Choosing Expected Shortfall over VaR in Basel III
Using Stochastic Dominance[*]

Chia-Lin Chang[1] ,Juan-Angel Jimenez-Martin[2],Esfandiar Maasoumi[3],

Michael McAleer[4, 5] [*] and Teodosio Perez-Amaral[6]

[1]Department of Applied Economics and Department of Finance

National Chung Hsing University, Taiwan

[2]Department of Quantitative Economics and Complutense

Institute of Economic Analysis Complutense (ICAE), University of Madrid Spain

[3]Department of Economics

Emory University, USA

[4]Department of Quantitative Finance

National Tsing Hua University, Taiwan

[5]Department of Mathematics and Statistics

University of Canterbury, New Zealand

[6]Department of Quantitative Economics and Complutense

Institute of Economic Analysis Complutense (ICAE), University of Madrid Spain

[*]michael.mcaleer@gmail.com

Abstract. Bank risk managers follow the Basel Committee on Banking Supervision (BCBS) recommendations that recently proposed shifting the quantitative risk metrics system from Value-at-Risk (VaR) to Expected Shortfall (ES). The Basel Committee on Banking Supervision (2013, p. 3) noted that: "a number of weaknesses have been identified with using VaR for determining regulatory capital requirements, including its inability to capture tail risk". The proposed reform costs and impact on bank balances may be substantial, such that the size and distribution of daily capital charges under the new rules could be affected significantly. Regulators and bank risk managers agree that all else being equal, a "better" distribution of daily capital charges is to be preferred. The distribution of daily capital charges depends generally on two sets of factors: (1) the risk function that is adopted (ES versus VaR); and (2) their estimated counterparts. The latter is dependent on what models are used by bank risk managers to provide for forecasts of daily capital charges. That is to say, while ES is known to be a preferable "risk function" based on its fundamental properties and greater accounting for the tails of alternative distributions, that same sensitivity to tails can lead to greater daily capital charges, which is the relevant (that is, controlling) practical reference for risk management decisions and observations. In view of the generally agreed focus in this field on the tails of non-standard distributions and low probability outcomes, an assessment of relative merits of estimated ES and estimated VaR is ideally not limited to mean variance considerations. For this reason, robust comparisons between ES and VaR will be achieved in the paper by using a Stochastic Dominance (SD) approach to rank ES and VaR.

---

1 Introduction

In the financial market industry, it is well known that the Basel III Accord requires that banks and other Authorized Deposit-taking Institutions (ADIs) communicate their daily risk forecasts to the appropriate monetary authorities at the beginning of each trading day, using one of a range of alternative financial risk models to forecast Value-at-Risk (VaR). Recently, the Basel Committee on Banking Supervision (BCBS) (2013) published a consultative document which presents the Basel Committee's initial proposals in regard to trading book capital requirement policies. A key element of the proposal is moving the quantitative risk metrics system from VaR to expected shortfall (ES), and decreasing the confidence level from 99% to 97.5%. The Basel Committee (2013, p. 3) observed that "a number of weaknesses have been identified in using Value-at-Risk (VaR) for determining regulatory capital requirements, including its inability to capture tail risk".

For this reason, the Basel Committee has considered using Expected Shortfall (ES) instead of VaR. ES is a coherent risk measure and has already become common in the insurance industry, although not yet in the banking industry. Artzner et al. (1997) proposed the use of ES to alleviate the problems inherent in VaR as: (1) ES considers losses beyond the VaR level and is shown to be sub-additive, whereas VaR disregards losses beyond the percentile and is not sub-additive; (2) although ES is mathematically superior to VaR in that it does not show "tail risk" and is a coherent risk measure in being subadditive, its practical implementation and greater computational requirements may pose operational challenges to financial firms.

Danielsson (2013) has examined the quantitative impact of such a proposal through analytical calculations, Monte Carlo simulations, and empirical results from observed data. He analyzed one of the key issues raised from this change in emphasis, namely that estimating ES conditional on VaR might be such that estimation and model risk for ES will be strictly higher than for VaR. Having found that 97.5% ES and 99% VaR are exactly the same for conditionally normal procedures, and a slightly greater ES than VaR for conditional Student-t, his analysis concluded that the 97.5% ES risk forecasts are generally more volatile than their 99% VaR counterparts.

Not only did prior methods for testing the main concerns of moving from VaR to ES focus on the first and second moments of these two risk distributions, but they also did not analyze the impact of such a measure on bank balances. The proposed reform costs and impact on bank balances may be substantial; in particular, the size and distribution of daily capital charges under the new rules may be considerably affected. Regulators and bank risk managers agree that all else being equal, a "better" distribution of daily capital charges is to be preferred. Distribution of daily capital charges depends on generally two sets of factors: (1) the risk function that is adopted (ES versus VaR), and (2) their estimated counterparts. The latter is dependent on what models are used by bank risk managers to provide forecasts of daily capital charges. That is to say, while ES is known to be a preferable "risk function" based on its fundamental properties and greater accounting for the tails of distributions, that same sensitivity to tails can lead to greater daily capital charges. Then the latter is the relevant (controlling) practical reference for risk management decisions and observations.

In view of the generally agreed focus in this field on the tails of non-standard distributions and low probability outcomes, an assessment of the relative merits of estimated ES and estimated VaR should not be limited to mean variance considerations. It will be consistent with the spirit of Basel III concerns with tail outcomes, to consider a more robust comparison, that is, robust to particular loss function (for example, mean and variance/quadratic), and underlying (unknown) distributions of the estimated daily capital charges. The former robustness is achieved by exploration of uniform

rankings over classes of loss functions, while the latter robustness is achieved by non-parametric estimation of the estimated daily capital charges distributions and resampling techniques.

Robust comparison between ES and VaR is achieved in this paper by using Stochastic Dominance (SD) methods and rankings. SD is a well-developed branch of "Decision Theory under Risk", with important applications in Economics, Finance, Portfolio Theory, and Financial Risk Management, among others. The SD approach has been regarded as one of the most useful tools to rank investment prospects when there are uncertainties (see, for example, Levy, 1992) as the ranking of the assets has been proven to be equivalent to utility maximization for the preferences of risk averters and risk lovers (see, for example, Quirk and Saposnik, 1962; Hanoch and Levy, 1969; Hammond, 1974; Stoyan, 1983; Li and Wong, 1999). Chang et al. (2015) used SD tests for choosing among several VaR forecasting models to analyze whether the daily capital charges produced by one model would stochastically dominate the daily capital charges produced by an alternative model.

The paper will analyse the optimality of the two standard risk measures, namely ES and VaR, with respect to the stochastic dominance relations induced by the sampling distribution of the daily capital charges produced by using either 97.5%-ES or 99%-VaR. As stated above, Basel III Accord requires that banks and other Authorized Deposit-taking Institutions (ADIs) communicate their daily risk forecasts to the appropriate monetary authorities at the beginning of each trading day, using one of a range of alternative financial risk models to forecast risk. The risk estimates from these models are used to determine the daily capital charges and associated capital costs of ADIs, depending in part on the number of previous violations, whereby realized losses exceed the estimated risk measure (for further details see, for example, Chang et al. (2011)). The objective or the risk manager is to minimize the overall cost associated with the risk measure chosen, which will be assumed to minimize the daily capital charges.

The SD approach uses as much information as possible from the probability distribution of daily capital charges. Such an approach can incorporate useful information about the likelihood of specific levels of daily capital charges, and provide decision makers with greater information about the risks associated with different options. Using the information contained in the distribution of daily capital charges, SD will provide the pairwise comparison of the two risk measures such that all risk bank managers whose utility functions belong to some set U will prefer one to another.

Seeking to choose between the two alternative risk measures, the paper will follow Weeks (1985), who defines a hypothetical utility-of-cost function. An appealing aspect of SD is that the criterion does not require a parameterized utility function, with a reliance on a general preference assumption. The bank manager's "utility-of-daily capital charges" function may be represented by a decreasing utility function with a negative second derivative, which reflects diminishing marginal utility of daily capital charges reduction associated with risk-averse decision-making (loss aversion). Risk taking for losses is not inconsistent with expected utility theory, provided the von Neumann-Morgenstern (1947) utility function possesses convex segments in the range of losses.

The paper will examine several standard models for forecasting VaR and ES, including GARCH, EGARCH, and GJR, paired with Gaussian and Student-t distributions. The results will likely show that the daily capital charges (DCC) produced using VaR (DCC_VaR) FSD (First-order Stochastic Dominance) the daily capital charges produced using ES (DCC_ES) for the four models and the two distributions. This would imply that the likelihood of a greater mean of DCC_ES is higher than the mean of DCC_VaR. In addition, given that FSD implies Second-order Stochastic Dominance (SSD), we would conclude that the uncertainty inherent in the estimated DCC_VaR is greater than the uncertainty in DCC_ES, at least in the case of extreme events, which is contrary to Danielsson's (2013) results. Given a utility-of-daily capital charges preference structure, a risk-averse manager should prefer VaR to ES as risk measurement because the expected utility of the daily capital charges produced is larger than the expected utility of the daily capital charges produced by the ES risk measure.

