ARTICLE IN PRESS

Journal of Econometrics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

A stochastic dominance approach to financial risk management strategies☆

Chia-Lin Chang [a,b], Juan-Ángel Jiménez-Martín [c], Esfandiar Maasoumi [d,*], Teodosio Pérez-Amaral [c]

[a] Department of Applied Economics, National Chung Hsing University, Taichung, Taiwan
[b] Department of Finance, National Chung Hsing University, Taichung, Taiwan
[c] Department of Quantitative Economics, Complutense University of Madrid, Spain
[d] Department of Economics, Emory University, USA

ARTICLE INFO

Article history:
Available online xxxx

JEL classification:
G32
G11
G17
C53
C22

Keywords:
Stochastic dominance
Value-at-Risk
Daily capital charges
Violation penalties
Optimizing strategy
Basel III Accord
VIX futures
Global financial crisis

ABSTRACT

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 risk models to forecast Value-at-Risk (VaR). The risk estimates from these models are used to determine the daily capital charges (DCC) and associated capital costs of ADIs, depending in part on the number of previous violations, whereby realized losses exceed the estimated VaR. In this paper we define risk management in terms of choosing sensibly from a variety of risk models and discuss the optimal selection of the risk models. Previous approaches to model selection for predicting VaR proposed combining alternative risk models and ranking such models on the basis of *average* DCC, or other quantiles of its distribution. These methods are based on the first moment, or specific quantiles of the DCC distribution, and supported by restrictive evaluation functions. In this paper, we consider robust uniform rankings of models over large classes of loss functions that may reflect different weights and concerns over different intervals of the distribution of losses and DCC. The uniform rankings are based on recently developed statistical tests of stochastic dominance (SD). The SD tests are illustrated using the prices and returns of VIX futures. The empirical findings show that the tests of SD can rank different pairs of models to a statistical degree of confidence, and that the alternative (recentered) SD tests are in general agreement.

1. Introduction

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). The risk estimates from these models are used to determine the daily capital charges (DCC) and associated capital costs of ADIs, depending in part on the number of previous violations, whereby realized losses exceed the estimated VaR (for further details see, for example, Chang et al., 2011).

In 1993 the Chicago Board Options Exchange (CBOE) introduced a volatility index, VIX (Whaley, 1993), which was originally designed to measure the market expectation of 30-day volatility implied by at-the-money S&P100 option prices. In 2003, together with Goldman Sachs, CBOE updated VIX to reflect a new way of measuring expected volatility, one that continues to be widely used by financial theorists.

The new VIX is based on the S&P500 Index, and estimates expected volatility by averaging the weighted prices of S&P500 puts and calls over a wide range of strike prices. Although many market participants considered the index to be a good predictor of short-term volatility, namely daily or intraday, it took several years for the market to introduce volatility products, starting with over-the-counter products, such as variance swaps and other financial derivatives. The first exchange-traded product, VIX futures, was

☆ The authors are most grateful to Michael McAleer for numerous insightful comments and guidance, and to two reviewers for very helpful comments and suggestions. For financial support, the first author wishes to thank the National Science Council, Taiwan, and the second and fourth authors acknowledge the Ministerio de Economía y Competitividad and Comunidad de Madrid, Spain.

* Corresponding author.
E-mail address: esfandiar.maasoumi@emory.edu (E. Maasoumi).

http://dx.doi.org/10.1016/j.jeconom.2015.02.032
0304-4076/

introduced in March 2004, and was followed by VIX options in February 2006. Both of these volatility derivatives are based on the VIX index as the underlying asset.

McAleer et al. (2013a,b,c) analyze, from a practical perspective, how the new market risk management strategies performed during the 2008–09 global financial crisis (GFC), and evaluate how the GFC affected the best risk management practices. These papers define risk management in terms of choosing appropriate financial targets, from a variety of financial risk models, and discuss the selection of optimal risk models. They forecast VaR using ten univariate conditional volatility models with different error distributions. Additionally, they analyze twelve new strategies based on combinations of the previous standard univariate model forecasts of VaR, namely: infimum (0th percentile), supremum (100th percentile), average, median and nine additional strategies based on the 10th through to the 90th percentiles. Such an approach is intended to select a robust VaR forecast, irrespective of the time period, that provides reasonable daily capital charges and number of violation penalties under the Basel II Accord. They found that the median is a GFC-robust strategy, in the sense that maintaining the same risk management strategy before, during and after the GFC leads to comparatively low daily capital charges and violation penalties under the Basel II Accord. Chang et al. (2011) apply a similar methodology for choosing the best strategy to forecast VaR for a portfolio based on VIX futures.

These prior methods focus on the first moment, or certain quantiles of the DCC distribution. Alternative criteria may consider mean–variance trade-offs, as in substantial areas of financial research, or general evaluation criteria that incorporate higher moments and quantiles of the underlying probability distributions. These will all provide appropriate "cardinal" and "complete" rankings of models and strategies, based on subjective valuations of different aspects, or parts of the DCC distribution. For instance, were DCC to be a Gaussian variate, mean–variance assessments would be strongly justified. This is not likely, however, and consensus on appropriate weighting and assessment functionals of non-Gaussian distributions has been elusive. It is of some importance to point out, that mean–variance type assessments are justified by a joint consideration of quadratic risk function, as well as the full characterization of the Gaussian case by the second moment. Absent a Gaussian setting, justification of a quadratic loss function itself becomes questionable. Why would we not be concerned with higher moments (when they exist), and often asymmetrical tail area behavior, especially when tail functions such as VaR are of interest?

A complementary robust alternative, is to seek weak uniform rankings over entire classes of evaluation functions, and based on nonparametric distributions of DCC. In this respect, Stochastic Dominance (SD) tests have been developed to test for statistically significant rankings of prospects. Assuming that F and G are the distribution functions of DCC produced by model 1 and model 2, respectively, model 1 first-order SD model 2, over the support of DCC, iff $F(\text{DCC}) \leq G(\text{DCC})$, with strict inequality over some values of DCC. This means that the model that produces G is dominant over all merely increasing evaluation functions since, at all quantiles, the probability that capital charges are smaller under G is greater than under F. In particular, the distribution F will have a higher median DCC than G. Similarly, each and every (quantile) percentile of the F distribution will be at a higher DCC level than the corresponding percentile of the G distribution. Consequently, model 2 will be preferred to model 1, to a statistical degree of confidence, on the basis of lower capital charges. Higher-order SD rankings reference further subclasses of evaluation functions, those that are increasing and concave, reflecting increasing risk aversion (see Sections 4–5 below).

In this paper we examine several standard models for forecasting VaRs, including GARCH, EGARCH, and, GJR, paired with Gaussian and Student-t distributions. The results show that the Gaussian distribution is preferred to Student-t in forecasting DCC. With the Student-t distribution, the EGARCH model provides a greater likelihood of higher DCC in comparison with GARCH and GJR. Using the Gaussian distribution to forecast DCC does not lead to either first- or second-order stochastic dominance. In respect of the CDF and integrated CDF, the basis of first- and second-order stochastic dominance testing, it seems that the higher expected DCC of GJR or GARCH may be compensated by lower risk compared with EGARCH.

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.

The remainder of the paper is organized as follows. Section 2 describes briefly the Basel II Accord for computing daily capital charges. Section 3 presents alternative GARCH models to produce daily capital charges. In Section 4 the definition, notation and properties of stochastic dominance are presented. Section 5 introduces the data, describes the block bootstrapping method to simulate time series, and illustrates the application of stochastic dominance to enhance financial risk management strategies of banks. Section 6 presents the main results. Section 7 gives some concluding comments.

2. Forecasting value-at-risk and daily capital charges

In this section, which follows McAleer et al. (2013a,b,c) closely, we introduce the calculation of daily capital charges (DCC) as a basic criterion for choosing between risk models. The Basel II Accord stipulates that daily capital charges (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:

$$\text{DCC}_t = \sup\left\{-(3+k)\,\overline{\text{VaR}}_{60},\ -\text{VaR}_{t-1}\right\} \tag{1}$$

where

$\text{DCC}_t =$ daily capital charges,
$\underline{\text{VaR}}_t = \hat{Y}_t - z_t \cdot \hat{\sigma}_t$, the Value-at-Risk for day t,
$\overline{\text{VaR}}_{60} =$ mean VaR over the previous 60 working days,
$\hat{Y}_t =$ estimated return at time t,
$z_t =$ 1% critical value of the distribution of returns at time t,
$\hat{\sigma}_t =$ estimated risk (or square root of volatility) at time t,
$0 \leq k \leq 1$ is the Basel II violation penalty (see Table 1).

It is well known that the formula given in Eq. (1) 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 (1As 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

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

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; McAleer et al., 2010).

As has been stated elsewhere, in calculating the number of violations, 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.

VaR refers to the lower bound of a confidence interval for a (conditional) mean, that is, a "worst case scenario on a typical day". If interest lies in modeling the random variable, Y_t, it could be decomposed as follows:

$$Y_t = E(Y_t|F_{t-1}) + \varepsilon_t. \tag{2}$$

This decomposition states that Y_t comprises a predictable component, $E(Y_t|F_{t-1})$, which is the conditional mean, and a random component, ε_t. The variability of Y_t, and hence its distribution, is determined by the variability of ε_t. If it is assumed that ε_t follows a conditional distribution, $\varepsilon_t \sim D(\mu_t, \sigma_t^2)$, where μ_t and σ_t are the time-varying conditional mean and standard deviation of ε_t, respectively, these can be estimated using a variety of parametric, semi-parametric or non-parametric methods.

