Available online at www.sciencedirect.com
ScienceDirect
Journal of Banking & Finance 32 (2008) 1404–1415
Journal of BANKING & FINANCE
www.elsevier.com/locate/jbf

Backtesting trading risk of commercial banks using expected shortfall ☆

Woon K. Wong *,[1]

Investment Management Research Unit, Cardiff Business School, Aberconway Buidling, Colum Drive, Cardiff CF10 3EU, United Kingdom

Received 8 January 2007; accepted 19 November 2007
Available online 19 February 2008

Abstract

This paper uses saddlepoint technique to backtest the trading risk of commercial banks using expected shortfall. It is found that four out of six US commercial banks have excessive trading risks. Monte Carlo simulation studies show that the proposed backtest is very accurate and powerful even for small test samples. More importantly, risk managers can carry out the proposed backtest based on any number of exceptions, so that incorrect risk models can be promptly detected before any further huge losses are realized.

JEL classification: G32

Keywords: Value-at-Risk; Expected shortfall; Backtesting; Saddlepoint technique

1. Introduction

Since Value-at-Risk (VaR) was sanctioned by the Basle Committee in 1996 for market risk capital requirement through the so called internal models, VaR has become the standard measure for financial market risk. Major commercial banks and other financial entities such as hedge funds disclose VaR on a regular basis. Also, risk management consulting firms and software systems have been set up to meet the demand generated by the VaR-based risk management. Despite the popularity of VaR, many commercial banks and other financial institutions suffered severe trading losses during the fall of 1998. This prompted a report by the BIS Committee on the Global Financial System in 1999, in which it is remarked that "*last autumn's events were in the 'tails' of distributions and that VaR models were useless for measuring and monitoring market risk.*" Such risk in the use of VaR is termed the "tail risk" by the literature; see Yamai and Yoshiba (2005). Given the role of commercial banks as principal dealers in the ever growing over-the-counter derivatives markets, their trading accounts have grown rapidly and become progressively more complex. Though most of them are hedged, during times of extreme market movements, correlations between various derivatives and their hedged counterparts may break down, resulting in the failure of VaR models in measuring and monitoring market risk; see for example Jorion (2000). On the other hand, Longin and Solnik (2001) and Campbell et al. (2002) provide evidence of increased correlation in bear markets, with reduced diversification and more heavy capital losses. Thus it is important to use a risk measure that takes into account the extreme losses beyond VaR, and expected shortfall (ES) proposed by Artzner et al. (1997, 1999) is an alternative risk measure that can remedy this shortcoming of VaR.

Essentially, VaR is a quantile measure which, at a given a confidence level, describes the loss that can occur over a given period due to exposure to market risk. ES is the

☆ This paper was reviewed and accepted while Prof. Giorgio Szego was the Managing Editor of The Journal of Banking and Finance and by the past Editorial Board.
* Tel.: +44 29 20875079; fax: +44 29 20874419.
E-mail address: wongwk3@cardiff.ac.uk
[1] This paper was partially completed when the author was with the Department of Banking and Finance in Tamkang University, Taiwan.

0378-4266/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2007.11.012

expected loss conditional on the loss being above the VaR level. Despite the fact that ES is a coherent risk measure whereas VaR is not subadditive,[2] the former is absent in Basel II. One main reason for this is that the backtest of ES is harder than that of VaR; see Kerkhof and Melenberg (2004, footnote 8). For instance, the test statistics of existing ES backtests, namely the censored Gaussian approach of Berkowitz (2001) and the functional delta approach of Kerkhof and Melenberg (2004), rely on large sample for convergence to the limiting distributions. This is undesirable because test samples tend to be small in practice and regulation requires backtest to be carried out on past 250 observations. In this paper, I propose to use saddlepoint or small sample asymptotic technique to compute under the null hypothesis the required p-value based on sample expected shortfall. The advantage of the saddlepoint technique is that the required p-value can be calculated *accurately for any given number of exceptions*[3] under the null hypothesis. This is of great significance because the Capital Accord stipulates VaR at 99% level, which implies that exceptions rarely occur in practice. Monte Carlo simulation study confirms that the proposed technique is not only very accurate but also very powerful even for small samples.

