PERFORMANCE AND CONSERVATISM OF MONTHLY FHS VAR: AN INTERNATIONAL INVESTIGATION [§]

Stéphane Chrétien[¤]
Frank Coggins[*]

This draft: September 2008

Abstract

This study examines sixteen models of monthly Value-at-Risk (VaR) for three equity indices with an emphasis on the filtered historical simulation (FHS) technique. We investigate the importance of historical simulation versus a parametrized approach, the presence of filter versus a static modeling of the return distribution, the choice of GARCH versus RiskMetrics conditional variances and the use of monthly versus daily data sampling frequencies. Tests for unconditional and conditional coverage and for independence show that two daily GARCH-type FHS models perform the best. The most conservative daily FHS model, an asymmetric GARCH specification, indicates that the CRSP value-weighted index, the DAX index and the NIKKEI 225 index have a 5% probability of a respective loss averaging at least 6.9%, 8.7% and 9.3% of their value over one month.

JEL classifications: G11, G23

Keywords: VaR models with filtered historical simulation, GARCH models, Unconditional and conditional coverage tests, Conservatism tests

[§] The authors thank Paul Gallant, Éric Maillé, Alexandre Roy, Mélissa Tremblay and seminar participants at Sherbrooke University for their comments, and Alexandre Fortier and Sébastien Rousseau for their excellent research assistance. Stéphane Chrétien gratefully acknowledges the financial support from the *Institut de finance mathématique de Montréal* and the Faculty of Business Administration at Laval University. Frank Coggins gratefully acknowledges the financial support of the Faculty of Business at Sherbrooke University. The authors are associated researchers at CIRPÉE.
[¤] Department of Finance and Insurance, Faculty of Business Administration, Laval University, Pavillon Palasis-Prince, 2325 rue Terrasse, Quebec City (Québec), Canada, G1V 0A6. Email: Stéphane.Chrétien@fas.ulaval.ca. Phone: (418) 656-2131 ext. 3380.
[*] Département de finance, Faculté d'administration, Université de Sherbrooke, 2500 Boulevard Université, Sherbrooke (Québec), J1K 2R1, Canada. Email: Frank.Coggins@usherbrooke.ca. Phone: (819) 821-8000, ext. 65156.

PERFORMANCE AND CONSERVATISM OF MONTHLY FHS VAR: AN INTERNATIONAL INVESTIGATION

Abstract

This study examines sixteen models of monthly Value-at-Risk (VaR) for three equity indices with an emphasis on the filtered historical simulation (FHS) technique. We investigate the importance of historical simulation versus a parametrized approach, the presence of filter versus a static modeling of the return distribution, the choice of GARCH versus RiskMetrics conditional variances and the use of monthly versus daily data sampling frequencies. Tests for unconditional and conditional coverage and for independence show that two daily GARCH-type FHS models perform the best. The most conservative daily FHS model, an asymmetric GARCH specification, indicates that the CRSP value-weighted index, the DAX index and the NIKKEI 225 index have a 5% probability of a respective loss averaging at least 6.9%, 8.7% and 9.3% of their value over one month.

1. INTRODUCTION

In the context of highly volatile and sometimes crisis-prone financial markets, the Value-at-Risk (hereafter VaR) measure has become an important risk management instrument for numerous organizations[1]. Conceptually simple, the VaR corresponds to a loss that should only be exceeded with a given target probability on a given time horizon. By focusing on the left tail of the return distribution, the VaR provides an intuitive measure of the downside risk of an investment.

Following the increased need for reliable quantitative risk management tool, researchers in the last few decades have developed a large number of VaR approaches. For example, Kuester, Mittnik and Paolella (2006) give a list which includes approaches based on mixture of distributions, extreme value theory, quantile regression, regime switching, realized volatility, option-implied volatility and stochastic volatility. Among the most promising approaches is the filtered historical simulation (hereafter FHS) technique introduced by Barone-Adesi, Bourgoin and Giannopoulos (1998) and Barone-Adesi, Giannopoulos and Vosper (1999). The FHS technique is a semiparametric method that forecasts the mean and variance of returns through a parametric specification and uses the percentile of the standardized returns in order to calculate the VaR. The goal of this study is to formally investigate the out-of-sample performance of sixteen models of monthly VaR for three equity indices with an emphasis on the FHS technique. We make three contributions to the literature.

Our first contribution is to provide a better understanding of the relevance of four important methodological features of FHS VaR models through our selection of models. Specifically, our choice of models highlights 1- the use of historical simulation, where the realized return distribution is assumed to be representative of the one expected over the VaR horizon, versus a parametrized approach, where an analytical VaR is computed using a Normal distribution, a Student-t distribution or a Cornish-Fisher approximation; 2- the presence of filter, where the specification of time variation in returns is required, resulting in a conditional VaR, versus a static

[1] VaR applications include the disclosure of risk to officers and shareholders of corporations, the allocation of resources and performance evaluation in companies, the risk management of institutional portfolios like pension or investment funds, the calculation of the legal capital requirements of financial institutions regulated by the Basle II agreement, the risk management of the market positions of brokers and abitragists, etc [Jorion (2006)].

modeling of the distribution of returns; 3- the choice of the GARCH-type conditional volatilities of Engle (1982), Bollerslev (1986) and Glosten, Jagannathan and Runkle (1993), which have been documented to perform well in the academic literature[2], versus the JP Morgan's RiskMetrics-type conditional specification, which is particularly popular in the industry; 4- the use of a monthly sampling frequency, which corresponds to VaR horizon investigated, versus the compounding of daily observations, which could provide more information for precise econometric estimation on the dynamics of the return distribution.

Our second contribution is the examination of the monthly VaR horizon and international equity market risk relevant to institutional portfolio management. Specifically, the market risk we are interested is reflected in the monthly returns on the American CRSP value-weighted index, the German DAX index and the Japanese NIKKEI 225 index over the last 50-plus year. Given that the risk of positions in brokerage firms must be measured daily and the risk associated with the capital requirement of financial institutions, according to the Basle II Committee, needs to be measured over ten days, the literature generally evaluates the VaR daily or over ten days. In contrast to daily returns, monthly returns follow a distribution with less asymmetry and fat tails, are less autocorrelated, and result in a smaller number of observations. Our unique focus on the monthly horizon relevant to longer-term investment allows us to investigate the importance of these characteristics for the performance of VaR models.

Our third contribution is the application of formal performance and conservatism statistical tests for two target probabilities (1% and 5%) to provide a more complete comparison than generally reported. Specifically, we apply the unconditional coverage test, the independence test and the conditional coverage test of Christoffersen (1998), which measure the ability of VaR models to comply with two conditions. First, the proportion of VaR violation, which refer to an event where the ex post index loss exceeds the ex ante VaR measure, should be on average equal to the theoretical target probability. Second, a VaR violation should not be predictable using available information. In particular, the proportions of VaR violation when there is and when there is not a VaR violation in the previous period should be on average the same. We also evaluate the

[2] For a review of the literature on GARCH models and a discussion of their performance, see Bollerslev, Chou and Kroner (1992), Engle and Ng (1993) and Hansen and Lunde (2005).

conservatism of the best performing VaR models by proposing a ranking coincidence test. The most conservative model produces consistently the highest risk measure, which is more prudent.

There is a growing number of studies on VaR models, but there is no consensus model adequate for all financial assets, sample frequencies, performance tests, target probabilities and sub-periods. On the FHS technique, Hull and White (1998), Barone-Adesi, Giannopoulos and Vosper (1999, 2002), Christoffersen and Gonçalves (2005), Pritsker (2006), Bao, Lee and Saltoglu (2006), Kuester, Mittnik and Paolella (2006), and Angelidis, Benos and Degiannakis (2007) show that this approach performs relatively well and argue that it is among the most promising, but in various contexts and with different samples than ours[3]. Comparative studies that do not consider the FHS technique include Beder (1996), Hendricks (1996), Alexander and Leigh (1997), Pritsker (1997), Mittnik and Paolella (2000), Sarma, Thomas and Shah (2003), Angelidis, Benos and Degiannakis (2004), Brooks, Clare, Dalle Molle and Persand (2005), and So and Yu (2006), but they do not examine the monthly VaR horizon. Furthermore, only Sarma, Thomas and Shah (2003), Angelidis, Benos and Degiannakis (2004), and Kuester, Mittnik and Paolella (2006) apply the conditional coverage test of Chirstoffersen (1998) to an extensive number of VaR models and none of the studies examine the conservatism test.