The VaR threshold for Y_t can be calculated as:

$$VaR_t = E(Y_t|F_{t-1}) - \alpha\sigma_t, \tag{3}$$

where α is the critical value from the distribution of ε_t to obtain the appropriate confidence level. It is possible for σ_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).

Some recent empirical studies (see, for example, Berkowitz and O'Brien, 2001; Gizycki and Hereford, 1998; Pérignon et al., 2008) have indicated that financial institutions tend to overestimate their market risks in disclosures to the appropriate regulatory authorities, which can imply a costly restriction to the banks trading activity. ADIs may prefer to report high VaR numbers to avoid the possibility of regulatory intrusion. This conservative risk reporting suggests that efficiency gains may be feasible. In particular, as ADIs have effective tools for the measurement of market risk, while satisfying the qualitative requirements, ADIs could conceivably reduce daily capital charges by implementing a context-dependent market risk disclosure policy. McAleer (2009) and McAleer et al. (2010) discuss alternative approaches to optimize VaR and daily capital charges.

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

3. Models for forecasting VaR

ADIs can use internal models to determine their VaR thresholds. There are alternative 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 as 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), and Caporin and McAleer (2012).

3.1. 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:

$$y_t = \varphi_1 + \varphi_2 y_{t-1} + \varepsilon_t, \quad |\varphi_2| < 1 \tag{4}$$

for $t = 1, \ldots, n$, where the shocks to returns are given by:

$$\begin{aligned} \varepsilon_t &= \eta_t\sqrt{h_t}, \quad \eta_t \sim iid(0, 1) \\ h_t &= \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}, \end{aligned} \tag{5}$$

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).

3.2. EGARCH

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

$$\log h_t = \omega + \alpha\left|\frac{\varepsilon_{t-1}}{h_{t-1}}\right| + \gamma\frac{\varepsilon_{t-1}}{h_{t-1}} + \beta\log h_{t-1}, \quad |\beta| < 1 \tag{6}$$

where the parameters α, β and γ have different interpretations from those in the GARCH(1, 1) and GJR(1, 1) models.

EGARCH captures asymmetries differently from GJR. The parameters α and γ 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). Unlike GJR, EGARCH can accommodate leverage, depending on the restrictions imposed on the size and sign parameters, though leverage is not guaranteed.

As noted in McAleer et al. (2007), there are some important differences between EGARCH and the previous two models, as follows: (i) EGARCH is a model of the logarithm of the conditional variance, which implies that no restrictions on the parameters are required to ensure $h_t > 0$; (ii) moment conditions are required for the GARCH and GJR models as they are dependent on lagged unconditional shocks, whereas EGARCH does not require moment conditions to be established as it depends on lagged conditional shocks (or standardized residuals); (iii) Shephard (1996) observed that $|\beta| < 1$ is likely to be a sufficient condition for consistency of

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

QMLE for EGARCH(1, 1) (see also the caveats given in McAleer and Hafner, 2014); (iv) as the standardized residuals appear in Eq. (6), $|\beta| < 1$ would seem to be a sufficient condition for the existence of moments; and (v) in addition to being a sufficient condition for consistency, $|\beta| < 1$ is also likely to be sufficient for asymptotic normality of the QMLE of EGARCH(1, 1).

3.3. 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 effect of negative shocks (or downward movements in daily returns) of equal magnitude. In order to accommodate asymmetric behavior, Glosten et al. (1992) proposed a model (hereafter GJR), for which GJR(1, 1) is defined as follows:

$$h_t = \omega + (\alpha + \gamma I(\eta_{t-1}))\varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{7}$$

where $\omega > 0,\ \alpha \geq 0,\ \alpha + \gamma \geq 0,\ \beta \geq 0$ are sufficient conditions for $h_t > 0$, and $I(\eta_t)$ is an indicator variable defined by:

$$I(\eta_t) = \begin{cases} 1, & \varepsilon_t < 0 \\ 0, & \varepsilon_t \geq 0 \end{cases} \tag{8}$$

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 γ. 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).

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.

4. Stochastic dominance[1]

The objective is to evaluate each of the alternative conditional volatility models with respect to the DCC function. Observe that each model will yield different values of DCC because they will produce different VaR forecasts. The stochastic dominance concept is applied to determine which model should be used to produce the lowest DCC for a given investment, while not taking into account the number of violations since the primary purpose of the analysis is to assist risk managers in choosing among alternative models. Below we briefly describe the SD tests that are used in this paper.

[1] This section closely follows Donald and Hsu (2013), and Linton, Maasoumi, and Whang (2007).

4.1. Definitions and hypothesis formulation

Let X and Y be two random variables with cumulative distribution functions (CDF) F_X and F_Y, respectively. For first-order stochastic dominance (SD1), Y SD1 X, if $F_Y(z) \leq F_X(z)$ for all $z \in R$. Let $W_U(F)$ denote an evaluation function of the form $W_U(F) = \int U(z)dF(z)$, where F *is* the distribution of an underlying variable, and U is any "utility" function. SD1 is defined over monotonically increasing utility functions, that is, $W_U(F_Y) \geq W_U(F_X)$ for all $U(z)$ such that $U'(z) \geq 0$.

The technical assumptions for the underlying statistical theory include the following:

Assumption 4.1. 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 $z = \bar{z}$.

(see Linton et al., 2005 (hereafter LMW), Linton et al., 2010 and Donald and Hsu, 2013 for further details).

In order to test if Y SD1 X, Donald and Hsu (2013) formulate their hypotheses as:

$$H_0 : F_Y(z) \leq F_X(z) \quad \text{for all } z \in Z, \tag{9}$$

$$H_1 : F_Y(z) > F_X(z) \quad \text{for some } z \in Z. \tag{10}$$

This is different from LMW, who provide a simultaneous test of either Y SD1 X or X SD1 Y.

Assumption 4.2. 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. Some authors deal with independent samples and observations. LMW allow dependent time series and possibly dependent X and Y.
2. M is a function of N satisfying that $M(N) \to \infty$ and $N/(N + M(N)) \to \lambda \sim (0, 1)$ when $N \to \infty$.

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

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

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

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

Let Ψ_{h_2} denote a zero mean Gaussian process with covariance kernel equal to $h_2 \in H_2$, where H_2 denotes the collection of all covariance kernels on $Z \times Z$. For F_X and F_Y satisfying Assumption 4.1, and let $h_2^{X,Y}$ denote the covariance kernel on $Z \times Z$ such that:

$$h_2^{X,Y}(z_1, z_2) = \lambda \cdot F_X(z_1)(1 - F_X(z_2)) + (1 - \lambda) \cdot F_Y(z_1)(1 - F_Y(z_2)) \quad \text{for } z_1 \leq z_2 \tag{11}$$

with λ *defined* in Assumption 4.2. Then,

$$\sqrt{NM/(N+M)}(\hat{F}_{Y,M}(z) - \hat{F}_{X,N}(z) - (F_Y(z) - F_X(z))) \Rightarrow \Psi_{h_2 X,Y}$$

(see, for example, Linton et al., 2005, or Donald and Hsu, 2013). A typical limiting result is:

1. *Under* H_0, $\hat{S}_N \xrightarrow{D} \sup_{z \in Z} \Psi_{h_2 X,Y}$

2. *Under* H_1, $\hat{S}_N \xrightarrow{D} \infty$.

Several approaches for resampling and subsampling implementation of SD tests have been proposed. Some methods simulate the limiting Gaussian process, in the spirit of Barrett and Donald (2003), using the Multiplier method, bootstrap with separate samples, or bootstrap with combined samples. Simulated processes weakly converge to the same process as the limit process, conditional on the sample path with probability approaching 1.

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

4.2. Re-centering functions

These simulation methods do not work when the data are weakly dependent, as for time series samples in this paper. In these cases one has to appeal to either the subsampling technique of Linton et al. (2005), or a variant of the block bootstrap. Donald and Hsu (2013) provide a comparative examination of these alternatives.

Donald and Hsu (2013) and LMW (2005) propose re-centering methods introduced by Hansen (2005) to construct critical values for the 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). The re-centering function proposed by Donald and Hsu (2013) is conditioned to apply on the null space:

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

For $\alpha \leq 1/2$, let

$$\hat{c}^{bb}_{\eta,N} = \max\{\tilde{c}^{bb}_N, \eta\},$$
$$\tilde{c}^{bb}_N = \sup\left\{c | P^u\left(\sup_{z\in Z}\sqrt{N}\left(\hat{D}^{bb}_N(z) + \hat{\mu}N(z)\right) \leq c\right) \leq 1-\alpha\right\}$$

where $\hat{D}^{bb}_N(z)$ is the "S" statistic computed with the bth resample/subsample block. If the decision rule is to reject the null hypothesis, $H_0 : F_Y(z) \leq F_X(z)$ for all $z \sim Z$ when $\hat{S}_{N>} \succ \hat{c}^{bb}_{\eta,N}$, then the corresponding test has the same size properties as in the independent random samples case.

