© *The Journal of Risk and Insurance*, 2009, Vol. 76, No. 3, 683-707
DOI: 10.1111/j.1539-6975.2009.01309.x

Control and Out-of-Sample Validation of Dependent Risks

Christian Gourieroux
Wei Liu

Abstract

This article introduces a framework to determine and allocate capital reserves to multiple dependent business lines, with or without overall reserve level constraints. The proposed methodology emphasizes the role of the loss function in the validation criterion and its conditional interpretation. Univariate and multivariate examples are discussed in detail.

Introduction

The current regulation introduced in Finance by Basel I and Basel II (or in Insurance by Solvency II) has at least three limitations. The first limitation is the choice of the value at risk (VaR) to measure the risk and to fix the required capital. Indeed, the VaR accounts for the probability of a loss, but not for the expected magnitude of this loss (as the TailVaR), or for the possibility of extreme losses (as distortion risk measures). The second limitation is the lack of coherency between the suggested computation of the VaR as a conditional quantile and the *ex post* validation of this computation, which uses unconditional (historical) hedging errors. Finally, the computation of the VaRs are performed independently on different business lines, without taking into account the possible dependence of the risks between business lines.

This article proposes solutions to the two last limitations by focusing on the (implicit) decision criterion of the internal or external validator. For expository purpose, we follow the practice of measuring the risk by means of the VaR. In the "Conditional Risk Control" section, we focus on a single business line, that is, on a single risk. We introduce the decision criterion that underlies the choice of a VaR and discuss the criterion value when a suboptimal control variate is selected. This criterion is chosen because it represents a natural way of measuring the goodness of a candidate model fitting the observed data in the quantile context and, thus, suits our problem well. The criterion value can be decomposed into its minimal attainable value plus

Christian Gourieroux is with CREST and University of Toronto. Wei Liu is with Fitch Solutions. Wei Liu can be contacted via e-mail: leo.liu@utoronto.ca. The authors gratefully acknowledge financial support of NSERC Canada and the chair AXA/Risk Foundation: "Large Risks in Insurance."

a term measuring the lack of optimality. In a perspective of measuring the out-of-sample performances of a proposed control variate, we analyze the joint dynamic of these two components. The case of several dependent business lines is considered in the "Dependent Risks" section. We introduce a decision criterion function, which extends naturally the criterion considered for a single business line. By optimizing this criterion, with or without overall reserve level constraints, we get the capital allocations of reserves by business lines. The "Conclusion" section concludes. Proofs are presented in the appendices.

CONDITIONAL RISK CONTROL

It is emphasized in the recent literature (see e.g., Hansen, 2005; Giacomini and White 2006) that the "out-of-sample comparison of predictive ability has to be based on inference about conditional expectations of forecast and forecast errors rather than their unconditional expectations." The same remark applies for the validation by regulators of the approaches followed by banks and insurance companies to fix their reserves. This section discusses this principle for the computation of the VaR and the analysis of its out-of-sample performances.

The Underlying Criterion

The approach is based on the underlying criterion followed by regulators to derive the VaR, that is a conditional quantile of the future portfolio value. Let us denote by y_t the opposite of the portfolio value at date t, by I_t the information available at time t, which includes at least the current and lagged values of y, and by z_t the level of reserve, which can be considered as a control variate. This level is proposed by the bank according to information I_t. The standard decision criterion is

$$\Psi(t, z_t) = E_t\left[\alpha(y_{t+1} - z_t)^+ + (1-\alpha)(y_{t+1} - z_t)^-\right], \tag{1}$$

where E_t denotes the conditional expectation given I_t, $y^+ = \max(y, 0)$, $y^- = \max(-y, 0)$, and α is the announced risk level. This decision criterion corresponds to an asymmetric loss function (written in \$). Two types of errors can arise. The amount of reserve z_t can be too small, implying a loss $y_{t+1} - z_t$; it can also be too large, which implies an amount of money $z_t - y_{t+1}$ not used for profitable investment. The criterion is weighting both types of errors, with more weights to avoid the first type of error from the prudential (regulator) point of view.[1]

Let us denote f_t and F_t the conditional pdf and cdf of y_{t+1}, respectively. It is well known (see e.g., Koenker, 2005), that the optimal value of the control variate

$$\hat{z}_t = \arg\min_{z_t \in I_t} E_t\left[\alpha(y_{t+1} - z_t)^+ + (1-\alpha)(y_{t+1} - z_t)^-\right], \tag{2}$$

[1] Since the quantile concerns the loss and profit variable, α is large in practice. Typically, it is fixed at 95 percent, that is, at a level of 5 percent when the problem is written in terms of profit and loss and concerns a portfolio of stocks. It has been fixed to some other values as 99 percent, or even 99.97 percent for some special credit derivatives.

is the conditional quantile at risk level α:

$$\hat{z}_t = F_t^{-1}(\alpha), \tag{3}$$

which explains why function (1) constitutes a natural choice of decision criterion for VaR models.

However, there exists a large number of decision criteria providing the conditional quantile as the optimal control variate. It has been shown in Gourieroux and Monfort (1995, vol. 1, section 8.5.9) that these criteria can be written as

$$\Psi(t, z_t) = E_t\big[\alpha\big[A(y_{t+1}) - A(z_t)\big]^+ + (1-\alpha)\big[A(y_{t+1}) - A(z_t)\big]^-\big], \tag{4}$$

where A is an increasing function. Thus, we get as many criteria compatible with the VaR as admissible choices of risk level α and distortion function A. The standard criterion (1) corresponds to the identity distortion function $A(y) = y$. By choosing an increasing convex distortion function, we would have more focus on extreme losses. We will not discuss the distortion function in the following sections and essentially consider direct extensions of standard criterion (1).

Optimality Versus Suboptimality

The optimal solution of problem (1), that is the conditional VaR, is clearly defined and is used as a benchmark in the following sections to discuss some properties and interpretations. However, the VaRs computed by banks or insurance companies differ generally from this optimal conditional VaR and their suboptimal methodologies can vary in time. This can be explained by the difficulty in taking into account the relevant information I_t, by misspecified dynamic models and by errors in estimating a conditional quantile when the model is specified and the information is available. For instance, Basel II proposes two approaches, called standard approach and advanced approach, respectively. The standard approach (i.e., the suboptimal approach) allows for VaR computed with a limited information. Typically a VaR computed by historical simulation neglects all information on the current market environment. Since the proposed control variate z_t are not necessarily optimal, it is important to analyze the magnitude and effect of suboptimality on the validation criterion.

For any control variate z_t not necessarily optimal, the criterion value can be decomposed as

$$\Psi(t, z_t) = \Psi(t, \hat{z}_t) + \big[\Psi(t, z_t) - \Psi(t, \hat{z}_t)\big], \tag{5}$$

where the second component of the right-hand side measures the lack of optimality.

(i) For the standard objective function, where A is the identity function, the lack of optimality is equal to (see Appendix A)

$$c(t, z_t) = \int_{z_t}^{\hat{z}_t} \big[F_t(\hat{z}_t) - F_t(y)\big] dy \tag{6}$$

$$= \begin{cases} E_t[(y_{t+1} - z_t)\mathbb{1}_{(z_t < y_{t+1} < \hat{z}_t)}], & \text{if } \hat{z}_t > z_t, \\ E_t[(z_t - y_{t+1})\mathbb{1}_{(\hat{z}_t < y_{t+1} < z_t)}], & \text{if } \hat{z}_t \leq z_t. \end{cases} \tag{7}$$

In particular, when the selected control variate is close to the optimal one $z_t \simeq \hat{z}_t$, we get the expansion:

$$c(t, z_t) \simeq \frac{1}{2} f_t(\hat{z}_t)(\hat{z}_t - z_t)^2 = \frac{1}{2} f_t\left[F_t^{-1}(\alpha)\right](\hat{z}_t - z_t)^2.$$

The quantity $1/f_t[F_t^{-1}(\alpha)]$ is equal to the derivative of the conditional quantile with respect to risk level α. This is a special case of delta-VaR: $delta - VaR_t(\alpha) = \frac{\partial}{\partial\alpha} F_t^{-1}(\alpha)$, which measures the sensitivity of the conditional VaR to level α (see e.g., Garman, 1997; Gourieroux et al., 2000). Thus, we get

$$c(t,\, z_t) \simeq \frac{1}{2} \frac{(\hat{z}_t - z_t)^2}{delta - VaR_t(\alpha)}. \tag{8}$$

The above expression shows that the local effect of the lack of optimality is linked with the granularity adjustment introduced in the literature, when the optimal control variate is computed by assuming an infinitely large credit portfolio, whereas the actual portfolio includes a large, but finite, number of credits (see e.g. Wilde, 2001; Martin and Wilde, 2003; Gordy and Lütkebohmert, 2007).

(ii) The first component of decomposition (5) is equal to

$$\begin{aligned} \Psi(t, \hat{z}_t) &= \alpha P_t\,[y_{t+1} > \hat{z}_t]\, E_t\left[(y_{t+1} - \hat{z}_t)^+ \mid y_{t+1} > \hat{z}_t\right] \\ &\quad + (1-\alpha) P_t\,[y_{t+1} < \hat{z}_t]\, E_t\left[(y_{t+1} - \hat{z}_t)^- \mid y_{t+1} < \hat{z}_t\right] \\ &= \alpha(1-\alpha)\left[TailVaR^+(\alpha) + TailVaR^-(\alpha)\right], \end{aligned} \tag{9}$$

where the TailVaR measures the "expected loss" when a loss occurs. This expression highlights the importance of the TailVaR in the validating criterion even if the control variate is suboptimal. Even if the optimal control is applied, the regulator needs to follow up both the VaR for the control variate, and the TailVaR, for the objective function.