Our empirical results highlight only two VaR models not rejected at the 95% confidence level with regard to each test, i.e. the two models with historical simulation using daily GARCH-type filters (GARCH(1,1) and asymmetrical GARCH(1,1)). These two FHS models, which generate the most volatile VaR measures, are satisfactory for the three performance tests of Christoffersen (1998), at the 1% and 5% target probabilities of VaR violation, and for the three equity indices. They thus provide adequate values of the market risk at monthly horizon for institutional portfolios. The conservatism tests suggest that the asymmetrical GARCH model is the most conservative of the two. Its VaR results indicate that the CRSP value-weighted index has 5% and 1% probabilities to

[3] Hull and White (1998) examine daily VaR for the returns on ten exchange rates and five stock indexes from 1988 to 1998. Barone-Adesi, Giannopoulos and Vosper (1999, 2002) present VaR with horizons from one to ten days for the returns on futures, options and swaps from 1994 to 1997. Christoffersen and Gonçalves (2005) look at estimation risk for one-day VaR with a simulation study using resampling methods. Pritsker (2006) estimates 10-day VaR for the returns on the UK pound/US dollar exchange rate from 1973 to 1997. Bao, Lee and Saltoglu (2006) compute daily VaR for the stock returns on five Asian countries from 1996 to 1999. Kuester, Mittnik and Paolella (2006) examine daily VaR for the returns on the NASDAQ index from 1971 to 2001. Angelidis, Benos and Degiannakis (2007) estimate daily VaR for the returns on the DJ Euro Stoxx large and small capitalization indices from 1987 to 2005.

lose on average 6.9% and 12.2% of its value over one month, respectively, whereas the DAX index shows corresponding losses of 8.7% and 15.3%, and the NIKKEI 225 index shows corresponding losses of 9.3% and 18.2%.

An examination of our results in terms of the methodological features of the FHS technique, the importance of the specific characteristics of monthly returns, and the application of formal performance tests lead to the following observations. First, the kurtosis in monthly equity returns, which is relatively small compare to the one in daily returns, is still sufficiently important that fat distribution tails need to be considered in the VaR specification. For example, the parametric VaR models relying on a Normal distribution fare worse than the ones based on a Student-*t* distribution in the unconditional coverage tests, which show that all rejected models underestimate the frequency of extreme losses. The historical simulation technique is generally able to account for the fat tails through its use of the realized return distribution.

Second, an adequate specification of the volatility dynamics is important for the success of monthly VaR models. The unconditional VaR models, which put relatively little emphasis on recent returns, are generally underperforming the conditional VaR models with respect to the independence tests, as they greatly underestimate the frequency of consecutive VaR violations. The independence tests also reveal that the RiskMetrics conditional volatility specification has more difficulty than the GARCH specification, although all the FHS models examined adjust relatively well to risk variations predictable from its immediate past.

Third, the use of daily rather monthly data in monthly FHS model improves the performance in the three tests of Christoffersen (1998), especially at the 1% target probability. This improvement is related to a better estimation of the conditional volatilities as the larger number of daily observations allows more precise estimates of the ARCH, GARCH and asymmetry effects than the smaller monthly sample.

The next section describes the sixteen monthly VaR models considered. The third section outlines the performance and conservatism tests. The fourth section discusses the data. The fifth section provides the empirical results whereas the last section concludes.

2. THE MONTLHY VAR MODELS

This section presents the montlhy VaR models. The first subsection presents the parametric and historical simulation models estimated with monthly data whereas the second subsection describes the estimation of montlhy VaR with filtered historical simulation using daily data.

2.1 MONTHLY VAR MODELS USING MONTHLY DATA

2.1.1 Parametric VaR

We define the monthly parametric VaR for portfolio *p* at the period *T*+1 by one of three functions according to whether the returns could be characterized by a Normal or Student-*t* distributions or a Cornish-Fisher approximation:

$$VaR_{p,T+1}^{Par-N} = -\mu_{p,T+1} + \alpha \cdot \sigma_{p,T+1}, \qquad \textit{Equation 1a}$$

$$VaR_{p,T+1}^{Par-t} = -\mu_{p,T+1} + \sqrt{\frac{d-2}{d}} \cdot t^{-1}(d) \cdot \sigma_{p,T+1}, \qquad \textit{Equation 1b}$$

$$VaR_{p,T+1}^{Par-t} = -\mu_{p,T+1} + CF^{-1} \cdot \sigma_{p,T+1}, \qquad \textit{Equation 1c}$$

where $\mu_{p,T+1}$ represents the monthly expected return of portfolio *p* on the VaR horizon (*T*+1) and $\sigma_{p,T+1}$ is the standard deviation of monthly returns of portfolio *p* on the VaR horizon. The equation *1a* implies that the return $R_{p,T+1}$ is a normally distributed random variable with a mean $\mu_{p,T+1}$ and a variance $\sigma^2_{p,T+1}$ [$R_{p,T+1} \sim N(\mu_{p,T+1}, \sigma^2_{p,T+1})$]. To account for the fat distribution tails generally observed in financial returns, we also consider in equation *1b* a parametric VaR that assumes that the standardized error term of portfolio *p* follows a Student-*t* distribution where *d* corresponds to the degrees of freedom and is equal to $6/(E(R_{p,T+1} - \mu_{p,T+1})^4 / \sigma^4_{p,T+1}) + 4$. To allow for skewness and kurtosis, we also consider in equation *1c* a parametric VaR that can be approximated by the Cornish-Fisher quantile[4]. Thus, α, $t^{-1}(d)$ and CF^{-1} represent the number of standard deviations associated with the target probability (*pr*) of a VaR violation according to the Normal or Student-*t* distributions or the Cornish-Fisher approximation, respectively.

2.1.2 VaR with Historical Simulation

Without explicitly parametrizing the distribution, this approach assumes that the historical return distribution is representative of the expected return distribution. For the probability *pr* of a VaR violation, the VaR with historical simulation is obtained by drawing the 100*pr* percentile of the distribution of the historical error terms[5] $[[\varepsilon_{p,T+1-\tau}^{pseudo}]_{\tau=1}^{T}]$, i.e. the historical deviations from the expected return. It can be written as follows:

$$VaR_{p,T+1}^{HS} = -\mu_{p,T+1} - Percentile\left\{[\varepsilon_{p,T+1-\tau}^{pseudo}]_{\tau=1}^{T}, 100\,pr\right\} \qquad \textit{Equation 2}$$

2.1.3 Specification of the First and Second Moments

We study monthly VaR with unconditional and conditional first and second moments.

2.1.3.1 Unconditional Specifications

If the portfolio returns are identically and independently distributed (i.i.d.), an unconditional specification of the first and second moments calculated from *t* to *T* leads to adequate estimates of the mean and volatility of returns at *T*+1. Then, we can describe unconditional parametric VaR models based on the Normal distribution (equation 3a), the Student-*t* distribution (equation 3b), the Cornish-Fisher approximation (equation 3c) as well as the unconditional VaR with historical simulation (equation 4) as follows:

---

[4] With $CF^{-1} = \alpha + \frac{S}{6}\left[\alpha^2 - 1\right] + \frac{K}{24}\left[\alpha^3 - 3\alpha\right] - \frac{S^2}{36}\left[2\alpha^3 - 5\alpha\right]$ and where *S* and *K* are respectively the skewness and the excess kurtosis coefficients.

[5] For some methodologies, these error terms are called pseudo-shocks.

$$VaR_{p,T+1}^{UncPar-N} = -\mu_{p,T+1} + \alpha\sigma_{p,T+1}, \qquad \text{Equation 3a}$$

$$VaR_{p,T+1}^{UncPar-t} = -\mu_{p,T+1} + \sqrt{\tfrac{d-2}{d}}\, t^{-1}(d)\sigma_{p,T+1} \qquad \text{Equation 3b}$$

$$VaR_{p,T+1}^{UncPar-CF} = -\mu_{p,T+1} + CF^{-1}\sigma_{p,T+1} \qquad \text{Equation 3c}$$

$$VaR_{p,T+1}^{HSInc} = -Percentile\left\{\left\{R_{p,T+1-\tau}\right\}_{\tau=1}^{T}, 100\,pr\right\}, \qquad \text{Equation 4}$$

where $\mu_{p,T+1} = \frac{1}{T}\sum_{t=1}^{T}\left(R_{p,t}\right)$, $\sigma_{p,T+1} = \sqrt{\sum_{t=1}^{T}\left(R_{p,t} - \mu_{p,T+1}\right)^2 /(T-1)}$. We thus obtain three unconditional parametric VaR, Normal, Student-*t* and Cornish-Fisher, by computing the historical mean and standard deviation of returns, and also the degrees of freedom in the case of the Student-*t* distribution and the skewness and excess kurtosis coefficients in the case of the Cornish-Fisher approximation. We furthermore determine the unconditional VaR with historical simulation by drawing from the 100*pr* percentile of the historical return distribution. We study the performance of the models estimated with data samples of the last five or fifteen years. Hereafter, we identify them as follows: Uncond. Normal, Uncond. Student-t, Uncond. Cornish-Fisher and Unconditional.