Three tests proposing different resampling procedures for estimating the critical values of a extended Kolmogorov-Smirnov (KS) test for SD will be used: (i) Barret and Donald (BD) (2003), who propose a standard bootstrap simulation method to mimic the asymptotic null distribution of the least favourable case (LFC); (ii) Linton, Maasoumi and Whang (2005) (LMW), who estimate the critical values using the sub-sampling method proposed in Politis and Romano (1994), which allows for general dependence among the prospects, and for observations not to be i.i.d., so that the critical values for this test do not rely on the (LFC); and (iii) the Donald and Hsu (2013) test that extends Hansen's (2005) recentering method to obtain critical values for the KS test, which increases the power properties compared with the unadjusted test mounted at the composite boundary of the null and alternative spaces, the so-called Least Favorable Case (LFC).

The remainder of the paper is organized as follows: Section 2 describes VaR and ES risk measures and how to produce daily capital charges. In Section 3 the definition, notation and properties of SD are presented along with a brief description of the SD tests. Section 4 introduces the data, describes the block bootstrapping method to simulate time series, and presents alternative conditional volatility models to produce VaR and ES in order to obtain the daily capital charges. Section 5 presents the main results. Section 6 gives some concluding comments.

2 Forecasting Value-at-Risk, Expected Shortfall and Daily Capital Charges

In this section we introduce the definitions and explain the forecasting of Value-at-Risk (VaR) and Expected Shortfall (ES). In addition, we describe how to compute Daily Capital Charges (DCC) under VaR and ES as a basic criterion for choosing between risk measures. The Basel II Accord stipulates that DCC must be set at the higher of the previous day's VaR or the average VaR over the last 60 business days, multiplied by a factor $(3+k)$ for a violation penalty, where a violation occurs when the actual negative returns exceed the VaR forecast negative returns for a given day. Although regulators have not released the DCC expression under the ES risk measure, it is fair to assume that it will consist in changing VaR to ES.

2.1 Value-at-Risk

VaR refers to the lower bound of a confidence interval for a (conditional) mean, that is, a "worst case scenario on a typical day". The VaR for a given confidence level $q \in (0,1)$ and time t is given by the smallest number y_q such that the lost Y_{t+1} at time $t+1$ will fall below y_q with probability q:

$$VaR_t^q = \inf \left\{ y_q \in \Re : P\left(Y_{t+1} \le y_q\right) \ge q \right\} = \inf \left\{ y_q \in \Re : P\left(Y_{t+1} > y_q\right) \le 1-q \right\}. \quad (1)$$

Thus, VaR is a quantile of the distribution of the loss function, and q is usually taken to be in the range [0.9, 1). For example, the Basel II accord refers to the "99%-VaR". Sometimes the level of significance or coverage rate, $\alpha = 1-q$, is used instead. If the random variable, Y_t, is normally distributed with mean μ_{t+1} and standard deviation σ_{t+1}, for $q \in (0,1)$, the VaR of Y_t is given by:

$$VaR_{t+1}^q \left(Y_t\right) = \mu_{t+1} + \sigma_{t+1}\Phi^{-1}\left(q\right), \quad (2)$$

where Φ is the cumulative distribution function of a standard normal variable.

If instead the normalized random variable $\tilde{Z}_t = \left(Y_t - \mu_t\right)/\sigma_t$ has a standardized t-distribution with $v > 2$ degrees of freedom, that is, the Student t distribution with mean 0 and variance 1, the VaR of $Y_t = \mu_t + \sigma_t\tilde{Z}_t$ would be given as:

$$VaR_{t+1}^q = \mu_{t+1} + \sigma_{t+1} \sqrt{v^{-1}(v-2)}\, t_v^{-1}(q),$$

(3)

where $t_v^{-1}(q)$ is the q quantile of the *standard* Student-t distribution. Since quantiles translate under monotonic transformations, the q quantile of the standardized Student-t distribution (mean 0 and variance 1) with v degrees of freedom is given as $t_v^{-1}(q)\sqrt{v^{-1}(v-2)}$.

In addition to these parametric VaR calculations, we will include the analysis based on the non-parametric historical VaR, which does not have to make an assumption about the parametric form of the distribution of the returns. The $100q$% historical VaR is the q quantile of the random variable, Y_t.

2.2. Expected Shortfall

The expected shortfall at level q is the expected value at time t of the loss in the next period, Y_{t+1}, conditional on the loss exceeding VaR_t^q :

$$ES_{t+1}^q = E_t\left[Y_{t+1} | Y_{t+1} > VaR_{t+1}^q\right].$$

(4)

For $q \in (0,1)$, the expected shortfall for a normally distributed random variable, $Y_t \sim N(\mu_t, \sigma_t^2)$, is given as:

$$ES_{t+1}^q = \mu_{t+1} + \sigma_{t+1} \frac{\phi\left(\Phi^{-1}(q)\right)}{1-q},$$

(5)

where ϕ is the density of a standard normal variable.

If instead the normalized random variable $\tilde{Z}_t = (Y_t - \mu_t)/\sigma_t$ has a *standardized* t-distribution with $v > 2$ degrees of freedom, then the expected shortfall of e_t is given by:

$$ES_{t+1}^q = \mu_{t+1} + \sigma_{t+1} \frac{f*_v\left(t*_v^{-1}(q)\right)}{1-q} \frac{\left(v-2+\left(t*_v^{-1}(q)\right)^2\right)}{q},$$

(6)

where $t*_v^{-1}(q)$ denotes the q quantile of the *standardized* Student t distribution (that is, with zero mean and unit variance) having v degrees of freedom, and $f*_v\left(t*_v^{-1}(q)\right)$ is the value of its density function at that point. The standardized Student t density function is given as:

$$f*_v(x) = \left((v-2)\pi\right)^{-1/2} \Gamma\left(\frac{v}{2}\right)\Gamma\left(\frac{v+1}{2}\right)\left(1+(v-2)^{-1} x^2\right)^{-(1+v)/2},$$

(7)

where the *gamma function,* Γ, is an extension of the factorial function to non-integer values (see Alexander, 2009, p. 130).

Analogous to the VaR calculations, for the calculation of ES it is possible for S_t to be replaced by alternative estimates of the conditional standard deviation in order to obtain an appropriate VaR (for useful reviews of theoretical results for conditional volatility models, see Li et al. (2002) and McAleer (2005), where several univariate and multivariate, conditional, stochastic and realized volatility models are discussed).

In the historical VaR model the ES can be estimated directly, simply by taking the average of all the losses in the tail above the VaR.

2.3. Forecasting Daily Capital Charges

In this section,which follows McAleer et al. (2013a, b, c) closely, we introduce the calculation of DCC. The Basel II Accord stipulates that DCC must be set at the higher of the previous day's VaR or the average VaR over the last 60 business days, multiplied by a factor $(3+k)$ for a violation penalty, where a violation occurs when the actual negative returns exceed the VaR forecast negative returns for a given day. Assuming that the new risk measure might be ES, we can generalize the DCC expression changing VaR to RiskM, which is the risk measure that can be either VaR or ES for day t,

$$DCC_t = \sup\left\{-(3+k)\overline{RiskM}_{60},\ -RiskM_{t-1}\right\}$$

(8)

where

DCC_t = daily capital charges,

$\overline{RiskM}_{60}$ = mean RiskM over the previous 60 working days,

$0 \leq k \leq 1$ is the Basel II violation penalty (see Table 1).

Table 1
Basel Accord Penalty Zones

Zone	Number of Violations	K
Green	0 to 4	0.00
Yellow	5	0.40
	6	0.50
	7	0.65
	8	0.75
	9	0.85
Red	10+	1.00
Note: The number of violations is given for 250 business days. The penalty structure under the Basel II Accord is specified for the number of violations and not their magnitude, either individually or cumulatively.		

It is well known that the formula given in equation (8) is contained in the 1995 amendment to Basel I, while Table 1 appears for the first time in the Basel II Accord in 2004. The multiplication factor (or penalty), *k,* depends on the central authority's assessment of the ADI's risk management practices and the results of a simple backtest. It is determined by the number of times actual losses exceed a particular day's VaR forecast (see Basel Committee on Banking Supervision).

As stated in a number of previous papers (see, for example, McAleer et al. (2013a, b, c)), the minimum multiplication factor of 3 is intended to compensate for various errors that can arise in model implementation, such as simplifying assumptions, analytical approximations, small sample biases and numerical errors that tend to reduce the true risk coverage of the model (see Stahl (1997)). Increases in the multiplication factor are designed to increase the confidence level that is implied by the observed number of violations at the 99% confidence level, as required by regulators (for a detailed discussion of VaR, as well as exogenous and endogenous violations, see McAleer (2009) and McAleer et al. (2010)).

In calculating the number of violations, it is well known that ADIs are required to compare the forecasts of VaR with realized profit and loss figures for the previous 250 trading days. In 1995, the 1988 Basel Accord (Basel Committee on Banking Supervision (1988)) was amended to allow ADIs to use internal models to determine their VaR thresholds (Basel Committee on Banking Supervision (1995)). However, ADIs that propose using internal models are required to demonstrate that their models are sound. Movement from the green zone to the red zone arises through an excessive number of violations. Although this will lead to a higher value of *k*, and hence a higher penalty, violations will also tend to be associated with lower daily capital charges. It should be noted that the number of violations in a given period is an important, though not the only, guide for regulators to approve a given VaR model.

