



Working Paper 09-24
Economic Series (14)
March 2009

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Analytical Value-at-Risk and Expected Shortfall under Regime Switching^{*}

Abderrahim Taamouti[†]
Universidad Carlos III de Madrid
March 30, 2009

Abstract

It is well known that the use of Gaussian models to assess financial risk leads to an underestimation of risk. The reason is because these models are unable to capture some important facts such as heavy tails and volatility clustering which indicate the presence of large fluctuations in returns. An alternative way is to use regime-switching models, the latter are able to capture the previous facts. Using regime-switching model, we propose an analytical approximation for multi-horizon conditional Value-at-Risk and a closed-form solution for conditional Expected Shortfall. By comparing the Value-at-Risks and Expected Shortfalls calculated analytically and using simulations, we find that the both approaches lead to almost the same result. Further, the analytical approach is less time and computer intensive compared to simulations, which are typically used in risk management.

JEL Classification: C22, C19, G11, G19

Keywords: Regime switching; probability distribution; Value-at-Risk; Expected Shortfall; analytical approximation; closed-form solution; simulation; multi-horizon.

^{*} The author thank Taouk Bouezmarni, Jean-Marie Dufour, Xiaodong Lin, Nour Meddahi, Jeroen Rombouts, Chen Ying, and the participants to the Conference on Volatility and High Frequency Data at Chicago (2007) for their helpful comments. This paper was a part of paper entitled .Risk Measures and Portfolio Optimization under Regime Switching Models. Financial support from the Spanish Ministry of Education through grants SEJ 2007-63098 is acknowledged.

[†] Economics Department, Universidad Carlos III de Madrid. Adresse and e-mail: Departamento de Economía Universidad Carlos III de Madrid Calle Madrid, 126 28903 Getafe (Madrid) España. TEL: +34-91 6249863; FAX: +34-91 6249329; e-mail: ataamout@eco.uc3m.es.

1 Introduction

Since the seminal work of Hamilton (1989, *Econometrica*), Markov switching models have been increasingly used in financial time-series econometrics because of their ability to capture some key features, such as heavy tails, volatility clustering, and mean reversion in asset returns [see Cecchetti and al. (1990), Pagan and Schwert (1990), Turner and al. (1989), Gray (1996), Hamilton (1988), and Timmermann (2000), among others]. In this paper, we use these popular models to provide an analytical approximation for multi-horizon conditional Value-at-Risk (hereafter VaR) and derive a closed-form solution for Expected Shortfall (hereafter ES).

VaR has become the most widely used technique to measure and control market risk. It is a quantile measure that quantifies risk for financial institutions and measures the worst expected loss over a given horizon (typically a day or a week) at a given statistical confidence level (typically 1% or 5%). Different methods exist to calculate VaR under different models of risk factors. Generally, there is a trade-off between the simplicity of the calculation method and realism of the assumptions in the risk factor model: As we allow the latter to capture more stylized effects, the calculation method becomes more complex. Under the assumption that returns follow elliptical conditional distribution, one can show that the VaR is given by a simple analytical formula [see for example RiskMetrics (1995) and Bauer (2000)]. However, when we relax this assumption, the analytical calculation of VaR becomes complicated and people tend to use computer intensive simulation based methods. An alternative measure of financial risk is ES defined by the conditional expectation of loss given that the loss is beyond the VaR level. This paper proposes an analytical approximation for conditional Value-at-Risk and a closed-form solution for conditional Expected Shortfall under more realistic assumptions using regime-switching.

The issue of VaR calculation under regime-switching has been considered at least by Billio and Pelizzon (2000) and Guidolin and Timmermann (2006). Billio and Pelizzon (2000) use a switching volatility model to forecast the distribution of returns and calculate the VaR of both single assets and linear portfolios. Comparing the calculated VaR with the variance-covariance approach and *GARCH*(1,1) models, they find that VaR values under switching regime models are preferable to both approaches. Guidolin and Timmermann (2006) examine the term structure of VaR under different econometric approaches, including multivariate regime switching, and they find that bootstrap and regime switching models are best overall for VaR levels of 5% and 1%, respectively. However, to our knowledge, no analytical method has been proposed to calculate the VaR and ES under regime-switching.

In this paper, we derive a closed-form solution for multi-horizon conditional ES under regime switching model. We use the same approach as Cardenas et al. (1997), Rouvinez (1997), and Duffie and Pan (2001) to provide an analytical approximation for multi-horizon conditional VaR. We first

use Fourier inversion method to compute the probability distribution function of multi-horizon portfolio returns. Thereafter, we use an efficient numerical integration step, designed by Davies (1973, 1980), to approximate the infinite integral in the inversion formula and make the calculation of VaR feasible. To account for conditional information and compute the conditional VaR and ES, we use the Hamilton filter. However, our derivations of VaR and ES are made under the assumption that the error terms in the Markov switching model are i.i.d. Consequently, the dependence in our framework is due to the mean and volatility dependence, thus we ignore the dependence of high-order moments. This could be a limitation relative to simulations based calculation of VaR and ES, which may allow for non-i.i.d. errors.

By comparing the Value-at-Risks and Expected Shortfalls calculated analytically and using simulations, we find that the both approaches lead to almost the same results. Further, the analytical approach is less time and computer intensive compared to simulations, which are typically used in risk management.

The remainder of this paper is organized as follows. In Section 2, we introduce some notations and define our model. In Section 3, we derive the multi-horizon conditional VaR and ES of linear portfolio. In Section 4, we compare the simulation and analytical calculations of VaR and ES. We conclude in Section 5.

2 Framework

We assume that there are n risky assets in the economy. We denote by $r_t = (r_{1t}, r_{2t}, \dots, r_{nt})^\top$ the vector of risky assets returns. We consider the following notations:

$$\zeta_t = \begin{cases} (1, 0, 0, \dots, 0)^\top & \text{when } s_t = 1 \\ (0, 1, 0, \dots, 0)^\top & \text{when } s_t = 2 \\ \dots & \\ (0, 0, 0, \dots, 1)^\top & \text{when } s_t = N \end{cases}$$

where s_t is a stationary and homogenous Markov chain with N states, and we define the information sets

$$\begin{aligned} J_t &= \sigma(r_\tau, \zeta_\tau, \tau \leq t) = \sigma(r_\tau, s_\tau, \tau \leq t), \\ I_t &= \sigma(r_\tau, \tau \leq t). \end{aligned}$$

It is well known that [see, e.g., Hamilton (1994, page 679)]

$$E[\zeta_{t+h} | J_t] = P^h \zeta_t, \quad h \geq 1, \tag{1}$$

where P is a transition probability matrix

$$P = [p_{ij}]_{1 \leq i, j \leq N}, \quad p_{ij} = Pr(s_{t+1} = j | s_t = i), \quad \sum_{j=1}^N p_{ij} = 1, \quad \forall i \in \{1, \dots, N\}. \tag{2}$$