2.1.3.2 Conditional Specifications

Conditional modeling can be described in two stages. The first stage determines the specification of the conditional expected return of portfolio *p* ($\mu_{p,T+1}$). Ljung-Box tests on the autocorrelation of monthly returns discriminate among ARMA models the one which best characterizes the portfolio returns. We also examine the explanatory power of the ARMA models for various VaR horizons and for the full sample of data[6]. In the second stage, we check for the presence of autocorrelation in the squared error terms using Ljung-Box and ARCH tests. In the presence of autocorrelation, we

[6] For details on ARMA models, see Chap. 3 of Hamilton (1994). In light of our results and similar to Busse (2001), we evaluate the first conditional moment with a MA(1) model, i.e. the process $R_{p,T+1} = c_{p,k} + \phi_{p,k}\varepsilon_{p,k,T} + \varepsilon_{p,k,T+1}$ with $\varepsilon_{p,k,T+1} \sim N(0, \sigma_{p,k,T+1}^2)$, where $c_{p,k}$ is a constant, $\phi_{p,k}$ captures the autocorrelation in the error terms, and $\sigma_{p,k,T+1}^2$ represents the conditional error term variance. The index *k* distinguishes between the various conditional variance specifications.

estimate the conditional variance using either the RiskMetrics exponential weighting, GARCH(1,1) or GJR-GARCH(1,1) models.

The RiskMetrics (RM) exponential weighting model specifies the variance at T+1 as a weighted average of the squared return and variance at T:

$$\sigma^2_{p,RM,T+1} = (1-\lambda)R^2_{p,T} + \lambda\sigma^2_{p,RM,T} \,, \qquad \textit{Equation 5}$$

where the parameter λ is lower than one. As this parameter moves away from one, the variance at T+1 puts more emphasis on the squared return at T and less emphasis on all the other squared returns[7]. In this study, we assume that λ = 0.97, one of the values suggested by RiskMetrics for the estimation of the second moments of monthly returns.

The GARCH(1,1) model of Bollerslev (1986) measures the conditional variance as follows:

$$\sigma^2_{p,GARCH,T+1} = \omega + \alpha\varepsilon^2_{p,GARCH,T} + \beta\sigma^2_{p,GARCH,T} \,, \qquad \textit{Equation 6}$$

where ω is a constant related to the unconditional variance, α is the parameter of the ARCH effect and captures the link between the variance at T+1 and the squared error term at T [ε^2_T] and β represents the GARCH effect as it measures the persistence of the previous squared error terms on the conditional variance [Engle (1982), Bollerslev (1986) and Chou (1988)].

The GJR-GARCH(1,1) model, proposed by Glosten, Jagannathan and Runkle (1993), considers the asymmetrical effect of the positive and negative error terms on the conditional variance:

$$\sigma^2_{p,GJR,T+1} = \omega + \alpha\varepsilon^2_{p,GJR,T} + \beta\sigma^2_{p,GJR,T} + \gamma I^-_{p,GJR,T}\,\varepsilon^2_{p,GJR,T} \,, \qquad \textit{Equation 7}$$

[7] For more details on the conditional RiskMetrics specification, see Jorion (2006), Christoffersen (2003) or the technical documentation of JP Morgan on RiskMetrics.

The parameter γ measures the asymmetrical effect of the negative error terms since I^- is a binary variable that takes a value of one if the error term is negative and zero otherwise.

We can rewrite the equation for the unconditional parametric VaR models using the conditional mean and variance specifications to obtain the conditional parametric VaR models:

$$VaR_{p,T+1}^{Par,k} = -\mu_{p,k,T+1} + \alpha\sigma_{p,k,T+1}, \qquad \textit{Equation 8}$$

for k = RM, GARCH(1,1) or GJR-GARCH(1,1). $\mu_{p,k,T+1}$ is the conditional expected return of portfolio p at T+1. In the case of the RiskMetrics VaR model, the conditional expected return is constant, while in the case of the two GARCH-type VaR models, we use a MA(1) model[8]. $\sigma_{p,k,T+1}$ is the standard deviation of returns at T+1 from one of the three conditional variance specifications. Finally, we consider parametric VaR models with a MA(1)-GJRGARCH(1,1) specification for the first two moments of the return distribution and either a Student-*t* distribution to account for fat tails or a Cornish-Fisher approximation for both the skewness and kurtosis. Specifically, we replace α by $t^{-1}(d)$ or CF^{-1} in the equation above. Hereafter, we identify these conditional parametric VaR models as follows: RiskMetrics, MA(1)-GARCH(1,1), MA(1)-GJRGARCH(1,1), MA(1)-GJRGARCH(1,1)-*t*(*d*) and MA(1)-GJRGARCH(1,1)-CF.

Similarly, we can modify the equation for the VaR model with historical simulation to account for the mean and variance dynamics to obtain the VaR models with filtered historical simulation:

$$VaR_{p,T+1}^{HS,k} = -\mu_{p,k,T+1} - Percentile\left\{\left[z_{p,k,T+1-\tau} \cdot \sigma_{p,k,T+1}\right]_{\tau=1}^{T}, 100pr\right\}, \qquad \textit{Equation 9}$$

for k = RM, GARCH(1,1) or GJR-GARCH(1,1). The variable $z_{p,k,T+1-t}$ represents the standardized error term at T+1-t, i.e. $\left[\frac{\varepsilon_{p,k,T+1-t}}{\sigma_{p,k,T+1-t}}\right]$, for t = 1, …, T. The VaR with filtered historical simulation is

[8] Specifically, $\mu_{p,k,T+1} = c_{p,k} + \phi_{p,k}\varepsilon_{p,k,T}$ where $\phi_{p,k} = 0$ for k = RM.

thus a function of the standardized error term related to the target probability (*pr*) of a return exceeding the VaR, multiplied by the estimate of the standard deviation at the VaR horizon (T+1). By adding this pseudo-shock to the conditional expected return, it results in a conditional VaR with historical simulation for each of the three variance specifications. Hereafter, these three VaR models with historical simulation are denoted RiskMetrics, MA(1)-GARCH(1,1)-MD and MA(1)-GJRGARCH(1,1)-MD, where MD refers to the use of monthly data. The next section outlines monthly VaR models with historical simulation using daily data.

2.2 MONTLHY VAR MODELS WITH FILTERED HISTORICAL SIMULATION USING DAILY DATA

Barone-Adesi, Giannopoulos and Vosper (2002) propose (but do not implement) VaR models with historical simulation using GARCH-type filter on returns measured at a higher frequency than the VaR horizon. In this spirit, we compute monthly VaR with filtered historical simulation using a sample of T monthly pseudo-returns, each simulated from N_t daily returns (N_t working days in month t) with an ARMA-GARCH filter[9]. Steps in the estimation of monthly VaR models with this methodology can be summarized in the following way.

First, we obtain standardized error terms from an ARMA-GARCH regression with the 3900 previous daily returns, or about fifteen years of data preceding the VaR evaluation date. This model can be written as follows:

$$R_{p,j} = u_{p,k,j} + \varepsilon_{p,k,j}\,, \qquad \textit{Equation 10}$$

where $\varepsilon_{p,k,j} \sim N(0, \sigma^2_{p,k,j})$, for j = 1,…, 3900 and for k = GARCH(1,1) or GJR-GARCH(1,1).

The next step consists in randomly drawing with replacement the i^{th} of the N_t standardized error terms of the month $\left[z_{p,k,i} = \frac{\varepsilon_{p,k,j}}{\sigma_{p,k,j}} \right]$ from the sample of observations, i.e. for j between 1 and 3900

[9] Giannopoulos (2003) also suggests using daily returns at the monthly horizon, but for the specific purpose of forming daily observations of overlapping monthly returns when the number of months in the period studied is insufficient.

and for each specification k. The i^{th} of the N_t filtered daily returns ($R^{'}_{p,k,i}$) of the month is then a function of the daily conditional expected return ($\mu_{p,k,i}$) and the daily conditional standard deviation ($\sigma_{p,k,i}$), which are recomputed at each working day i. For each specification k, this i^{th} filtered daily return is calculated as follows:

$$R^{'}_{p,k,i} = \mu_{p,k,i} + z_{p,k,i} \cdot \sigma_{p,k,i}, \quad \text{for } i = 1, \ldots, N_t. \qquad \textit{Equation 11}$$

The filtered monthly return of portfolio p at period t ($R^{'}_{p,k,t}$) is then determined by compounding the N_t daily returns, where N_t varies according to the month and year. The monthly return ($R^{'}_{p,k,t}$) then becomes one of the T observations of the return distribution leading to the VaR estimation. More specifically, the last step consists in evaluating the VaR using the distribution of the T generated monthly returns, with T = 1000 in this study, as follows:

$$VaR^{HSDay,k}_{p,T+1} = -Percentile\left\{[R^{'}_{p,k,t}]^{T}_{\tau=1}, 100pr\right\}, \qquad \textit{Equation 12}$$

for k = GARCH(1,1) or GJR-GARCH(1,1). The VaR is thus determined by drawing from the distribution of filtered returns the return associated with the target probability (pr) of a VaR violation. Hereafter, we denote these two VaR models as MA(1)-GARCH(1,1)-DD and MA(1)-GJRGARCH(1,1)-DD. The next section discusses the performance and conservatism tests.