Dynamic of the Decision Criterion

Let us now study the consequence on the criterion of any proposed methodology to compute the control variate. More precisely, for any methodology z_t, we can consider the joint dynamics of $\Psi(t, \hat{z}_t)$ and $c(t,\, z_t)$.

As an illustration, let us assume that the dynamics of y_{t+1} is such that

$$y_{t+1} = m_t + \sigma_t u_{t+1}, \tag{10}$$

where m_t, σ_t are functions of the past and (u_t) is a sequence of independent and identically distributed error terms with $E(u_t) = 0$. This specification includes the ARMA-GARCH model, where the information set corresponds to the current and lagged values of y, but also the stochastic mean and volatility model, where the information set also includes the current and lagged values of the factors driving the stochastic parameters. Let us denote by Q the quantile function of the standardized error u, and $H(x) = E(u - x)^+$. The conditional quantile is

$$\hat{z}_t = m_t + \sigma_t Q(\alpha). \tag{11}$$

It follows that

$$\begin{aligned}\Psi(t, \hat{z}_t) &= \alpha E_t \left[(y_{t+1} - \hat{z}_t)^+ \right] + (1 - \alpha) E_t \left[(y_{t+1} - \hat{z}_t)^- \right] \\ &= \alpha \sigma_t E_t \left[(u_{t+1} - Q(\alpha))^+ \right] + (1 - \alpha) \sigma_t E_t \left[(u_{t+1} - Q(\alpha))^- \right] \\ &= \sigma_t \delta(\alpha), \end{aligned} \tag{12}$$

where

$$\delta(\alpha) = H\left[Q(\alpha)\right] + (1 - \alpha) Q(\alpha). \tag{13}$$

Let us now assume that the VaR proposed by the bank is computed by the unconditional quantile associated with a risk level α (or equivalently by historical simulation based on a wide window). Then $z_t = z$, say, is constant overtime and is in particular independent of the environment (but a function of risk level α). The value of the criterion becomes

$$\begin{aligned}\Psi(t, z) &= \alpha E_t \left[(y_{t+1} - z)^+ \right] + (1 - \alpha) E_t \left[(y_{t+1} - z)^- \right] \\ &= E_t \left[(y_{t+1} - z)^+ \right] - (1 - \alpha) \left(E_t[y_{t+1}] - z \right) \\ &= \sigma_t H \left[\tfrac{z - m_t}{\sigma_t} \right] - (1 - \alpha)(m_t - z). \end{aligned} \tag{14}$$

The correcting component is

$$c(t, z_t) = \sigma_t \left[H\left(\frac{z - m_t}{\sigma_t} \right) - H\left(Q(\alpha) \right) \right] - (1 - \alpha) \left[m_t + \sigma_t Q(\alpha) - z \right], \tag{15}$$

and its local expansion becomes (see Appendix B):

$$c(t, z_t) = \frac{1}{2\sigma_t} g\left[Q(\alpha) \right] \left[m_t + \sigma_t Q(\alpha) - z \right]^2, \tag{16}$$

where g is the pdf of the error term u.

This example shows that

1. The minimal attainable criterion value is a stochastic function of the scale factor σ_t only (see Equation (12));
2. The lack of optimality of the unconditional VaR depends on both location and scale factors m_t, σ_t (see Equation (15)); therefore, the two components of the decision criterion have dependent joint dynamics;
3. The same remark applies to the relative lack of optimality $c(t, z_t)/\Psi(t, \hat{z}_t)$;
4. The optimal criterion value is highly volatile, since it is proportional to σ_t. This is a consequence of the standard choice of a fixed risk level α by validators instead of a path dependent risk level α_t.

The question of fixed versus path dependent risk level has already been considered in the VaR literature for cyclical effects. In a point-in-time (PIT) approach, the banks have to analyze the risk without correcting for cyclical effect, and thus, the whole treatment of the cycle is left to the regulator. In the through-the-cycle (TTC) approach the banks analyze a risk corrected for cycle, and in some sense the treatment of the cycle is left to the bank. Basel II decided to ask for banks' actual risk, which is the PIT approach, but the regulator has unfortunately fixed a constant level α, keeping the same α in expansion period, or in crisis period, for instance. In our framework, the question of fixed versus market relatively dependent risk level α has not yet been carefully considered by regulators.

As an illustration, let us consider the stochastic volatility model:

$$r_t = -y_t = 0.5\sigma_t^2 + \sigma_t u_t, \tag{17}$$

where the volatility σ_t^2 follows an autoregressive gamma (ARG) process with parameters $\rho = 0.96$, $\delta = 3.6$, $c = 7.34.10^{-3}$ (see Gourieroux and Jasiak, 2006). The historical parameters ρ, δ, c are such that the ARG volatility process matches the stationary mean, variance and first-order autocorrelation of the discretely sampled Cox–Ingersoll–Ross process estimated by Andersen et al. (2002). The dynamic model is an extension of the standard Ball and Roma (1994) model, including a risk premium. We provide in Figure 1 a joint simulated path of (r_t, σ_t^2), $t = 1, \ldots, 250$, corresponding to a solid line and dashed line, respectively.

This path is used to estimate the historical (unconditional) distribution of the return and to deduce its (estimated) unconditional quantile at $1 - \alpha = 5$ percent. This unconditional quantile is equal to $\hat{z} = 2.31$. Figure 2displays the evolution of the conditional quantile $\hat{z}_t$ (solid line) and compares it with the unconditional quantile (dashed line) $\hat{z}$.[2] The simulated path of r_t (shorter dashed line) is also plotted to reveal the relative locations of the estimated quantiles. The unconditional quantile is above the average of the conditional quantiles. This represents the cost of not taking into account the largest information when evaluating the risk.[3]

[2] The conditional quantile is computed with the value of the parameters used in the simulation study. In practice these values are unknown and have to be estimated, which induces an additional uncertainty.

[3] The interpretation in terms of cost of information seems natural by analogy with what is known for the variance: the unconditional variance is always larger than the average of the

FIGURE 1
Joint Simulated Path of (r_t, σ_t^2)

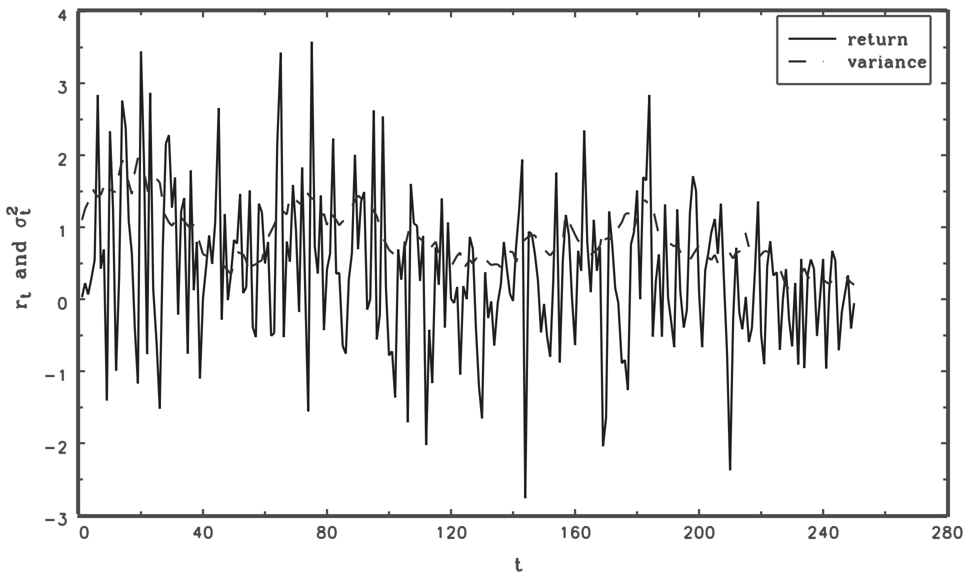
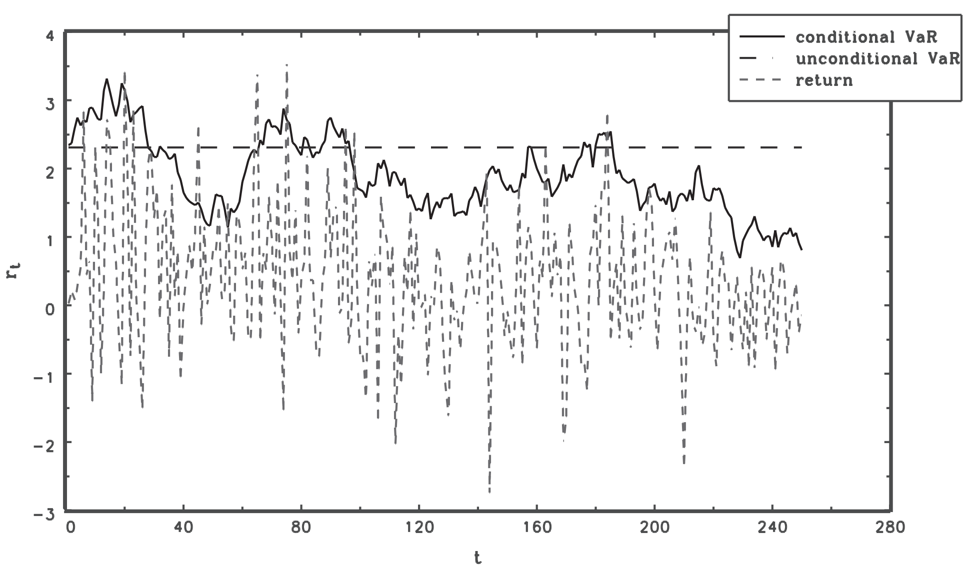


FIGURE 2
Conditional and Unconditional Quantiles



conditional variances by the variance analysis equation. However, a similar result does not exist for quantiles, and there are likely examples where unconditional quantiles are smaller than the average of conditional quantiles.