3. Stochastic Dominance and Risk Measures

The purpose of the paper is to evaluate the optimality of 97.5%-ES and 99%-VaR with respect to the stochastic dominance relations induced by the sampling distribution of the Daily Capital Charges produced by using both risk measures. Notice that each measure will yield different values of DCC. The stochastic dominance concept is applied to determine which risk measure should be used to maximize the expected utility of a risk averse bank manager. The objective is to rank DCC, which are treated as costs/losses for banks, according to the expected value of the utility-of-DCC function introduced below in this section. In addition, we define first- and second-order stochastic dominance and its relation to the decision making process of risk seeking for losses agents. Finally, we briefly describe the SD tests that are used in this paper.

3.1. Daily Capital Charges (DCC) and evaluation framework: Stochastic Dominance

SD rules have been shown to offer superior and more efficient criteria on which to base investment decisions than the criteria derived from the traditional strategies based on first and second moments. In this paper we use SD and utility function theory to compare DCC produced by VaR or ES. We focus on the behaviour of the decision maker when the potential DCC is used as the variable of interest. Following Weeks (1985), who defines a utility-of-cost function (attractively, SD criteria do not require a parameterized utility function, but rather they rely on general preference assumption), the bank manager's utility-of-DCC function may be represented by *U (DCC)*, assuming *U´< 0* and *U´´< 0*, where *U´* is the first derivative and *U´´* is the second derivative, which reflects diminishing marginal utility of DCC reduction associated with risk-averse decision making (loss aversion, individuals seem to feel more pain from losses than from equivalent gains).

Hershey and Schoemaker (1980) investigated risk loving in the domain of losses. Risk taking for losses is not inconsistent with EU theory provided the von Neumann-Morgenstern (1947) utility function possesses convex segments in the range of losses.

Consequently, the definition of first- and second-order stochastic dominance applied to risk loving-decision makers is defined as follows:

Definition 1: Consider two distributions of DCC, DCC1 and DCC2, characterized by the cumulative distribution functions, with F and G as their respective distribution functions defined on *[a, b]*, we can define that DCC1 First-order stochastic Dominates DCC2, denotes DCC1 $\succeq_{FSD}$ DCC2, if and only if either:

a) $E\big[u(DCC_1)\big] \geq E\big[u(DCC_2)\big]$ for all u with strict inequality for some u; *or*

b) $F(x) \geq G(x)$ for all x with strict inequality for some x.

Definition 2: DCC1 Second order stochastic Dominates DCC2, denotes DCC1 $\succeq_{SSD}$ DCC2, if and only if either:

c) $E\big[u(DCC_1)\big] \geq E\big[u(DCC_2)\big]$ for all u with strict inequality for some u; *or*

d) $\int_{-\infty}^{x} F(t)\,\mathrm{d}t \geq \int_{-\infty}^{x} G(t)\,\mathrm{d}t$ for all x with strict inequality for some x. With strict inequality for some x. Having loss aversion, individuals will often be willing to accept a gamble with uncertainty and an expected loss than a guaranteed loss of the same amount.

3.2. Test statistics and critical values[1]

Three tests proposing different resampling procedures for estimating the critical values of a extended Kolmogorov-Smirnov (KS) test for SD are used, namely: (i) Barret and Donald (BD) (2003), who propose a standard bootstrap simulation method to mimic the asymptotic null distribution of the *least favourable case* (LFC)*;* (ii) Linton, Maasoumi and Whang (2005) (LMW), who estimate the critical values using the subsampling method proposed in Politis and Romano (1994), which allow for general dependence amongst the prospects, and for observations not to be i.i.d., such that the critical values for this test does not rely on the (LFC); and (iii) the Donald and Hsu (2013) test that extends Hansen's (2005) recentering method to obtain critical values for the KS test, which increases power properties compared with the unadjusted test mounted at the composite boundary of the null and alternative spaces, the so-called Least Favorable Case (LFC).

Let *DCC1* and *DCC2* be two random variables with cumulative distribution functions (CDF), F_X and F_Y, respectively. We are interested in knowing if DCC1 first stochastic dominates DCC2 that according to definition 1 corresponds to $F_Y(z) \leq F_X(z)$ for all $z \in R$. The technical assumptions required for the underlying statistical theory include the following (see Linton, Maasoumi and Whang (2005) (hereafter LMW), Linton, Song and Whang (2010) and Donald and Hsu (2013) for further details):

Assumption 3.1:

1. $Z = [0, \bar{z}]$, *where* $\bar{z} < \infty$.

2. F_X *and* F_Y *are continuous functions on Z such that* $F_X(z) = F_Y(z) = 0$ *iff z = 0, and* $F_X(z) = F_Y(z) = 1$ *iff* $z = \bar{z}$.

---

[1]This section is based on Donald and Hsu (2013), and Linton, Maasoumi, and Whang (2007).

The first part of assumption 3.1 requires F_X and F_Y to be continuous on Z. The second part rules out the cases where $F_X(z) = F_Y(z) = 0$

Assumption 3.2*:*

1. $\{X_i\}_{i=1}^N$ and $\{Y_i\}_{i=1}^M$ *are samples from distributions with CDF's F_X and F_Y , respectively. It is possible to deal with independent samples and observations. Linton, Maasoumi and Whang (2005) allow dependent time series, and possibly dependent X and Y.*

2. *M is a function of N such that $M(N) \to \infty$ and $N/(N + M(N)) \to \lambda \sim (0, 1)$ when $N \to \infty$.*
Assumption 3.2 requires that N and M grow at the same rate. In order to test whether X FSD Y, the hypothesis of interest can now be stated as in McFadden (1989):

$$H_0 : F_Y\left(z\right) \le F_X\left(z\right) \text{ for all } z \in Z, \qquad (9)$$

$$H_1 : F_Y\left(z\right) > F_X\left(z\right) \text{ for all } z \in Z, \qquad (10)$$

Under the null, alternative X would provide a higher probability of lowest DCCs.

The CDF's F_X and F_Y are estimated by empirical CDFs:

$$\hat{F}_{X,N}(z) = \frac{1}{N}\sum_{i=1}^N 1(X_i \le z), \ \hat{F}_{Y,M}(z) = \frac{1}{M}\sum_{i=1}^N 1(Y_i \le z),$$

where $1(\cdot)$ denotes the indicator function. The Kolmogorov-Smirnov test statistic is given by:

$$\hat{S}_N = \sqrt{\frac{NM}{N+M}} \sup_{z \in Z} (\hat{F}_{Y,N}(z) - \hat{F}_{X,M}(z)) .$$

Regarding higher-order SD, Davidson and Duclos (2000) offer a very useful characterization of any SD order and tests. They define $D_k^{(s)}(z) = \int_{-\infty}^z D_k^{(s-1)}(t)dt, k = Y, X,$ where $D_k^{(1)}(z) = F_k(z).$ Then X stochastically dominates Y at order s, if $D_y^{(s)}(z) \le D_x^{(s)}(z)$ for all z with strict inequality for some z. In order to test if *X SD2 Y*, we can formulate the null and alternative hypotheses as:

$$H_0 : D_y^{(2)}(z) \le D_x^{(2)}(z) \text{ for all } z \in Z, \qquad (11)$$

$$H_1 : D_y^{(s)}(z) > D_x^{(2)}(z) \text{ for all } z \in Z, \qquad (12)$$

Under the null, over the entire range of DCC outcomes, risk adverse bank managers would prefer alternative X because it provides a greater utility of DCC.

The corresponding test statistic is defined as:

$$\hat{S}_{2,N} = \sqrt{\frac{NM}{N+M}} \sup_{z \in Z} \left(\hat{D}_{Y,N}^2(z) - \hat{D}_{X,M}^2(z)\right),$$

where a natural estimator of $D_k^{(2)}(z)$ is:

$$\hat{D}_k^{(2)}(z) = \int_0^z (z-t) d\hat{F}_k(t) = \frac{1}{N} \sum_{i=1}^N (z-t_i) I(t_i \leq z) = \frac{1}{N} \sum_{i=1}^N (z-t_i)_+$$

Under the above assumptions, which we adopt here, these tests are known to be asymptotically Gaussian. The asymptotic approximation is also known to be too poor to provide reliable empirical guidance. In response, many resampling techniques and approaches have been proposed in the literature. We implement three of these suggested resampling techniques in this paper.

3.2.A.- Barret and Donald (BD) (2003)

BD apply three different techniques for simulation to mimic the asymptotic null distribution in the LFC: the multiplier method and bootstrap with both separate and combined samples. In this paper, the critical values $\hat{q}_N^{BD}$ are computed using the bootstrap with separate samples. Draw a random sample of size N from $\{X_1, \ldots, X_N\}$ and sample size M from $\{Y_1, \ldots, Y_M\}$ to form $\hat{D}_{k,N}^{(s)BD}(z) = \int_{-\infty}^z \hat{D}_k^{(s-1)BD}(t) dt$, where $\hat{D}_k^{(1)BD}(z) = \hat{F}_k^{DB}(z), k = X, Y$.

Define $\hat{S}_{YX}^s(z) = \left(\hat{D}_{Y,N}^{(s)BD}(z) - \hat{D}_{Y,N}^s(z) \right) - \left(\hat{D}_{X,M}^{(s)BD}(z) - \hat{D}_{X,M}^s(z) \right)$ and let P^u denote the conditional probability measure given the observed sample. Let α be the significance level, then $\hat{q}_N^{BD}$ can be computed as:

$$\hat{q}_N^{BD} = \sup \left\{ q \mid P^u \left(\sqrt{\frac{NM}{N+M}} \sup_{z \in Z} \hat{S}_{YX}^s(z) \leq q \right) \leq 1 - \alpha \right\}.$$

(13)

$\hat{q}_N^{BD}$ is bounded away from zero in probability.