4.3. More on weakly dependent data

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 $H_0 : F_Y(z) \leq F_X(z)$ for all $z \sim Z$, $\hat{S}_N \xrightarrow{D} \sup_{z\in Z}\Psi_{h_2}(z)$, where

$$h_2(z_1, z_2) = \lim_{N\to\infty} Cov\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N (1(Y_i \leq z_1) - 1(X_i \leq z_1) - F_Y(z_1) + F_X(z_1)), \frac{1}{\sqrt{N}}\sum_{i=1}^N (1(Y_i \leq z_2) - 1(X_i \leq z_2) - F_Y(z_2) + F_X(z_2))\right) \quad (12)$$

which is the long-run covariance kernel function. In order to simulate Ψ_{h_2}, Donald and Hsu (2013) propose the blockwise bootstrap as in LMW (2005) because the multiplier method and the bootstrap methods do not account for the weak dependence of the data. Then under the same conditions as in LMW, we have $\sqrt{N}D^{bb}_N(\cdot) \xRightarrow{P} \Psi_{h_2}(\cdot)$ where h_2 is defined in (12).

We adopt the Donald and Hsu recentering procedure and critical values as they are less conservative under the null and at least as powerful under the alternative. To appreciate the role played by the choice of critical values here, consider the critical value in either of the multiplier method (mp), bootstrap with separate samples (bs), or bootstrap with combined samples (bc), given below for $k = mp$, bs and bc is as follows:

$$\hat{q}^k_N = \sup\left\{q | P^u\left(\sqrt{\frac{NM}{N+M}}\sup_{z\in Z}\hat{D}^k_N(z) \leq q\right) \leq 1-\alpha\right\}. \quad (13)$$

The critical value $\hat{q}^k_N$ is bounded away from zero in probability. Since η can be chosen to be arbitrarily small, we can assume that $\eta \prec \hat{q}^k_N$, which implies that $\hat{c}^k_{\eta,N} = \max\{\tilde{c}^k_{,N}, \eta\} \leq \hat{q}^k_N$ given that $\tilde{c}^k_{\eta,N} \leq \hat{q}^k_N$. Thus, Donald and Hsu (2013) are able to show that, given Assumptions 4.1 and 4.2, and $\alpha \leq 1/2$:

$$P(\hat{S} > \hat{q}^k_N) \leq P(\hat{S} > \hat{c}^k_{\eta,N}), \quad \text{for } k = mp, bs \text{ and } bc.$$

This would imply that the power and size of these tests are never smaller than those of BD (2003).

4.4. Linton, Maasoumi and Whang's subsampling test

LMW (2005) estimate the critical value by the subsampling method proposed in Politis and Romano (1992). 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 $\{X_1, \ldots, X_N\}$:

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

A random draw denoted by $\{X^b_1, \ldots, X^b_s\}$ from X_s would be a random sample of size s without replacement from the original data. Let $\hat{F}^b_{X,s}$ be the empirical CDF based on the random draw, $\{X^b_1, \ldots, X^b_s\}$. Define $\hat{F}^b_{Y,s}$ similarly. Let s_N and s_M denote the subsampling sizes for the X and Y samples, respectively. The subsampling critical value $\hat{c}^s_N$ is given by

$$\hat{c}^s_N = \sup\left\{c | P^u\left(\sqrt{\frac{s_N s_M}{s_N + s_M}}\sup_{z\in Z}(\hat{F}^b_{Y,s_M}(z) - \hat{F}^b_{Y,s_N}(z)) \leq c\right) \leq 1-\alpha\right\}.$$

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.

The first part is standard for the subsampling method. The second part requires that the subsample sizes from the two samples grow at the same rate and that the limit of the ratio of the subsample sizes be the same as that of the original samples. This condition is important if, for example, $s_N/(s_N + s_M) \to \lambda_s \neq \lambda$, so that, under the null hypothesis:

$$\sqrt{\frac{s_N s_M}{s_N + s_M}}\sup_{z\in Z^*}(\hat{G}^b_{s_M}(z) - \hat{F}^b_{Y;s_N}(z)) \xrightarrow{D} \sup_{z\in Z^*}\sqrt{\lambda_s}B_G \circ G(z) - \sqrt{1-\lambda_s}B_F \circ F(z) \quad (14)$$

conditional on the sample(s) with probability one (denoted as D→p). When $\lambda s \neq \lambda$, in the limit, the left-hand side of (14) may not be a good approximation to the limiting null distribution of the original test statistic.

The limiting distribution theory in LMW (2005) covers weakly stationary, dependent samples, with certain mixing conditions, such as in our application. In addition, they allow for the prospects that are ranked to be estimated functions, rather than original series described above. If the estimators involved in these estimated functions permit certain expansions, as described in LMW (2005), Assumption 2, Section 3.1, the limiting distribution theory is preserved with re-centering. The DCC functions being ranked here are certainly estimated functions of the data. The

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

conditional means and variance estimators in the VaR forecasts are generally consistent and admit the type of expansions assumed in LMW (2005). However, DCC is a maximal function of two functions of past VaRs and, to the extent that this may lead to discontinuities, or non-smoothness, in the estimated empirical CDFs, the accuracy of the underlying limiting distributions may be affected. It is known that the subsampling method is valid for many non-standard cases of this kind, but the same may not be true of bootstrap methods. A more careful technical examination of this issue is beyond the scope of this paper.

5. Data and implementation of tests

5.1. Data description

The data used for estimation and forecasting are closing daily prices (settlement prices) for the 30-day maturity CBOE VIX volatility index futures (ticker name VX). They were obtained from the Thomson Reuters-Data Stream Database for the period 26 March 2006–29 November 2013 (2526 observations). The settlement price is calculated by the CBOE as the average of the closing bid and ask quote so as to reduce the noise due to any microstructure effects. The contracts are cash settled on the Wednesday 30 days prior to the third Friday on the calendar month immediately following the month in which the contract expires. The underlying asset is the VIX index that was originally introduced by Whaley (1993) as an index of implied volatility on the S&P100. In 2003 the new VIX based on the S&P500 index was introduced.

VIX is a measure of the implied volatility of 30-day S&P500 options. It is independent of an option pricing model and is calculated from the prices of the front month and next-to-front month S&P500 at-the-money and out-the-money call and put options. The level of VIX represents a measure of the implied volatilities of the entire smile for a constant 30-day to maturity option chain. VIX is quoted in percentage points (for example, 30.0 VIX represents an implied volatility of 30.0%). In order to invest in VIX, an investor can take a position in VIX futures or VIX options.

Although VIX represents a measure of the expected volatility of the S&P500 over the next 30-days, the prices of VIX futures are based on the current expectation of what the expected 30-day volatility will be at a particular time in the future (on the expiration date). Although the VIX futures should converge to the spot at expiration, it is possible to have significant disparities between the spot VIX and VIX futures prior to expiration. Fig. 1 shows the daily VIX futures index together with the 30-day maturity VIX futures closing prices. VIX has a correlation (0.96) with the 30-day maturity VIX futures. VIX futures prices tend to show significantly lower volatility than VIX, which can be explained by the fact that VIX futures must be priced in a manner that reflects the mean reverting nature of VIX. For the whole sample, the standard deviation is 9.99 for VIX and 8.58 for VIX futures prices.

If P_t denotes the closing prices of the VIX futures contract at time t, the returns at time t (R_t) are defined as:

$$R_t = 100^* \log\left(P_t/P_{t-1}\right). \tag{15}$$

Fig. 2 shows the daily VIX futures returns, and the descriptive statistics for the daily returns are given in Table 2. The total number of observations is 2526 so that, after computing the VIX futures returns, there are 2525 observations. The histogram in Table 2 includes all the returns observations.

The returns to the VIX futures are driven by changes in expectations of implied volatility. Fig. 3 shows the histograms for the daily returns, together with the theoretical Gaussian and Student-t probability density functions and a kernel density estimator. The Student-t density fits the returns distributions better than does its Gaussian counterpart.

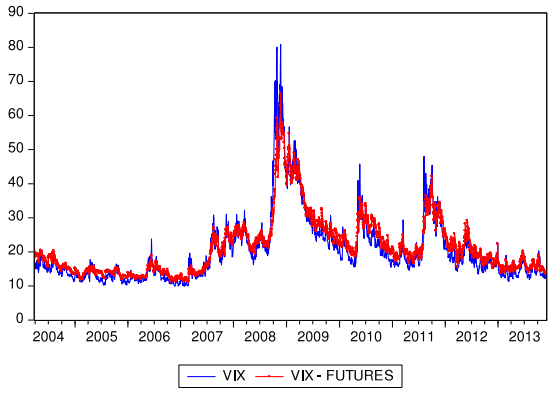


Fig. 1. VIX and 30-day maturity VIX futures closing prices 26 march 2004–29 November 2013.

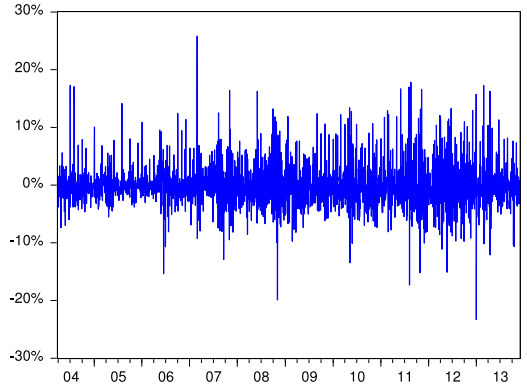


Fig. 2. 30-day maturity VIX futures returns 26 march 2004–29 November 2013.