FIGURE 3
Joint Evolution of $\Psi(t, \hat{z}_t), \Psi(t, z)$, and $c(t, z_t)$

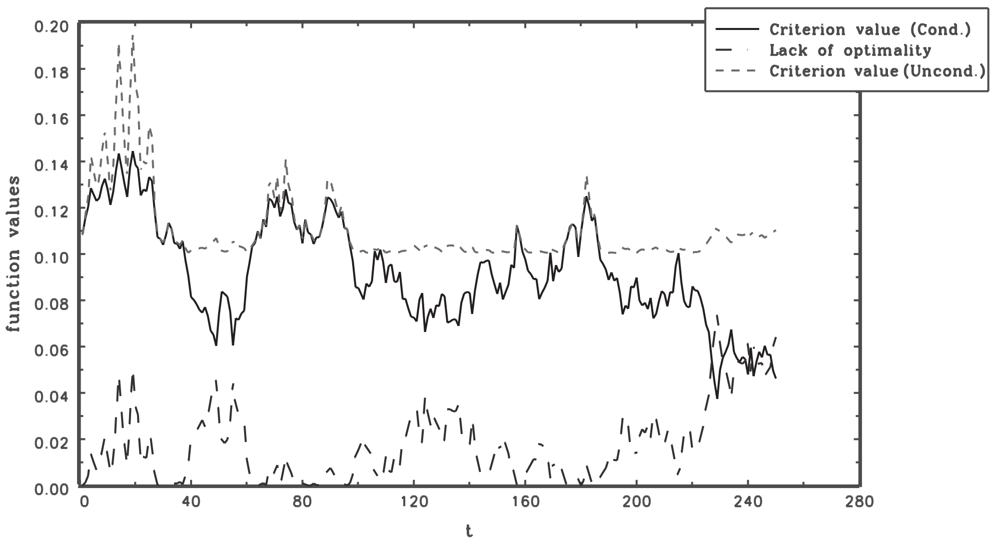


TABLE 1
Summary Statistics

Series	Mean	Variance	Skewness	Kurtosis
$\psi(t, \hat{z}_t)$	0.0926	0.0005	0.0762	2.4029
$c(t, z_t)$	0.0178	0.0003	1.0883	3.5836
$\psi(t, z)$	0.1104	0.0003	2.6492	10.8276

We observe the clustering of reserve levels deduced from the underlying volatility clustering. Moreover, the observations exceed the conditional and unconditional VaRs by a similar number of times, even though the VaR estimates are quite different during large and small volatility periods. This suggests a limitation of the current validation based on the number of exceedances only. Figure 3 provides the joint dynamics of the value of the optimal criterion (solid line), the value of the suboptimal criterion (shorter dashed line), and the measure of suboptimality (dashed line), evaluated when unconditional quantile is used. By definition, the criterion value computed with the unconditional quantile is above the criterion value computed with the conditional one. The optimal criterion value exhibits a pattern similar to the pattern of conditional VaR in Figure 2, since both processes are proportional to the volatility.

For a suboptimal control, both components of the objective function have to be analyzed jointly. We provide in Table 1 the historical summary statistics of the series, that are their historical mean, variance, skewness, and kurtosis.

The distribution of the loss function $\Psi(t, z)$, when the suboptimal unconditional quantile is used, features strong positive skewness and very fat tails. The fat right

TABLE 2
Autocorrelations

	Series	$\psi(t,\hat{z}_t)$	$c(t,z_t)$	$\psi(t,z)$	Series	$\psi(t,\hat{z}_t)^2$	$c(t,z_t)^2$	$\psi(t,z)^2$
$\rho(0)$	$\psi(t,\hat{z}_t)$	1.0000	−0.7024	0.6730	$\psi(t,\hat{z}_t)^2$	1.0000	−0.5363	0.7420
	$c(t,z_t)$	−0.7024	1.0000	0.0539	$c(t,z_t)^2$	−0.5363	1.0000	0.0772
	$\psi(t,z)$	0.6730	0.0539	1.0000	$\psi(t,z)^2$	0.7420	0.0772	1.0000
$\rho(1)$	$\psi(t,\hat{z}_t)$	0.9562	−0.6681	0.6465	$\psi(t,\hat{z}_t)^2$	0.9561	−0.5136	0.7008
	$c(t,z_t)$	−0.6725	0.9064	−0.0083	$c(t,z_t)^2$	−0.5193	0.8856	0.0170
	$\psi(t,z)$	0.6429	0.0034	0.8975	$\psi(t,z)^2$	0.6973	0.0273	0.8759
$\rho(2)$	$\psi(t,\hat{z}_t)$	0.9122	−0.6397	0.6136	$\psi(t,\hat{z}_t)^2$	0.9095	−0.4952	0.6514
	$c(t,z_t)$	−0.6489	0.8108	−0.0778	$c(t,z_t)^2$	−0.5109	0.7891	−0.0496
	$\psi(t,z)$	0.6044	−0.0558	0.7777	$\psi(t,z)^2$	0.6417	−0.0248	0.7306
$\rho(3)$	$\psi(t,\hat{z}_t)$	0.8657	−0.6006	0.5877	$\psi(t,\hat{z}_t)^2$	0.8630	−0.4713	0.6161
	$c(t,z_t)$	−0.6156	0.7337	−0.1115	$c(t,z_t)^2$	−0.4898	0.7069	−0.0757
	$\psi(t,z)$	0.5719	−0.0794	0.7054	$\psi(t,z)^2$	0.6030	−0.0556	0.6523
$\rho(4)$	$\psi(t,\hat{z}_t)$	0.8265	0.5512	0.5822	$\psi(t,\hat{z}_t)^2$	0.8289	−0.4304	0.6159
	$c(t,z_t)$	−0.5707	0.6928	−0.0929	$c(t,z_t)^2$	−0.4553	0.6569	−0.0535
	$\psi(t,z)$	0.5621	−0.0491	0.7184	$\psi(t,z)^2$	0.5996	−0.0214	0.6906
$\rho(5)$	$\psi(t,\hat{z}_t)$	0.7988	−0.5131	0.5806	$\psi(t,\hat{z}_t)^2$	0.8048	−0.3997	0.6220
	$c(t,z_t)$	−0.5364	0.6795	−0.0606	$c(t,z_t)^2$	−0.4301	0.6455	−0.0219
	$\psi(t,z)$	0.5571	−0.0066	0.7502	$\psi(t,z)^2$	0.6037	0.0171	0.7464
$\rho(10)$	$\psi(t,\hat{z}_t)$	0.6082	−0.4021	0.4263	$\psi(t,\hat{z}_t)^2$	0.5942	−0.3417	0.4364
	$c(t,z_t)$	−0.4448	0.4699	−0.1492	$c(t,z_t)^2$	−0.3770	0.5445	−0.0973
	$\psi(t,z)$	0.3796	−0.0591	0.4436	$\psi(t,z)^2$	0.4076	−0.0459	0.4185
$\rho(15)$	$\psi(t,\hat{z}_t)$	0.4237	−0.2383	0.3302	$\psi(t,\hat{z}_t)^2$	0.4124	−0.2166	0.3183
	$c(t,z_t)$	−0.3565	0.2752	−0.2140	$c(t,z_t)^2$	−0.3058	0.3453	−0.1515
	$\psi(t,z)$	0.2239	−0.0322	0.2545	$\psi(t,z)^2$	0.2471	−0.0412	0.2278
$\rho(25)$	$\psi(t,\hat{z}_t)$	0.0286	−0.0163	0.0215	$\psi(t,\hat{z}_t)^2$	0.0064	0.0601	−0.0077
	$c(t,z_t)$	−0.0884	0.0626	−0.0580	$c(t,z_t)^2$	−0.1541	0.0636	−0.0841
	$\psi(t,z)$	−0.1145	0.0907	−0.0685	$\psi(t,z)^2$	−0.1011	0.1283	−0.0740

tail arises since the unconditional VaR estimate tends to underestimate the TailVaR when the volatility is large. This underestimation of TailVaR is amplified by the volatility level as specified by Equation (14). This effect is significantly smaller when the conditional quantile is used, and the kurtosis is four times smaller.

Cross-autocorrelations of the series and their squares are provided in Table 2. $\rho(k)$ denotes the kth-order autocorrelation matrix. For instance,

$$\rho(3) = corr\left([\psi(t,\hat{z}_t),c(t,z_t),\psi(t,z)]',[\psi(t-3,\hat{z}_{t-3}),c(t-3,z_{t-3}),\psi(t-3,z)]'\right).$$

By construction, the serial correlations on the optimal criterion value $\Psi(t, \hat{z}_t)$ and their squares are equal to the serial correlations of σ_t and σ_t^2. Thus, it is not surprising to observe the long memory effect corresponding to the high persistence of volatility. A smaller persistence can be observed on the criterion value and its squares, when unconditional quantile is used as control variate.

Comparing the Methodologies Proposed to Compute the Reserves

The aim of this section is to compare different methodologies $z_{1,t}$, $z_{2,t}$, say, used to compute the reserve corresponding to a same risky portfolio and a same announced risk level. This question is important for the implementation of new Basel II regulation, which is based on a sequential checking of what is proposed by the banks. Loosely speaking, the regulator will receive, from some given department of a bank (i.e., for a given portfolio), proposed VaR levels $z_{1,t}$, $t = 1, \ldots, T$, for the past and current periods with some incomplete information on the followed methodology. The regulator will check if the VaR levels $(z_{1,t})$ are compatible with its validation criterion. If they are not compatible, the regulator will ask for new computations $z_{2,t}$, $t = 1, \ldots, T$ and will check if there is a real improvement.

Two methodologies cannot be ranked unambiguously in general; indeed, methodology 1 can be preferred in some environments (such as expansion periods or high-volatility periods), whereas methodology 2 can be preferable in other environments (such as recession periods, or low-volatility periods).