3.2.B.- Linton, Maasoumi and Whang (2005)

BD simulation methods do not work well enough when the data are weakly dependent, as for time series samples that are used in this paper. In these cases, one has to appeal to either the subsampling technique of LMW, or a variant of the block bootstrap. Donald and Hsu (2013) provide a comparative examination of these alternatives

LMW estimate the critical value by the subsampling method proposed by Politis and Romano (1994) that allows the series to be mutually dependent over time. Let $\{(X_i, Y_i)\}_{i=1}^N$ be a strictly stationary time series sequence with joint distribution function F_{XY} on Z^2 and marginal CDF's, F_X and F_Y, respectively. Suppose that Assumption 1 of LMW holds. Then under the null hypothesis that H_0: $D^s_Y(z) \leq D^s_X(z)$ for all $z \sim Z$, the SD tests defined earlier are asymptotically Gaussian.

Donald and Hsu (2013) introduce LMW's test with a modification that allows for different sample sizes. For $s \geq 1$, let X_s denote the collection of all of the subsets of size s of $\{X_1, ...,X_N\}$:

$$X_s \equiv \left\{ \{X_{r_1}, \ldots, X_{r_s}\} \mid \{r_1, \ldots, r_s\} \subseteq \{1, \ldots, N\} \right\}.$$

A random draw denoted by $\{X_1^b, \ldots, X_s^b\}$ from X_s would be a random sample of size s without replacement from the original data. Let $\hat{F}_{X,s}^b$ be the empirical CDF based on the random draw, $\{X_1^b, \ldots, X_s^b\}$. Define $\hat{F}_{Y,s}^b$ similarly and let s_N and s_M denote the subsampling sizes for the X and Y samples, respectively, and define $\hat{D}_{YX}^{(s)b}(z) = \hat{D}_{Y,N}^{(s)b}(z) - \hat{D}_{X,M}^{(s)b}(z),$ where $\hat{D}_{k,N}^{(s)b}(z) = \int_{-\infty}^{z} \hat{D}_k^{(s-1)b}(t)\,dt,$ where $\hat{D}_k^{(1)b}(z) = \hat{F}_k^b(z), k = Y, X.$ Then the subsampling critical value $\hat{c}_N^{(s)LMW}$ for any s pre-specified order is given by:

$$\hat{c}_N^{(s)LMW} = \sup\left\{ c \mid P^u\left(\sqrt{\frac{s_N s_M}{s_N + s_M}} \sup_{z\in Z}\hat{D}_{XY}^{(s)b} \leq c\right) \leq 1 - \alpha\right\}.$$

Assume that:

1. $s_N \to \infty$, $s_M \to \infty$, $s_N/N \to 0$ and $s_M/M \to 0$ as $N \to \infty$.
2. $s_N/(s_N + s_M) \to \lambda$, where λ is defined in Assumption 4.2.

These subsampling SD tests are known to be asymptotically Gaussian under these assumptions, and provide consistent tests. The limiting distribution theory in LMW covers weakly stationary, dependent samples, with certain mixing conditions, such as in our applications. In addition, they allow for the prospects that are ranked to be estimated functions, rather than the original series described above. If the estimators involved in these functions permit certain expansions, as described in Linton, Maasoumi and Whang (2005), Assumption 2, Section 3.1, the limiting distribution theory will be preserved with re-centering.

3.2.C.- Donald and Hsu (2013) and Re-centering Functions.

Donald and Hsu (2013) and Linton, Maasoumi and Whang (2005) propose re-centering methods introduced by Hansen (2005) to construct critical values for Kolmogorov-Smirnov type tests. This approach provides a test with improved size and power properties compared with the unadjusted test mounted at the composite boundary of the null and alternative spaces, the so-called Least Favorable Case (LFC).

For a negative number a_N, Donald and Hsu (2013) define the re-centering function $\hat{\mu}_N(z)$ as:

$$\hat{\mu}_N(z) = \left(\hat{F}_{Y,M}(z) - \hat{F}_{X,N}(z)\right) \cdot \mathbb{1}\left(\sqrt{N}\left(\hat{F}_{Y,M}(z) - \hat{F}_{X,N}(z)\right) < a_N\right).$$

See Donald and Hsu (2013) for further details of recentering functions. The blockwise bootstrap is used to compute the critical values based on the sum of the simulated processes and the recentering function. Let $\hat{D}_k^{(1)bb}(z) = \hat{F}_k^{bb}(z), k = Y, X$ denote the empirical CDF computed with the b-th resample/subsample block. Define $\hat{D}_{YX}^{(s)bb}(z) = \hat{D}_{Y,N}^{(s)bb}(z) - \hat{D}_{X,M}^{(s)bb}(z),$ where $\hat{D}_{k,N}^{(s)bb}(z) = \int_{-\infty}^{z}\hat{D}_k^{(s-1)bb}(t)\,dt.$ Then for $\alpha < \frac{1}{2}$:

$$\hat{c}_{\eta,N}^{(s)bb} = \max\{\tilde{c}_N^{(s)bb}, \eta\},$$

$$\tilde{c}_N^{(s)bb} = \sup\left\{ c \mid P^u\left(\sup_{z \in Z} \sqrt{N}\left(\hat{D}^{(s)bb}{}_N(z) + \hat{\mu}_N(z)\right) \le c\right) \le 1 - \alpha\right\},$$

where η is an arbitrarily small positive number. If the decision rule is to reject the null hypothesis, H_0: $D^{(s)}{}_Y(z) \le D^{(s)}{}_X(z)$ for all $z \sim Z$ when $\hat{S}_{sN} > \hat{c}_{\eta,N}^{(s)bb}$, then the corresponding test has the same size properties as in the independent random samples case.

4. Data and Implementation of Tests

In this section, we describe the data used together with the block bootstrapping procedure for simulating the time series of the stock prices. In addition, for computing 97.5%-ES and 99%-VaR, the conditional variances must be estimated. We use three different univariate conditional volatility models that are also described in this section.

4.1. Data description

The data used for estimation and forecasting are the closing daily prices for Standard and Poor's Composite 500 Index (S&P500), which were obtained from the Thomson Reuters-Datastream database for the period 1 January 1999 to 26 June 2014, giving 4040 observations.

The returns at time t (R_t) are defined as:

$$R_t = \log\left(P_t / P_{t-1}\right), \qquad (14)$$

where P_t is the market price.

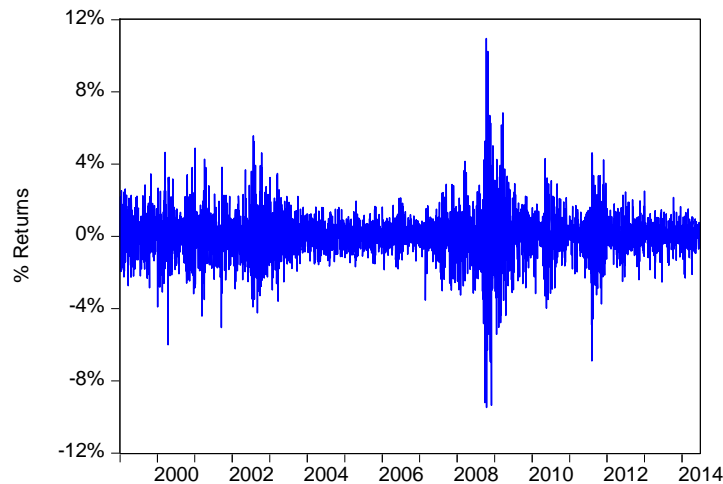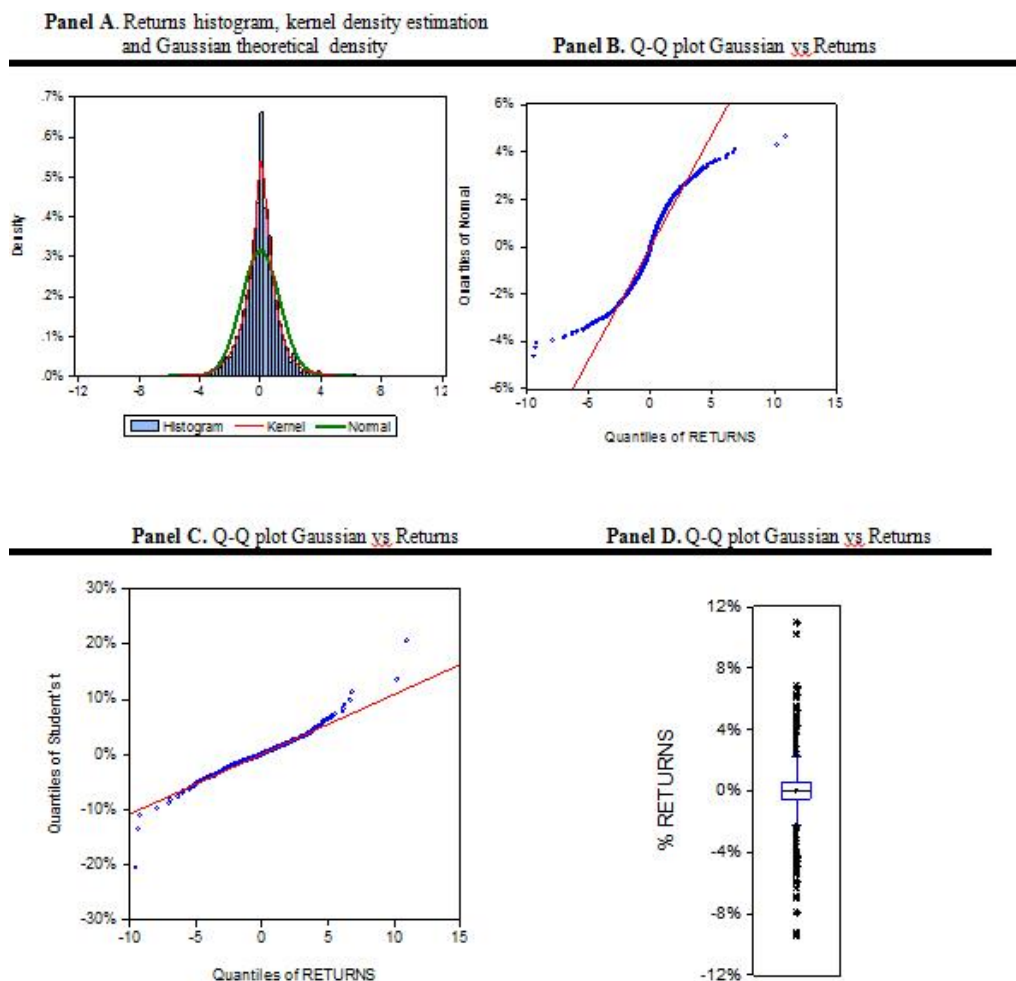**Figure 1**