The probability p_{ij} of which regime is in operation at time $t+1$ depends on the past only through the most recent value s_t .¹ We assume that the Markov chain is stationary with an ergodic distribution Π , $\Pi \in \mathbb{R}^N$, i.e.

$$E[\zeta_t] = \Pi = (\pi_1, \dots, \pi_N)^\top. \quad (3)$$

Observe that

$$P^h \Pi = \Pi, \forall h. \quad (4)$$

In what follows, we suppose that r_t follows a multivariate Markov switching model

$$\begin{aligned} r_{t+1} &= \mu \zeta_t + \Sigma(\zeta_t) \varepsilon_{t+1}, \quad \varepsilon_{t+1} \text{ i.i.d. } \sim \mathcal{N}(0, I_n), \\ E \left[(\Sigma(\zeta_t) \varepsilon_{t+1}) (\varepsilon_{t+1} \Sigma(\zeta_t))^\top \mid J_t \right] &= \Sigma(\zeta_t) I_n \Sigma(\zeta_t)^\top = \Omega(\zeta_t), \end{aligned} \quad (5)$$

where I_n is an $n \times n$ identity matrix and

$$\mu = \begin{pmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1N} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2N} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nN} \end{pmatrix}, \quad \Omega(\zeta_t) = \begin{pmatrix} \omega_{11}^\top \zeta_t & \omega_{12}^\top \zeta_t & \dots & \omega_{1n}^\top \zeta_t \\ \omega_{21}^\top \zeta_t & \omega_{22}^\top \zeta_t & \dots & \omega_{2n}^\top \zeta_t \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \omega_{n1}^\top \zeta_t & \omega_{n2}^\top \zeta_t & \dots & \omega_{nn}^\top \zeta_t \end{pmatrix},$$

μ_{ij} , for $i = 1, \dots, n$ and $j = 1, \dots, N$, is the mean return of an asset i at state j , and ω_{il} , for $i, l = 1, \dots, n$, is the vector of covariances between assets i and l at the N states. The processes $\{s_t\}$ and $\{\varepsilon_t\}$ are assumed jointly independent.

Finally, we adopt the notations, $\forall U = (u_1, u_2, \dots, u_N)^\top \in \mathbb{R}^N$,

$$A(U) = \text{Diag}(\exp(u_1), \exp(u_2), \dots, \exp(u_N)) P. \quad (6)$$

3 Value-at-Risk and Expected Shortfall under regime switching

In this section, we propose an analytical approximation for conditional VaR under regime switching. We use the conditional characteristic function and a standard Fourier-inversion formula [see for example Gil-Pelaez (1951)] to derive the conditional distribution function, from which the conditional VaR is immediate. Further, we provide a closed-form solution for conditional ES.

3.1 Simple returns

We consider the portfolio allocation between the n risky assets. The portfolio's return at time $t+h$ is given by:

$$r_{p,t+h} = \sum_{i=1}^n \alpha_i^h r_{it+h} = \alpha_h^\top r_{t+h}, \quad (7)$$

¹The assumption of fixed transition probabilities p_{ij} could be relaxed. Many models with time-varying transition probabilities have been considered [see Diebold, Rudebusch and Sichel (1993), Diebold, Ohanian and Berkowitz (1994), Filardo (1994), Lahiri and Wang (1994), Durland and McCurdy (1994), among others].

where r_{t+h} follows multivariate Markov switching model (5) and $\alpha_h = (\alpha_1^h, \alpha_2^h, \dots, \alpha_n^h)^\top$ is the vector representing the weights attributed to each risky asset in the portfolio.

To compute the conditional VaR of the portfolio's return $r_{p,t+h}$ we proceed in two steps. Firstly, we derive the conditional distribution function of $r_{p,t+h}$ by inverting its conditional characteristic function using a standard Fourier-inversion formula [see Gil-Pelaez (1951)]. Secondly, we compute the conditional VaR of $r_{p,t+h}$ by inverting numerically its conditional distribution function using an efficient numerical integration step designed by Davies (1980); for related work, see also Imhof (1961), Bohmann (1961, 1970), and Davies (1973).

We can show [see Taamouti (2008)] that the conditional characteristic function of simple returns $r_{p,t+h}$ is given by, $\forall u \in \mathbb{R}$ and $h \geq 2$,

$$E[\exp(iur_{p,t+h}) | J_t] = e^\top A \left(iu\mu^\top \alpha_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \alpha_{l_1}^h \alpha_{l_2}^h \omega_{l_1 l_2} \right) P^{h-2} \zeta_t, \quad (8)$$

where $i = \sqrt{-1}$, the matrix $A(\cdot)$ is defined in (6), and e denotes the $N \times 1$ vector whose components are all equal to one. The characteristic function (8) depends on the state variable ζ_t . In practice, the current state variable ζ_t is not observable. Using the observable information set I_t , the law of iterated expectations yields

$$E[\exp(iur_{p,t+h}) | I_t] = e^\top A \left(iu\mu^\top \alpha_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \alpha_{l_1}^h \alpha_{l_2}^h \omega_{l_1 l_2} \right) P^{h-2} \Pi_t, \quad \forall u \in \mathbb{R} \quad (9)$$

where

$$\Pi_t = (Pr(s_t = 1 | I_t), \dots, Pr(s_t = N | I_t))^\top.$$

Notice that, Equation (9) is a complex function and using Euler's formula it can be written as follows:

$$E[\exp(iur_{p,t+h}) | I_t] = e^\top [A_1(u) + i A_2(u)] P^{h-1} \Pi_t, \quad \forall u \in \mathbb{R} \quad (10)$$

where, for any $u \in \mathbb{R}$

$$A_1(u) = \begin{bmatrix} \exp\left(\frac{-u^2}{2} (\alpha_h^\top \Omega_1 \alpha_h)\right) \cos(u (\alpha_h^\top \mu_1)) & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & \exp\left(\frac{-u^2}{2} (\alpha_h^\top \Omega_N \alpha_h)\right) \cos(u (\alpha_h^\top \mu_N)) \end{bmatrix},$$

$$A_2(u) = \begin{bmatrix} \exp\left(\frac{-u^2}{2} (\alpha_h^\top \Omega_1 \alpha_h)\right) \sin(u (\alpha_h^\top \mu_1)) & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & \exp\left(\frac{-u^2}{2} (\alpha_h^\top \Omega_N \alpha_h)\right) \sin(u (\alpha_h^\top \mu_N)) \end{bmatrix}. \quad (11)$$

μ_j and Ω_j , for $j = 1, \dots, N$, are the vector of mean and variance-covariance matrix of returns at state j , respectively.