However, it is possible to develop tractable diagnostic concerning properties valid for all environments, for instance to test the hypothesis

H_0: Methodology 1 is better than methodology 2 for any environment.

The null hypothesis is

$$H_0\colon \Psi(t, z_{1,t}) \leq \Psi(t, z_{2,t}), \text{ for any value of the conditioning variable,}$$

and involves an infinite number of conditional inequality moment restrictions. This hypothesis can be written equivalently in terms of unconditional inequality moment restrictions by introducing instruments g_t as

$$\begin{aligned} H_0\colon\ & E\left\{g_t\left[\alpha(y_{t+1} - z_{1,t})^+ + (1-\alpha)(y_{t+1} - z_{1,t})^-\right]\right\} \\ & \leq E\left\{g_t\left[\alpha(y_{t+1} - z_{2,t})^+ + (1-\alpha)(y_{t+1} - z_{2,t})^-\right]\right\}, \\ & \text{for any positive function } g_t \text{ depending on information } I_t. \end{aligned}$$

At this step, let us mention that the standard diagnostics, proposed in the literature (see, e.g., Diebold and Mariano, 1995; Harvey et al., 1997, 1998; Harvey and Newbold, 2000; or the survey by Stock and Watson, 2005) or used by practitioners, validators, or regulators, are based on the unconditional loss function only, that is, select $g_t = 1$

as the single instrumental variable.[4] The recent literature has emphasized the importance of selecting several appropriate instruments. For instance, Gourieroux and Jasiak (2008) considered lagged values of the loss function, such as

$$g_{j,k,t} = \alpha(y_{t+1-k} - z_{j,t-k})^+ + (1-\alpha)(y_{t+1-k} - z_{j,t-k})^-, \quad j = 1,2.$$

By considering the single unitary instrument, we focus on the historical average loss. By choosing lagged values of the objective function as instruments, we consider also the possibility of loss clustering, that is of a large cumulated loss on consecutive periods of time. When a finite number of instruments $g_{1,t}, \ldots, g_{K,t}$, say, are selected, the hypothesis concerns K positivity restrictions on unconditional moments. The tests of inequality restrictions have been well studied in the literature (Kudô, 1963; Gourieroux et al., 1980, 1982; Wolak, 1987; Wolak, 1991). Asymptotically most powerful procedures can be based on Wald approach described below.

The procedure is as follows:

1. Approximate each unconditional moment by its sample counterpart. Namely, the following moment:

$$\begin{aligned}\theta_k = E\big[g_{k,t}\big\{\alpha\left[(y_{t+1} - z_{2,t})^+ - (y_{t+1} - z_{1,t})^+\right] \\ + (1-\alpha)\left[(y_{t+1} - z_{2,t})^- - (y_{t+1} - z_{1,t})^-\right]\big\}\big], \quad k = 1,\ldots,K,\end{aligned}$$

   is estimated by:

$$\begin{aligned}\hat{\theta}_{k,T} = \frac{1}{T}\sum_{t=1}^{T-1} g_{k,t}\big\{\alpha\left[(y_{t+1} - z_{2,t})^+ - (y_{t+1} - z_{1,t})^+\right] \\ + (1-\alpha)\left[(y_{t+1} - z_{2,t})^- - (y_{t+1} - z_{1,t})^-\right]\big\}, \quad k = 1,\ldots,K.\end{aligned}$$

2. Estimate the asymptotic variance–covariance matrix $\hat{\Omega}_T$ of $\hat{\theta}_T = \left(\hat{\theta}_{1,T},\ldots,\hat{\theta}_{K,T}\right)'$, by the sample variance of the vector of sample averages, given by

$$\hat{\Omega}_T = \hat{\Gamma}_T(0) + \sum_{\tau=-L_T}^{L_T} \omega_T(\tau)\hat{\Gamma}_T(\tau),$$

   where $\hat{\Gamma}_T(\tau)$ is the sample counterpart of the autocovariance matrix at lag τ of the multivariate process $g_t\{\alpha[(y_{t+1} - z_{2,t})^+ - (y_{t+1} - z_{1,t})^+] + (1-\alpha)$ $[(y_{t+1} - z_{2,t})^- - (y_{t+1} - z_{1,t})^-]\}, L_T$ defines the bandwidth and $\omega_T(\tau)$ a weighting scheme. The bandwidth and weighting schemes have to be chosen to ensure the consistency of $\hat{\Gamma}_T(\tau)$ and its finite sample positiveness (see, e.g., Andrews, 1991).

[4] Direct comparisons of the proposed and optimal VaR have the same drawback (see, e.g., Mancini and Trojani, 2005).

3. Compute the inequality constrained estimator $\hat{\theta}_T^0$ of θ as the solution of

$$\hat{\theta}_T^0 = \underset{\theta:\theta_k \geq 0, k=1,\ldots,K}{\operatorname{argmin}} (\hat{\theta}_T - \theta)' \hat{\Omega}_T^{-1} (\hat{\theta}_T - \theta).$$

4. Compute the test statistics as the optimal value of the objective function

$$W_T = \left(\hat{\theta}_T - \hat{\theta}_T^0\right)' \hat{\Omega}_T^{-1} \left(\hat{\theta}_T - \hat{\theta}_T^0\right).$$

5. Reject the null hypothesis H_0 if $W_T > c$, where the critical value with type-I error 5 percent is the 95 percent quantile of a mixture of chi-square distributions $\sum_{k=0}^{K} \pi_k \chi^2(k)$, where the weights depend on the pattern of the Ω matrix. For instance, if $K = 2$, we have

$$\pi_0 = \frac{1}{2\pi} \cos^{-1}\left(\rho(\theta_1, \theta_2)\right), \pi_1 = \frac{1}{2}, \quad \text{and } \pi_2 = \frac{1}{2} - \pi_0,$$

where $\rho(\theta_1, \theta_2)$ is the correlation between θ_1 and θ_2.

The weight π_k is the probability that the K-dimensional vector $\hat{\theta}_T^0$ has exactly k positive elements under the null hypothesis.[5] Thus, a Monte-Carlo procedure can be designed to approximate these weights. First, one generates a K dimensional Gaussian series (10,000 observations, say) with mean 0 and variance–covariance matrix $\hat{\Omega}_T$. Then, π_k is estimated by the frequency of exactly k positive components (out of K components). Finally, the p-value is computed by substituting these estimated weights into the mixture of chi-square distributions.

The above test procedure requires the observations on $z_{1,t}, z_{2,t}, t = 1, \ldots, T$, but does not demand the detailed methodology, which has been used by the banks to get z_1 and z_2. Thus, the regulator does not have to check if $z_{1,t}$ has really been computed as announced by the bank, and thus to replicate all banks' computations, which will be impossible in practice.

As an illustration of this procedure, let us continue the example from the "Dynamic of the Decision Criterion" section, with the suboptimal unconditional quantile $z_{1,t} = \hat{z}$ and the optimal conditional quantile $z_{2,t} = \hat{z}_t$. The test procedure can be applied with the following instruments:

Case I: $g_{1t} = 1;$

Case II: $g_{1t} = 1,\ g_{2t} = |r_t - \hat{z}|;$

Case III: $g_{1t} = 1,\ g_{3t} = |r_t - \hat{z}_{t-1}|;$

Case IV: $g_{1t} = 1,\ g_{2t} = |r_t - \hat{z}|, g_{3t} = |r_t - \hat{z}_{t-1}|.$

[5] Wolak (1987, section 4) provides an excellent summary on how the weights (π_k) can be calculated. Closed form formulas are available for $K \leq 4$; Monte-Carlo methods are necessary for higher order.

TABLE 3
Test Results

Cases	W_T	c	Results
I	13.0939	2.7056	Y
II	14.7373	4.9471	Y
III	14.4964	4.7363	Y
IV	15.1493	5.9920	Y

Notes: The Wald test statistics are provided in column 2. Column 3 lists the critical values. When $K \leq 2$, we obtain the values of c by using the closed form expressions of the distribution of W_T and a grid search numerical algorithm. Otherwise, the weights π_k are obtained through a Monte-Carlo procedure. The letter "Y" means that the test concludes that the estimated optimal procedure $\hat{z}_t$ is better than the suboptimal procedure $\hat{z}$.

Case I is the standard practice of counting the historical number of exceedances. Case II (respectively, Case III) includes also an instrument to capture clustering effect for the suboptimal control (respectively, optimal control). Case IV gathers all instruments. The testing procedures are applied for $T = 250$ (i.e., 1 year of daily data) and $\hat{\Omega}_T$ is computed with $L_T = 5$. The results of the test are given in Table 3, where Y (Yes) means that the test concludes that the estimated optimal VaR $\hat{z}_t$ is better than the suboptimal VaR $\hat{z}$. When $K \leq 2$, the critical values c are computed by using the closed form expressions provided by Gourieroux et al. (1982). For Case IV, the weights π_k are obtained through a Monte-Carlo procedure. As expected, the optimal conditional VaR is proved to be better than the suboptimal unconditional VaR in all cases. The comparison of the methodologies based on conditional and unconditional VaR is less clear if we follow the standard validation practice, which counts the number of exceedances only. Indeed in our example the numbers of times that the observed returns exceed the estimated VaRs are similar for both models.

The example above has just been given to illustrate the testing procedure, since in practice both competing control variates $z_{1,t}$, $z_{2,t}$ are suboptimal and the validator has only partial knowledge on the way they have been computed.

DEPENDENT RISKS

This section considers the joint analysis of two business lines. The method accounts for the possible dependence between the extreme risks associated with the two lines. First, we consider this question by focusing on an appropriate criterion function, and, as a by-product, derives a notion of bidimensional quantile. Then, we discuss the reserve allocation when the overall reserve level is fixed. The validation approach introduced in the "Conditional Risk Control" section can be easily extended to this framework.