3. PERFORMANCE AND CONSERVATISM TESTS

We apply three likelihood ratio tests using the interval forecast method developed by Christoffersen (1998) to evaluate the ability of the VaR models to meet the target probabilities of VaR violation. According to these tests, a VaR model should meet two conditions. Firstly, the proportion of VaR violation should be on average equal to the theoretical target probability pr. The following unconditional coverage test (see also the binomial evaluation method of Kupiec, 1995) examines this hypothesis:

$$LR_{unc} = 2Log\left[\frac{\left(1-\frac{n_1}{n_0+n_1}\right)^{n_0}\left(\frac{n_1}{n_0+n_1}\right)^{n_1}}{(1-pr)^{n_0}(pr)^{n_1}}\right] \sim \chi^2(1), \qquad \textit{Equation 13}$$

where n_1 and n_0 are the number of VaR violations and non-violations, respectively, and $\frac{n_1}{n_0+n_1}$ represents the empirical probability of VaR violation. If this test results in a rejection, then the VaR model is biased as it produces an incorrect proportion of VaR violation.

Secondly, a VaR violation should not be predictable using available information. In particular, the proportions of VaR violation when there is and when there is not a VaR violation in the previous period should be on average the same. The following independence test examines this hypothesis:

$$LR_{ind} = 2Log\left[\frac{\left(1-\frac{n_{01}}{n_{00}+n_{01}}\right)^{n_{00}}\left(\frac{n_{01}}{n_{00}+n_{01}}\right)^{n_{01}}\left(1-\frac{n_{11}}{n_{10}+n_{11}}\right)^{n_{10}}\left(\frac{n_{11}}{n_{10}+n_{11}}\right)^{n_{11}}}{\left(1-\frac{n_1}{n_0+n_1}\right)^{n_0}\left(\frac{n_1}{n_0+n_1}\right)^{n_1}}\right] \sim \chi^2(1), \qquad \textit{Equation 14}$$

where n_{01} and n_{00} are the number of VaR violations and non-violations following a non-violation, respectively, and $\frac{n_{01}}{n_{00}+n_{01}}$ represents the empirical probability of VaR violation following a non-violation, and where n_{11} and n_{10} are the number of VaR violations and non-violations following a violation, respectively, and $\frac{n_{11}}{n_{10}+n_{11}}$ represents the empirical probability of VaR violation following a violation. If this test results in a rejection, then a VaR violation is predictable from the immediate past and is thus not purely random.

Lastly, a VaR model should meet *jointly* the two preceding conditions. The following conditional coverage test examines this hypothesis:

$$LR_{cond} = LR_{unc} + LR_{ind} \sim \chi^2(2). \qquad \textit{Equation 15}$$

If this test results in a rejection, then the cases of VaR violations are not simultaneously independent and of a proportion corresponding to the target probability.

In addition, we examine a ranking coincidence test to determine if, among the best performing VaR models with respect to the tests of Christoffersen (1998), some VaR models are more conservative than others. The model with the highest VaR is considered the most conservative since it indicates the highest risk, which translates into the most prudent risk measure. The null hypothesis is that the VaR of two different models are ranked in a purely random way. The following conservatism test examines this hypothesis by using the index of coincidence developed by Friedman (1920):

$$IC = 2M\left[(\overline{R}_{VaR1} - 1{,}5)^2 + (\overline{R}_{VaR2} - 1{,}5)^2\right] \sim \chi^2(1), \qquad \textit{Equation 16}$$

where M is the number of observations common to both VaR models, $\overline{R}_{VaR1} = \frac{1}{M}\sum_{m=1}^{M} R_{VaR1,m}$ is the average rank of the VaR1 model compared to the VaR2 model, $R_{VaR1,m} = 1$ $\left(R_{VaR1,m} = 2\right)$ if the VaR1 model is more (less) conservative than the VaR2 model for observation m, and $\overline{R}_{VaR2} = 3 - \overline{R}_{VaR1}$ is the average rank of the VaR2 model. If this test results in a rejection, then the VaR model with the highest average rank is the most conservative. The next section describes the data.

4. DATA

We study the performance and conservatism of sixteen monthly VaR models for three stock indices, the American CRSP value-weighted index, the German DAX index and the Japanese NIKKEI 225 index. The monthly series of the CRSP, DAX and NIKKEI indices begin in January 1950, January 1965 and February 1960, respectively, and end in January 2008. Their daily series go from July 1st, 1963, January 5th, 1965, and January 4th, 1960, respectively, and end on January 31st, 2008, for the Japanese index. The data source is the web site of Prof. Kenneth French for the CRSP Index and Datastream for the two other series. In each month, we use a moving window of the previous fifteen years for the VaR estimation. Our study thus obtains a maximum of 498 montlhy VaR for each model[10].

[10] For the unconditional VaR models, we also estimate the models based on the previous five years of monthly data.

Table 1 presents descriptive statistics of the daily and monthly index returns. At the monthly frequency, the German index has a higher standard deviation and kurtosis than the two other indices. Based on this result, we anticipate larger monthly VaR values for the DAX index. For all the series, the worst return occurs during the crash of October 1987. The skewness and kurtosis coefficients suggest that the historical daily and monthly returns do not follow a Normal distribution as they show fat tails. The Jarque-Bera tests indicate that the normality hypothesis is rejected at the 99% confidence level. The VaR models that do not account for these characteristics should have difficulty in passing the coverage tests of Christoffersen (1998).

[Please insert Table 1 here]

We also reject at the 99% confidence level the hypothesis that the daily and monthly returns are serialy independent. The Q^2-tests for five lags show that the squared returns are autocorrelated, so that the return variances are predictable. We furthermore reject at the 99% confidence level the hypothesis of no autocorrelation in daily returns, but are not able to reject the hypothesis for monthly returns. This result is partially explained by the presence of microstructure elements in high frequency returns. Given the significant autocorrelation in squared monthly returns, conditional VaR models should outperform their unconditional counterpart.

Overall, comparing the descriptive statistics of the daily and monthly returns, monthly returns follow a distribution with less asymmetry and fat tails, are less autocorrelated, and result in a smaller number of observations. These characteristics could lead to best performing monthly VaR models different from the best performing daily VaR models documented in the literature. The next section discusses our empirical results.

5. EMPIRICAL RESULTS ON THE VAR MODELS

This section presents the empirical results. The first section shows summary statistics of the VaR models. The second section gives the results of the performance tests of Chirstoffersen (1998). For the best performing models with regard to the performance tests, we discuss in the third section their conservatism on a pairwise basis.

5.1 DESCRIPTIVE STATISTICS

Table 2 presents summary statistics of the VaR estimated for different indices (CRSP, DAX and NIKKEI), models (nine parametric VaR models and seven VaR models with historical simulation) and target probabilities (5% and 1%)[11]. As anticipated from the descriptive statistics of the monthly returns, the DAX index is riskier than the other indices according to the VaR. Also, the conditional approach leads to more volatile VaR values than the unconditional approach. This higher volatility is caused by the larger sensitivity to recent returns in the conditional models, whereas the unconditional models consider the effect of past returns over a long sample. The VaR models showing the highest average VaR values and the highest VaR volatility use filtered historical simulation with daily data or is based on a conditional Student-*t* distribution approach. [MA(1)-GARCH(1,1)-DD, MA(1)-GJRGARCH(1,1)-DD or MA(1)-GJRGARCH(1,1)-t(d)][12].

[Please insert Table 2 here]

5.2 PERFORMANCE TESTS

5.2.1 Unconditional Coverage Tests

Table 3 gives the results of the unconditional coverage test LR_{unc} . This test evaluates the ability of the VaR models to meet the target probabilities of VaR violation. Overall, six models are never rejected at the 95% confidence level: the two models based on the Student-*t* distribution, the two unconditional models with historical simulation, as well as the two models with historical simulation using a daily filter. At the 5% target probability, seven models show empirical probabilities that are significantly too high for the CRSP index, but there is only one rejected model for each of the two other indices. At the 1% target probability, a majority of models are rejected for all indices. The proportions of VaR violation for the rejected models are generally

[11] We do not estimate the unconditional VaR with historical simulation using a five-year estimation window at the 1% target probability since the 1st percentile of 60 observations is not sufficiently informative.
[12] We do not report the parameter estimates of the models due to limited space and the fact that the results are typical of the existing literature. For example, we find ARCH, GARCH and asymmetry (GJRGARCH) effects in the error terms that are generally significant and more important in daily than monthly data.

between 7% and 10% at the 5% target probability, and between 2% and 4% at the 1% target probability. The rejected models thus underestimate the frequency of the extreme losses.