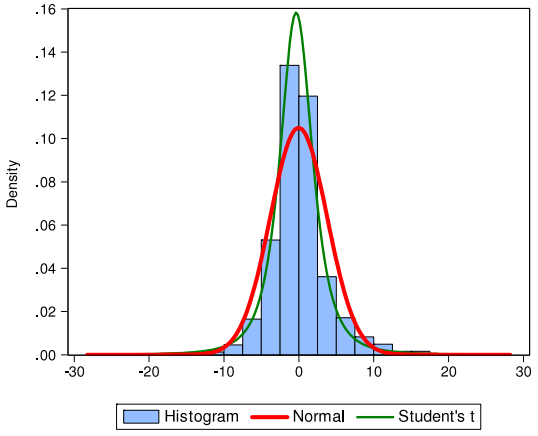


Fig. 3. Histogram, normal and Student-t distributions 30-day maturity VIX futures returns 26 march 2006–29 November 2013.

Regarding the returns volatility, several measures of volatility are available in the literature. In order to gain some intuition, we adopt the measure proposed in Franses and van Dijk (1999), who define the true volatility of returns as:

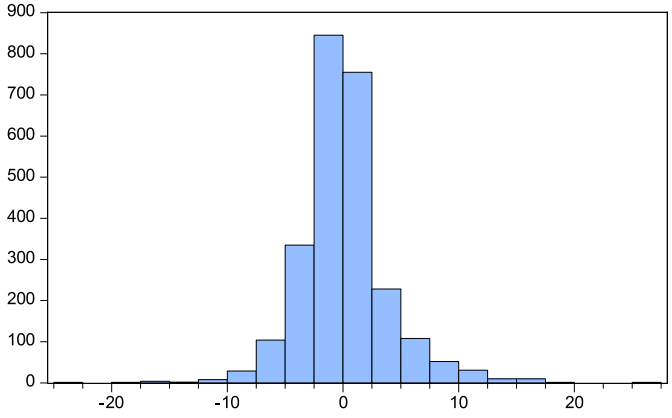
$$V_t = \left[\left(R_t - E\left(R_t|F_{t-1}\right)\right)^2\right]^{1/2}, \tag{16}$$

where F_{t-1} is the information set at time $t-1$.

Fig. 4 presents V_t in Eq. (16) as "volatilities". The series exhibit clustering that should be captured by an appropriate time series model. Until January 2007, a month before the first reports of subprime losses, the volatility of the series seems to be stable. The volatility reached an all time peak on 27 February 2007, when it climbed to 26% (the mean (median) for the entire sample is

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

Table 2
30-day maturity VIX futures returns histogram and descriptive statistics for 26 march 2004–29 November 2013.



Series: RETURNS	
Sample 26/03/2004 29/11/2013	
Observations 2525	
Mean	-0.014896
Median	-0.273598
Maximum	25.81866
Minimum	-23.32617
Std. Dev.	3.800655
Skewness	0.697327
Kurtosis	7.350578
Jarque-Bera	2195.970
Probability	0.000000

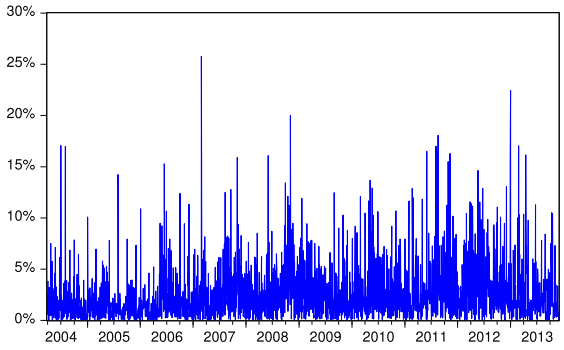


Fig. 4. Volatility of 30-day maturity VIX futures returns 26 march 2004–29 November 2013.

2.62(1.78)), as the US equity market had its worst day in four years. Then it remained above historic levels, but the VIX futures volatility increases again after August 2008, due in large part to the worsening global credit environment, with a maximum again on 3 November 2008. Then the volatility remained low until the news about the sovereign debt crisis in the Euro zone created another spike in volatility in the first week of May 2010, when the VIX futures reached 35 with a high volatility in returns. Finally, at the end of September 2011 we observe another maximum of 42 because of the 2011 US debt ceiling episode.

5.2. Block bootstrapping and subsampling

In order to test for SD rankings between risk models using different conditional volatility models for forecasting VaR, we implement the Circular Block Bootstrapping (CBB) method developed in Politis and Romano (1992) for resampling the VIX futures through the MFE toolbox of Kevin Sheppard. 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 $\{X_i : i = 1, \ldots, n\}$, and let $l \in \{1, \ldots, n\}$ 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:

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

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

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

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 $i_n = i \mod n$ 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 $X_1, \ldots, X_n$ 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 DCC (VaR) is estimated for each drawn sample, thus generating the bootstrap (subsample) distribution of the test statistics.

5.3. Daily capital charges (DCC) and evaluation framework: stochastic dominance

The primary objective is to evaluate each of the alternative conditional volatility models with respect to the DCC function. Each model will imply different values of DCC as they will produce different VaR forecasts. The stochastic dominance concept is applied in this context to determine which model should be used for producing a lower DCC for a given amount invested.[2] The main point of the analysis is to help risk managers to choose between alternative models.

In analyzing alternative risk management strategies, McAleer et al. (2013b) forecast VaR using ten univariate conditional volatility models with different error distributions. Additionally, they analyze twelve new strategies based on combinations of the previous standard univariate forecasts of VaR, namely: infimum (0th percentile), supremum (100th percentile), average, median and nine additional strategies based on the 10th through to the 90th percentiles. This was intended to select a robust VaR forecast, irrespective of the time period, that provides reasonable daily capital charges and number of violation penalties under the Basel Accord. They found that the median is a GFC-robust strategy, in the sense that maintaining the same risk management strategy before, during and after the GFC leads to comparatively low daily capital charges and violation penalties under the Basel Accord. Chang et al.

[2] Subsequent analysis would take into account if each model not only provides lower DCCs, but also if the model satisfies the requirement of keeping the number of violations away from the red zone (see Table 1). Looking at Table 3 in Chang et al. (2011), the number of violations for all the strategies and periods analyzed in that paper remains below 5. In this paper we use the same asset and models, which means that all the models remain within the Basel II green zone, thereby avoiding the possibility of regulatory intrusion.

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

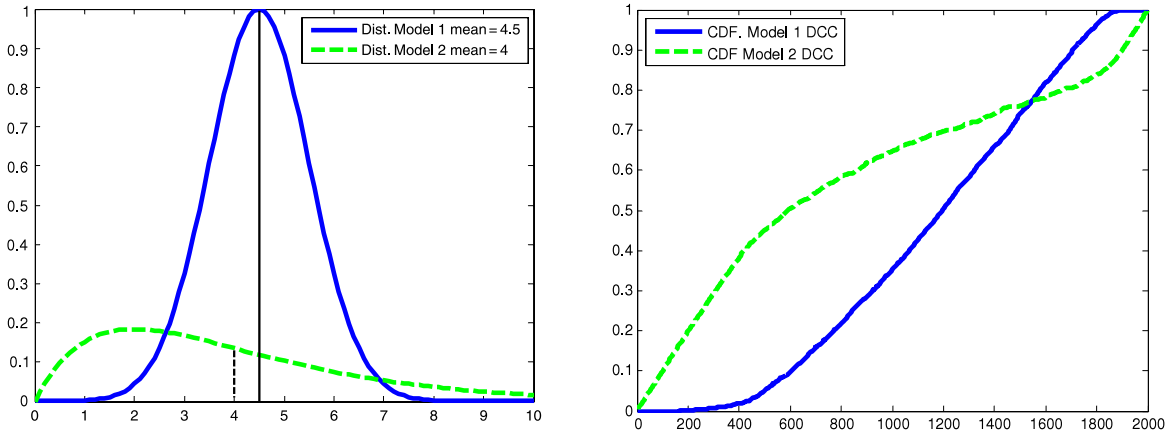


Fig. 5. Probability density functions and CDFs of DCC models. Note: in the left panel, the solid line is for model 1, a Gaussian density with mean 4.5 and standard deviation 1. The dashed line for model 2 depicts the density of a Chi-square with 4 degrees of freedom. The right panel depicts the CDFs for models 1 (Gaussian) and 2 (Chi-square).

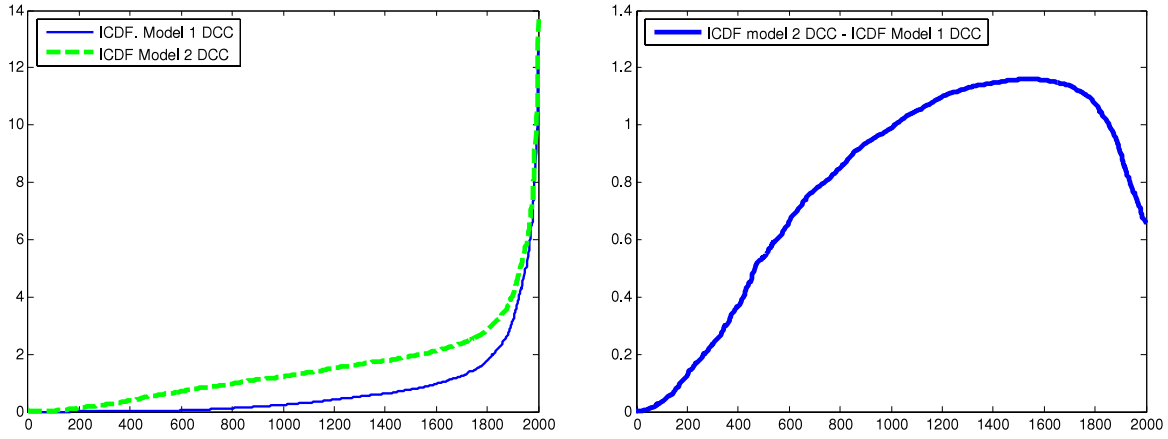


Fig. 6. DCC and ICDF for models 1 and 2 and their difference. Note: in the left panel, the solid line represents the integrated cumulative distribution function (ICDF) for a Gaussian variable with mean 4.5 and standard deviation 1. The dashed line for model 2 depicts the ICDF of a 4-degrees of freedom Chi-square variable. The right panel depicts the difference between the ICDFs given in the left panel.

