT} = (hT)^{-1/2} \sum_{t=1}^{T} \bar{V}_{it}$ with $\bar{V}_{it} = \bar{W}_{it} - E[\bar{W}_{it}]$. Since $V[(Th)^{-1/2} \sum_{t=1}^{T} \bar{W}_{it}] \to 0$ as $D \to \infty$, uniformly in large T, we only have to show that the vector $\bar{S}_T = (\bar{S}_{iT}, \bar{S}_{jT})'$ has, for fixed D, an asymptotic normal distribution. First we have

$$h^{-1}E[\bar{V}_{it}\bar{V}_{jt}] \to pE[Y_iY_j\mathbb{I}_{|Y_i| \le D}\mathbb{I}_{|Y_j| \le D} | a'Y < q(a, p)] \int K^2(z) dz.$$

Second, the continuity of $G_i(u)$ and $H_{ij}(u)$ implies the continuity of $\overline{G}_i(u) = E[Y_i \mathbb{I}_{|Y_i| \le D} | a' Y < u]$ and $\overline{H}_{ij}(u) = E[Y_i Y_j \mathbb{I}_{|Y_i| \le D} | \mathbb{I}_{|Y_j| \le D} | a' Y < u]$. Hence the central limit theorem of Lemma 7.1 in Robinson (1983), valid for bounded \overline{V}_{it} , may be applied to \overline{S}_T , and we get the stated normality result since $\overline{G}_i(u) \to G_i(u)$, $\overline{H}_{ij}(u) \to H_{ij}(u)$ by letting D tend to infinity.

(b) We continue with the proof of the second normality result. We have the expansion

$$P[a' Y < q(a, p)] = p = \int_{-\infty}^{\hat{q}(a, p)} [(1, a' Y_t); u] du$$

=
$$\int_{-\infty}^{q(a, p)} [(1, a' Y_t); u] du + [(1, a' Y_t); \bar{q}(a, p)](\hat{q}(a, p) - q(a, p)),$$

from which we get

$$(\hat{q}(a, p) - q(a, p)) = \frac{1}{[(1, a'Y_t); \bar{q}(a, p)]} \left(p - \int_{-\infty}^{q(a, p)} [(1, a'Y_t); u] \, du \right).$$

We also have

$$\sqrt{Th}\left(p-E\left[\int_{-\infty}^{q(a,p)}[(1,a'Y_t);u]\,du\right]\right)\to 0.$$

To establish the asymptotic normality of $\sqrt{Th}(p - \int_{-\infty}^{q(a,p)} [(1, a'Y_t); u] du)$, we build $S_T = (hT)^{-1/2} \sum_{t=1}^{T} V_t$ with $V_t = K_t - E[K_t]$, and $K_t = \int_{-\infty}^{q(a,p)} K((u - a'Y_t)/h) du$, as above. We may here directly apply the central limit theorem of Lemma 7.1 in Robinson (1983) since V_t is bounded:

$$S_T \to N\left(0, p \int K^2(z) dz\right).$$

We deduce then the second normality result from the convergence of $[(1, a' Y_t); \bar{q}(a, p)]$ to f(q(a, p)).

(c) Eventually, since $\frac{\partial \hat{I}(\bar{q}(a,p))}{\partial u} = [(Y_t, a'Y_t); \bar{q}(a, p)]$ converges to $E[Y' \mathbb{I}_{a'Y=q(a,p)}]$, we obtain the final stated normality results for $\hat{m}(a, p)$ and $\hat{m}^{(1)}(a, p)$ after computation of the covariance terms.

REFERENCES

- ARTZNER, P., F. DELBAEN, J. M. EBER, and D. HEATH (1999): Coherent Measures of Risk, *Math. Finance* 9, 203–228.
- BIERENS, H. (1985): Kernel Estimators of Regression Functions; in Advances in Econometrics, T. BEWLEY, ed. Cambridge: Cambridge University Press, 99–144.
- Bosq, D. (1998): *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*. Lecture Notes in Statistics. New York: Springer-Verlag.

DENAULT, M. (2001): Coherent Allocation of Risk Capital, J. Risk 4, 7-21.

- DOUKHAN, P. (1994): *Mixing: Properties and Examples*. Lecture Notes in Statistics. New York: Springer-Verlag.
- DUFFIE, D., and J. PAN (1997): An Overview of Value at Risk, J. Derivatives 4, 7-49.

- EMBRECHTS, P., C. KLUPPELBERG, and T. MIKOSCH (1997): *Modelling Extremal Events: For Finance and Insurance*. New York: Springer-Verlag.
- GARMAN, M. (1997): Taking VaR to Pieces, RISK, 10, 70-71.
- GOURIÉROUX, C., J. P. LAURENT, and O. SCAILLET (2000): Sensitivity Analysis of Values at Risk, J. Empirical Finance 7, 225–245.
- JORION, P. (1997): Value at Risk: The New Benchmark for Controlling Market Risk. Chicago: Irwin.
- MARKOWITZ, H. (1952): Portfolio Selection, J. Finance 7, 77-91.
- MCNEIL, A. (1997): Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory, *ASTIN Bull.* 27, 117–137.

ROBINSON, P. (1983): Nonparametric Estimators for Time Series, J. Time Series Anal. 4, 185-207.

TASCHE, D. (1999): Risk Contributions and Performance Measurement. Diss TU München.

Copyright of Mathematical Finance is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.