[Please insert Table 3 here]

A comparative analysis of the results for the parametric VaR models reveals that those relying on the Student-*t* distribution, which has fat tails, perform better than those based on the Normal distribution. For the VaR models with historical simulation, the use of daily rather than monthly data improves the performance, especially at the 1% target probability. This improvement is related to the better estimation of the conditional volatilities with 3700 daily returns than with 180 monthly returns. The daily data allow more precise estimates of the ARCH, GARCH and asymmetry (GJRGARCH) effects.

5.2.2 Independence Tests

Whereas the preceding section analyzes the proportions of VaR violation for the full sample, the independence test LR_{ind} of Christoffersen (1998) checks if the proportions of VaR violation are significantly different depending on whether there is or is not a VaR violation in the previous period. Table 4 reports the proportions of VaR violation in the period following a VaR violation, and the significance of the independence test on whether these proportions are different from the ones following a non-violation.

[Please insert Table 4 here]

The independence tests show that numerous VaR models obtain empirical probabilities of VaR violation in the period following a VaR violation greater than 10%, thus greatly underestimating the frequency of consecutive extreme losses. The highest proportions generally belong to the unconditional models. At the 5% target probability, all six unconditional models, but only two of the ten conditional models, are rejected for at least two of the three indices. In particular, none of the VaR models with filtered historical simulation are rejected. The joint consideration of fat distribution tails and conditional volatility dynamics is responsible for the success of these models.

At the 1% target probability, we do not reject any model for the CRSP index and only reject three VaR models for both other indices: the two unconditional parametric models using a Normal distribution and 5 or 15-year estimation windows and the unconditional model with historical simulation. The independence test may lack power as almost no consecutive return belonging to the 1% extremity of the left distribution tail occurs in the available sample[13]. Overall, the VaR models with filtered historical simulation using a daily filter are the only models that they have not been rejected by any test yet.

5.2.3 Conditional Coverage Tests

The conditional coverage test LR_{cond} of Christoffersen (1998) jointly examines the proportion of VaR violation for the full sample and the independence of VaR violations for two consecutive periods. Table 5 provides the results. For each target probability and at the 90% confidence level, three models are not rejected for all three indices. At the 5% target probability, these models are the unconditional parametric model using a Normal distribution with a 5-year estimation window and the two VaR models with historical simulation and a daily filter [(MA(1)-GARCH(1,1)-DD and MA(1)-GJRGARCH(1,1)-DD)]. At the 1% target probability, they are the parametric GJRGARCH model with a Student-*t* distribution and the two VaR models with historical simulation and a daily filter. The RiskMetrics VaR model with historical simulation also performs relatively well, as it is only rejected for the CRSP index at the 5% target probability and for the DAX index at the 1% target probability.

[Please insert Table 5 here]

As an estimate of the loss incurred beyond the value announced by the VaR, Table 5 also reports the average deviation between the realized return and the VaR when all models simultaneously have a VaR violation[14]. The unconditional parametric VaR model with a Student-*t* distribution is the model that obtains consistently among the smallest average losses. However, this model is

[13] On the issue of power, see also Christoffersen (1998), Lopez (1998) and Christoffersen and Pelletier (2004).
[14] The average deviation is not available for the NIKKEI index at the 1% target probability as there is no observation when there is a VaR violation for all models.

rejected in the conditional coverage tests for the three indices. As expected, the two VaR models never rejected by the conditional coverage tests (the models with historical simulation and a daily filter) show small average deviations. For example, at the 5% target probability, the average deviations for the MA(1)-GARCH(1,1)-DD and MA(1)-GJRGARCH(1,1)-DD models are -6.00% and -5.75%, respectively, while the deviations of the other models average -7.18%.

In summary, not only the two VaR models with daily filtered historical simulation perform well in all the conditional coverage tests, but they are the only ones never rejected by any of the three performance tests of Christoffersen (1998) at the 95% confidence level. Our empirical evidence thus supports the use of these two models because they simultaneously present an adequate mean level of risk and adjust quickly to risk variations predictable from their immediate past. The next section determines which of the most performing models is the most conservative. For each target probability, we study the conservatism of the VaR models that are never rejected at the 90% confidence level in the conditional coverage tests.

5.3 CONSERVATISM TESTS

The conservatism test *IC* examines the null hypothesis that the VaR of two different models are ranked in a purely random way. In a rejection, the model with the highest VaR is considered the most conservative because it suggests a more prudent risk level. Conservatism is only an interesting characteristic for VaR models that perform well with respect to the tests previously discussed. In this section, we thus compare the three best performing VaR models identified previously for each target probability. At the 5% target probability, we study the conservatism of the parametric unconditional model using a Normal distribution and a 5-year estimation window and the two VaR models with historical simulation and a daily filter [(MA(1)-GARCH(1,1)-DD and MA(1)-GJRGARCH(1,1)-DD)]. At the 1% target probability, we apply the conservatism test to the parametric GJRGARCH model with a Student-*t* distribution and the two VaR models with daily filtered historical simulation. These models are the only ones never rejected at the 90% confidence level in the conditional coverage tests of Christoffersen (1998). Table 6 has the test results as well as the proportion of observations where a model is more conservative than another.

[Please insert Table 6 here]

The results in Table 6 reject the hypothesis that the ranking between the best performing VaR models is randomly drawn in thirteen of the eighteen cases tested. At the 5% target probability, the most conservative model is the MA(1)-GJRGARCH(1,1)-DD, as it shows significantly higher-ranked risk measures for two of the three indices. This model obtains the highest VaR measures for 68% of the CRSP index observations and 59% of the NIKKEI index observations compare to the MA(1)-GARCH(1,1)-DD model, which is the second most conservative model. The unconditional parametric model using a Normal distribution obtains significantly lower-ranked risk measures.

At the 1% target probability, the most conservative model is the parametric GJRGARCH model with a Student-*t* distribution for two of the three indices. This model gets the highest-ranked risk measures with proportions of 58% for the CRSP index and 75% for the DAX index compare to the second most conservative model, the MA(1)-GJRGARCH(1,1)-DD model. Our results also indicate that the MA(1)-GJRGARCH(1,1)-DD model is significantly more conservative than the MA(1)-GARCH(1,1)-DD model for the three indices.

Overall, the MA(1)-GJRGARCH(1,1)-DD model is to be the most conservative model among the two VaR models with daily filtered historical simulation. This finding highlights the role of the asymmetrical effect in the conditional variance in terms of conservatism. In the implementation of a monthly risk management program based on a threshold VaR, our results suggest that the portfolio manager can benefit from using the MA(1)-GJRGARCH(1,1)-DD VaR model because it is one of the most prudent among the best performing models. The next section concludes.

6. CONCLUSION

This article examines the performance and conservatism of sixteen monthly VaR models to estimate the risk on the American CRSP value-weighted index, the German DAX index and the Japanese NIKKEI 225 index. We study three parametric VaR models that assume a Normal distribution and are based on three different conditional variance specifications, i.e. RiskMetrics, GARCH(1,1) and GJRGARCH(1,1), and two parametric VaR models with a GJRGARCH(1,1) specification and a Student-*t* distribution or a Cornish-Fisher approximation. We also study three VaR models with filtered historical simulation using the three above conditional variance specifications [Hull and White (1998) and Barone-Adesi, Giannopoulos and Vosper (2002)]. The last two conditional VaR models use historical simulation with either a GARCH(1,1) or GJRGARCH(1,1) specifications, but with daily data rather than monthly data [Barone-Adesi, Giannopoulos and Vosper (1999, 2002)]. The other VaR models, either parametric using Normal or Student-*t* distributions, or a Cornish-Fisher approximation, or with historical simulation, are unconditional and are evaluated with samples of either five years or fifteen years of historical data. We estimate the VaR models at the 1% and 5% target probabilities of VaR violation.

The coverage and independence tests of Christoffersen (1998) reveal that only the monthly VaR models with daily filtered historical simulation and the GARCH(1,1) or GJRGARCH(1,1) volatility specifications are never rejected by any of the tests at the 95% confidence level. These two VaR models, which generate the most volatile values, obtain an adequate expected proportion of VaR violation, and do not present an abnormal probability of a VaR violation immediately after another one. They thus best succeed in capturing the fat tails of the return distribution and adapting to the changing market conditions. Among these two models, the conservatism tests indicate that the model with the GJRGARCH(1,1) specification is the most conservative, thus providing the most prudent measure of risk.

The parametric VaR models using a Normal distribution have statistically higher than expected proportions of VaR violation, especially at the 1% target probability. The bad performance of these models is partly explained by the inability of the Normal distribution to capture the fat tails of the index return distribution. The unconditional VaR models, which put relatively little emphasis on recent returns, are generally underperforming the conditional VaR models with respect to the independence tests. Specifically, the unconditional VaR models have a larger-than-expected propensity to obtain two consecutive VaR violations, suggesting that they do not adjust quickly to the dynamics of financial market risk.

7. REFERENCES

Alexander, C. (2001). "A Primer on the Orthogonal GARCH Model," Working paper, University of Reading.

