VALUE AT RISK AND EXPECTED SHORTFALL: A COMPARATIVE ANALYSIS OF PERFORMANCE IN NORMAL AND CRISIS MARKETS

by

Peter Mak
Bachelor of Education, University of Ottawa 2007
Bachelor of Science in Engineering, Queen's University 2006

and

Qi Meng
Bachelor of Science, University of Waterloo 2009

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN FINANCE

In the Master of Science in Finance Program
of the
Faculty
of
Business Administration

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name: **Peter Mak, Qi Meng**

Degree: **Master of Science in Finance**

Title of Project: **Comparative Study of Value at Risk verse Expected Shortfall Base on Empirical Research in Normal and Stressed Market Conditions**

Supervisory Committee:

____

Dr. Andrey Pavlov
Senior Supervisor
Academic Chair, MSc in Finance

____

Dr. Phil Goddard
Second Reader
Instructor

Date Approved: ______________________________

Abstract

In this paper, we compare two risk measures, Value at Risk (VaR) and Expected Shortfall (ES) in their ability to capture risk associated with tail thickness. We test them under both normal and stressed market conditions using historical daily return data for capital-weighted stock indices from major markets around the world. Both log-returns (serially correlated) and ARMA-GARCH residuals (not serially correlated) are tested.

We find that expected shortfall is better at capturing tail risk than VaR under all market conditions; the improvement is more dramatic in normal conditions and in serially correlated data. However, in crisis conditions, the stochastic component of returns shows that both risk measures contain very high tail risk. We recommend that practitioners and regulators using VaR consider switching to expected shortfall to be better prepared for extreme negative events.

Keywords: Value at Risk, VaR, Expected Shortfall, Crisis, Extreme Value Theory

Acknowledgements

We would like to thank Dr. Andrey Pavlov and Dr. Phil Goddard for their guidance, insight and suggestions. We also appreciate the work of Dr. Peter Klein and Suzanne Yim in facilitating the final project process.

Thanks also to Angie Leung and Joyce Yue for their support and patience as we worked on this project.

Table of Contents

List of Figures

List of Tables

1: Introduction

1.1 Overview

The objective of this study is to build upon the pioneering work of Artzner et al. (1997, 1999) on coherent risk measures, which take into account the most extreme returns of a security, as a substitute for value at risk (VaR), which only considers a specific percentile of the returns distribution. We specifically consider expected shortfall, a well-known coherent risk measure.

Yamai and Yoshiba (2002b, 2005) extend this concept by applying extreme value theory to measure tail thickness. By analysing foreign currency exchange rates during the Asian financial crisis, they establish that over this period both value at risk and expected shortfall have tail risk (risk due to fat tails, see Section 3.2), but that expected shortfall generally has less than value at risk.

However, the above study does not make any comparison between typical market behaviour and stressed market behaviour. It is important to know whether or not expected shortfall offers a significant improvement over VaR under stressed conditions compared to typical conditions.

This study therefore extends the existing body of work using similar methodology to that introduced by Yamai and Yoshiba (2002b, 2005) on differentiated normal and stressed stock index returns data. Data from the 2007-2008 financial crisis and from a more typical period shortly before it are used to establish normal and crisis regimes of returns behaviour. The distributions in each regime are analysed to determine the tail risk of both expected shortfall and value at risk.

Furthermore, the work of Yamai and Yoshiba (2005) is expanded by applying their methodology to the stochastic component of returns through the application of an ARMA/GARCH model. This is done to provide insight into the ability of VaR and expected shortfall to capture truly unpredictable tail risk, removing the effects of autocorrelation.

We find that VaR is outperformed by expected shortfall in all conditions in terms of the capture of tail risk. However, the outperformance in the crisis regime is not as extreme as expected, where both measures have significant tail risk. This is found to be especially true of the stochastic component of returns, where both VaR and expected shortfall are found to have very significant tail risk. Indeed, much of expected shortfall's superior performance in the crisis regime is eliminated when applied to data that are not auto correlated.

1.2 Background

Value at risk (VaR) is a risk measurement that describes the riskiness of the return of a portfolio. It is widely used across the financial industry, both by participants and by regulators. Its popularity is owed to its simplicity: it is reported as a single number that represents potential losses with some confidence level in raw dollar terms. As a result, it is easy to understand and intuitive to manage and regulate. Nevertheless, VaR has two major shortcomings, both of which are side effects of the way in which it is calculated. Its value is based completely on the losses at the specified confidence level, while no weight whatsoever is assigned to losses beyond that confidence level.

The main concern with the use of VaR as a risk measure is that it does not respond to losses exceeding the confidence level. As a result, it works well when applied to normally distributed returns, but it cannot capture the risk associated with the shape of the distribution beyond the confidence level. In particular, if a portfolio's returns distribution has thick tails representing low-probability, extreme-magnitude losses, VaR cannot accurately convey the riskiness of the portfolio's tails.