S&P500 Returns

1 January 1999 – 24 June 2014

Figure 1 shows the S&P500 returns. The extremely high positive and negative returns are evident from September 2008 onward, after the Lehman Brothers bankruptcy, and have continued well into 2009. Then, in spring, 2010, the European debt crisis began, with the European Union together with the International Monetary Fund providing 110 million Euros to Greece that became unable to borrow from the market. Greece required a second bailout in mid-2011. Thereafter, Ireland and Portugal also received bailouts in November 2010 and May 2011, respectively. Higher volatility in the S&P500 returns is observed during these periods. Regarding the descriptive statistics, the median (0.022) is above the mean (0.012) and the range is between 11% and -9.5%, with a standard deviation of 1.27. S&P500 returns show negative skewness (-0.17) and high kurtosis (10.99), which would seem to indicate the existence of extreme observations and non-Gaussianity.

Figure 2 shows several graphs that provide valuable information for identifying the returns probability distribution. Panel A displays the empirical histogram, together with the density function of the Gaussian distribution and a kernel density estimate of the distribution that show fatter tails than normal and some slight asymmetry. Panels B and C exhibit two theoretical quantile-quantile plots (QQ-plot) comparing the quantiles of the S&P500 returns with the quantiles of both a fitted normal, Panel B, and Student-t, Panel C, distributions. For the Gaussian case, the QQ-plot does not lie on a straight line, overall on the tails, supporting the non-normality of returns. According to the QQ-plots, the Student-t distribution seems to fit the observed data better than does the Gaussian distribution. Finally, Panel D displays a boxplot that summarizes the returns distributions showing the extreme observation mentioned above.

Figure 2

SP500 Returns distribution analysis
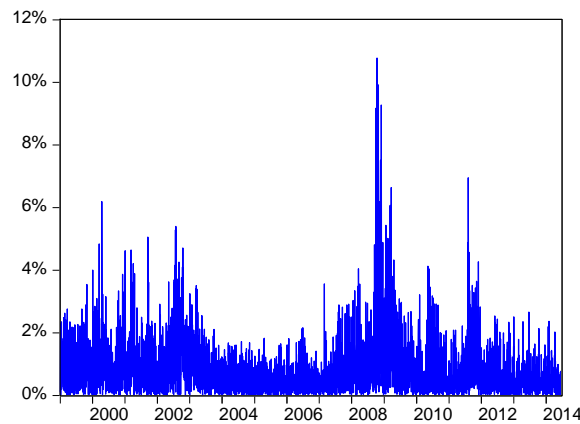
1 January 1999 – 24 June 2014

Figures 1 and 2 show that stock markets have been working under stress during the last seven years. Traditional risk measurement, specifically VaR, might not work properly under these extreme price fluctuations. The BIS Committee on the Global Financial System discussed the shortcomings of VaR for measuring and monitoring market risk when many such events are taking place in the tails of the distributions. VaR that suffers tail risk only measures the distribution quantile, and disregards the extreme loss beyond the VaR level, ignoring important information regarding the tails of the distribution. Expected shortfall might be a more appropriate tool for risk monitoring under stress circumstances. As Yamai and Yoshiba (2005) state, expected shortfall has no tail risk under more lenient conditions than VaR.

Fat tails might be explained by clusters of volatility that seem to appear in Figure 1. A closer examination of the volatility of returns using a measure proposed in Franses and van Dijk (1999) is given as:

$$V_t = \left(R_t - E\left(R_t \mid F_{t-1} \right) \right)^2 , \tag{15}$$

where F_{t-1} is the information set at time *t*-1, highlights the volatility clustering (see Figure 3).

Figure 3

Volatility of S&P500 Returns

1 January 1999 – 24 June 2014

4.2. Block bootstrapping

In order to increase the power of the three tests used in this analysis, we will use block bootstrapping for simulating 500 time series of the S&P500 returns for the 3000-observation rolling window that will be used for producing a total of 500 one-step-ahead 97.5%-ES and 99%-VaR forecasts. We implement the Circular Block Bootstrapping (CBB) method developed in Politis and Romano (1992) for resampling the S&P500 through the MFE toolbox of Sheppard (2013). The block bootstrap is widely used for implementing the bootstrap with time series data. It consists of dividing the data into blocks of observations and sampling the blocks randomly, with replacement.

In the CBB, let the data consist of observations $\{X_i : i = 1, ..., n\}$, and let $l \in \{1, ..., n\}$ and $b \geq 1$ denote the length and the number of blocks, respectively, such that $lx\,b \leq n$. Let *n* and *m* be the initial data size and the bootstrap sample size, $m \leq n$ and *k* the number of blocks chosen. CBB consists of dividing the time series into *b* blocks of consecutive observations denoted by:

$$B_i = \left(X_{(i-1)l+1}, ..., X_{il} \right), \ i = 1, ..., n.$$

A random sample of k blocks, $k \geq 1$, $B_1^*, ..., B_k^*$ is selected with replacement from $B_1^*, ..., B_k^*$. Joining the k blocks with $m = k \times l$ observations, the bootstrap sample is given as:

$$\left(X_1^*, ..., X_l^*, ..., X_{(k-1)l+1}^* ..., X_l^* \right).$$

The CBB procedure is based on wrapping the data around a circle and forming additional blocks using the "circularly defined" observations. For $i \geq n$, it is defined that $X_1 = X_{i_n}$, where $i_n = i \bmod n$ and $X_0 = X_n$. The CBB method resamples overlapping and periodically extended blocks of length l. Notice that each X_i appears exactly l times in the collection of blocks and, as the CBB resamples the blocks from this collection with equal probability, each of the original observations $X_1, ..., X_n$ receives equal weight under the CBB. This property distinguishes the CBB from previous methods, such as the non-overlapping block bootstrap of Carlstein (1992). Note that and 97.5%-ES and 99%-VaR are estimated for each drawn sample, thereby generating the bootstrap (subsample) distribution of the test statistics.

The next section describes several volatility models that are widely used to forecast the 1-day ahead conditional variances and VaR thresholds for the parametric cases.

4.3. Models for Forecasting VaR

ADIs can use internal models to determine their VaR thresholds. There are alternative univariate time series models for estimating conditional volatility. In what follows, we present several well-known conditional volatility models that can be used to evaluate strategic market risk disclosure, namely GARCH, GJR and EGARCH, with Gaussian and Student-*t* distributions. These univariate models are chosen because they are widely used in the literature. For an extensive discussion of the theoretical properties of several of these models see, for example, Ling and McAleer (2002a, b, 2003a), McAleer (2005), Caporin and McAleer (2012), McAleer and Hafner (2014), and McAleer (2014).

4.3.A.- GARCH

For a wide range of financial data series, time-varying conditional variances can be explained empirically through the autoregressive conditional heteroskedasticity (ARCH) model, which was proposed by Engle (1982). When the time-varying conditional variance has both autoregressive and moving average components, this leads to the generalized ARCH(p,q), or GARCH(p,q), model of Bollerslev (1986). It is very common in practice to impose the widely estimated GARCH(1,1) specification in advance.

Consider the stationary AR(1)-GARCH(1,1) model for daily returns, y_t:

$$y_t = \varphi_1 + \varphi_2 y_{t-1} + \varepsilon_t, \qquad |\varphi_2| < 1 \tag{16}$$

For $t = 1, ..., n$, where the shocks to returns are given by:

$$\varepsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid(0,1)$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{17}$$

and $\omega > 0, \alpha \geq 0, \beta \geq 0$ are sufficient conditions to ensure that the conditional variance $h_t > 0$. The stationary AR(1)-GARCH(1,1) model can be modified to incorporate a non-stationary ARMA(*p,q*) conditional mean and a stationary GARCH(*r,s*) conditional variance, as in Ling and McAleer (2003b). Tsay (1987) shows that $\alpha > 0$ in the derivation of the GARCH model.