(2011) apply this model selection criterion for choosing the best risk management strategy when dealing with VIX futures.

For each criterion above there corresponds an implied evaluation, or weighting function, over the quantiles of the distribution. SD ranking may provide "uniform" evaluations based on large classes of loss functions, and/or distribution functions. To a statistical degree of confidence, SD rankings may provide for robust risk model selection and reporting. This method offers the advantage of always being consistent with expected utility maximization, while not requiring a specific loss function.

As an example, consider the probability density functions of DCC from two alternative models 1 and 2, as shown in Fig. 5. Based on an illustrative example using 1000 simulated observations, Figs. 5 and 6 depict probability density functions, empirical CDFs and ICDFs, where one distribution is Chi-squared and the other is Gaussian. The empirical CDFs, $\hat{F}_X$ and $\hat{F}_Y$, for any two variables, X and Y, that are used in the SD tests, are plotted in Figs. 5–13, and are computed as follows: (i) a vector z containing all the distinct X and Y values is created, and is sorted in the ascending order; and (ii) each point j in $\hat{F}_X(\hat{F}_Y)$ is calculated as the number of observations less than or equal to observation j in vector z, divided by the total number of observations of X (Y). The number of points on the horizontal axes in Figs. 5–6, as determined by the number of distinct values of X and Y, show the numbers of observations for $\hat{F}_X$ and $\hat{F}_Y$.

In the left panel of Fig. 5, the bold line represents model 1, which has a Gaussian distribution function with mean 4.5 and standard deviation 1. The dotted line is for model 2 and depicts the density of a Chi-square with 4 degrees of freedom. If we made the choice according to previous criteria based only on first moments, this would imply choosing model 2 as the mean DCC from model 2 is lower than that from model 1. However, in the right panel in Fig. 5, which depicts the CDF for models 1 (Gaussian) and 2 (Chi-square), we see that the DCC values of model 2 are more uncertain, with a greater probability of either small (the CDF for model 2 is initially above that for model 1) or large DCC (the CDF for model 2 lies below that of model 1 for high DCC values). As these CDF functions cross, first-order stochastic dominance cannot be established. However, given the high crossing point, second order SD is likely.

Second-order dominance is prospective, and we can test for it based on the cumulative areas under the CDFs (ICDF), as shown in the right panel of Fig. 6. As the difference in ICDF model 2-ICDF model (1) is always positive, this implies that model 1 second-order dominates model 2, and will be preferred by a risk averse risk manager, whatever their cardinal (concave) loss function. Choosing model 2 would imply a greater risk during periods of turmoil and greater uncertainty.

This is quite important since safe practice based on a specific risk assessment function, such as mean–variance, or any particular quantile, may be an artifact of the particular functional employed, and how that functional places different weights on different quantiles, as well as how different quantiles are compared. This is a question of *increasing* risk aversion, embodied by, or absent from, different loss functions.

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

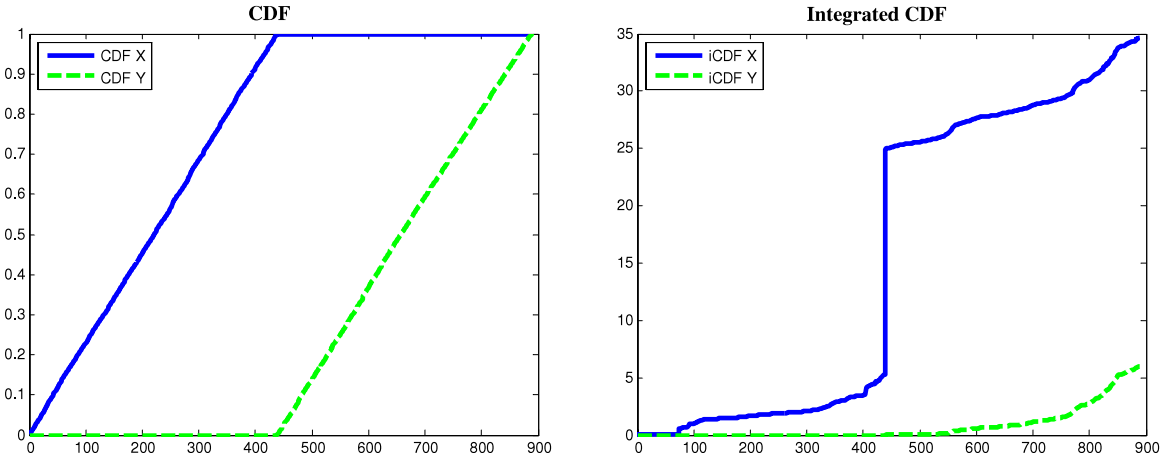


Fig. 7. Alternative A the solid line is GARCH-t and the dashed line is GARCH–Gaussian. Note: in the left panel, the solid line is the CDF of DCC produced by a GARCH model assuming a Student-t distribution of VIX futures returns. The dashed line is the CDF of a DCC produced by a GARCH model assuming a Gaussian distribution of VIX futures 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.

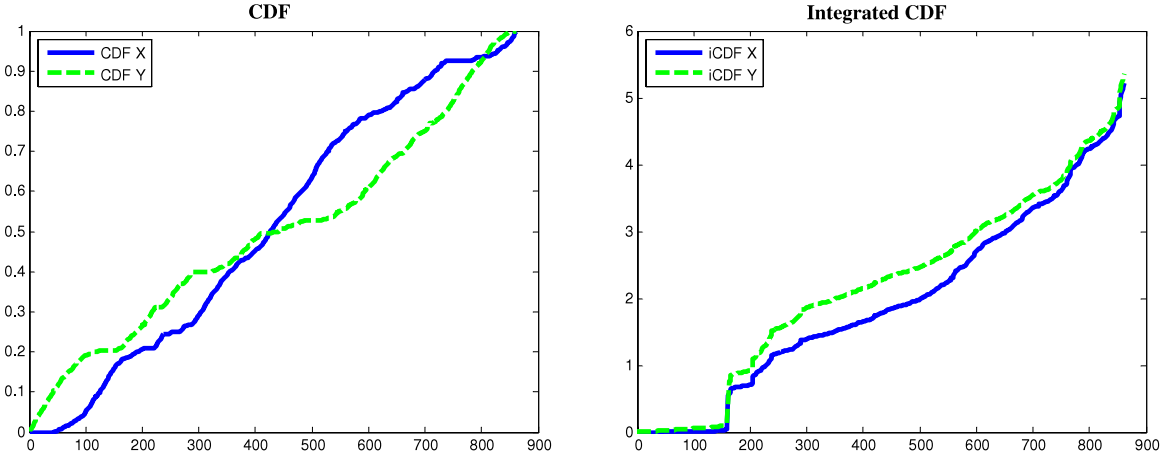


Fig. 8. Alternative D. The solid line is GARCH–Gaussian and the dashed line is EGARCH–Gaussian. Note: in the left panel, the solid line is the CDF of DCC produced by a GARCH model assuming a Gaussian distribution of VIX futures returns. The dashed line is the CDF of a DCC produced by an EGARCH model assuming a Gaussian distribution of VIX futures returns. In the right panel, the two lines depict the integrated cumulative distribution functions (ICDF) of the CDFs shown in the left panel.

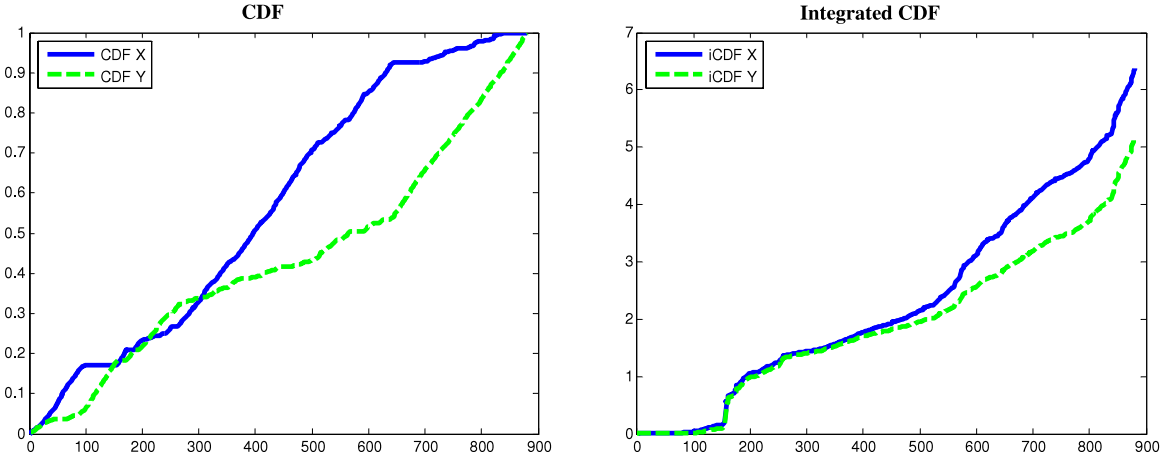


Fig. 9. Alternative E. The solid line is GARCH–Gaussian and the dashed line is GJR–Gaussian. Note: in the left panel, the solid line is the CDF of DCC produced by a GARCH model assuming a Gaussian distribution of VIX futures returns. The dashed line is the CDF of a DCC produced by a GJR model assuming a Gaussian distribution of VIX futures 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.