Now, we derive the conditional probability distribution function of $r_{p,t+h}$. Given the conditional characteristic function of $r_{p,t+h}$ from (9), a standard Fourier-inversion formula [see Gil-Pelaez (1951)] implies that, for $r_p \in \mathbb{R}$,

$$Pr(r_{p,t+h} < r_p | I_t) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im}(E[\exp(iur_{p,t+h}) | I_t] \exp(-iur_p))}{u} du, \quad (12)$$

where $\text{Im}(\cdot)$ denotes the imaginary part of a complex number. Using equations (10)-(12), we can show that the conditional distribution function of $r_{p,t+h}$, evaluated at r_p , for $r_p \in \mathbb{R}$, is given by:

$$Pr(r_{p,t+h} < r_p | I_t) = \frac{1}{2} - \frac{1}{\pi} e^\top \int_0^\infty \frac{\bar{A}_2(u, r_p)}{u} du P^{h-1} \Pi_t, \quad (13)$$

where, for any $u \in \mathbb{R}$

$$\bar{A}_2(u, r_p) = \begin{bmatrix} \exp\left(\frac{-u^2}{2} (\alpha_h^\top \Omega_1 \alpha_h)\right) \sin(u(\alpha_h^\top \mu_1 - r_p)) & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & \exp\left(\frac{-u^2}{2} (\alpha_h^\top \Omega_N \alpha_h)\right) \sin(u(\alpha_h^\top \mu_N - r_p)) \end{bmatrix}. \quad (14)$$

VaR is a quantile measure that quantifies risk and measures the worst expected loss over a given horizon h (typically a day or a week) at a given statistical confidence level α (typically 1% or 5%). Considering that conditional Value-at-Risk, say $VaR_t^\alpha(r_{p,t+h})$, is a positive quantity, we have

$$Pr(r_{p,t+h} < -VaR_t^\alpha(r_{p,t+h}) | I_t) = \frac{1}{2} - \frac{1}{\pi} e^\top \int_0^\infty \frac{\bar{A}_2(u, VaR_t^\alpha(r_{p,t+h}))}{u} du P^{h-1} \Pi_t, \quad (15)$$

where $\bar{A}_2(u, VaR_t^\alpha(r_{p,t+h}))$ is defined by the right-hand side of Equation (14) where we replace the constant r_p with $-VaR_t^\alpha(r_{p,t+h})$. The conditional VaR of $r_{p,t+h}$ can be calculated by inverting the conditional distribution function (15). However, inverting analytically (15) is not feasible, for reasons explained below, and a numerical solution (hereafter analytical approximation) is required.

Proposition 1 *The conditional VaR of a portfolio's simple returns $r_{p,t+h}$ with coverage probability α , denoted $VaR_t^\alpha(r_{p,t+h})$, is the solution of the following equation:*

$$e^\top \int_0^\infty \frac{\bar{A}_2(u, VaR_t^\alpha(r_{p,t+h}))}{u} du P^{h-1} \Pi_t - \left(\frac{1}{2} - \alpha\right) \pi = 0, \quad (16)$$

where $\bar{A}_2(u, VaR_t^\alpha(r_{p,t+h}))$ is defined by the right-hand side of Equation (14) where we replace the constant r_p with $-VaR_t^\alpha(r_{p,t+h})$.

Proposition 1 results from Equation (15). Thus, the conditional VaR of $r_{p,t+h}$ can be obtained by solving numerically the equation:

$$f(\text{VaR}_t^\alpha(r_{p,t+h})) = e^\top \int_0^\infty \frac{\bar{A}_2(u, \text{VaR}_t^\alpha(r_{p,t+h}))}{u} du P^{h-1}\Pi_t - \left(\frac{1}{2} - \alpha\right)\pi = 0. \quad (17)$$

The function $f(\text{VaR}_t^\alpha(r_{p,t+h}))$ can be rewritten in the following form:

$$f(\text{VaR}_t^\alpha(r_{p,t+h})) = -\pi [Pr(r_{p,t+h} < -\text{VaR}_t^\alpha(r_{p,t+h}) | I_t) - \alpha]. \quad (18)$$

Using the properties of the probability distribution function [monotonically increasing, $\lim_{x \rightarrow -\infty} Pr(r_{p,t+h} < x) = 0$, and $\lim_{x \rightarrow +\infty} Pr(r_{p,t+h} < x) = 1$] we can show that (17) admits a unique solution. Another way to calculate the conditional VaR of $r_{p,t+h}$ is to consider the following optimization problem:

$$\widehat{\text{VaR}_t^\alpha(r_{p,t+h})} = \arg \min_{\text{VaR}_t^\alpha(r_{p,t+h})} \left[e^\top \int_0^\infty \frac{\bar{A}_2(u, \text{VaR}_t^\alpha(r_{p,t+h}))}{u} du P^{h-1}\Pi_t - \left(\frac{1}{2} - \alpha\right)\pi \right]^2. \quad (19)$$

The following algorithm shows how to estimate the conditional portfolio's VaR using Hamilton filter:

1. Estimate the vector of unknown parameters

$$\theta = \left(\text{vec}(\mu)^\top, \text{vech}(\Omega_1)^\top, \dots, \text{vech}(\Omega_N)^\top, \text{vec}(P)^\top \right)^\top \quad (20)$$

using the maximum-likelihood method [see Hamilton (1994, pages 690-696)]. In Equation (20), “*vec*” denotes the column stacking operator and “*vech*” is the column stacking operator that stacks the elements on and below the diagonal only.

2. Estimate the conditional probabilities of regimes,

$$\Pi_t = \hat{\xi}_{t|t} = (Pr(s_t = 1 | I_t), \dots, Pr(s_t = N | I_t))^\top,$$

by iterating on the following pair of equations [see Hamilton (1994)]:

$$\hat{\xi}_{t|t} = \frac{\left(\hat{\xi}_{t|t-1} \odot \eta_t \right)}{e^\top \left(\hat{\xi}_{t|t-1} \odot \eta_t \right)}, \quad (21)$$

$$\hat{\xi}_{t+1|t} = P \hat{\xi}_{t|t}, \quad (22)$$

where, for $t = 1, \dots, T$,

$$\eta_t = \begin{bmatrix} (2\pi)^{-\frac{T}{2}} | \Omega_1^{-1} |^{\frac{1}{2}} \exp \left\{ (r_t - \mu_1)^\top \Omega_1^{-1} (r_t - \mu_1) \right\} \\ (2\pi)^{-\frac{T}{2}} | \Omega_2^{-1} |^{\frac{1}{2}} \exp \left\{ (r_t - \mu_2)^\top \Omega_2^{-1} (r_t - \mu_2) \right\} \\ \dots \\ (2\pi)^{-\frac{T}{2}} | \Omega_N^{-1} |^{\frac{1}{2}} \exp \left\{ (r_t - \mu_N)^\top \Omega_N^{-1} (r_t - \mu_N) \right\} \end{bmatrix},$$

T is the sample size and the symbol \odot denotes element-by-element multiplication. Given a starting value $\hat{\xi}_{1|0}$ and the estimator $\hat{\theta}^{MV}$ of the vector θ , one can iterate on (21) and (22) to compute the values of $\hat{\xi}_{t|t}$ and $\hat{\xi}_{t+1|t}$ for each date t in the sample. Hamilton (1994, pages 693-694) suggests several options for choosing the starting value $\hat{\xi}_{1|0}$. One approach is to set $\hat{\xi}_{1|0}$ equal to the vector of unconditional probabilities of regimes Π . Another option is to set $\hat{\xi}_{1|0} = \rho$, where ρ is a fixed $N \times 1$ vector of nonnegative constants summing to unity, such as $\rho = N^{-1}e$. Alternatively, ρ can be estimated by maximum likelihood, along with θ , subject to the constraint that $e^\top \rho = 1$ and $\rho_j \geq 0$ for $j = 1, 2, \dots, N$.