4.3.B. - GJR

In the symmetric GARCH model, the effects of positive shocks (or upward movements in daily returns) on the conditional variance, h_t, are assumed to be the same as the effects of negative shocks (or downward movements in daily returns) of equal magnitude. In order to accommodate asymmetric behaviour, Glosten, Jagannathan and Runkle (1992) proposed a model (hereafter GJR), for which GJR(1,1) is defined as follows:

$$h_t = \omega + (\alpha + \gamma I(\eta_{t-1}))\varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{18}$$

where $w > 0, a \geq 0, a + g \geq 0, b \geq 0$ are sufficient conditions for $h_t > 0$, and $I(\eta_t)$ is an indicator variable defined by:

$$I(\eta_t) = \begin{cases} 1, & \varepsilon_t < 0 \\ 0, & \varepsilon_t \geq 0 \end{cases} \tag{19}$$

as η_t has the same sign as ε_t. The indicator variable differentiates between positive and negative shocks, so that asymmetric effects in the data are captured by the coefficient g. For financial data, it is expected that $\gamma \geq 0$ because negative shocks have a greater impact on risk than do positive shocks of similar magnitude. The asymmetric effect, γ, measures the contribution of shocks to both short run persistence, $\alpha + \gamma/2$, and to long run persistence, $\alpha + \beta + \gamma/2$.

Although GJR permits asymmetric effects of positive and negative shocks of equal magnitude on conditional volatility, the special case of leverage, whereby negative shocks increase volatility while positive shocks decrease volatility (see Black (1976) for an argument using the debt/equity ratio), cannot be accommodated, in practice (for further details on asymmetry versus leverage in the GJR model, see Caporin and McAleer (2012)). McAleer (2014) showed that $\alpha > 0$ and $\gamma > 0$ in the derivation of the GJR model.

4.3.C.- EGARCH
An alternative model to capture asymmetric behaviour in the conditional variance is the Exponential GARCH, or EGARCH(1,1), model of Nelson (1991), namely:

$$logh_t = \omega + \alpha \left| \frac{\varepsilon_{t-1}}{h_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{h_{t-1}} + \beta logh_{t-1}, \quad |\beta| < 1, \tag{20}$$

where the parameters α, b and g have different interpretations from those in the GARCH(1,1) and GJR(1,1) models presented above.

EGARCH captures asymmetries differently from GJR. The parameters α and g in EGARCH(1,1) represent the magnitude (or size) and sign effects of the standardized residuals, respectively, on the conditional variance, whereas α and $\alpha + \gamma$ represent the effects of positive and negative shocks, respectively, on the conditional variance in GJR(1,1). As in the case of GJR, EGARCH cannot accommodate leverage (further details can be found in McAleer et al. (2007) and McAleer (2014)). McAleer and Hafner (2014) showed that $\alpha > 0$ and $\gamma > 0$ in the derivation of the EGARCH model, which prevents leverage from consideration in the EGARCH model.

In the empirical analysis, the three conditional volatility models given above are estimated under the following distributional assumptions on the conditional shocks: (1) Gaussian and (2) Student-*t*, with estimated degrees of freedom. As the models that incorporate the *t* distributed errors are estimated by QMLE, the resulting estimators are consistent and asymptotically normal, so they can be used for estimation, inference and forecasting.

5. Empirical Results

Let DCC_ES and DCC_ VaR be the DCC produced using the 97.5%-ES and 99%-VaR, respectively. Based on definitions 1 and 2, if DCC_VaR first-order stochastically dominates DCC_ES, then the DCC_VaR will involve higher probability of lower DCCs than the latter. Similarly, if DCC_VaR distribution dominates the DCC_ES distribution stochastically at second order, it would imply that a risk averse risk manager would prefer the DCC_VaR distribution because it provides a greater utility of costs. In essence, stochastic dominance preference ordering rule would divide the two alternative risk measurements to compute DCC into the efficient set of undominated alternatives and the inefficient set of dominated alternatives. The expected utility of those alternatives in the efficient set is larger than the expected utility of those in the inefficient set.

According to definitions 1 and 2, graphically, DCC_VaR would dominate DCC_ES when its cumulative distribution function is above the CDF of DCC_ES. SSD implies that the difference between the integral of the cumulative distribution functions (ICDF) (area under CDFs) DCC_VaR and ICDF DCC_ES is always positive for every level of probability.

Neither ES nor VaR is observed; they have to be estimated, so we proceed as follows:

(1) We use a 3000-observations long rolling window (from 1 January 1999 to 1 July 2010, around 75% of the total number of available observations) for smoothing out spikes of volatility during the Global Financial Crisis, while estimating the conditional volatility models for producing one-step-ahead 97.5%-ES and 99%-VaR forecasts.

(2) In order to obtain the *empirical distribution of DCC* using both risk measures, block bootstrapping is used for simulating 500 time series of the S&P500 returns for the 3000-observations rolling window chosen in step 1 that will be used for producing a total of 500 one-step-ahead 97.5%-ES and 99%-VaR forecasts.

(3) Steps 1 and 2 are then repeated for the 1040 days remaining in the total sample (from 2 July 2010 until 26 June 2014), each time leaving out the first observation and adding a new observation at the end. This procedure yields a 500×1040 matrix for computing the cumulative distribution functions needed for testing SD.

Table 2
Rejection Rates for First-order SD Tests

	Gaussian			Student-t				
Design	**GARCH**	**EGARCH**	**GJR**	**GARCH**	**EGARCH**	**GJR**		**HS**
BB	0.0001	0.0112	0.0010	0.0000	0.0000	0.0000		0.2002
BD	0.0000	0.0092	0.0000	0.0000	0.0000	0.0000		0.1848
LMW	0.0061	0.0112	0.0031	0.0000	0.0000	0.0000		0.1848

Note:Rejection rates are from three different tests, namely Donald and Hsu (2013) (BB), Barrett and Donald (2003) (BD), and Linton, Maasoumi and Whang (2005) (LMW) for the null hypothesis: H_0: DCC_VaR FSD DCC_ES, where DCC_VaR and DCC_ES denote the DCC produced using ES and VaR risk measurements, respectively. Forecast risk measures are produced using two probability distributions, Gaussian (left panel of the table) and Student t, and three conditional volatility models, stated in the first row of each table. In the last column, VaR and ES are computed using the Historical Simulation (HS) procedure.

Table 3
Rejection Rates for Second-order SD Tests

	Gaussian			Student-t			HS
Design	**GARCH**	**EGARCH**	**GJR**	**GARCH**	**EGARCH**	**GJR**	**HS**
BB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
LMW	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Rejection rates are from three different tests, namely Donald and Hsu (2013) (BB), Barrett and Donald (2003) (BD), and Linton, Maasoumi and Whang (2005) (LMW) for the null hypothesis: H_0: DCC_VaR SSD DCC_ES, where DCC_VaR and DCC_ES denote the DCC produced using ES and VaR risk measurements, respectively. Forecast risk measures are produced using two probability distributions, Gaussian (left panel of the table) and Student t, and three conditional volatility models, stated in the first row of each table. In the last column, VaR and ES are computed using the Historical Simulation (HS) procedure.
.

Table 2 presents rejection rates from three different tests, namely Donald and Hsu (2013) (BB), Barrett and Donald (2003) (BD), and Linton, Maasoumi and Whang (2005) (LMW), for the null hypothesis: *H_0*: DCC_VaR FSD DCC_ES. Forecast risk measures, needed for computing DCCs, are

produced using two probability distributions, Gaussian (left panel of the table) and Student-t, and three conditional volatility models given in the first row of each table. In addition, historical simulation (HS) results are shown in the analysis. Table 3 shows rejection rates for SSD tests of BB, BD and LMW for the same distributions and volatility models. The *p*-values for the blockwise bootstraps method are approximated based on 200 replications, and the *p*-values for the subsampling method are approximated based on the 176 possible subsamples.
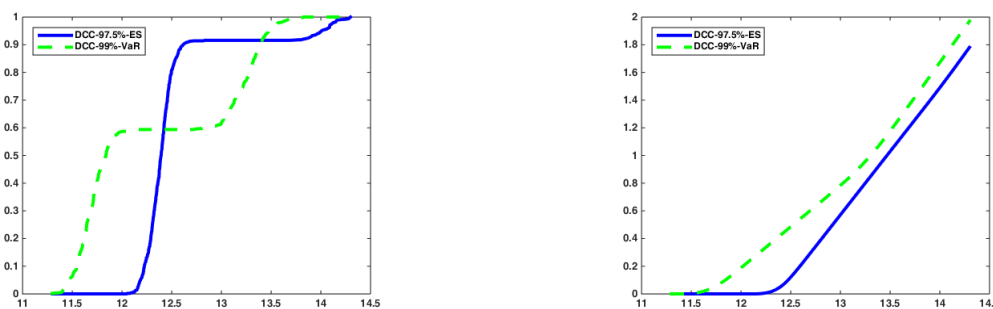
The significance level is set to 5%. For example, in Table 2, under the Gaussian distribution and using GARCH to produce DCC_97.5%-ES and DCC_99%-VaR forecasts, the BB test obtained a 0.0001% rejection rate for the null hypothesis that DCC_VaR FSD DCC_*VaR*. Following Donald and Hsu (2013), when implementing the blockwise bootstrap, the block sizes are set to 12 and the subsample size is set to 25.