The way that SD may be used to choose the best risk management strategy is as follows. For notational simplicity, we write *Model* 1 $\overset{FSD}{\succeq}$ *Model* 2 and *Model* 1 $\overset{SSD}{\succeq}$ *Model* 2 whenever Model 1 dominates Model 2 according to FSD and SSD, respectively. Let Y and X be the DCC produced using model 1 and model 2, respectively. Based on the previous definition, if Y first-order stochastically dominates X, then Y will involve higher DCC than X in the sense that it has a higher probability of producing higher values of

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

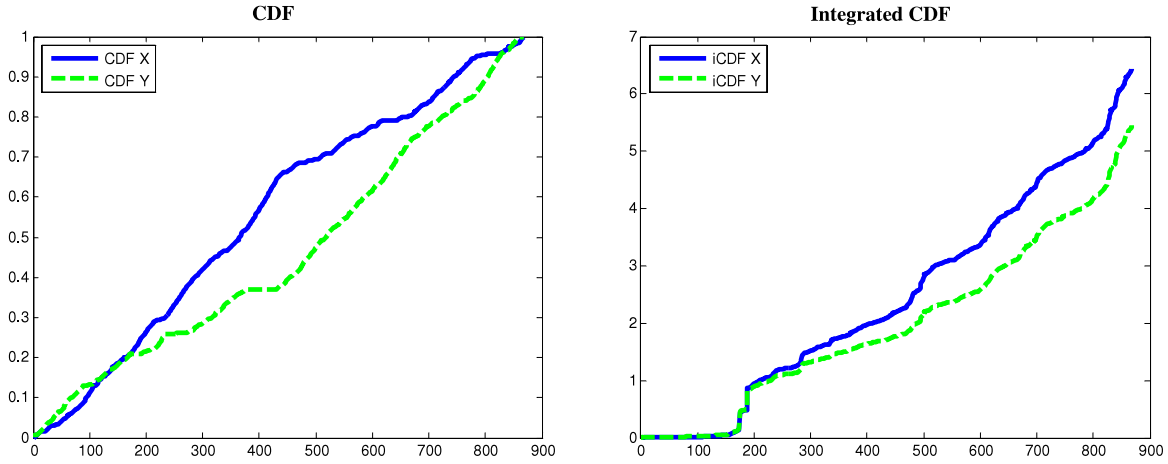


Fig. 10. Alternative F. The solid line is EGARCH–Gaussian and the dashed line is GJR–Gaussian. Note: in the left panel, the solid line is the CDF of DCC produced by an EGARCH model assuming a Gaussian distribution of VIX futures returns. The dashed line is the CDF of a DCC produced by a GJR model assuming a Gaussian distribution of VIX futures returns. In the right panel, the solid and dashed lines depict the integrated cumulative distribution functions (ICDFs) of the CDFs shown in the left panel.

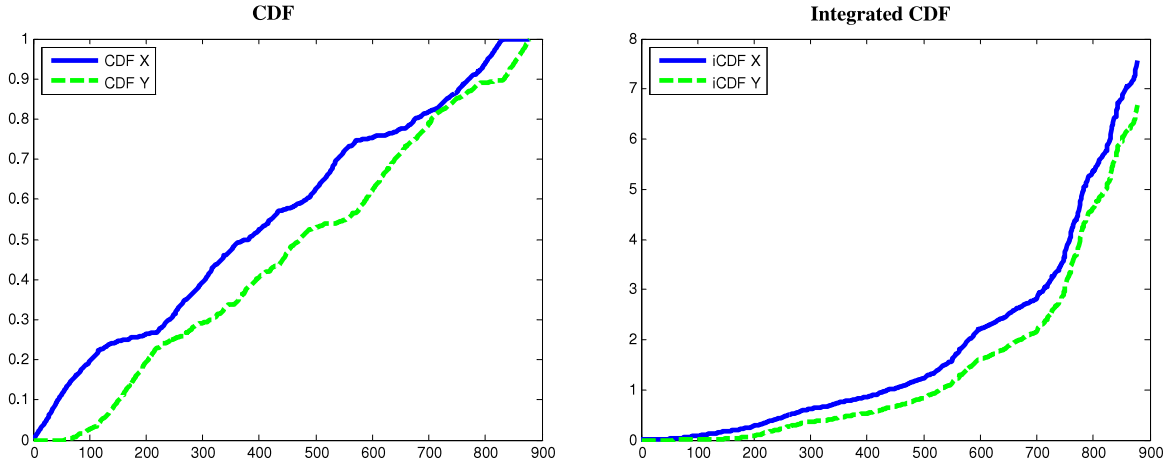


Fig. 11. Alternative G. The solid line is GARCH-*t* and the dashed line is EGARCH-*t*. Note: in the left panel, the solid line is the CDF of DCC produced by a GARCH model assuming a Student-*t* distribution of VIX futures returns. The dashed line is the CDF of DCC produced by an EGARCH model assuming a Student-*t* distribution of VIX futures 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.

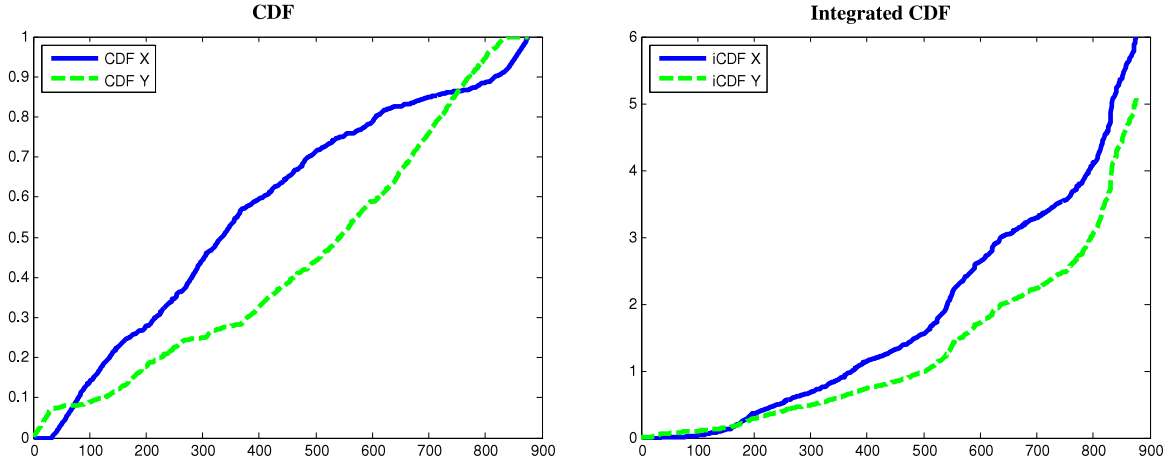


Fig. 12. Alternative H. The solid line is GARCH-*t* and the dashed line is GJR-*t*. Note: in the left panel, the solid line is the CDF of DCC produced by a GARCH model assuming a Student-*t* distribution of VIX futures returns. The dashed line is the CDF of a DCC produced by a GJR model assuming a Student-*t* distribution of VIX futures 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.

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

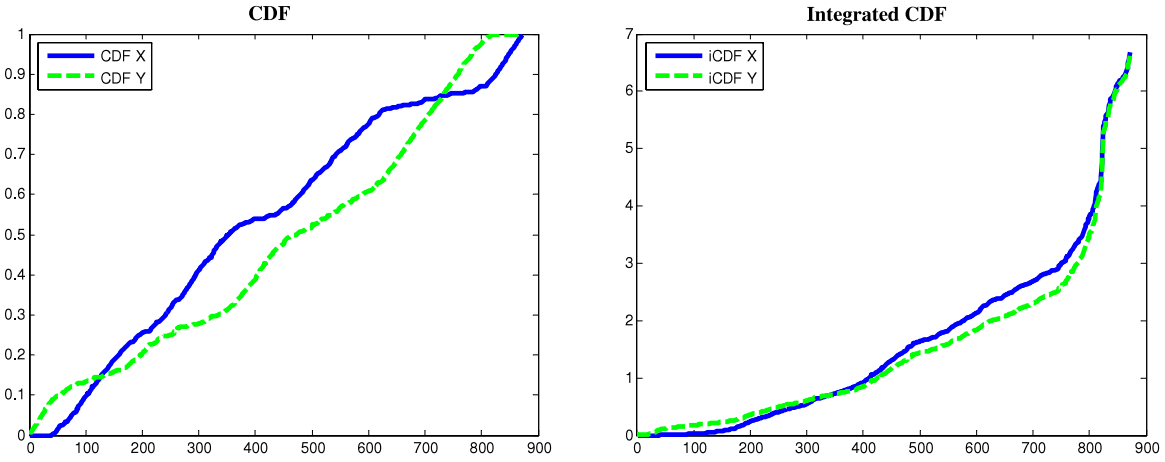


Fig. 13. Alternative I. The Blue line is EGARCH-t and the green line is GJR-t. Note: In the left panel, the solid line is the CDF of a DCC produced by an EGARCH model assuming a Student-t distribution of VIX futures returns. The dashed line is the CDF of a DCC produced by a GJR model assuming a Student-t distribution of VIX futures returns. In the right panel, the solid and dashed lines depict the integrated cumulative distribution functions (ICDFs) of the CDFs shown in the left panel.