The Decision Criterion

Let us now consider two business lines with loss and profit values $y_{j,t}$, $j = 1, 2$, and denote $z_{j,t}$, $j = 1, 2$ the associated amount of reserves. The criterion function can be generalized to:

$$\begin{aligned}\Psi(t, z_{1,t}, z_{2,t}) = E_t \big[& \alpha_{11}(y_{1,t+1} - z_{1,t})^{+}(y_{2,t+1} - z_{2,t})^{+} \\ & + \alpha_{12}(y_{1,t+1} - z_{1,t})^{+}(y_{2,t+1} - z_{2,t})^{-} + \alpha_{21}(y_{1,t+1} - z_{1,t})^{-} \\ & \times (y_{2,t+1} - z_{2,t})^{+} + \alpha_{22}(y_{1,t+1} - z_{1,t})^{-}(y_{2,t+1} - z_{2,t})^{-} \big], \end{aligned} \tag{18}$$

where $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$ are positive weights. By introducing the cross-products, the loss function accounts for all types of joint losses on the two business lines.

Remark 1: If $\alpha_{11} = \alpha^2, \alpha_{12} = \alpha_{21} = \alpha(1 - \alpha), \alpha_{22} = (1 - \alpha)^2$, the criterion becomes:

$$\begin{aligned}\Psi(t, z_{1,t}, z_{2,t}) = E_t \big\{ & \big[\alpha(y_{1,t+1} - z_{1,t})^{+} + (1 - \alpha)(y_{1,t+1} - z_{1,t})^{-}\big] \\ & \times \big[\alpha(y_{2,t+1} - z_{2,t})^{+} + (1 - \alpha)(y_{2,t+1} - z_{2,t})^{-}\big]\big\}, \end{aligned} \tag{19}$$

or, equivalently,

$$\begin{aligned}\Psi(t, z_{1,t}, z_{2,t}) = {} & E_t \big[\alpha(y_{1,t+1} - z_{1,t})^{+} + (1 - \alpha)(y_{1,t+1} - z_{1,t})^{-}\big] \\ & E_t \big[\alpha(y_{2,t+1} - z_{2,t})^{+} + (1 - \alpha)(y_{2,t+1} - z_{2,t})^{-}\big] + Cov_t \big[\alpha(y_{1,t+1} - z_{1,t})^{+} \\ & + (1 - \alpha)(y_{1,t+1} - z_{1,t})^{-}, \alpha(y_{2,t+1} - z_{2,t})^{+} + (1 - \alpha)(y_{2,t+1} - z_{2,t})^{-}\big]. \end{aligned} \tag{20}$$

The first component is the product of criterion values for separate analysis of the two business lines, whereas the second term emphasizes the effect of the possible dependence between risks.

Remark 2: As in the "Conditional Risk Control" section, alternative decision criteria can be introduced by weighting differently the losses according to their magnitudes. For instance, we could consider the criterion:

$$\begin{aligned} & \Psi(t, z_{1,t}, z_{2,t}; A_1, A_2) \\ & = E_t\big\{\alpha_{11} \left[A_1(y_{1,t+1}) - A_1(z_{1,t})\right]^{+} \left[A_2(y_{2,t+1}) - A_2(z_{2,t})\right]^{+} \\ & \quad + \alpha_{12} \left[A_1(y_{1,t+1}) - A_1(z_{1,t})\right]^{+} \left[A_2(y_{2,t+1}) - A_2(z_{2,t})\right]^{-} \\ & \quad + \alpha_{21} \left[A_1(y_{1,t+1}) - A_1(z_{1,t})\right]^{-} \left[A_2(y_{2,t+1}) - A_2(z_{2,t})\right]^{+} \\ & \quad + \alpha_{22} \left[A_1(y_{1,t+1}) - A_1(z_{1,t})\right]^{-} \left[A_2(y_{2,t+1}) - A_2(z_{2,t})\right]^{-} \big\}, \end{aligned}$$

where A_1, A_2 are two increasing functions.

The Optimal Control Variates

The optimal control variates are solutions of the first-order conditions (note that the objective function is convex); we get

$$\begin{cases} \dfrac{\partial \Psi}{\partial z_1}(t, \hat{\hat{z}}_{1,t}, \hat{\hat{z}}_{2,t}) = 0, \\ \dfrac{\partial \Psi}{\partial z_2}(t, \hat{\hat{z}}_{1,t}, \hat{\hat{z}}_{2,t}) = 0, \end{cases} \tag{21}$$

where

$$\begin{aligned} \frac{\partial \Psi}{\partial z_1}(t, z_{1,t}, z_{2,t}) = {} & E_t \left\{ -\mathbb{1}_{y_{1,t+1} > z_{1,t}} \left[\alpha_{11}(y_{2,t+1} - z_{2,t})^+ + \alpha_{12}(y_{2,t+1} - z_{2,t})^- \right] \right\} \\ & + E_t \left\{ \mathbb{1}_{y_{1,t+1} < z_{1,t}} \left[\alpha_{21}(y_{2,t+1} - z_{2,t})^+ + \alpha_{22}(y_{2,t+1} - z_{2,t})^- \right] \right\}, \end{aligned}$$

and a similar expression for $\frac{\partial \Psi}{\partial z_2}(t, z_{1,t}, z_{2,t})$.

If $y_{1,t}$ and $y_{2,t}$ are independent and the weights satisfy $\alpha_{11} = \alpha^2, \alpha_{12} = \alpha_{21} = \alpha(1 - \alpha), \alpha_{22} = (1 - \alpha)^2$ (see Remark 1), then Equations (21) are equivalent to the first-order conditions of two univariate optimization problems. For instance,

$$\frac{\partial \Psi}{\partial z_1}(t, \hat{\hat{z}}_{1,t}, \hat{\hat{z}}_{2,t}) = 0$$

will lead to

$$\alpha E_t \left\{ -\mathbb{1}_{y_{1,t+1} > z_{1,t}} \right\} + (1 - \alpha) E_t \left\{ \mathbb{1}_{y_{1,t+1} < z_{1,t}} \right\} = 0,$$

which is the first-order condition associated with problem (2). In fact we have the following result (see proof in Appendix C):

Proposition 1: *The solutions $\hat{\hat{z}}_{1,t}, \hat{\hat{z}}_{2,t}$ of system (21) are equal to optimal values $\hat{z}_{1,t}, \hat{z}_{2,t}$ computed per single business line for any independent risks $y_{1,t}$, $y_{2,t}$ if and only if: $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = 0$.*

If $y_{1,t}$ and $y_{2,t}$ are dependent and the weights satisfy $\alpha_{11} = \alpha^2, \alpha_{12} = \alpha_{21} = \alpha(1 - \alpha)$, $\alpha_{22} = (1 - \alpha)^2$, the criterion function is easily concentrated with respect to $z_{1,t}$ (resp. $z_{2,t}$) and can be solved recursively. Indeed, let us assume that $z_{2,t}$ is given. We have

$$\begin{aligned} & \Psi(t, z_{1,t}, z_{2,t}) \\ & = E_t \big\{ (y_{1,t+1} - z_{1,t})^+ E\big[\alpha^2 (y_{2,t+1} - z_{2,t})^+ + \alpha(1 - \alpha)(y_{2,t+1} - z_{2,t})^- \,\big|\, y_{1,t+1}, I_t \big] \\ & \quad + (y_{1,t+1} - z_{1,t})^- E\big[\alpha(1 - \alpha)(y_{2,t+1} - z_{2,t})^+ + (1 - \alpha)^2 (y_{2,t+1} - z_{2,t})^- \,\big|\, y_{1,t+1}, I_t \big] \big\} \\ & = E_t \left\{ \delta(y_{1,t+1}, t) \left[\alpha (y_{1,t+1} - z_{1,t})^+ + (1 - \alpha)(y_{1,t+1} - z_{1,t})^- \right] \right\}, \end{aligned}$$

where

$$\delta(y_{1,t+1}, t) = E_t\left[\alpha(y_{2,t+1} - z_{2,t})^+ + (1-\alpha)(y_{2,t+1} - z_{2,t})^- \mid y_{1,t+1}, I_t\right]. \tag{22}$$

Thus, $\Psi(t, z_{1,t}, z_{2,t})$ is proportional to

$$E_t^{Q_t^1}\left[\alpha(y_{1,t+1} - z_{1,t})^+ + (1-\alpha)(y_{1,t+1} - z_{1,t})^-\right],$$

where Q_t^1 denotes the probability deduced from the historical one by applying a change of density proportional to $\delta(y_{1,t+1}, t)$. This modified probability depends on the methodology $z_{2,t}$. Since a similar approach can be used to concentrate with respect to $z_{2,t}$, the optimal strategies can be derived recursively. At step p, the strategies are $z_{1,t}^{(p)}, z_{2,t}^{(p)}$, say. They are used to derive the two modified probabilities $Q_t^{1,p}$ and $Q_t^{2,p}$. Then $z_{1,t}^{(p+1)}, z_{2,t}^{(p+1)}$ are the α-quantiles of the modified conditional probabilities $Q_t^{1,p}$ and $Q_t^{2,p}$.

Thus, even when the weights satisfy $\alpha_{11} = \alpha^2$, $\alpha_{12} = \alpha_{21} = \alpha(1-\alpha)$, $\alpha_{22} = (1-\alpha)^2$, the distributions have to be changed for computing the quantile, and the optimal strategy does not correspond to the standard choice of VaR performed separately for each business line.