Summarizing, the main results are given as follows:

(1) The BB, BD and LMW tests in Table 2 show that DCC_VaR, assuming both Gaussian and Student-t distributions and for every conditional volatility model, first-order stochastically dominates DCC_ES. This is equivalent to the statement that using VaR for producing DCC provides a higher probability of obtaining lower DCCs than using ES. Therefore, the expected utility of the risk manager would be higher using VaR instead of ES. Nonetheless, DCC-VaR does not FSD DCC-ES when using Historical Simulation. Rejection rates are greater than 0.05. Figure 4 shows CDFs and ICDFs for observation 3558 (21 August, 2012) under HS. As CDFs shown in Figure 4 cross, first degree stochastic dominance cannot be shown. DCC_VaR model has a much smaller risk of higher costs at levels of dcc lower than 12% and greater than 13.5%, nonetheless the risk of higher costs for dcc level between 12 and 13.5% is higher. Thus, the risk manager must next test for second degree stochastic dominance by plotting the difference in the areas under de CDFs at all daily capital charges levels. This is shown in the right panel in Figure 4. In this example, the area between the CDFs of the DCC_VaR and the DCC_ES is always greater than or equal to zero for all possible outcomes. Therefore the DCC_VaR second order dominates the DCC_ES, even though the uncertainty associated with the DCC_VaR is greater than the associated with DCC_ES.

Figure 4

CDF and ICDF for DCC_97.5%-ES and DCC_99%-VaR

Note: In the left panel, the solid line is the CDF of DCC_97.5%-ES and the dashed line is the CDF of DCC_99%-VaR produced by historical simulation. In the right panel, the solid and dashed lines depict the integrated cumulative distribution functions (ICDF) of the CDFs shown in the left panel. These are the empirical distributions of DCC_99%-ES and DCC_99%-VaR of observation 3559 (21 August, 2012).
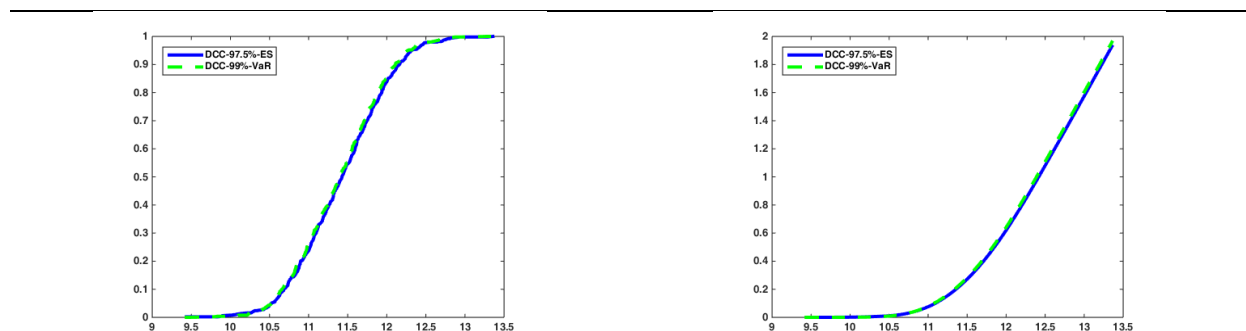
(2) As FSD implies SSD, then DCC_VaR SSD DCC_ES for all the conditional volatility models and distribution errors used in the paper. Therefore, using VaR for producing DCC would be chosen by risk managers who prefer less DCC to more and have loss aversion, who will often be willing to accept a gamble with uncertainty and an expected loss than a guaranteed loss of the same amount.

Contradictory objectives might exist between regulators and risk bank managers: on the one hand, regulators would prefer 97.5%-ES to the non-coherent 99%-VaR as a risk measure because it only provides the amount that is at risk with a particular probability. It does not suggest how much is at risk at twice that probability, or at half that probability, as it only tells part of the risk condition. On the other hand, following the SD analysis in this paper, VaR would be the preferred risk measure of risk managers. In addition, it seems that, contrary to the analysis of Danielsson (2013), which was based only on a comparison of standard deviations, VaR turns out to be a stochastically different measure of tail-risk when compared with ES.

(3) The fact that Student-t has heavier tails than the Gaussian distribution explains why CDFs and ICDFs are very close under Gaussianity, while this is not the case when the Student-t is used. This can be seen in Figure 5 assuming the Gaussian distribution, and Figure 6 under Student-t. Cumulative Distribution Functions (CDF) in the left panel and the integrated CDFs (ICDF) in the right panel are for observation 3558 (21 August, 2012). These figures represent DCC_VaR and DCC_ES produced using 99%-VaR and 97.5%-ES when GARCH is used. The outcomes of the SD tests for the Gaussian distribution, showing that DCC_VaR FSD DCC_ES, are illustrative in the light of previous results (see Danielsson, 2013), in which 97.5%-ES and 99%-VaR had similar statistical properties (namely mean and standard deviation), thereby making it difficult to uncover any inherent empirical differences.
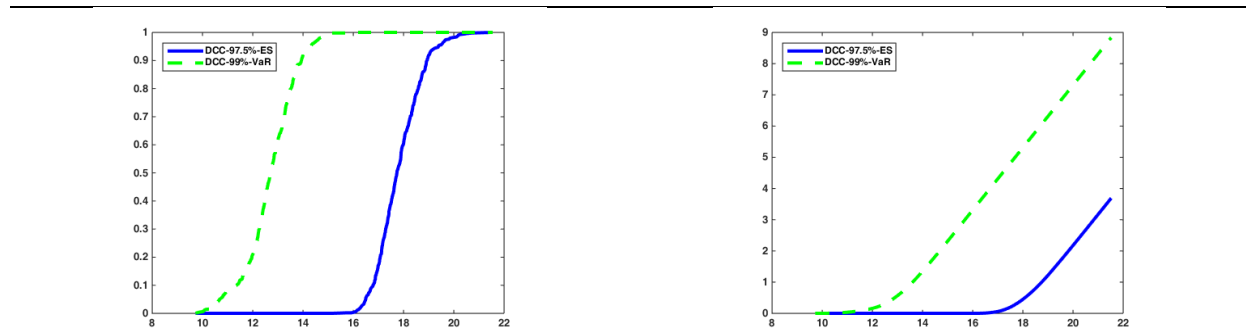
Figure 5

CDF and ICDF for DCC_97.5%-ES and DCC_99%-VaR

Note: In the left panel, the solid line is the CDF of DCC_97.5%-ES and the dashed line is the CDF of DCC_99%-VaR produced by a EGARCH model assuming a Gaussian distribution of S&P500 returns. In the right panel, the solid and dashed lines depict the integrated cumulative distribution functions (ICDF) of the CDFs shown in the left panel. These are the empirical distributions of DCC_99%-ES and DCC_99%-VaR of observation 3559 (21 August, 2012).

Figure 6

CDF and ICDF for DCC_97.5%-ES and DCC_99%-VaR

Note: In the left panel, the solid line is the CDF of DCC_97.5%-ES and the dashed line is the CDF of DCC_99%-VaR produced by a EGARCH model assuming a Student-t distribution of S&P500 returns. In the right panel, the solid and dashed lines depict the integrated cumulative distribution functions (ICDF) of the CDFs shown in the left panel. These are the empirical distributions of DCC_99%-ES and DCC_99%-VaR of observation 3559 (21 August, 2012).

6.Conclusions

This paper proposed a robust comparison between ES and VaR using the SD ranking that is not limited to mean variance considerations. It is consistent with the spirit of Basel III concerns with tail outcomes, and considers a more general comparison that is robust to particular loss functions (for example, mean and variance/quadratic), and underlying (unknown) exact distributions of the estimated DCC. While VaR is still in great use, the ES may be a preferable risk measure (Dowd, 2005). Specifically, ES considers losses beyond the VaR level and is shown to be sub-additive, while VaR disregards losses beyond the percentile and is not sub-additive.

However, the calculation of ES will almost always involve greater computation than VaR. In this paper, we have presented easily-computed expressions for evaluating the ES for two distributions that are commonly used for modelling asset returns, namely the Gaussian and Student-t. For other more complicated cases, obtaining ES might be more challenging.

The paper evaluates the optimality of 97.5%-ES and 99%-VaR with respect to the stochastic dominance relations induced by the sampling distribution of the Daily Capital Charges produced by using both risk measures. SD provides the pairwise comparison of the two risk measures such that risk bank managers whose utility functions belong to some set U will prefer one to another. Stochastic dominance ordering is theoretically superior to statistical moment rules (for example, mean-variance analysis). The SD approach uses as much information as possible from the DCC probability distribution.

Stochastic dominance incorporates alternative risk attitudes (such as risk aversion and loss aversion) of decision makers more easily than other decision ranking methods, and provides more robust solutions. It was found that, under any risk manager utility-of-cost function such that less DCC is preferred to more and showing loss aversion, DCC_VaR stochastically dominated DCC_ES, so that DCC_VaR is preferable to DCC_ES in an expected utility sense. Loss aversion may reflect risk managers who will often be more willing to accept an expected loss than a guaranteed loss of the same amount.

According to previous results, some conclusions arise that might shed light on the Basel Committee dilemma of moving from 99%-VaR to 97.5%-ES. First, the null hypothesis of dominance of the 97.5%-ES by the 99%-VaR cannot be rejected. Hence, using VaR for producing DCC would be chosen by risk managers who prefer less DCC to more and have loss aversion. SD is even more easily perceived for fat-tailed conditional distributions. Secondly, the 97.5%-ES not only accounts for the tail-risk, but also provides a more stable measurement of risk, and being less sensitive to extreme observations, it might be the preferred option by regulators.