3. Given $\hat{\theta}^{MV}$ and Π_t , the portfolio's conditional VaR with coverage probability α can be obtained by solving the optimization problem in (19). Notice that, in practice an exact solution for (19) is not feasible, since the integral $\int_0^\infty \frac{\bar{A}_2(u, VaR_t^\alpha(r_{p,t+h}))}{u} du$ is difficult to evaluate. The latter can be approximated using the results from Imhof (1961), Bohmann (1961, 1970), and Davies (1973, 1980). These authors propose a numerical approximation for distribution function using characteristic function. The proposed approximation introduces two types of errors: discretization and truncation errors. Davies (1973), provides a criterion to control for discretization error, and Davies (1980) suggests three different bounds to control for truncation error. More details about how to approximate numerically the distribution function using the characteristic function can be found in Duffie and Pan (2001).

An alternative measure to assess financial risk is Expected Shortfall, say $ES_t^\alpha(r_{p,t+h})$, defined by the conditional expectation of loss given that the loss is beyond the VaR level:

$$ES_t^\alpha(r_{p,t+h}) = E_t[r_{p,t+h} \mid r_{p,t+h} \leq -VaR_t^\alpha(r_{p,t+h})],$$

where $E_t[\cdot]$ defines the expectation conditional on I_t . The following proposition provides a closed-form solution for the conditional ES of simple returns $r_{p,t+h}$.

Proposition 2 *The conditional ES of a portfolio's simple returns $r_{p,t+h}$ with coverage probability α , denoted $ES_t^\alpha(r_{p,t+h})$, is given by:*

$$ES_t^\alpha(r_{p,t+h}) = \frac{1}{\alpha} e^\top \left[R_1(VaR_t^\alpha(r_{p,t+h})) - \frac{1}{\sqrt{2\pi}} R_2(VaR_t^\alpha(r_{p,t+h})) \right] P^{h-1} \Pi_t, \quad (23)$$

where

$$\begin{aligned} R_1(VaR_t^\alpha(r_{p,t+h})) &= \text{Diag} \left(\alpha_h^\top \mu_1 \Phi \left(- \left(\frac{VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_1}{\sqrt{(\alpha_h^\top \Omega_1 \alpha_h)}} \right) \right), \dots, \alpha_h^\top \mu_N \Phi \left(- \left(\frac{VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_N}{\sqrt{(\alpha_h^\top \Omega_N \alpha_h)}} \right) \right) \right), \\ R_2(VaR_t^\alpha(r_{p,t+h})) &= \text{Diag} \left(\sqrt{(\alpha_h^\top \Omega_1 \alpha_h)} \exp \left(- \frac{1}{2} \frac{(VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_1)^2}{(\alpha_h^\top \Omega_1 \alpha_h)} \right), \dots, \sqrt{(\alpha_h^\top \Omega_N \alpha_h)} \exp \left(- \frac{1}{2} \frac{(VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_N)^2}{(\alpha_h^\top \Omega_N \alpha_h)} \right) \right), \end{aligned}$$

$\Phi(\cdot)$ is the standard normal distribution function. See proof in Appendix A.

The calculation of conditional ES given by Equation (23) does not require any numerical approximation as in case of VaR. In practice, to compute the conditional ES we need to run steps 1 and 2 of the above algorithm. Once we get the estimates of θ and Π_t and for a known $VaR_t^\alpha(r_{p,t+h})$, we evaluate the standard normal distribution functions $\Phi(\cdot)$ and plug-in the formula (23) to get an estimate of the conditional ES of $r_{p,t+h}$.

3.2 Extension to aggregated returns

We extend the discussion on the analytical calculation of VaR to the case of aggregated returns. To compute the VaR of aggregated returns we follow the same steps of Section 3.1 [see Paragraph 2 of Section 3.1]. Unfortunately, a closed-form solution for the ES of aggregated returns may not be tractable for reasons explained below. Now, consider h periods ahead aggregated returns:

$$r_{t:t+h} = \sum_{k=1}^h r_{t+k}, \quad (24)$$

where r_{t+k} follows multivariate Markov switching model (5). The portfolio's aggregated returns is given by:

$$r_{p,t:t+h} = \bar{\alpha}_h^\top r_{t:t+h}, \quad (25)$$

where $\bar{\alpha}_h = (\bar{\alpha}_1^h, \bar{\alpha}_2^h, \dots, \bar{\alpha}_n^h)^\top$ is the vector of weights attributed to each risky asset in the portfolio.

We can show [see Taamouti (2008)] that the conditional characteristic function of $r_{p,t:t+h}$ is given by, $\forall u \in \mathbb{R}$ and $h \geq 2$,

$$\begin{aligned} E[\exp(iur_{p,t:t+h}) | J_t] &= e^\top \left(A \left(iu\mu^\top \bar{\alpha}_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \bar{\alpha}_{l_1}^h \bar{\alpha}_{l_2}^h \omega_{l_1 l_2} \right) \right)^{h-1} \\ &\quad \times \exp \left(\left(iu\mu^\top \bar{\alpha}_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \bar{\alpha}_{l_1}^h \bar{\alpha}_{l_2}^h \omega_{l_1 l_2} \right)^\top \zeta_t \right) \zeta_t, \end{aligned}$$

where the matrix $A(\cdot)$ is defined in (6) and e denotes the $N \times 1$ vector whose components are all equal to one. The law of iterated expectations yields,

$$E[\exp(iur_{p,t:t+h}) | I_t] = e^\top \left(A \left(iu\mu^\top \bar{\alpha}_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \bar{\alpha}_{l_1}^h \bar{\alpha}_{l_2}^h \omega_{l_1 l_2} \right) \right)^{h-1} D(u) \Pi_t, \quad \forall u \in \mathbb{R} \quad (26)$$

where, for any $u \in \mathbb{R}$

$$D(u) = \text{Diag} \left(\exp \left(iu\bar{\alpha}_h^\top \mu_1 - \frac{u^2}{2} \left(\bar{\alpha}_h^\top \Omega_1 \bar{\alpha}_h \right) \right), \dots, \exp \left(iu\bar{\alpha}_h^\top \mu_N - \frac{u^2}{2} \left(\bar{\alpha}_h^\top \Omega_N \bar{\alpha}_h \right) \right) \right).$$