Reserve Allocation Under Overall Reserve Level Constraint

The external regulator controls the risk on some aggregated business line. For internal control, the management of the bank has to allocate this overall reserve among the different business lines. Let us consider a global line composed of two business lines $y_{1,t}, y_{2,t}$. The disaggregate reserves are $z_{1,t}, z_{2,t}$, whereas the overall reserve is $z_t = z_{1,t} + z_{2,t}$. The optimization problem becomes

$$\begin{cases} \min_{z_{1,t}, z_{2,t}} \Psi(t, z_{1,t}, z_{2,t}), \\ \text{s.t. } z_{1,t} + z_{2,t} = z_t. \end{cases} \tag{23}$$

The restriction on the total reserve can be solved to get the new optimization problem

$$\begin{aligned} \min_{z_{1,t}} E_t\big[&\alpha_{11}\left(y_{1,t+1} - z_{1,t}\right)^+ \left(y_{2,t+1} - z_t + z_{1,t}\right)^+ \\ &+ \alpha_{12}\left(y_{1,t+1} - z_{1,t}\right)^+ \left(y_{2,t+1} - z_t + z_{1,t}\right)^- \\ &+ \alpha_{21}\left(y_{1,t+1} - z_{1,t}\right)^- \left(y_{2,t+1} - z_t + z_{1,t}\right)^+ \\ &+ \alpha_{22}\left(y_{1,t+1} - z_{1,t}\right)^- \left(y_{2,t+1} - z_t + z_{1,t}\right)^- \big]. \end{aligned} \tag{24}$$

The solutions of this problem will be denoted by $\hat{\hat{z}}_{j,t}(z_t, \alpha), j = 1,2$, where $\alpha = (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})$. Thus, $\hat{\hat{z}}_{j,t}(z_t, \alpha)$ solves

$$\begin{cases} \dfrac{\partial \Psi}{\partial z_1}\left(t, \hat{\hat{z}}_{1,t}(z_t, \alpha), z_t - \hat{\hat{z}}_{1,t}(z_t, \alpha)\right) = 0, \\ \hat{\hat{z}}_{2,t}(z_t, \alpha) = z_t - \hat{\hat{z}}_{1,t}(z_t, \alpha), \end{cases} \tag{25}$$

where

$$\begin{aligned} \frac{\partial \Psi}{\partial z_1}(t, z_{1,t}, z_t - z_{1,t}) = E_t \Big\{ \big[& \mathbb{1}_{y_{1,t+1} > z_{1,t}} \mathbb{1}_{y_{2,t+1} > z_t - z_{1,t}} \alpha_{11} - \mathbb{1}_{y_{1,t+1} > z_{1,t}} \mathbb{1}_{y_{2,t+1} < z_t - z_{1,t}} \alpha_{12} \\ & - \mathbb{1}_{y_{1,t+1} < z_{1,t}} \mathbb{1}_{y_{2,t+1} > z_t - z_{1,t}} \alpha_{21} + \mathbb{1}_{y_{1,t+1} < z_{1,t}} \mathbb{1}_{y_{2,t+1} < z_t - z_{1,t}} \alpha_{22} \big] \\ & \times (y_{1,t+1} - y_{2,t+1} + z_t - 2 z_{1,t}) \Big\}. \end{aligned}$$

Comparison of the Methodologies

By comparing the standard approach and the new methodologies introduced in sections titled "The Optimal Control Variates" and "Reserve Allocation Under Overall Reserve Level Constraint," we can distinguish three ways of fixing the reserves $z_{1,t}$, $z_{2,t}$.

Method 1: Fix a risk level α^* and compute the disaggregate reserve levels by $\hat{z}_{j,t}(\alpha^*) = F_{j,t}^{-1}(\alpha^*)$, where $F_{j,t}$ is the conditional cdf of $y_{j,t}$.

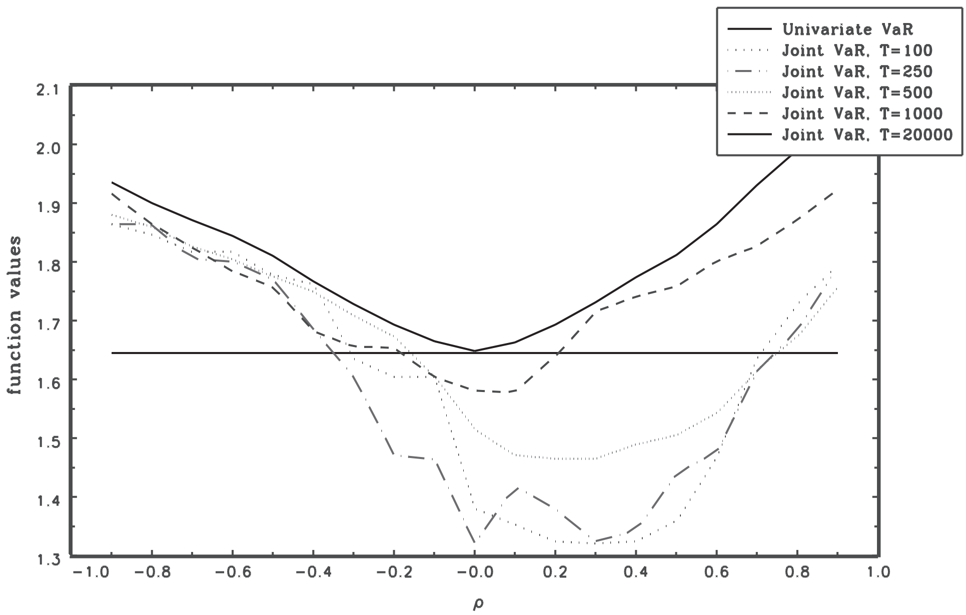
Method 2: Apply the optimization problem (18) without restrictions on $z_{1,t}, z_{2,t}$. We get solutions $\hat{\hat{z}}_{j,t}(\alpha), j = 1, 2$, say, which depend on $\alpha = (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})$.

Method 3: Apply a two-step approach. First fix a risk level α^* for the aggregated line and compute $\hat{z}_t(\alpha^*) = F_t^{-1}(\alpha^*)$, where F_t is the conditional cdf of the aggregate line $y_t = y_{1,t} + y_{2,t}$. Then, look for solutions of the constrained optimization problem, that are $\hat{\hat{z}}_{j,t}\left[\hat{z}_t(\alpha^*), \alpha\right], j = 1, 2$.

The three approaches above depend on different risk level parameters, which are α^* for method 1, α for method 2, and (α^*, α) for method 3. They are difficult to compare in a general setting, especially since it is difficult to choose in a coherent way the aggregate and disaggregate risk levels, α^* and α, respectively. We discuss below the reserve allocation for specific weighting functions and independent and identically distributed Gaussian returns. The independent and identically distributed Gaussian framework has been retained for illustration purpose for the following reasons:

1. It corresponds to the basic mean-variance framework and thus it is natural to first consider how the one-dimensional Gaussian VaR are extended.
2. It is useful to study how the various notions of bivariate quantiles depend on the expected return and their volatility–covolatility.

FIGURE 4
Disaggregate Reserve Level, Function of ρ



In practice, the methodology will be applied to dynamic models compatible with stylized facts such as fat tail, high volatility persistence, but also path dependent risk factors.

(i) **Unconstrained methods**

As mentioned before, Methods 1 and 2 coincide if $\alpha_{11} = (\alpha^*)^2, \alpha_{22} = (1-\alpha^*)^2, \alpha_{12} = \alpha_{21} = \alpha^*(1-\alpha^*)$, and if the disaggregate risks $y_{1,t}$ and $y_{2,t}$ are independent. In the sequel, we select these weights, but do not assume independent risks. To highlight the differences between methods 1 and 2, let us consider the case of independent and identically distributed Gaussian variables

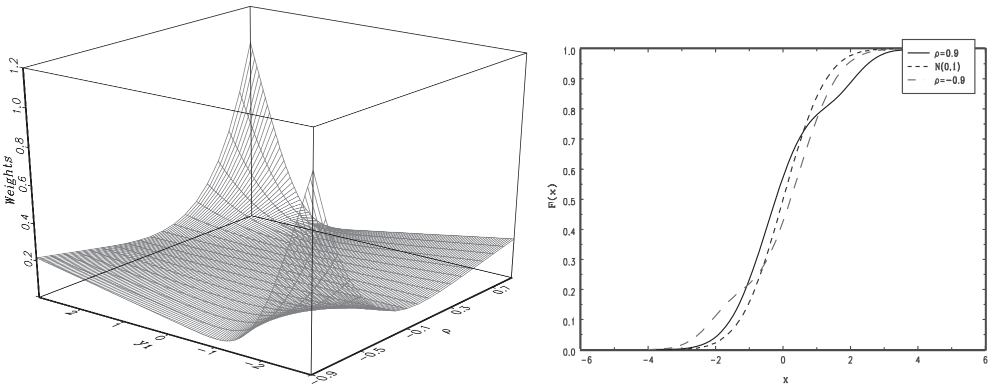
$$(y_{1,t}, y_{2,t})' \sim N\left[\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right].$$

It is equivalent to write $y_{j,t} = m_j + \sigma_j u_{j,t},\ j = 1, 2$, where

$$(u_{1,t}, u_{2,t})' \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right].$$

It is easily checked that the reserves computed with $y_{j,t}$ for both methods 1 and 2 are deduced from the reserves computed with $u_{j,t},\ j = 1, 2$ by a same affine transformations $u_j \to m_j + \sigma_j u_j$. Thus, without loss of generality, we can assume $m_1 = m_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$; then, the solutions satisfy $z_{1,t}(\alpha) = z_{2,t}(\alpha)$ for both methods. We provide in Figure 4 the pattern of the disaggregate

FIGURE 5
Value of Weights as Function of (ρ, y_1) and Modified cdf When $\rho = \pm\,0.9$



reserve level as function of correlation ρ, for $\alpha^* = 95$ percent. The reserve level corresponding to $\rho = 0$ corresponds to the standard derivation business line per business line. The horizontal line 1.64 represents the 95th percentile of the standard normal distribution. In order to reduce the effect of sampling errors, we choose a sample size equal to 20,000 when performing the simulation and computation. We refer this as the true pattern, displayed by the solid curve. For comparison, finite sample patterns are also provided with $T = 100, 250, 500$, and 1,000.