Current VaR backtesting as stipulated by the Basle Committee focuses on checking number of violations of the VaR level. While it is advantageous for such backtesting procedure to be applicable to any distribution, it ignores valuable information conveyed by the sizes of losses exceeding the VaR level, and thus lacks statistical test power. Indeed, the report by the Basle Committee for Banking Supervision (1996b, page 5) acknowledges that, "*tests of this type are limited in their power to distinguish an accurate model from an inaccurate model...This limitation has been a prominent consideration in the design of the framework presented here, and should also be prominent among the considerations of national supervisors in interpreting the results of a bank's backtesting program.*" Therefore, various methods of ES backtesting, noticeably by Berkowitz (2001) and Kerkhof and Melenberg (2004), have been proposed. Though parametric assumptions for the null distributions are made in the ES backtests, considerable powers have been gained.[4] The proposed backtest of ES using saddlepoint technique, however, have several advantages over these two existing approaches. First, as mentioned above, both existing ES backtests rely on asymptotic test statistics that might be inaccurate when sample size is small. If the Basle Committee were to employ the backtest of ES to determine the risk capital requirement, such lack of accuracy might be unacceptable because banks may be penalized based on incorrect inference. Second, not only Monte Carlo simulation study shows that the saddlepoint backtest has correct size, it is also more powerful than the two existing ES backtests. Finally, perhaps the most important of all from a practical risk management view point, the proposed saddlepoint technique allows one to compute the required p-value conditional on the number of exceptions instead of the size of the full sample. This means that it is possible to detect failure of a risk model based on just one or two exceptions before any more data are observed. This is extremely useful since prompt action is often required in order to avert extreme financial losses due to market risk.

The proposed backtest of ES is used to test for non-normal or excessive trading risk at the six commercial banks studied by Berkowitz and O'Brien (2002). For the sample banks, internal VaR is calculated using structural model which takes into consideration all positions in the trading portfolio. Since the saddlepoint technique is conditional on number of realized exceptions but not on sample size, ES backtests based on the bank's internal VaR model are carried out for the full sample period from January 1998 to March 2000 as well as for the three months sub-sample period from August 1998 to October 1998 (when the financial markets were highly volatile). In both cases, four out of six commercial banks' trading accounts are found to have sample ES that are significantly larger than a normal null hypothesis would suggest. This serves as a warning for risk modeling based on normal distribution which is quite common in practice; see for example Lotz and Stahl (2005).

This paper is organized as follows: Section 2 introduces the sample ES statistic for backtesting and the saddlepoint technique to calculate the p-value of the proposed backtest is derived in Section 3. Section 4 presents the Monte Carlo simulation results whereas Section 5 applies the proposed backtest to the trading accounts at six large banks in US. Section 6 briefly illustrates how the proposed ES backtesting could help risk management to be more responsive. Finally, Section 7 concludes.

2. Backtesting expected shortfall for tail risk

In this section, some preliminary definitions of VaR and ES are first provided, followed by the definition of the sample ES statistic to be used in backtesting. Difficulties of existing methods for backtesting ES and the motivation for the use of small sample asymptotic technique are explained.

2.1. Some preliminary definitions

Let R be the random profit or loss of a portfolio over a holding period. For simplicity, R is assumed to have a

[2] A coherent risk measure is subadditive (sum of individual risks cannot be smaller than risk of sums), homogenous (larger positions are associated with greater risk), monotonous (systematic lower returns imply greater risk), and translation invariant (investing in risk-free asset will lower the risk of portfolio). See, for example, Artzner et al. (1997, 1999) and Szego (2002) for more details.

[3] In this paper, an exception always means an exception or exceedance of VaR. For brevity, similar presumption is also made for 'violation' and 'breach' of VaR.

[4] For a distribution-free test, readers are referred to Section 4.4.3 in McNeil et al. (2005).