The characteristic function (26) is expressed in terms of the observable information set I_t . Using Euler's formula, the function (26) can be written as follows:

$$E[\exp(iur_{p,t:t+h}) | I_t] = e^\top [D_1(u) + i D_2(u)] \Pi_t, \quad \forall u \in \mathbb{R}$$

where, for any $u \in \mathbb{R}$

$$D_1(u) = \text{Re} \left(\left(A \left(iu\mu^\top \bar{\alpha}_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \bar{\alpha}_{l_1}^h \bar{\alpha}_{l_2}^h \omega_{l_1 l_2} \right) \right)^{h-1} D(u) \right),$$

$$D_2(u) = \text{Im} \left(\left(A \left(iu\mu^\top \bar{\alpha}_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \bar{\alpha}_{l_1}^h \bar{\alpha}_{l_2}^h \omega_{l_1 l_2} \right) \right)^{h-1} D(u) \right).$$

$\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real and imaginary parts of a complex matrix, respectively.

Using Gil-Pelaez's (1951) inversion formula, the conditional distribution function of $r_{p,t:t+h}$, evaluated at \bar{r}_p , for $\bar{r}_p \in \mathbb{R}$, is given by:

$$\text{Pr}(r_{p,t:t+h} < \bar{r}_p \mid I_t) = \frac{1}{2} - \frac{1}{\pi} e^\top \int_0^\infty \frac{\bar{D}_2(u, \bar{r}_p)}{u} du \Pi_t,$$

where, for any $u \in \mathbb{R}$

$$\bar{D}_2(u, \bar{r}_p) = \text{Im} \left\{ \exp(-iu\bar{r}_p) A \left(iu\mu^\top \bar{\alpha}_h - \frac{u^2}{2} \sum_{1 \leq l_1, l_2 \leq n} \bar{\alpha}_{l_1}^h \bar{\alpha}_{l_2}^h \omega_{l_1 l_2} \right)^{h-1} D(u) \right\}. \quad (27)$$

An explicit expression for matrix $\bar{D}_2(u, \bar{r}_p)$ is not easy to compute when the horizon h is large and this makes it hard to get a closed-form solution for ES. However, for a given short horizon h , one can calculate this expression and get an analytical formula for ES.

Proposition 3 *The conditional VaR of a portfolio's aggregated returns $r_{p,t:t+h}$ with coverage probability α , denoted $\text{VaR}_t^\alpha(r_{p,t:t+h})$, is the solution of the following equation:*

$$e^\top \int_0^\infty \frac{\bar{D}_2(u, \text{VaR}_t^\alpha(r_{p,t:t+h}))}{u} du \Pi_t - \left(\frac{1}{2} - \alpha\right)\pi = 0,$$

where $\bar{D}_2(u, \text{VaR}_t^\alpha(r_{p,t:t+h}))$ is defined by the right-hand side of Equation (27) where we replace the constant \bar{r}_p with $-\text{VaR}_t^\alpha(r_{p,t:t+h})$.

In practice, to calculate the conditional VaR of $r_{p,t:t+h}$, we follow the same algorithm described in section (3.1). Given an estimator $\hat{\theta}^{MV}$ of the vector

$$\theta = \left(\text{vec}(\mu)^\top, \text{vech}(\Omega_1)^\top, \dots, \text{vech}(\Omega_N)^\top, \text{vec}(P)^\top \right)^\top \quad (28)$$

and the conditional probabilities of regimes Π_t , the conditional VaR of $r_{p,t:t+h}$ with coverage probability α can be obtained by solving the following optimization problem:

$$\widehat{\text{VaR}}_t^\alpha(r_{p,t:t+h}) = \arg \min_{\text{VaR}_t^\alpha(r_{p,t:t+h})} \left[e^\top \int_0^\infty \frac{\bar{D}_2(u, \text{VaR}_t^\alpha(r_{p,t:t+h}))}{u} du \Pi_t - \left(\frac{1}{2} - \alpha\right)\pi \right]^2, \quad (29)$$

where the integral $\int_0^\infty \frac{\bar{D}_2(u, \text{VaR}_t^\alpha(r_{p,t:t+h}))}{u} du$ can be approximated using the results from Imhof (1961), Bohmann (1961, 1970), and Davies (1973, 1980) [see Section (3.1)].

In the next section, we compare the analytical and simulation calculations of VaR and ES using a multivariate regime switching model estimated from a real data.

4 Analytical versus simulation calculations of VaR and ES

In this section, we compare the analytical and simulation calculations of VaR and ES under regime switching. Our comparison is based on the following multivariate regime switching model:²

$$\begin{aligned} \text{State 1: } r_{t+1} &= \begin{pmatrix} 0.0096 \\ 0.0010 \end{pmatrix} + \Sigma_1 \varepsilon_{t+1} \quad \text{Var} [\Sigma_1 \varepsilon_{t+1}] = \Omega_1 = \begin{pmatrix} 0.0006 & -0.0003 \\ -0.0003 & 0.0009 \end{pmatrix}, \\ \text{State 2: } r_{t+1} &= \begin{pmatrix} -0.005 \\ -0.0003 \end{pmatrix} + \Sigma_2 \varepsilon_{t+1} \quad \text{Var} [\Sigma_2 \varepsilon_{t+1}] = \Omega_2 = \begin{pmatrix} 0.0025 & 4.5265 \cdot 10^{-5} \\ 4.5265 \cdot 10^{-5} & 0.0029 \end{pmatrix}, \end{aligned} \quad (30)$$

where $\varepsilon_{t+1} \sim \mathcal{N}(0, I)$, and the transition probability matrix is given by:

$$P = \begin{bmatrix} 0.96 & 0.126 \\ 0.04 & 0.874 \end{bmatrix}. \quad (31)$$

The parameter values in the model (30)-(31) are obtained by estimating a two-regime Markov switching model using a real data. The latter consists of monthly observations on S&P composite index and 10-years Government Bond from January 1958 to December 2006 for a total of 588 observations. The returns are computed using the standard continuous compounding formula. We applied the Hamilton filter to recuperate the conditional probabilities Π_t [see Section 3.1] and compute the conditional VaR and ES.

To compute analytically the conditional VaR, we apply the algorithm described in Section 3.1.³ We calculate the VaR and ES for different portfolios constructed by considering a number of different investment strategies: (A) 100% stock; (B) 75% stock and 25% bond; (C) 50% stock and 50% bond; (D) 25% stock and 75% bond; and (E) 100% bond. Using model (30)-(31), the analytical and simulated values of VaR of aggregated returns (25) and ES of simple returns (7) at horizons $h = 1, 2, 3, 4, 5$ are presented in Tables 1 and 2, respectively.