The true pattern suggests that, when estimating jointly, one tends to get higher reserve levels than that estimated independently if $\rho \neq 0$. This is due to the behavior of the weighting function defined in (22). As illustrated in Figure 5, for a nonzero correlation ρ, the weighting function takes large values at left and right tails of y_1 and small values in the middle. As a consequence, the modified cdf has fatter tails than the marginal cdf. This is illustrated by the graph at the right side of Figure 5, where the cdf of the standard normal (the shorter dashed line), and the modified cdf with $\rho = \pm\,0.9$ (dashed and solid lines) are plotted.

Errors due to small sample size can be substantial. It is seen in Figure 4 that, for sample sizes less than 500 (i.e., 2 years), the reserve level could be largely underestimated. This underestimation is diminished when sample size rises up to 1,000 (i.e., 4 years).

(ii) **Constrained method**

Let us now consider method 3. Due to the two-step approach, the reserves computed with $y_{j,t}$, $j = 1, 2$ are no longer deduced by linear transformations of the reserves computed with $u_{j,t}$, $j = 1, 2$. To highlight this effect, let us consider returns

$$(y_{1,t}, y_{2,t})' \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix}\right],$$

with a distribution depending on the correlation ρ and the ratio of variances σ^2.

FIGURE 6
Disaggregate Reserve Level as Function of ρ With Aggregate Reserve Constraint

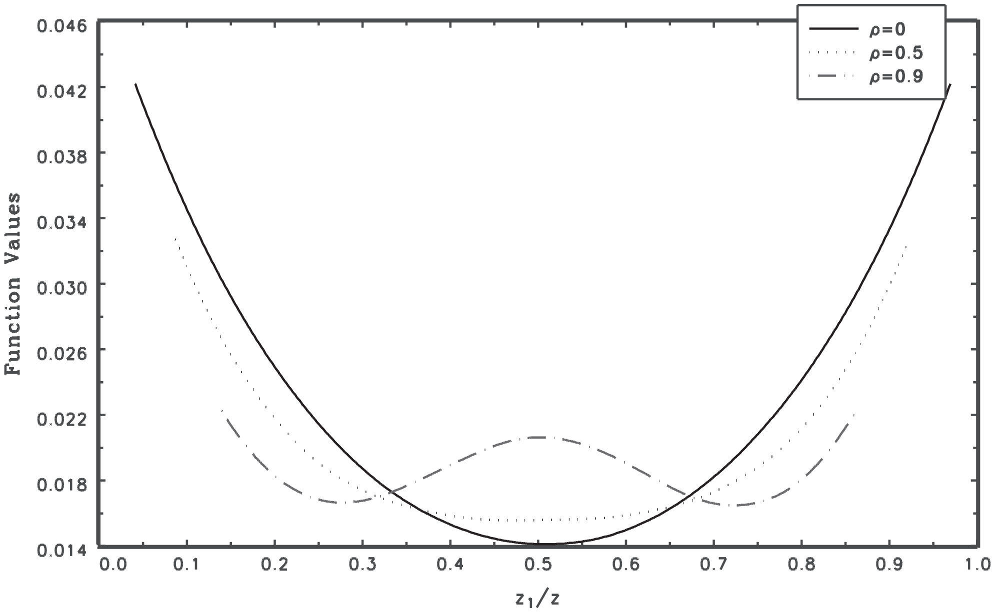


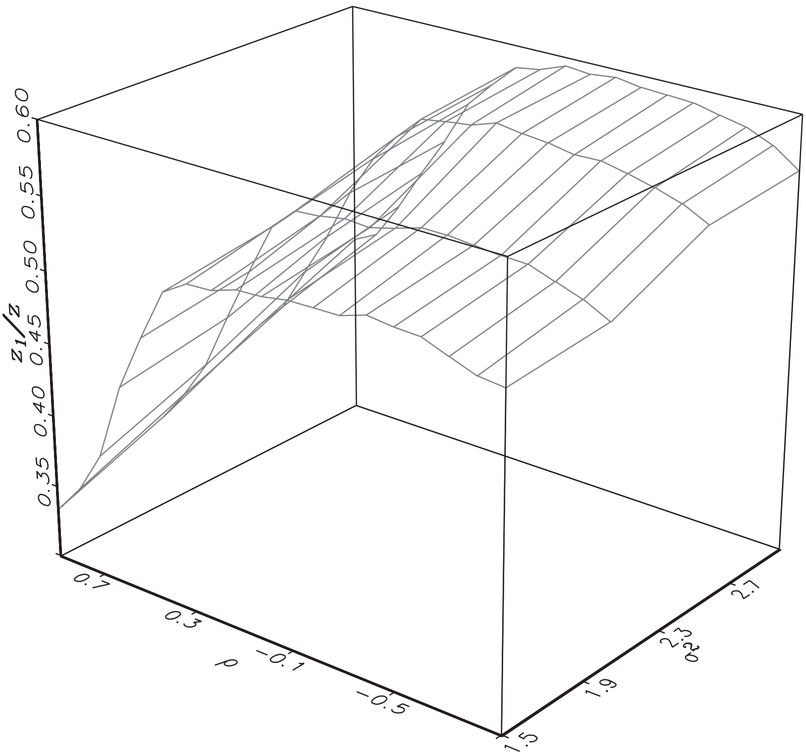
Figure 6 shows a symmetric situation (i.e., $\sigma = 1$), where the objective function values are plotted against the proportion $\hat{\hat{z}}_{1,t}[\hat{z}_t(\alpha^*),\alpha]/\hat{z}_t(\alpha^*)$, for $\rho = 0, 0.5, 0.9$. The reserve level allocated to asset 1 is roughly one-half of the aggregate reserve level for small and medium correlations (say, $\rho < 0.6$). However, if the correlation is large, it is optimal to assign more reserve to one asset than the other. To illustrate this formally, let us consider the limiting case where $\rho = 1$ (i.e., $y_{1,t} = y_{2,t}$ for all $t = 1, \ldots, T$). The FOC (25) is satisfied if $z^*_{1,t} = z^*_{2,t} = z_t/2$. Evaluating at the same value of $z^*_{j,t}$, $j = 1, 2$, the second-order condition (SOC) is given by

$$\begin{aligned}\frac{\partial^2 \Psi}{\partial z_1^2}\left(t, z^*_{1,t}, z_t - z^*_{1,t}\right) &= -2E_t\left\{\left[\mathbb{1}_{y_{1,t+1}>z^*_{1,t}}\mathbb{1}_{y_{2,t+1}>z_t-z^*_{1,t}}\alpha_{11} + \mathbb{1}_{y_{1,t+1}<z^*_{1,t}}\mathbb{1}_{y_{2,t+1}<z_t-z^*_{1,t}}\alpha_{22}\right]\right\} \\ &= -2\left(\alpha_{11} Pr\left[y_{1,t} > \frac{z_t}{2}\right] + \alpha_{22} Pr\left[y_{1,t} < \frac{z_t}{2}\right]\right) < 0. \qquad (26)\end{aligned}$$

The FOC and SOC imply that a local maximum of the criterion value is achieved when $z^*_{1,t}/z_t = 0.5$. Thus, the minimal value of the criterion function, if it exists, must be a different solution $(\tilde{z}^*_{j,t}, j = 1, 2)$ to the FOC (25), where $\tilde{z}^*_{1,t} \neq \tilde{z}^*_{2,t} = z_t - \tilde{z}^*_{1,t}$. If $\tilde{z}^*_{1,t} < \tilde{z}^*_{2,t}$, this implies:

$$\alpha_{11} Pr\left[y_{1,t} > \tilde{z}^*_{2,t}\right] - \alpha_{12} Pr\left[\tilde{z}^*_{1,t} \leq y_{1,t} \leq \tilde{z}^*_{2,t}\right] + \alpha_{22} Pr\left[y_{1,t} < \tilde{z}^*_{1,t}\right] = 0. \qquad (27)$$

FIGURE 7
Relative Allocation as Function of ρ and σ^2



Given condition (27), the SOC evaluated at $(\tilde{z}^*_{j,t},\ j = 1,2)$ is given by

$$\frac{\partial^2 \Psi}{\partial z_1^2}\left(t, \tilde{z}^*_{1,t}, z_t - \tilde{z}^*_{1,t}\right) = \left(z_t - 2\tilde{z}^*_{1,t}\right)\left[\alpha_{22} f\left(\tilde{z}^*_{1,t}\right) - \alpha_{11} f\left(z_t - \tilde{z}^*_{1,t}\right)\right], \tag{28}$$

where f denotes the pdf of y_1 (resp. y_2). In general, a positive SOC requires the following relationship of (α_{11}, α_{22}, f, and z_t):

$$\alpha_{22} f\left(\tilde{z}^*_{1,t}\right) - \alpha_{11} f\left(z_t - \tilde{z}^*_{1,t}\right) > 0.$$

Since $z_{1,t}$ and $z_{2,t}$ are interchangeable, the minimal point is not unique.

We report, in Figure 7, the relative optimal allocation $\hat{\tilde{z}}_{1,t}/\hat{z}_t$ as function of ρ and σ^2. The data are generated with ρ ranging from −0.9 to 0.9 and $\sigma^2 = \{1.5, 2, 2.5, 3\}$. When the value of correlation is small (say less than 0.5), larger reserves are allocated to the riskiest business line. However, the reverse is true for strongly and positively correlated business lines. In both cases, the relative reserve allocation to the riskiest business line increases with the risk level.