Table 3a
Definition of pairs of alternative models to be tested for stochastic dominance.

Alternative	GARCH Gaussian	GARCH Student-t	EGARCH Gaussian	EGARCH Student-t	GJR Gaussian	GJR Student-t
A	X	Y				
B			X	Y		
C					X	Y
D	X		Y			
E	X				Y	
F			X		Y	
G		X		Y		
H		X				Y
I				X		Y
J	Y		X			
K	Y				X	
L			Y		X	
M		Y		X		
N		Y				X
O				Y		X

Notes (1): the first column denotes 14 alternatives of 30 permutations of the models considered: GARCH, EGARCH, GJR, with two different distributions, Gaussian and Student-t, of VIX futures returns. FSD and SSD are tested for each alternative using three different tests, BB, BD, and LMW for the null hypothesis: H_0: Y SD X.
(2). For example, in alternative A, X represents the empirical distribution of DCC produced using a GARCH model assuming a Gaussian distribution of the VIX futures returns. Alternative Y represents the empirical distribution of DCC produced using a GARCH model assuming a Student-t distribution of the VIX futures returns. Failing to reject H_0 would imply that DCC produced by option Y SD option X, so that, under option Y there is a higher chance of producing a higher DCC than under option Y.

DCC. Therefore, Y must be associated with a higher DCC than X if both X and Y require the same initial investment. In this context, $Model\ 1 \overset{FSD}{\succeq} Model\ 2$ would mean that the risk manager would prefer Model 2 because the probability is higher that the bank will have to set aside less money for covering losses.

In summary, the decision rule would be as follows:

Decision rule			
H_0: Y FSD X	**If do not rejectH_0: YdominatesX**	DCC of Model 1 is likely to be higher than of Model 2	**Risk manager prefers Model 2 to Model 1**
Y = DCC of Model 1 X = DCC of Model 2			

Graphically, FSD exists when the cumulative distribution functions do not intersect; if they do cross, then the FSD is statistically indeterminate. This is the case shown in Fig. 5, where the right box includes CDFs for models 1 and 2 from the previous example. As the two CDFs cross, first-order stochastic dominance cannot be established. From Fig. 5, a significant amount of information in the probability distribution would be lost examining only the first or second moments. For example, for low values of DCC, model 1 provides a higher probability of having higher values of DCC, while for high values of DCC, a higher probability would be more likely under model 2.

We can test for second-order stochastic dominance (SSD) in order to account for risk aversion, over increasing and concave utility functions. Model 1 SSD model 2 when the area under the cumulative distribution for model 1 is less than the corresponding area for model 2, at every point on the support. In Fig. 6, model 1 SDD model 2, with the graph on the right showing the difference in areas, which is always greater than zero over the entire range of possible values of DCC. This means that model 2 yields higher DCC for all but one set of circumstances. If this case is not too severe, SSD indicates that model 1 is better than model 2. This testing strategy may also be extended to consider cases in which a limited part of the support might be of particular interest.

6. Results

Table 3a describes the different alternatives analyzed in the paper. Tables 3b and 3c present rejection rates from three different

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

Table 3b
Rejection rates for first-order SD tests.

				←						→					
				Gaussian			Student-t			Gaussian			Student		
Design	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
BB(12)	0.00	0.00	0.00	0.53	0.50	0.47	0.02	0.49	0.68	0.50	0.51	0.55	0.87	0.28	0.14
BD(12)	0.00	0.00	0.00	0.45	0.43	0.42	0.01	0.45	0.67	0.42	0.46	0.47	0.87	0.25	0.11
BB(24)	0.00	0.00	0.00	0.35	0.30	0.32	0.01	0.27	0.51	0.34	0.34	0.37	0.62	0.15	0.07
BD(24)	0.00	0.00	0.00	0.33	0.28	0.29	0.00	0.25	0.50	0.30	0.31	0.34	0.62	0.14	0.05
LMW	0.00	0.00	0.00	0.28	0.21	0.26	0.02	0.25	0.52	0.28	0.27	0.28	0.18	0.13	0.05

Table 3c
Rejection rates for second-order SD tests.

				←						→					
				Gaussian			Student-t			Gaussian			Student		
Design	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
BB(12)	0.00	0.00	0.00	0.44	0.38	0.37	0.00	0.40	0.67	0.40	0.46	0.48	1.00	0.19	0.05
BD(12)	0.00	0.00	0.00	0.41	0.36	0.34	0.00	0.39	0.67	0.37	0.43	0.45	1.00	0.19	0.05
BB(24)	0.00	0.00	0.00	0.36	0.34	0.31	0.01	0.30	0.55	0.33	0.39	0.39	0.97	0.14	0.05
BD(24)	0.00	0.00	0.00	0.35	0.33	0.29	0.00	0.30	0.44	0.32	0.39	0.38	0.97	0.14	0.03
LMW	0.00	0.00	0.00	0.29	0.26	0.30	0.06	0.24	0.44	0.31	0.31	0.31	0.94	0.09	0.02

tests, namely Donald and Hsu (2013) (BB), Barret and Donald (BD) (2003), and Linton et al. (2005) (LMW) for the null hypothesis: H_0: Y SD X. For example, in alternative 1, failing to reject H_0 would imply that DCC produced by GARCH with a Student-t distribution (Y) FSD the DCC produced by GARCH with a Gaussian distribution (X). Following Donald and Hsu (2013), when implementing the blockwise bootstrap, the block sizes are set to 12 and 24. The subsample size is set at 25. 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 176 possible subsamples. The significance level is set at 5%.

(1) With 524-VaR forecasts, we omit the first 60 days for computing DCC values, which leaves a 464-observation DCC vector for computing the cumulative distribution functions that appear in Figs. 7–13. For alternatives A, B and C in Table 3b, the BB and BD tests clearly show that the DCC values, assuming the Student-t distribution, first-order dominate Gaussian distributions for all three GARCH models. As these results imply a higher likelihood of higher DCC under the Student-t distribution, they are preferred to the Gaussian distribution.

Fig. 7 shows the Cumulative Distribution Function (CDF) and the integrated CDF (ICDF) for alternative A. Regarding CDF in the left panel, dashed line, DCC produced by GARCH for the Student-t distribution lies ahead of the solid line, DCC produced by GARCH assuming the Gaussian distribution. The right panel in Fig. 7 shows that DCC produced by GARCH assuming the Student-t distribution second-order dominates DCC produced by GARCH assuming the Gaussian distribution, as the dashed line is always below the solid one. This means that GARCH assuming the Gaussian distribution would be preferred by a financial risk manager.

(2) Assuming the Gaussian distribution, alternatives D, E and F, the three tests conclude that neither EGARCH nor GJR *FSD* GARCH, and GJR does not FSD EGARCH. Figs. 8–10 show the CDFs and ICDFs for the alternatives for one of the 500 daily simulations for DCC. In each case, Y CDF (dashed line) crosses X CDF (solid line), so that we do not find FSD. Not having found FSD, we check for SSD, but again, the hypothesis is rejected in all three cases. The rejection rates for SSD are given in Table 3c. Even though we cannot establish SD of any order, it is worth examining Fig. 8 to shed light on the richness of the information provided by the DCC probability distributions. For low values of DCC, GARCH provides a higher likelihood of higher DCCs (the solid line lies above the dashed line), but this is reversed for high values of DCC. The right panel shows the dashed line (EGARCH) lies above the blue line (GARCH), and the difference between them (dashed minus solid) is always positive. This means that the DCC produced by GARCH SSD the DCC produced by EGARCH. GARCH would be preferred to EGARCH for forecasting DCC as the higher expected DCC of GARCH for low DCC values can be compensated for a lower degree of uncertainty.

A similar intuition can be drawn from Figs. 9–10 for alternative F comparing DCC probability distributions produced by GARCH (solid) and GJR (dashed), and by EGARCH (solid) and GJR (dashed), respectively. Even though these tests cannot establish either first- or second-order stochastic dominance for these particular simulations, the CDFs nevertheless cross. For example, the left panel in Fig. 10 shows that the dashed line generally lies ahead of the solid line. However, in the right panel it seems that the higher expected DCC produced by GJR is compensated by lower uncertainty (dashed line, beneath solid line in the right panel), making this strategy more desirable. Therefore, based on Figs. 8 and 10, GARCH and GJR would be preferred to EGARCH for forecasting DCC (for the Gaussian case).

(3) Assuming the Student-t distribution, EGARCH FSD GARCH (alternative G) and EGARCH SSD GJR (alternative O). Neither GJR FSD GARCH (alternative H) nor GARCH *FSD* GJR. Fig. 11 for alternative G shows that the CDF of DCC produced by EGARCH lies ahead of the solid line, CDF of DCC produced by GARCH. This implies a higher likelihood of high DCC values under EGARCH, in which case GARCH would be preferred. The left panel in Fig. 12 shows the solid CDF (GJR) underneath the blue CDF (GARCH) at low quantiles, then crosses the dashed one, stays above during some quantiles, and then returns below the dashed CDF. Accordingly, we do not find either first or second-order SD. Fig. 13 shows

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

the last case, EGARCH versus GJR using the Student-*t* distribution (alternatives I and O). Even though the tests in Table 3b show that EGARCH FSD GJR (alternative O), the selected case[3] for Fig. 13 does not provide clear evidence in favor of the hypothesis.