The second column of Table 1 shows the 1% VaR calculated using 100000 simulations and the third and fourth columns present the 1% VaR calculated using an analytical approximation under two different approximation errors 10^{-3} and 10^{-5} , respectively. From this table we draw the following conclusions. First, there is almost no difference between the analytical and simulated values of 1% VaR at horizons under consideration; this is true for all the investment strategies that we consider. Second, it seems that investing only in stock (100% stock) or bond (100% bond) is more risky, since for these portfolios the VaRs are higher. Finally, the best investment strategy, in terms of reducing risk, is the one that corresponds to 50% stock and 50% bond.

²We also consider other univariate models and our conclusion, that there is almost no difference between simulation and analytical solutions of VaR, does not change. The results are available from the author upon request.

³The details about how to control for the discretization and truncation errors in the numerical approximation of the integral $\int_0^\infty \frac{\bar{D}_2(u, \text{VaR}_t^\alpha(r_{p,t:t+h}))}{u} du$ and a GAUSS code for calculation of ES and VaR are available from the author upon request.

Table 2 compares the analytical and simulation calculations of 1% ES for different investment strategies that we discussed above. The second column shows the 1% ES calculated using 100000 simulations and the third column presents the 1% ES calculated using the analytical formula of Proposition 2. From this table we draw the following conclusions. First, there is almost no difference between the analytical and simulated values of 1% ES. Second, as we find for VaR, it is more risky to invest only in stock or bond. Further, the best investment strategy, in terms of reducing risk, is the one that corresponds to 50% stock and 50% bond.

Table 1: Simulation and analytical approximation of 1% VaR of a portfolio's aggregated returns $r_{p,t:t+h}$

Horizon	Simulation	Analytical Approximation	
		Approx Error = 10^{-3}	Approx Error = 10^{-5}
<i>(A) 100% Stock</i>			
1	0.1067	0.1071	0.1062
2	0.1517	0.1606	0.1521
3	0.1853	0.1861	0.1861
4	0.2137	0.2127	0.2138
5	0.2380	0.2370	0.2390
<i>(B) 75% Stock 25% Bond</i>			
1	0.0858	0.0887	0.0851
2	0.1204	0.1220	0.1210
3	0.1476	0.1519	0.1480
4	0.1696	0.1669	0.1701
5	0.1882	0.1846	0.1892
<i>(C) 50% Stock 50% Bond</i>			
1	0.0783	0.0821	0.0774
2	0.1099	0.1116	0.1112
3	0.1342	0.1386	0.1360
4	0.1535	0.1505	0.1541
5	0.1706	0.1663	0.1715
<i>(D) 25% Stock 75% Bond</i>			
1	0.0883	0.0902	0.0889
2	0.1227	0.1242	0.1233
3	0.1490	0.1527	0.1501
4	0.1707	0.1692	0.1710
5	0.1889	0.1861	0.1901
<i>(E) 100% Bond</i>			
1	0.1104	0.1109	0.1094
2	0.1542	0.1620	0.1549
3	0.1868	0.1877	0.1877
4	0.2135	0.2126	0.2138
5	0.2367	0.2359	0.2387

Note: This table presents the simulated and analytical approximation values of 1% VaR of aggregated returns given by Equation (25) under model (30)-(31) and at horizons $h = 1, 2, 3, 4, 5$, where the horizon lengths are in monthly units. The calculated 1% VaRs correspond to different portfolios constructed by considering a number of different investment strategies: (A) 100% stock; (B) 75% stock and 25% bond; (C) 50% stock and 50% bond; (D) 25% stock and 75% bond; and (E) 100% bond. The number of simulations used to compute the 1% VaRs is 100000.

Table 2: Analytical and simulation calculations of 1% ES of a portfolio's simple returns $r_{p,t+h}$

Horizon	Simulation	Analytical
(A) 100% Stock		
1	-0.1128	-0.1125
2	-0.1110	-0.1130
3	-0.1036	-0.0991
4	-0.1031	-0.1097
5	-0.1036	-0.0973
(B) 75% Stock 25% Bond		
1	-0.0893	-0.0927
2	-0.0888	-0.0943
3	-0.0822	-0.0792
4	-0.0825	-0.0812
5	-0.0835	-0.0811
(C) 50% Stock 50% Bond		
1	-0.0815	-0.0812
2	-0.0826	-0.0820
3	-0.0748	-0.0749
4	-0.0750	-0.0754
5	-0.0754	-0.0761
(D) 25% Stock 75% Bond		
1	-0.0921	-0.0930
2	-0.0924	-0.0914
3	-0.0850	-0.0848
4	-0.0861	-0.0837
5	-0.0855	-0.0859
(E) 100% Bond		
1	-0.1156	-0.1219
2	-0.1151	-0.1204
3	-0.1048	-0.1191
4	-0.1062	-0.1169
5	-0.1078	-0.1192

Note: This table presents the analytical and simulated values of 1% ES of simple returns given by Equation (7) under model (30)-(31) and at horizons $h = 1, 2, 3, 4, 5$, where the horizon lengths are in monthly units. The calculated 1% ES correspond to different portfolios constructed by considering a number of different investment strategies: (A) 100% stock; (B) 75% stock and 25% bond; (C) 50% stock and 50% bond; (D) 25% stock and 75% bond; and (E) 100% bond. The number of simulations used to compute the 1% ES is 100000.

In Figure 1, we calculate 100 values of 1% ES using the analytical formula in Proposition 2 and 10000 and 100000 simulations. In this figure, we assume that we are at time t and we compute the 1% ES of simple returns given by Equation (7) at horizons $t + 1, t + 2, \dots, t + 100$, where the horizon lengths are in monthly units. To get a conditional ES we use the conditional probabilities Π_t [see Section 3.1]. The computational time of computing these values is presented in Table 3.⁴

⁴We use GAUSS for the analytical and simulation calculations of VaR and ES. Some characteristics of the computer hardware employed are:

- (1) Memory (RAM): 3582 MB;
- (2) Processor: intel(R) Core (TM)2 Duo CPU T7500 @ 2.20GHz 2.20 GHz;
- (3) System type: 32-bit Operating System.

From the latter, we see that using the analytical formula requires less than 1 second, whereas if we use 100000 simulations the computational time is about 3 hours, 32 minutes, and 7 seconds. For 10000 simulations, the computational time decreases to 21 minutes and 18 seconds. However, Figure 2 shows that using 10000 simulations may overestimate the values of the 1% ES.⁵ Thus, we need a very large number of simulations in order to get a good approximation for the tails of the distribution of returns and a large number of simulations requires several hours.