Conclusion

A standard validation criterion of the reserve levels proposed by a bank or an insurance company considers the difference between the frequency of exceedance of the VaR and the desired risk level. This criterion is not very informative, since it leads to accept some level of reserves on an historical basis, that is, without distinguishing the environment, such as period of low and high volatility. In this article, we consider the decision criterion underlying the VaR, which accounts for the level of loss beyond the VaR. This criterion can be directly extended to allocate reserves between multiple dependent business lines with or without restrictions on the overall reserve level. This extension shows the importance of considering seriously the loss function, that is, the validation criterion. The appropriate validation criteria depend on various arguments, that are the weights α, the distortion functions A, the instruments g, function of the information set I_t. In a coherent validation, the validator (i.e., the regulator) would have to announce his validation criterion and possibly how the criterion is updated according to the environment. Clearly, the arguments α, A, g are very powerful instruments of economic policy, and can provide a flexible global management of risk on financial markets.

Appendix A: The Component for Lack of Optimality

i) We have

$$\begin{aligned}\Psi(t, z_t) - \Psi(t, \hat{z}_t) &= \alpha(\hat{z}_t - z_t) + E_t\left[(z_t - y_{t+1})\mathbb{1}_{z_t > y_{t+1}}\right] \\ &\quad - E_t\left[(\hat{z}_t - y_{t+1})\mathbb{1}_{\hat{z}_t > y_{t+1}}\right].\end{aligned}$$

Let us consider the case $\hat{z}_t > z_t$. We get

$$\begin{aligned}\Psi(t, z_t) - \Psi(t, \hat{z}_t) &= \alpha(\hat{z}_t - z_t) + E_t\left[(z_t - y_{t+1})\mathbb{1}_{z_t > y_{t+1}} - (\hat{z}_t - y_{t+1})\mathbb{1}_{z_t > y_{t+1}}\right] \\ &\quad - E_t\left[(\hat{z}_t - y_{t+1})\mathbb{1}_{z_t < y_{t+1} < \hat{z}_t}\right] \\ &= \alpha(\hat{z}_t - z_t) + (z_t - \hat{z}_t)P_t\left[z_t > y_{t+1}\right] - E_t\left[(\hat{z}_t - y_{t+1})\mathbb{1}_{z_t < y_{t+1} < \hat{z}_t}\right] \\ &= (\hat{z}_t - z_t)E\left(\mathbb{1}_{z_t < y_{t+1} < \hat{z}_t}\right) - E_t\left[(\hat{z}_t - y_{t+1})\mathbb{1}_{z_t < y_{t+1} < \hat{z}_t}\right] \\ &= E_t\left[(y_{t+1} - z_t)\mathbb{1}_{z_t < y_{t+1} < \hat{z}_t}\right].\end{aligned}$$

A similar computation can be done when $z_t > \hat{z}_t$.

ii) We have

$$
\begin{aligned}
E_t\left[(y_{t+1}-z_t)\mathbb{1}_{z_t<y_{t+1}<\hat{z}_t}\right] &= \int_{z_t}^{\hat{z}_t}(y-z_t)d\,F_t(y) \\
&= (y-z_t)F_t(y)\Big]_{z_t}^{\hat{z}_t} - \int_{z_t}^{\hat{z}_t} F_t(y)dy \\
&= (\hat{z}_t-z_t)F_t(\hat{z}_t) - \int_{z_t}^{\hat{z}_t} F_t(y)dy \\
&= \int_{z_t}^{\hat{z}_t}\left[F_t(\hat{z}_t)-F_t(y)\right]dy.
\end{aligned}
$$

iii) Finally, if $z_t \simeq \hat{z}_t$, we get

$$
\int_{z_t}^{\hat{z}_t}\left[F_t(\hat{z}_t)-F_t(y)\right]dy \simeq \int_{z_t}^{\hat{z}_t} f_t(\hat{z}_t)(\hat{z}_t-y)dy \simeq \frac{1}{2}f_t(\hat{z}_t)(\hat{z}_t-z_t)^2.
$$

APPENDIX B: WEIGHT CORRECTION

Let us assume that $y_{t+1}=m_t+\sigma_t\,u_{t+1}$, where the error term has a path independent distribution with pdf g and cdf G. The conditional pdf and cdf of y_{t+1} are respectively given by:

$$
f_t(y)=\frac{1}{\sigma_t}g\left(\frac{y-m_t}{\sigma_t}\right), \quad F_t(y)=G\left(\frac{y-m_t}{\sigma_t}\right).
$$

We deduce

$$
f_t\left[F_t^{-1}(\alpha)\right]=f_t\left[m_t+\sigma_t G^{-1}(\alpha)\right]=\frac{1}{\sigma_t}g\left[G^{-1}(\alpha)\right].
$$

APPENDIX C: PROOF OF PROPOSITION 1

Proof: Let us consider the first-order condition $\frac{\partial\Psi}{\partial z_1}(t,\hat{\hat{z}}_{1,t},\hat{\hat{z}}_{2,t})=0$. We get

$$
\begin{aligned}
&-P_t(y_{1,t+1}>\hat{\hat{z}}_{1,t})\left[\alpha_{11}E_t(y_{2,t+1}-\hat{\hat{z}}_{2,t})^+ + \alpha_{12}E_t(y_{2,t+1}-\hat{\hat{z}}_{2,t})^-\right] \\
&\quad + P_t(y_{1,t+1}\leq\hat{\hat{z}}_{1,t})\left[\alpha_{21}E_t(y_{2,t+1}-\hat{\hat{z}}_{2,t})^+ + \alpha_{22}E_t(y_{2,t+1}-\hat{\hat{z}}_{2,t})^-\right]=0,
\end{aligned}
$$

or equivalently:

$$
\begin{aligned}
&\left[-(1-\alpha_1)\alpha_{11}+\alpha_1\alpha_{21}\right]E_t(y_{2,t+1}-\hat{\hat{z}}_{2,t})^+ + \left[-(1-\alpha_1)\alpha_{12}+\alpha_1\alpha_{22}\right] \\
&\quad\times E_t(y_{2,t+1}-\hat{\hat{z}}_{2,t})^- = 0,
\end{aligned}
$$

where α_1 denotes the marginal risk level for risk $y_{1,t}$. Since the equality holds for any conditional distribution of $y_{2,t+1}$ with risk level α_2, it follows that

$$-(1-\alpha_1)\alpha_{11}+\alpha_1\alpha_{21}=0 \quad \text{and} \quad -(1-\alpha_1)\alpha_{12}+\alpha_1\alpha_{22}=0,$$

and

$$\alpha_{11}\alpha_{22}=\alpha_{21}\alpha_{12}. \qquad \blacksquare$$

REFERENCES

Andersen, T. G., L. Benzoni, and J. Lund, 2002, An Empirical Investigation of Continuous-Time Models for Equity Returns, *Journal of Finance*, 57(3): 1239–1284.

Andrews, D. W. K., 1991, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, 59(3): 817–858.

Ball, C., and A. Roma, 1994, Stochastic Volatility Option Pricing, *Journal of Financial and Quantitative Analysis*, 29: 589–607.

Diebold, F. X., and R. S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13: 253–263.

Garman, M., 1997, Taking VaR to Pieces, *Risk*, 10: 70–71.

Giacomini, R., and H. White, 2006, Tests of Conditional Predictive Ability, *Econometrica*, 74: 1545–1578.

Gordy, M. B., and E. Lütkebohmert, 2007, Granularity Adjustment for Basel II. Deutsche Bank, DP, 01/2007.

Gourieroux, C., A. Holly, and A. Monfort, 1980, Kuhn-Tucker, Likelihood Ratio and Wald Tests for Nonlinear Models with Inequality Constraints on the Parameters. Harvard Institute of Economic Research, Mimeographed Paper No. 770.

Gourieroux, C., A. Holly, and A. Monfort, 1982, Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters, *Econometrica*, 50: 63–80.

Gourieroux, C., and J. Jasiak, 2006, Autoregressive Gamma Process, *Journal of Forecasting*, 25: 129–152.

Gourieroux, C., and J. Jasiak, 2008, Dynamic Quantile Model, *Journal of Econometrics*, 147: 198–205.

Gourieroux, C., J. P. Laurent, and O. Scaillet, 2000, Sensitivity Analysis of Values at Risk, *Journal of Empirical Finance*, 7: 225–245.

Gourieroux, C., and A. Monfort, 1995, *Statistics and Econometric Models* Vol. 1 (Cambridge, UK: Cambridge University Press).

Hansen, P. R., 2005, A Test for Superior Predictive Ability, *Journal of Business and Economic Statistics*, 23: 365–380.

Harvey, D., S. Leybourne, and P. Newbold, 1997, Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting*, 13(2): 281–291.

Harvey, D., S. Leybourne, and P. Newbold, 1998, Tests for Forecast Encompassing, *Journal of Business and Economic Statistics*, 16: 254–259.

Harvey, D., and P. Newbold, 2000, Tests for Multiple Forecast Encompassing, *Journal of Applied Econometrics*, 15(5): 471–482.

Koenker, R., 2005, *Quantile Regression* (Cambridge, UK: Cambridge University Press).

Kudô, A., 1963, A Multivariate Analogue of the One-Sided Test, *Biometrika*, 50: 403–418.

Mancini, L., and F. Trojani, 2005, Robust Semiparametric Bootstrap Methods for Value-at-Risk Prediction Under GARCH-Type Volatility Processes, Working Paper, University of St. Gallen.

Martin, R., and T. Wilde, 2003, Unsystematic Credit Risk, *Risk Magazine*, 15: 123–128.

Stock, J., and M. Watson, 2005, Forecasting with Many Predictors. in: G. Elliott, C. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting* (North-Holland: Elsevier).

Wilde, T., 2001, Probing Granularity, *Risk Magazine*, 14: 103–106.

Wolak, F., 1987, An Exact Test for Multiple Inequality and Equality Constraints in the Linear Regression Model, *Journal of the American Statistical Association*, 82: 781–793.

Wolak, F., 1991, The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models, *Econometrica*, 59: 981–991.