Alexander, C., and G. T. Leigh. (1997). "On the Covariance Matrices Used in Value-at-Risk Models," *Journal of Derivatives* 4, 50-62.

Angelidis, T., A. Benos, and S. Degiannakis. (2004). "The Use of GARCH Models in VaR Estimation," *Statistical Methodology* 1, 105-128.

Angelidis, T., A. Benos, and S. Degiannakis. (2007). "A Robust VaR Model under Different Time Periods and Weighting Schemes," *Review of Quantative Finance and Accounting* 28, 187-201.

Beder, T. (1995). "VaR: Seductive But Dangerous," *Financial Analysts Journal* 51, 12-24.

Bao, Y., T.-H. Lee, and B. Saltoglu. (2006). "Evaluating Predictive Performance of Value-at-Risk Models in Emerging Markets: A Reality Check," *Journal of Forecasting* 25, 101-128.

Barone-Adesi, G., F. Bourgoin, and K. Giannopoulos. (1998). "Don't Look Back," *Risk* 11, 100-104.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper. (1999). "VAR without Correlations for nonlinear Portfolios of Derivative Securities," *Journal of Futures Market* 19, 583-602.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper. (2002). "Backtesting Derivative Portfolios with Filtered Historical Simulation (FHS)," *European Financial Management* 8, 31-58.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, 307-327.

Bollerslev, T., R. Y. Chou, and K. F. Kroner. (1992). "ARCH Modeling in Finance: A Selective Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52, 5-59.

Brooks, C., A. D. Clare, J. W. Dalle Molle, and G. Persand. (2005). "A Comparison of Extreme Value Theory Approaches for Determining Value at Risk," *Journal of Empirical Finance* 12, 339-352.

Busse, J. A. (2001). "Another Look at Mutual Fund Tournaments," *Journal of Financial and Quantitative Analysis* 36, 53-73.

Chou, R. Y. (1988). "Volatility Persistence and Stock Valuations: Some Empirical Evidence Using Garch," *Journal of Applied Econometrics* 3, 279-294.

Christoffersen, P. (1998). "Evaluating Interval Forecasts," *International Economic Review* 39, 841-862.

Christoffersen, P. (2003). Elements of Financial Risk Management, San Diego: Academic Press.

Christoffersen, P., and S. Gonçalves. (2005). "Estimation Risk in Financial Risk Management," *Journal of Risk* 7, 1-28.

Christoffersen, P., and D. Pelletier. (2004). "Backtesting Value-at-Risk: A Duration-Based Approach," *Journal of Financial Econometrics* 2, 84-108.

Ding, Z., C. W. J. Granger, and R.F. Engle. (1993). "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance* 1, 83-106.

Engle, R. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica* 50, 987-1008.

Engle, R. F., and V. K. Ng (1993) "Measuring and Testing the Impact of News on Volatility," *Journal of Finance* 48, 1749-1777.

Friedman, W. F. (1920) "The Index of Coincidence and its Application in Cryptography," *The Riverbank Publications*, Aegean Park Press, Laguna Hills, 22.

Giannopoulos, K. (2003). "VaR Modelling on Long Run Horizons," *Automation and Remote Control* 64, 1094-1100.

Glosten, L., R. Jagannathan, and D. Runkle. (1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance* 48, 1779-1801.

Hamilton, J. D. (1994). Time Series Analysis. Princeton: Princeton University Press.

Hansen, P. R., and A. Lunde. (2005). "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?," *Journal of Applied Econometrics* 20, 873-889.

Hendricks, D. (1996). "Evaluation of Value-at-Risk Models Using historical Data," *Federal Reserve Bank of New York Economic Policy Review* 2, 39-69.

Hull, J., and A. White. (1998). "Incorporating Volatility Updating into the Historical Simulation Method for VAR," *Journal of Risk* 1, 5-19.

Jorion, P. (2006). Value at Risk: The New Benchmark for Managing Financial Risk. New York: McGraw-Hill.

Kuester, K., S. Mittnik, and M. S. Paolella. (2006). "Value-at-Risk Prediction: A Comparison of Alternative Strategies," *Journal of Financial Econometrics* 4, 53-89.

Kupiec, P. (1995). "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives* 3, 73-84.

Lopez, J.A. (1998). "Methods for Evaluating Value-at-Risk Estimates," *FRBNY Economic Policy Review*, October, 119-124.

Mittnik, S., and M. S. Paolella. (2000). "Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates," *Journal of Forecasting* 19, 313-333.

Pritsker, M. (1997). "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time," *Journal of Financial Services Research* 12, 201-242.

Pritsker, M. (2006). "The Hidden Dangers of Historical Simulation," *Journal of Banking & Finance* 30, 561-582.

Sarma, M., S. Thomas, and A. Shah. (2003). "Selection of Value-at-Risk Models," *Journal of Forecasting* 22, 337-358.

So, M. K. P., and P. L. H. Yu. (2006). "Empirical Analysis of GARCH Models in Value at Risk Estimation," *Journal of International Financial Markets, Institutions and Money* 16, 180-197.

Table 1

Descriptive Statistics of the Equity Index Returns

The monthly data for the value-weighted CRSP (United States), DAX (Germany) and NIKKEI 225 (Japan) equity indices cover the periods from 01/1950 to 01/2008, from 01/1965 to 01/2008 and from 02/1960 to 01/2008, respectively. The daily data end on 01/31/2008 for the three indices, but start on 07/01/1963 for the CRSP index, on 01/05/1965 for the DAX index and on 01/04/1960 for the NIKKEI index. Panel A shows the number of observations, mean, standard deviation, maximum, minimum, skewness and kurtosis for the return series. Panel B presents the results of the Jarque-Bera tests on normality, and of the Q-tests on return autocorrelation and Q^2-tests on squared return autocorrelation for 5 lag returns (K=5). The *, ** and *** symbols indicate statistical significance at the 90%, 95% and 99% confidence levels, respectively.

Panel A: Summary Statistics

International Index	Number	Mean	Std Dev	Maximum	Minimum	Skewness	Kurtosis
Daily Data		(×100)					
CRSP	11263	0.043	0.009	0.087	-0.171	-0.719	20.138
DAX	10787	0.031	0.012	0.093	-0.128	-0.217	9.959
NIKKEI	11897	0.030	0.012	0.132	-0.149	-0.161	11.116
Monthly Data							
CRSP	697	0.010	0.042	0.166	-0.225	-0.476	5.013
DAX	517	0.007	0.056	0.214	-0.254	-0.357	5.031
NIKKEI	576	0.006	0.053	0.201	-0.192	-0.250	3.923

Panel B: Normality and Independence Tests

International Index	Jarque-Bera Test	Q-test (K=5)	Q^2 test (K=5)
Daily Data			
CRSP	138800.800***	155.560***	1 551.100***
DAX	21853.750***	26.192***	2 675.500***
NIKKEI	32701.620***	22.854***	1 500.400***
Monthly Data			
CRSP	144.061***	7.271	21.389***
DAX	99.852***	1.394	32.319***
NIKKEI	26.444***	3.498	62.479***

Table 2

Descriptive Statistics of the Montlhy VaRs

This table provides the mean and standard deviation of the monthly VaRs estimated for the CRSP index (United States), DAX index (Germany) and NIKKEI index (Japan) for each model. Panels A and B report the results for the target probabilities of VaR violation of 5% and 1%, respectively. The monthly VaR series using monthly data end in 01/2008 and start at the earliest in 01/1965 for the CRSP index, in 01/1980 for the DAX index and in 01/1975 for the NIKKEI index. For a description of the VaR models, see Section 2. For a description of the data, see Table 1.