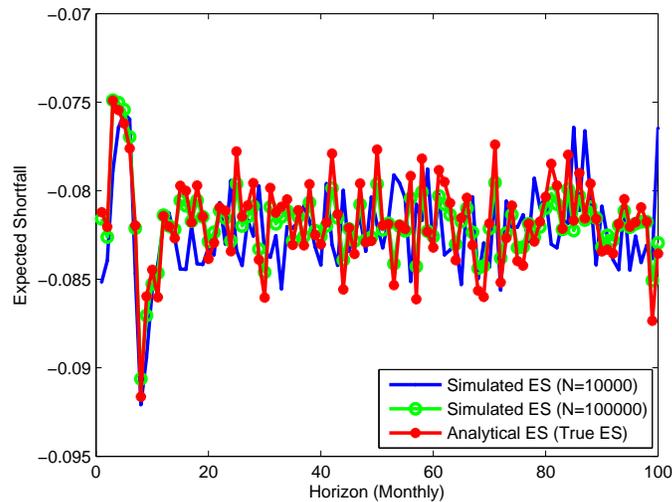


Figure 1: *This figure presents the analytical and simulated 100 values of 1% ES of simple returns given by Equation (7) under model (30)-(31) and using the analytical formula in Proposition 2 and different number of simulations (N). To compute these values, we assume that we are at time t and we compute the 1% ES of simple returns at horizons $t+1, t+2, \dots, t+100$, where the horizon lengths are in monthly units. The calculated values of 1% ES correspond to the investment strategy 50% stock and 50% bond.*

Table 3: Computational time of 100 values of 1% ES using analytical and simulation methods

	100 Analytical 1% ES	100 Simulated 1% ES	
		$N = 10000$	$N = 100000$
Computational Time	Less than 1 sec	21 min 18 sec	3 h 32 min 7 sec

Note: *This table presents the computational time of computing 100 values of 1% ES of simple returns (7) under model (30)-(31) and using the analytical and simulation methods. The calculated values of 1% ES correspond to the investment strategy 50% stock and 50% bond. N represents the number of simulations.*

⁵Similarly to Figure 1, in Figure 2 we calculate 5 values of 1% ES using the analytical formula in Proposition 2 and 10000, 50000, 100000, 500000, 1000000, and 3000000 simulations. In this figure, we assume that we are at time t and we compute the 1% ES of simple returns given by Equation (7) at horizons $t+1, t+2, \dots, t+5$, where the horizon lengths are in monthly units.

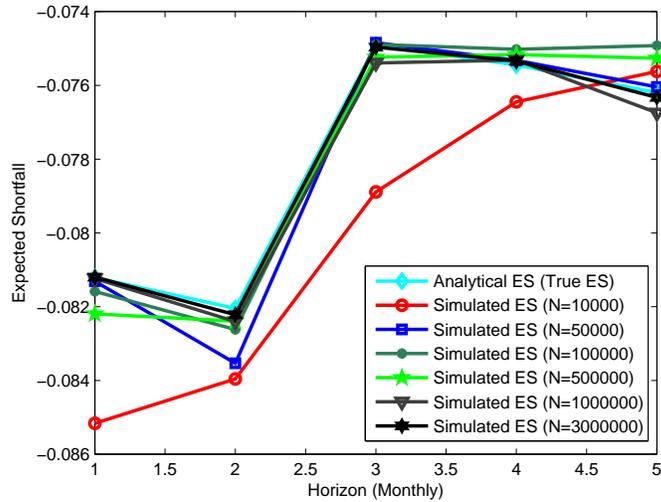


Figure 2: This figure presents the analytical and simulated 5 values of 1% ES of simple returns given by Equation (7) under model (30)-(31) and using the analytical formula in Proposition 2 and different number of simulations (N). To compute these values, we assume that we are at time t and we compute the 1% ES of simple returns at horizons $t + 1, t + 2, \dots, t + 5$, where the horizon lengths are in monthly units. The calculated values of 1% ES correspond to the investment strategy 50% stock and 50% bond.

5 Conclusion

We consider a regime switching model to capture important features such as heavy tails, persistence, and nonlinear dynamics in returns. These features are crucial to assess financial risk. Using Fourier inversion method, we propose an analytical approximation for multi-horizon Value-at-Risk and a closed-form solution for Expected Shortfall under regime-switching. By comparing the Value-at-Risks and Expected Shortfalls calculated analytically and using simulations, we find that the both approaches lead to almost the same result. Further, the analytical approach is less time and computer intensive compared to simulations, which are typically used in risk management.

References

- [1] Billio, M., Pelizzon, L., 2000. Value-at-Risk: a multivariate switching regime approach. *Journal of Empirical Finance* 7, 531-554.
- [2] Bauer, C., 2000. Value-at-Risk using hyperbolic distributions. *Journal of Economics and Business* 52, 455-467.
- [3] Bohmann, H., 1961. Approximate Fourier analysis of distribution functions. *Arkiv für Matematik* 4, 99-157.
- [4] Bohmann, H., 1970. A method to calculate the distribution function when the characteristic function is known. *Nordisk Tidskr. Informationsbehandling (BIT)* 10, 237-242.
- [5] Cardenas, J., Fruchard, E., Koehler, E., Michel, C., Thomazeau, I., 1997. VAR: One step beyond. *Risk* 10 (10), 72-75.
- [6] Cecchetti, S.G., Lam, P., Mark, N.C., 1990. Mean reversion in equilibrium asset prices. *American Economic Review* 80, 398-418.
- [7] Davies, R.B., 1973. Numerical inversion of a characteristic function. *Biometrika* 60 (2), 415-417.
- [8] Davies, R.B., 1980. The distribution of a linear combination of χ^2 random variables. *Applied Statistics* 29, 323-333.
- [9] Diebold, F.X., Rudebusch, G.D., Sichel, D.E., 1993. Further evidence on business cycle duration dependence. In Stock, J., and Watson, M. (eds.), *Business Cycles, Indicators, and Forecasting*, 255-280. Chicago: University of Chicago Press and NBER.
- [10] Diebold, F.X., Ohanian, L.E., Berkowitz, J., 1994. Dynamic equilibrium economies: A framework for comparing models and data. *Review of Economic Studies* 65, 433-451.
- [11] Duffie, D., and Pan, J., 2001. Analytical Value-at-Risk with Jumps and Credit Risk. *Finance and Stochastics* 5 (2), 155-180.
- [12] Durland, J.M., McCurdy, T.H., 1994. Duration dependent transitions in a Markov model of U.S. GNP growth. *Journal of Business and Economic Statistics* 12, 279-288.
- [13] Gil-Pelaez, J., 1951. Note on the inversion theorem. *Biometrika* 38, 481-482.
- [14] Gray, S.F., 1996. Modeling the conditional distribution of interest rates as a regime switching process. *Journal of Financial Economics* 42, 27-62.
- [15] Guidolin, M., Timmermann, A., 2006. Term structure of risk under alternative econometric specifications. *Journal of Econometrics* 131, 285-308.
- [16] Feller, W., 1966. *An introduction to probability theory and its applications*. New York: Wiley.
- [17] Filardo, A.J., 1994. Business cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics* 12, 299-308.
- [18] Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57 (2), 357-384.