The financial crisis has reinforced the need and highlighted the importance of a well-functioning regulatory framework for transparent and efficient markets with the right incentives. This paper, providing worthwhile information about the ES and VaR statistical properties, suggests that regulators should weight the advantages of moving from VaR to ES in terms of providing a more stable risk measure, but bearing in mind how this change can jeopardize bank profits. There is not a unique winner in terms of risk management. On the one hand, lacking the tail risk problem, the information provided by ES would not mislead investors but, on the other, ES implies a greater chance of larger DCC than using VaR.

An important caveat of the paper is that SD between ES and VaR can only be obtained for the class of utility function defined above that we consider might fit the bank risk manager preferences. One cannot determine the best risk measure with respect to other classes of utility functions. SD has not been used to identify the risk measure that is preferred under a regulator's utility function. The regulator choice of the risk measure using the SD approach would require further specification of a specific utility function. Note that in certain situations there may be no dominant risk measure.

Finally, while these results are obtained using a single asset and a limited range of models and distributions, they are likely to hold in a more general setting. The next step in this research agenda is to use a variety of assets, both from the USA and other countries, and also alternative univariate and multivariate risk models to ascertain the validity of these empirical results.

References

Alexander, C. (2009), Market Risk Analysis: Value-at-Risk Models, Volume 4, Wiley, New York.

Artzner, P., F. Delbaen, J.M.Eber, and D. Heath (1997), Thinking coherently, Risk, 10(11), 68-71.

Barrett, G., and S. Donald (2003), Consistent tests for stochastic dominance, Econometrica, 71, 71-104.

Basel Committee on Banking Supervision (1988), International Convergence of Capital Measurement and Capital Standards, BIS, Basel, Switzerland.

Basel Committee on Banking Supervision (1995), An Internal Model-Based Approach to Market Risk Capital Requirements, BIS, Basel, Switzerland

Basel Committee on Banking Supervision (2013), Consultative Document, Fundamental Review of the Trading Book: A revised Market Risk framework, BIS, Basel, Switzerland, http://www.bis.org/publ/bcbs265.pdf

Black, F. (1976), Studies of stock market volatility changes, in 1976 Proceedings of the American Statistical Association, Business & Economic Statistics Section, pp. 177-181.

Bollerslev, T. (1986), Generalised autoregressive conditional heteroskedasticity, Journal of Econometrics, 31, 307-327.

Caporin, M. and M. McAleer (2012), Model selection and testing of conditional and stochastic volatility models, in L. Bauwens, C. Hafner and S. Laurent (eds.), Handbook on Financial Engineering and Econometrics: Volatility Models and Their Applications, Wiley, New York, pp. 199-222.

Carlstein, E. (1992). Resampling Techniques for Stationary Time-Series: Some Recent Developments. New Directions in Time Series Analysis, 75–85.

Chang, C.-L., J-A. Jimenez-Martin, E. Maasoumi, and T. Pérez-Amaral (2015), A stochastic dominance approach to financial risk management strategies, Journal of Econometrics, 187(2), 472-485.

Chang, C.-L., J-A. Jimenez-Martin, M. McAleer, and T. Pérez-Amaral (2011), Risk management of risk under the Basel Accord: Forecasting value-at-risk of VIX futures, Managerial Finance, 37, 1088-1106.

Danielsson, J. (2013), The new market regulations, VoxEU.org, 28 November 2013, http://www.voxeu.org/article/new-market-risk-regulations.

Donald, S. and Y. Hsu (2013), Improving the power of tests of stochastic dominance. Mimeo. http://yuchinhsu.yolasite.com/resources/papers/powerful_SD.pdf

Dowd, K. (2005), Measuring Market Risk, Wiley, New York, 2nd edition.

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica, 50, 987-1007.

Fama, E. F. and R. Roll (1968), Some properties of symmetric stable distributions, Journal of the American Statistical Association, 63, 817-36.

Fama, E. F. and R. Roll (1971), Parameter estimates for symmetric stable distributions, Journal of the American Statistical Association, 66, 331-38.

Feller, W. (1969), An Introduction to Probability Theory and Its Applications, Volume 2, Wiley.

Franses, P.H. and D. van Dijk (1999), Nonlinear Time Series Models in Empirical Finance, Cambridge, Cambridge University Press.

Glosten, L., R. Jagannathan and D. Runkle (1992), On the relation between the expected value and volatility of nominal excess return on stocks, Journal of Finance, 46, 1779-1801.

Hammond, J.S. (1974), Simplifying the choice between uncertain prospects where preference is nonlinear, Management Science 20(7), 1047-1072.

Hanoch, G., Levy, H., 1969. The efficiency analysis of choices involving risk. Review of Economic studies, 36, 335-346.

Hansen, P.R. (2005), A test for superior predictive ability, Journal of Business and Economic Statistics, 23, 365-380.

Hershey, J.C., and Schoemaker, P.J.H. (1980), Risk taking and problem context in the domain of losses: an expected utility analysis, Journal of Risk and Insurance, 47(1), 111-132.

Levy, H. (1992), Stochastic dominance and expected utility: Survey and analysis, Management Science, 38(4), 555-593.

Li, W.K., S. Ling and M. McAleer (2002), Recent theoretical results for time series models with GARCH errors, Journal of Economic Surveys, 16, 245-269. Reprinted in M. McAleer and L. Oxley (eds.), Contributions to Financial Econometrics: Theoretical and Practical Issues, Blackwell, Oxford, 2002, pp. 9-33.

Li, C.K., Wong, W.K. (1999), A note on stochastic dominance for risk averters and risk takers. RAIRO Recherche Opérationnelle, 33, 509-524.

Ling, S. and M. McAleer (2002a), Stationariety and the existence of moments of a family of GARCH processes, Journal of Econometrics, 106, 109-117.

Ling, S. and M. McAleer (2002b), Necessary and sufficient moment conditions for the GARCH(r,s) and asymmetric power GARCH(r,s) models, Econometric Theory, 18, 722-729.

Ling, S. and M. McAleer (2003a), Asymptotic theory for a vector ARMA-GARCH model, Econometric Theory, 19, 278-308.

Ling, S. and M. McAleer (2003b), On adaptive estimation in nonstationary ARMA models with GARCH errors, Annals of Statistics, 31, 642-674.

Linton, O., E. Maasoumi and Y.J. Whang (2005), Consistent testing for stochastic dominance under general sampling schemes, Review of Economic Studies, 72, 735-765.

McAleer, M. (2005), Automated inference and learning in modelling financial volatility, Econometric Theory, 21, 232-261.

McAleer, M. (2009), The Ten Commandments for optimizing value-at-risk and daily capital charges, Journal of Economic Surveys, 23, 831-849.

McAleer, M. (2014), Asymmentry and leverage in conditional volatility models, Econometrics, 2(3), 145-150.

McAleer, M., F. Chan and D. Marinova (2007), An econometric analysis of asymmetric volatility: theory and application to patents, Journal of Econometrics, 139, 259-284.

McAleer, M. and C. Hafner (2014), A one line derivation of EGARCH, Econometrics, 2(2), 92-97.

McAleer, M., J.-A. Jimenez-Martin and T. Pérez-Amaral (2010), A decision rule to minimize daily capital charges in forecasting value-at-risk, Journal of Forecasting, 29, 617-634.

McAleer, M., J.-A. Jimenez-Martin and T. Pérez-Amaral (2013a), Has the Basel II Accord improved risk management during the global financial crisis?, North American Journal of Economics and Finance,26, 250-256.

McAleer, M., J.-A. Jimenez-Martin and T. Pérez-Amaral (2013b), GFC-robust risk management strategies under the Basel, International Review of Economics and Finance, 27, 97-111.

McAleer, M., J.-A. Jimenez-Martin and T. Pérez-Amaral (2013c), International evidence on GFC-robust forecasts for risk management under the Basel Accord, Journal of Forecasting, 32, 267–288.

Nelson, D.B. (1991), Conditional heteroskedasticity in asset returns: a new approach, Econometrica, 59, 347-370.

Politis, D.N. and J.P. Romano (1992), A Circular Block-resampling Procedure for Stationary Data,in R. Lepage and L. Billard (eds.), Exploring the Limits of Bootstrap, Wiley, New York, 263-270.

Quirk, J.P., and Saposnik, R. (1962), Admissibility and measurable utility functions. Review of Economic Studies 29, 140-146.

Sheppard, K. (2013), MFE Toolbox, https://www.kevinsheppard.com/MFE_Toolbox, 7 June, 2013.

Shiryaev, A.N. (1999), Essentials of Stochastic Finance, Facts, Models, Theory, World Scientific.

Stahl, G. (1997), Three cheers, Risk, 10, 67-69.

Stoyan, D. (1983), Comparison methods for queues and other stochastic models, New York, Wiley.

Tsay, R.S. (1987), Conditional heteroskedastic time series models, Journal of the American Statistical Association, 82, 590-604.

von Neumann, J. and Morgenstern, 0., Theory of Games and Economic Behaviour, Princeton University Press (2nd ed.), 1947.

Weeks, J. K. (1985), Stochastic dominance: A methodological approach to enhancing the conceptual foundations of operations management theory, Academy of Management Review, 10(1), 31-38.