VaR Models	CRSP		DAX		NIKKEI	
Panel A: Pr($R_p < -VaR_p$) = 5%	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Parametric VaR						
Uncond. Normal (5 years)	6.2%	1.6%	8.5%	2.7%	7.9%	3.1%
Uncond. Normal (15 years)	6.0%	0.8%	8.1%	1.0%	8.0%	1.9%
Uncond. Student-t	7.5%	1.0%	10.0%	1.3%	9.7%	2.2%
Uncond. Cornish-Fisher	5.3%	1.1%	7.5%	0.7%	7.5%	2.1%
MA(1)-GARCH(1,1)	6.0%	1.8%	8.4%	2.9%	8.0%	3.5%
MA(1)-GJRGARCH(1,1)	5.7%	2.8%	8.4%	2.4%	8.2%	3.8%
RiskMetrics	7.3%	1.2%	9.5%	2.2%	8.6%	2.2%
MA(1)-GJRGARCH(1,1)-t(d)	7.0%	3.4%	10.2%	2.9%	10.0%	4.5%
MA(1)-GJRGARCH(1,1)-CF	5.0%	2.4%	8.1%	2.3%	7.9%	3.7%
VaR with Historical Simulation						
Unconditional (5 years)	6.9%	1.8%	9.8%	3.7%	8.4%	3.5%
Unconditional (15 years)	6.0%	0.7%	7.9%	1.4%	7.7%	2.2%
MA(1)-GARCH(1,1)-MD	6.1%	2.0%	8.3%	3.2%	7.9%	3.7%
MA(1)-GJRGARCH(1,1)-MD	5.6%	2.7%	8.2%	2.7%	8.1%	3.9%
RiskMetrics	6.3%	1.5%	8.5%	2.4%	7.3%	2.2%
MA(1)-GARCH(1,1)-DD	6.8%	3.2%	8.6%	3.7%	9.2%	4.6%
MA(1)-GJRGARCH(1,1)-DD	6.9%	3.1%	8.7%	3.7%	9.3%	4.4%
Panel B: Pr($R_p < -VaR_p$) = 1%	Mean	SD	Mean	SD	Mean	SD
Parametric VaR						
Uncond. Normal (5 years)	9,1%	2,1%	12,4%	3,7%	11,4%	4,1%
Uncond. Normal (15 years)	8,9%	1,1%	11,8%	1,5%	11,5%	2,5%
Uncond. Student-t	13,0%	1,7%	17,0%	2,4%	16,3%	3,2%
Uncond. Cornish-Fisher	8,8%	2,1%	12,1%	1,0%	11,9%	2,6%
MA(1)-GARCH(1,1)	8,9%	2,5%	12,2%	4,2%	11,6%	4,9%
MA(1)-GJRGARCH(1,1)	8,5%	4,0%	12,2%	3,5%	11,9%	5,3%
RiskMetrics	10,3%	1,7%	13,4%	3,1%	12,2%	3,1%
MA(1)-GJRGARCH(1,1)-t(d)	12,1%	5,6%	17,2%	4,9%	16,7%	7,2%
MA(1)-GJRGARCH(1,1)-CF	8,3%	3,9%	12,6%	3,4%	12,5%	6,3%
VaR with Historical Simulation						
Unconditional (15 years)	9,7%	1,4%	13,8%	3,6%	13,5%	2,5%
MA(1)-GARCH(1,1)-MD	9,5%	2,9%	13,5%	5,3%	12,4%	5,0%
MA(1)-GJRGARCH(1,1)-MD	8,9%	4,6%	13,3%	4,5%	12,9%	5,4%
RiskMetrics	9,6%	2,1%	13,4%	3,8%	12,4%	4,1%
MA(1)-GARCH(1,1)-DD	11,4%	5,3%	14,7%	6,6%	17,0%	8,5%
MA(1)-GJRGARCH(1,1)-DD	12,2%	5,8%	15,3%	6,9%	18,2%	9,0%

Table 3

Unconditional Coverage Tests for the Montlhy VaR Models

This table presents the results of the unconditional coverage test proposed by Christoffersen (1998). It shows the number of estimated VaRs and the empirical probability of VaR violation for the CRSP index (United States), DAX index (Germany) and NIKKEI index (Japan) for each model. Panels A and B report the results for the target probabilities of VaR violation of 5% and 1%, respectively. The *, ** and *** symbols indicate that the likelihood ratio of the unconditional coverage test is statistically significant at the 90%, 95% and 99% confidence levels, respectively. For a description of the VaR models, see Section 2. For a description of the data, see Table 1.

VaR Models	CRSP		DAX		NIKKEI	
Panel A: Pr($R_p < -VaR_p$) = 5%	N	Empirical Prob of Exceeding	N	Empirical Prof of Exceeding	N	Empirical Prob of Exceeding
Parametric VaR						
Uncond. Normal (5 years)	517	6.770%*	337	5.638%	396	5.303%
Uncond. Normal (15 years)	517	5.996%	337	6.528%	396	4.545%
Uncond. Student-t	517	3.675%	337	4.748%	396	3.788%
Uncond. Cornish-Fisher	517	8.511%***	337	7.122%*	396	4.798%
MA(1)-GARCH(1,1)	506	7.115%**	320	5.625%	389	6.427%
MA(1)-GJRGARCH(1,1)	517	7.544%**	329	5.775%	384	5.208%
RiskMetrics	517	4.255%	337	5.045%	396	3.030%*
MA(1)-GJRGARCH(1,1)-t(d)	517	5.029%	329	4.559%	384	3.906%
MA(1)-GJRGARCH(1,1)-CF	517	9.478%***	329	6.383%	384	5.469%
VaR with Historical Simulation						
Unconditional (5 years)	517	5.029%	337	5.341%	396	4.293%
Unconditional (15 years)	517	6.383%	337	6.825%	396	5.051%
MA(1)-GARCH(1,1)-MD	506	6.917%*	320	5.938%	389	5.913%
MA(1)-GJRGARCH(1,1)-MD	517	7.737%***	329	6.383%	384	5.469%
RiskMetrics	517	6.383%	337	5.935%	396	5.808%
MA(1)-GARCH(1,1)-DD	355	3.662%	337	5.341%	397	3.778%
MA(1)-GJRGARCH(1,1)-DD	355	3.662%	337	4.748%	397	4.282%
Panel B: Pr($R_p < -VaR_p$) = 1%						
Parametric VaR						
Uncond. Normal (5 years)	517	2.321%***	337	3.264%***	396	1.768%
Uncond. Normal (15 years)	517	2.515%***	337	2.967%***	396	2.020%*
Uncond. Student-t	517	0.774%	337	1.187%	396	1.010%
Uncond. Cornish-Fisher	517	3.288%***	337	3.264%***	396	1.768%
MA(1)-GARCH(1,1)	506	2.569%***	320	2.500%**	389	2.828%***
MA(1)-GJRGARCH(1,1)	517	2.128%**	329	3.040%***	384	2.083%*
RiskMetrics	517	1.547%	337	2.374%**	396	2.020%*
MA(1)-GJRGARCH(1,1)-t(d)	517	0.967%	329	0.912%	384	0.521%
MA(1)-GJRGARCH(1,1)-CF	517	4.062%***	329	2.736%***	384	2.083%*
VaR with Historical Simulation						
Unconditional (15 years)	517	1.934%*	337	2.077%*	396	1.515%
MA(1)-GARCH(1,1)-MD	506	2.174%**	320	2.500%*	389	2.057%*
MA(1)-GJRGARCH(1,1)-MD	517	2.708%***	329	2.128%*	384	1.563%
RiskMetrics	517	1.741%	337	2.374%**	396	1.515%
MA(1)-GARCH(1,1)-DD	355	1.127%	337	1.484%	397	0.252%*
MA(1)-GJRGARCH(1,1)-DD	355	0.845%	337	1.484%	397	0.252%*

Table 4

Independence Tests for the Montlhy VaR Models

This table presents the results of the independence test proposed by Christoffersen (1998). It shows the number of estimated VaRs and the empirical probability of VaR violation following a VaR violation, represented by $N11/(N11+N10)$, for the CRSP index (United States), DAX index (Germany) and NIKKEI index (Japan) for each model. Panels A and B report the results for the target probabilities of VaR violation of 5% and 1%, respectively. The *, ** and *** symbols indicate that the likelihood ratio of the independence test is statistically significant at the 90%, 95% and 99% confidence levels, respectively. For a description of the VaR models, see Section 2. For a description of the data, see Table 1.