In summary, the Gaussian distribution is preferred to Student-*t* for forecasting DCC. EGARCH seems to provide a higher likelihood of higher DCC when compared with GARCH and GJR, when using the Student-*t* distribution. Using the Gaussian distribution for forecasting DCC does not lead to either first- or second-order stochastic dominance. However, on the basis of the CDF and integrated CDF, it seems that the higher expected DCC values of GJR or GARCH may be compensated by lower uncertainty than for EGARCH. These results lend support to the empirical findings in Table 5 of Chang et al. (2011), where it is shown that EGARCH provides the highest average DCC for all periods in comparison with GARCH and GJR. Moreover, Chang et al. (2011) also show that GJR always provides higher DCC values than GARCH. These results notwithstanding, the SD criterion does not seem to show any dominance between these two models.

7. Conclusions

In the spectrum of financial assets, VIX futures prices are a relatively new product. As with any financial asset, VIX futures are subject to risk. In this paper we analyzed the performance of a variety of strategies for managing the risk, through forecasting VaR, of VIX futures under the Basel II Accord.

The alternative strategies for forecasting VaR of VIX futures, and for managing financial risk under the Basel II Accord, are several univariate conditional volatility models, specifically GARCH, EGARCH and GJR, with each based on either the Gaussian and Student *t* distributions. The main criterion for choosing among the alternative strategies was minimizing average daily capital charges. In the paper we used a methodology based on stochastic dominance that permits partial ordering of strategies by accommodating the entire distribution of DCC values. This methodology provides a search for uniformly higher ranked volatility models, based on large classes of evaluation functions and the entire DCC distribution.

The main empirical findings of the paper are as follows:

1. The Gaussian models are generally preferred to their Student-*t* counterparts.
2. SD relations between DCC values produced by Gaussian models are generally not uniformly ranked. An analysis of CDFs and ICDF seems to show, however, that EGARCH provides DCC distributions with greater uncertainty, so the other models would be preferred. A lack of uniform rankings by SD also indicates that there exist special utility/evaluation functions that may provide complete, albeit subjective, rankings.
3. Within the class of Student-*t* distributions, EGARCH SD both GARCH and GJR, implying that EGARCH would be *uniformly* preferred to GARCH and GJR by a financial risk manager.
4. In general, a stochastic dominance criterion can be used to rank different models of VIX futures and distributions, as illustrated in the previous empirical results. Even in cases of no FSD and SSD, the tests provide additional information about the entire distribution over specific ranges.
5. The graphs of the CDFs of each pair of models allow a comparison globally for the whole distribution, and also locally for a given range of DCC values and probabilities. This allows more specific comparisons than previously afforded based on the mean and other moments of the relevant distributions.

In this paper we have not found an optimal model in the sense that it outperforms the other models during the whole sample period. On the other hand, we have restricted attention to a set of widely used, though not necessarily exhaustive set of forecasting models and distributions. This paper takes into account the number of violations as defined by the Basel Accord through the computation of DCC. Subsequent analysis would take into account if each model not only provides lower DCCs, but also satisfies the requirement of keeping the number of violations away from the red zone. For the VIX futures returns, all our models remain in the green zone.

The results of the paper suggest that further work is needed to compare, not only univariate models, but also combinations of models, such as based on the mean or median. This framework presented above should also be applied to a portfolio of assets to determine the usefulness of the stochastic dominance approach. This paper performed pairwise comparisons for a variety of models. The extension to comparisons among multivariate models is a topic for future research, as is a detailed analysis of the useful information that is contained in the CDF and ICDF.

References

Barrett, G., Donald, S., 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.
Berkowitz, J., O'Brien, J., 2001. How accurate are value-at-risk models at commercial banks?, Discussion Paper, Federal Reserve Board.
Black, F., 1976. Studies of stock market volatility changes. In: Proceedings of the American Statistical Association, Business & Economic Statistics Section, pp. 177–181.
Bollerslev, T., 1986. Generalised autoregressive conditional heteroskedasticity. J. Econometrics 31, 307–327.
Caporin, M., McAleer, M., 2012. Model selection and testing of conditional and stochastic volatility models. In: Bauwens, L., Hafner, C., Laurent, S. (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 Dir. Time Ser. Anal. 75–85.
Chang, C.-L., Jimenez-Martin, J.-A., McAleer, M., Pérez-Amaral, T., 2011. Risk management of risk under the Basel accord: forecasting value-at-risk of VIX futures. Manage. Finance 37, 1088–1106.
Donald, S., Hsu, Y., 2013. Improving the power of tests of stochastic dominance. mimeo. [details].
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.
Franses, P.H., van Dijk, D., 1999. Nonlinear Time Series Models in Empirical Finance. Cambridge University Press, Cambridge.
Gizycki, M., Hereford, N., 1998. Assessing the dispersion in banks' estimates of market risk: the results of a value-at-risk survey, Discussion Paper 1, Australian Prudential Regulation Authority.
Glosten, L., Jagannathan, R., Runkle, D., 1992. On the relation between the expected value and volatility of nominal excess return on stocks. J. Finance 46, 1779–1801.
Hansen, P.R., 2005. A test for superior predictive ability. J. Bus. Econom. Statist. 23, 365–380.
Li, W.K., Ling, S., McAleer, M., 2002. Recent theoretical results for time series models with GARCH errors. In: McAleer, M., Oxley, L. (Eds.), Contributions to Financial Econometrics: Theoretical and Practical Issues, Blackwell, Oxford, pp. 9–33. Reprinted from: Journal of Economic Surveys 16, 245-269 (2002).
Ling, S., McAleer, M., 2002a. Stationariety and the existence of moments of a family of GARCH processes. J. Econometrics 106, 109–117.
Ling, S., McAleer, M., 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., McAleer, M., 2003a. Asymptotic theory for a vector ARMA–GARCH model. Econometric Theory 19, 278–308.
Ling, S., McAleer, M., 2003b. On adaptive estimation in nonstationary ARMA models with GARCH errors. Ann. Statist. 31, 642–674.
Linton, O., Maasoumi, E., Whang, Y.-J., 2005. Consistent testing for stochastic dominance under general sampling schemes. Rev. Econom. Stud. 72, 735–765.
Linton, O., Song, K., Whang, Y.-J., 2010. An improved bootstrap test of stochastic dominance. J. Econometrics 154, 186–202.
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. J. Econ. Surv. 23, 831–849.

[3] Individual simulation number 250 of 500 simulations was chosen for illustrative purposes.

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032

McAleer, M., Chan, F., Marinova, D., 2007. An econometric analysis of asymmetric volatility: theory and application to patents. J. Econometrics 139, 259–284.
McAleer, M., Hafner, C., 2014. A one line derivation of EGARCH. Econometrics 2 (2), 92–97.
McAleer, M., Jimenez-Martin, J.-A., Pérez-Amaral, T., 2010. A decision rule to minimize daily capital charges in forecasting value-at-risk. J. Forecast. 29, 617–634.
McAleer, M., Jimenez-Martin, J.-A., Pérez-Amaral, T., 2013a. Has the Basel II Accord improved risk management during the global financial crisis? N. Am. J. Econ. Finance 26, 250–256.
McAleer, M., Jimenez-Martin, J.-A., Pérez-Amaral, T., 2013b. GFC-robust risk management strategies under the Basel. International Review of Economics and Finance 27, 97–111.
McAleer, M., Jimenez-Martin, J.-A., Pérez-Amaral, T., 2013c. International evidence on GFC-robust forecasts for risk management under the Basel Accord. J. Forcast. 32, 267–288.
Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347–370.
Pérignon, C., Deng, Z.-Y., Wang, Z.-J., 2008. Do banks overstate their value-at-risk? J. Bank. Finance 32, 783–794.
Politis, D.N., Romano, J.P., 1992. A circular block-resampling procedure for stationary data. In: Lepage, R., Billard, L. (Eds.), Exploring the Limits of Bootstrap. Wiley, New York, pp. 263–270.
Shephard, N., 1996. Statistical aspects of ARCH and stochastic volatility. In: Barndorff-Nielsen, O.E., Cox, D.R., Hinkley, D.V. (Eds.), Statistical Models in Econometrics, Finance and Other Fields. Chapman & Hall, London, pp. 1–67.
Stahl, G., 1997. Three cheers. Risk 10, 67–69.
Whaley, R.E., 1993. Derivatives on market volatility: Hedging tools long overdue. Journal of Derivatives 1, 71–84.

Further reading

Basel Committee on Banking Supervision, 1996. Supervisory Framework for the Use of Backtesting in Conjunction with the Internal Model-Based Approach to Market Risk Capital Requirements. BIS, Basel, Switzerland.
Basel Committee on Banking Supervision. 2006. International Convergence of Capital Measurement and Capital Standards, a Revised Framework Comprehensive Version. BIS, Basel, Switzerland.

Please cite this article in press as: Chang, C.-L., et al., A stochastic dominance approach to financial risk management strategies. Journal of Econometrics (2015), http://dx.doi.org/10.1016/j.jeconom.2015.02.032