- [19] Hamilton, J.D., 1988. Rational expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control* 12, 385-423.
- [20] Hamilton, J.D., 1994. *Time series analysis*. Princeton University Press.
- [21] Imhof, J.P., 1961. Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419-426.
- [22] Lahiri, K., Wang, J.G., 1994. Predicting cyclical turning points with leading index in a Markov switching model. *Journal of Forecasting* 13, 245-263.
- [23] Pagan, A., Schwert, G.W., 1990. Alternative models for conditional stock volatility. *Journal of Econometrics* 45, 267-90.
- [24] RiskMetrics, 1995. Technical document, JP Morgan. New York, USA.
- [25] Rouvinez, C., 1997. Going greek with VaR, *Risk* 10 (2): 57-65.
- [26] Taamouti, A., 2008. The risk-return trade-off under regime switching. Working Paper, Universidad Carlos III de Madrid.
- [27] Timmermann, A., 2000. Moments of Markov switching models. *Journal of Econometrics* 96, 75-111.
- [28] Turner, C.M, Startz, R., Nelson, C.R., 1989. A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics* 25, 3-22.

A Appendix: Proof of Propositions

Proof of Proposition 2.

$$\begin{aligned}
ES_t^\alpha(r_{p,t+h}) &= E_t[r_{p,t+h} \mid r_{p,t+h} \leq -VaR_t^\alpha(r_{p,t+h})] \\
&= \int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} r_p f_t(r_p \mid r_p \leq -VaR_t^\alpha(r_{p,t+h})) dr_p \\
&= \int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} r_p \frac{\sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \frac{1}{\sqrt{2\pi(\alpha_h^\top \Omega_j \alpha_h)}} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right)}{Pr(r_p \leq -VaR_t^\alpha(r_{p,t+h}) \mid I_t)} dr_p.
\end{aligned}$$

Since $Pr\left(r_p \leq -VaR_t^\alpha(r_{p,t+h}) \mid I_t\right) = \alpha$, we get

$$\begin{aligned}
& ES_t^\alpha(r_{p,t+h}) \\
&= \frac{1}{\alpha\sqrt{2\pi}} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{r_p}{\sqrt{(\alpha_h^\top \Omega_j \alpha_h)}} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p \\
&= \frac{-1}{\alpha\sqrt{2\pi}} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \left[\sqrt{(\alpha_h^\top \Omega_j \alpha_h)} \int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{-r_p}{(\alpha_h^\top \Omega_j \alpha_h)} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p \right] \\
&\quad - \frac{1}{\alpha\sqrt{2\pi}} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \left[\sqrt{(\alpha_h^\top \Omega_j \alpha_h)} \int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{\alpha_h^\top \mu_j}{(\alpha_h^\top \Omega_j \alpha_h)} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p \right] \\
&\quad + \frac{1}{\alpha\sqrt{2\pi}} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \left[\sqrt{(\alpha_h^\top \Omega_j \alpha_h)} \int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{\alpha_h^\top \mu_j}{(\alpha_h^\top \Omega_j \alpha_h)} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p \right] \\
&= \frac{-1}{\alpha\sqrt{2\pi}} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \sqrt{(\alpha_h^\top \Omega_j \alpha_h)} \left[\int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{(-r_p + \alpha_h^\top \mu_j)}{(\alpha_h^\top \Omega_j \alpha_h)} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p \right] \\
&\quad + \frac{1}{\alpha} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) (\alpha_h^\top \mu_j) \left[\int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{1}{\sqrt{2\pi}\sqrt{(\alpha_h^\top \Omega_j \alpha_h)}} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p \right].
\end{aligned}$$

Notice that

$$\int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{(-r_p + \alpha_h^\top \mu_j)}{(\alpha_h^\top \Omega_j \alpha_h)} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p = \exp\left(-\frac{1}{2} \frac{(VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) \quad (32)$$

and

$$\int_{-\infty}^{-VaR_t^\alpha(r_{p,t+h})} \frac{1}{\sqrt{2\pi}\sqrt{(\alpha_h^\top \Omega_j \alpha_h)}} \exp\left(-\frac{1}{2} \frac{(r_p - \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) dr_p = \Phi\left(-\left(\frac{VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_j}{\sqrt{(\alpha_h^\top \Omega_j \alpha_h)}}\right)\right), \quad (33)$$

where $\Phi(\cdot)$ is the standard normal distribution function. Thus, given (32) and (33), we have

$$\begin{aligned}
ES_t^\alpha(r_{p,t+h}) &= \frac{-1}{\alpha\sqrt{2\pi}} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) \sqrt{(\alpha_h^\top \Omega_j \alpha_h)} \exp\left(-\frac{1}{2} \frac{(VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_j)^2}{(\alpha_h^\top \Omega_j \alpha_h)}\right) \\
&\quad + \frac{1}{\alpha} \sum_{j=1}^N Pr(s_{t+h-1} = j \mid I_t) (\alpha_h^\top \mu_j) \Phi\left(-\left(\frac{VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_j}{\sqrt{(\alpha_h^\top \Omega_j \alpha_h)}}\right)\right).
\end{aligned}$$

$ES_t^\alpha(r_{p,t+h})$ can be written as follow

$$ES_t^\alpha(r_{p,t+h}) = \frac{1}{\alpha} e^\top \left[R_1 (VaR_t^\alpha(r_{p,t+h})) - \frac{1}{\sqrt{2\pi}} R_2 (VaR_t^\alpha(r_{p,t+h})) \right] P^{h-1} \Pi_t,$$

where

$$\begin{aligned}
R_1 (VaR_t^\alpha(r_{p,t+h})) &= \text{Diag} \left(\alpha_h^\top \mu_1 \Phi\left(-\left(\frac{VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_1}{\sqrt{(\alpha_h^\top \Omega_1 \alpha_h)}}\right)\right), \dots, \alpha_h^\top \mu_N \Phi\left(-\left(\frac{VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_N}{\sqrt{(\alpha_h^\top \Omega_N \alpha_h)}}\right)\right) \right), \\
R_2 (VaR_t^\alpha(r_{p,t+h})) &= \text{Diag} \left(\sqrt{(\alpha_h^\top \Omega_1 \alpha_h)} \exp\left(-\frac{1}{2} \frac{(VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_1)^2}{(\alpha_h^\top \Omega_1 \alpha_h)}\right), \dots, \sqrt{(\alpha_h^\top \Omega_N \alpha_h)} \exp\left(-\frac{1}{2} \frac{(VaR_t^\alpha(r_{p,t+h}) + \alpha_h^\top \mu_N)^2}{(\alpha_h^\top \Omega_N \alpha_h)}\right) \right),
\end{aligned}$$

$\Phi(\cdot)$ is the standard normal distribution function. See proof in Appendix. ■