VaR Models	CRSP		DAX		NIKKEI	
Panel A: $\Pr(R_p < -VaR_p) = 5\%$	N	$\frac{N11}{(N11+N10)}$	N	$\frac{N11}{(N11+N10)}$	N	$\frac{N11}{(N11+N10)}$
Parametric VaR						
Uncond. Normal (5 years)	517	11.765%	337	16.667%*	396	15.000%*
Uncond. Normal (15 years)	517	20.000%***	337	14.286%	396	17.647%**
Uncond. Student-t	517	21.053%***	337	20.000%**	396	21.429%**
Uncond. Cornish-Fisher	517	25.581%***	337	21.739%**	396	16.667%*
MA(1)-GARCH(1,1)	506	17.143%**	320	6.250%	389	8.333%
MA(1)-GJRGARCH(1,1)	517	5.263%	329	17.647%*	384	5.556%
RiskMetrics	517	14.286%*	337	18.750%**	396	9.091%
MA(1)-GJRGARCH(1,1)-t(d)	517	0.000%*	329	23.077%**	384	0.000%
MA(1)-GJRGARCH(1,1)-CF	517	6.250%	329	15.789%	384	5.263%
VaR with Historical Simulation						
Unconditional (5 years)	517	12.000%	337	17.647%*	396	18.750%**
Unconditional (15 years)	517	21.875%***	337	13.636%	396	21.053%**
MA(1)-GARCH(1,1)-MD	506	12.121%	320	11.765%	389	4.545%
MA(1)-GJRGARCH(1,1)-MD	517	5.128%	329	15.789%	384	5.263%
RiskMetrics	517	15.625%*	337	15.789%	396	13.636%
MA(1)-GARCH(1,1)-DD	355	0.000%	337	5.882%	397	0.000%
MA(1)-GJRGARCH(1,1)-DD	355	0.000%	337	6.667%	397	11.765%
Panel B: $\Pr(R_p < -VaR_p) = 1\%$						
Parametric VaR						
Uncond. Normal (5 years)	517	0.000%	337	20.000%**	396	16.667%*
Uncond. Normal (15 years)	517	7.692%	337	22.222%**	396	12.500%*
Uncond. Student-t	517	0.000%	337	0.000%	396	25.000%*
Uncond. Cornish-Fisher	517	11.765%	337	20.000%**	396	14.286%
MA(1)-GARCH(1,1)	506	0.000%	320	14.286%	389	0.000%***
MA(1)-GJRGARCH(1,1)	517	0.000%	329	12.500%	384	0.000%
RiskMetrics	517	0.000%	337	14.286%	396	14.286%*
MA(1)-GJRGARCH(1,1)-t(d)	517	0.000%	329	0.000%	384	0.000%
MA(1)-GJRGARCH(1,1)-CF	517	5.000%	329	14.286%	384	0.000%
VaR with Historical Simulation						
Unconditional (15 years)	517	10.000%	337	28.571%***	396	33.333%***
MA(1)-GARCH(1,1)-MD	506	0.000%	320	14.286%	389	0.000%
MA(1)-GJRGARCH(1,1)-MD	517	0.000%	329	20.000%*	384	0.000%
RiskMetrics	517	0.000%	337	14.286%	396	20.000%
MA(1)-GARCH(1,1)-DD	355	0.000%	337	0.000%	397	0.000%
MA(1)-GJRGARCH(1,1)-DD	355	0.000%	337	0.000%	397	0.000%

Table 5

Conditional Coverage Tests for the Monthly VaR Models

This table presents the results of the conditional coverage test proposed by Christoffersen (1998). It shows the average deviation between the return and the VaR when all models have a VaR violation, represented by $(\overline{R_p + VaR_p \mid R_p < -VaR_{p,i}})$, and the likelihood ratio of the conditional coverage test for the CRSP index (United States), DAX index (Germany) and NIKKEI index (Japan) for each model. Panels A and B report the results for the target probabilities of VaR violation of 5% and 1%, respectively. The *, ** and *** symbols indicate that the likelihood ratio of the conditional coverage test is statistically significant at the 90%, 95% and 99% confidence levels, respectively. For a description of the VaR models, see Section 2. For a description of the data, see Table 1.

VaR Models	CRSP		DAX		NIKKEI	
Panel A: Pr($R_p < -VaR_p$) = 5%	$(\overline{R_p + VaR_p \mid R_p < -VaR_{p,i}})$	Likelihood Ratio	$(\overline{R_p + VaR_p \mid R_p < -VaR_{p,i}})$	Likelihood Ratio	$(\overline{R_p + VaR_p \mid R_p < -VaR_{p,i}})$	Likelihood Ratio
Parametric VaR						
Uncond. Normal (5 years)	-7.498%	4.430	-8.206%	3.293	-6.371%	2.846
Uncond. Normal (15 years)	-7.148%	8.510**	-7.951%	3.239	-6.899%	4.497
Uncond. Student-t	-5.555%	10.810***	-6.054%	4.864*	-5.255%	7.738**
Uncond. Cornish-Fisher	-7.968%	23.873***	-8.679%	8.405**	-7.304%	3.786
MA(1)-GARCH(1,1)	-7.380%	8.571**	-8.486%	0.265	-7.104%	1.678
MA(1)-GJRGARCH(1,1)	-8.006%	6.631**	-7.798%	3.577	-6.908%	0.039
RiskMetrics	-5.908%	4.222	-6.719%	4.153	-5.250%	4.688*
MA(1)-GJRGARCH(1,1)-t(d)	-6.707%	2.756	-5.969%	6.000**	-5.288%	2.097
MA(1)-GJRGARCH(1,1)-CF	-8.644%	18.401***	-8.179%	3.446	-7.079%	0.174
VaR with Historical Simulation						
Unconditional (5 years)	-7.138%	2.093	-7.183%	3.634	-6.224%	5.384*
Unconditional (15 years)	-7.623%	11.125***	-8.040%	3.514	-6.922%	6.425**
MA(1)-GARCH(1,1)-MD	-7.592%	4.764*	-8.543%	1.436	-7.054%	0.730
MA(1)-GJRGARCH(1,1)-MD	-8.386%	7.651**	-7.853%	3.446	-6.858%	0.174
RiskMetrics	-7.155%	5.691*	-7.500%	3.120	-7.075%	2.491
MA(1)-GARCH(1,1)-DD	-6.468%	2.458	-6.863%	0.091	-4.657%	2.536
MA(1)-GJRGARCH(1,1)-DD	-5.909%	2.458	-6.523%	0.161	-4.830%	2.159
Panel B: Pr($R_p < -VaR_p$) = 1%						
Parametric VaR						
Uncond. Normal (5 years)	-14.872%	7.211**	-11.032%	15.439***	-	4.885*
Uncond. Normal (15 years)	-12.514%	9.446***	-11.385%	13.921***	-	5.369*
Uncond. Student-t	-8.249%	0.368	-7.312%	0.208	-	5.009*
Uncond. Cornish-Fisher	-11.299%	19.630***	-10.029%	15.439***	-	4.581
MA(1)-GARCH(1,1)	-13.011%	9.465***	-10.603%	7.166**	-	9.421***
MA(1)-GJRGARCH(1,1)	-15.649%	5.495*	-10.326%	10.406***	-	3.767
RiskMetrics	-11.621%	1.624	-8.711%	6.759**	-	5.620*
MA(1)-GJRGARCH(1,1)-t(d)	-12.902%	0.123	-5.960%	0.045	-	1.101
MA(1)-GJRGARCH(1,1)-CF	-14.776%	27.830***	-8.836%	8.662**	-	3.810
VaR with Historical Simulation						
Unconditional (15 years)	-14.630%	5.426*	-11.896%	10.909***	-	10.865***
MA(1)-GARCH(1,1)-MD	-11.430%	5.763*	-11.896%	7.166**	-	3.654
MA(1)-GJRGARCH(1,1)-MD	-11.303%	11.166***	-9.816%	6.185**	-	1.206
RiskMetrics	-14.990%	2.702	-9.193%	6.759**	-	4.569
MA(1)-GARCH(1,1)-DD	-11.860%	0.147	-9.405%	0.814	-	3.210
MA(1)-GJRGARCH(1,1)-DD	-11.626%	0.142	-11.033%	0.814	-	3.210

Table 6

Conservatism Tests for the Monthly VaR Models

This table presents the results of the conservatism test using the pairwise ranking coincidence test proposed by Friedman (1920). It shows the index of coincidence and, in parentheses, the proportion of observations where the model identified in the same row is more conservative than the model identified in the same column, for the CRSP index (United States), DAX index (Germany) and NIKKEI index (Japan). The tests are applied to the three best models with regard to the conditional coverage tests for each target probability. Panels A and B report the results for the target probabilities of VaR violation of 5% and 1%, respectively. The *, ** and *** symbols indicate that the index of coincidence of the conservatism test is statistically significant at the 90%, 95% and 99% confidence levels, respectively. For a description of the VaR models, see Section 2. For a description of the data, see Table 1.

Panel A: Pr($R_p < -VaR_p$) = 5%

	MA(1)-GARCH(1,1)-DD	MA(1)-GJRGARCH(1,1)-DD
CRSP index		
Uncond. Normal (5 years)	26.504***	71.214***
	[36.338%]	[27.606%]
MA(1)-GARCH(1,1)-DD		43.099***
		[32.394%]
DAX index		
Uncond. Normal (5 years)	0.074	0.667
	[49.258%]	[47.775%]
MA(1)-GARCH(1,1)-DD		3.635*
		[44.807%]
NIKKEI index		
Uncond. Normal (5 years)	24.252***	20.455***
	[37.626%]	[38.636%]
MA(1)-GARCH(1,1)-DD		13.423***
		[40.806%]

Panel B: Pr($R_p < -VaR_p$) = 1%

	MA(1)-GARCH(1,1)-DD	MA(1)-GJRGARCH(1,1)-DD
CRSP index		
MA(1)-GJRGARCH(1,1)-t(d)	38.561***	9.1521***
	[66.479%]	[58.028]
MA(1)-GARCH(1,1)-DD		44.014***
		[32.394%]
DAX index		
MA(1)-GJRGARCH(1,1)-t(d)	86.811***	74.921***
	[75.684%]	[73.860%]
MA(1)-GARCH(1,1)-DD		23.505***
		[36.795%]
NIKKEI index		
MA(1)-GJRGARCH(1,1)-t(d)	0.010	1.500
	[49.740%]	[46.875%]
MA(1)-GARCH(1,1)-DD		35.670***
		[35.013%]