The second major problem is that VaR is super-additive under certain conditions. This means that VaR may suggest that combining two portfolios together may result in greater total risk than the sum of the component portfolios' risk: it may fail to accurately capture the benefits of diversification. This happens because the weights assigned to losses beyond the VaR confidence level is a decreasing function of their probability (from 100% at the confidence level to 0% beyond it). It therefore fails to qualify as coherent risk measure defined by Artzner et al. (1997).

To remedy the problems inherent in VaR, Artzner et al. (1997) propose the use of expected shortfall (ES) as an improvement on VaR. Expected shortfall is the expected value of losses beyond the confidence level. Since expected shortfall assigns non-decreasing weights (actually, equal weights) to losses beyond the confidence level, it is always sub-additive and therefore also a coherent risk measure. Furthermore, since losses beyond the confidence level are taken into account in the calculation of ES, it should by definition have a superior ability to pick up any risk associated with returns distributions whose tails are not normal.

Returns on stock markets generally follow normal distributions, and as a result, under typical conditions, VaR is thought to be almost as effective as ES at capturing risk. However, the recent financial crisis highlights the importance of measuring the risk associated with non-normal returns and extreme tails. It is therefore desirable to confirm that ES does, in fact, outperform VaR in its ability to capture the risk associated with thicker tails.

To this end, the objective of this study is to compare VaR and expected shortfall by measuring their ability to pick up tail risk (see Section 3 for details on tail risk). We calculate return distributions from the daily values of major stock exchanges. We then analyze VaR and expected shortfall from these data series as functions of tail size to compare their abilities to pick up tail risk.

Our findings suggest that expected shortfall is more effective at capturing tail risk and therefore should replace VaR as the risk measure of choice for the post-crisis financial industry, where stress testing and worst-case scenarios are at the forefront of risk managers' thoughts. We also find that, under crisis conditions, both measures have significant tail risk, although expected shortfall has less than VaR.

The rest of the paper is organized as follows. Section 2 formally defines value at risk and expected shortfall. Section 3 defines tail index as a measurement of tail thickness and explains how it is calculated. Tail risk is also defined. Section 4 explains our method of determining VaR and expected shortfall of a data series. Section 5 describes the empirical data chosen and how it was prepared for testing. Section 6 presents the results of the tests and the conclusions drawn from them. Section 7 summaries the paper.

2: Risk Measure Definitions

We focus on two risk measures in this paper: value at risk (VaR) and expected shortfall (ES).

2.1 Value at Risk

Following convention, VaR represents the maximum loss that the portfolio will experience over some time interval, with some confidence level. For example, if portfolio A has a 1-month 95% VaR of $150 000, then one is 95% confident that over a given month the portfolio will not lose more than $150 000.

Mathematically, we denote the distribution of returns Z of the portfolio over the holding period T. VaR at the (100-X) % confidence level is then defined to be the upper X percentile of the loss distribution, as shown in Equation 1 below:

$$VaR_{X\%}(Z) = \sup\{z|P(Z \geq z) > X\%\} \qquad \text{Equation 1}$$

Where sup{z|A} is the upper limit of Z given event A as used in Artzner et al. (1999). In this paper, we analyse the daily returns of stock indices, so all distributions Z are discrete distributions, but without a loss of generality the same definition of VaR holds for continuous distributions as well.

2.2 Expected Shortfall

We define expected shortfall as proposed by Artzner et al. (1999). It is the conditional expectation of a portfolio's losses beyond VaR of the same confidence level over the same holding period. So, recalling portfolio A above with 95% one-month VaR of $150 000, if the same portfolio has a 95% ES of $300 000, then in the 5% of cases where losses exceed $150 000, expected shortfall tells us the expected loss is $300 000. Mathematically, expected shortfall is defined in Equation 2:

$$ES_{X\%}(Z) = E[Z|Z \geq VaR_{X\%}(Z)] \qquad \text{Equation 2}$$

Where E[Z|A] is the expectation of Z given event A. For a continuous distribution, expected shortfall is therefore the probability-weighted average of all losses beyond VaR. Given that our data is discrete and of equal weighting, the expected shortfall calculated is the simple average of all daily losses beyond VaR; numerically, the probability weighting is reflected in the frequency of losses at a given level.

Because of the way in which expected shortfall is calculated, larger data sets are generally required in order for it to achieve the same level of accuracy as VaR (Yamai and Yoshiba, 2002a). Specifically, Yamai and Yoshiba found that about 1000 data points are required in order for the accuracies of the two risk measures to converge for practical purposes. This is a significant limitation of expected shortfall, since it makes it less reliable than VaR for small data sets.

3: Tail index and tail risk

3.1 Tail index

We use extreme value theory to represent the tails of asset returns in this paper, following the same method as Yamai and Yoshida (2002b, 2005). In particular, we consider only the univariate case in this study.

A brief introduction to extreme value theory is needed to explain our study of the tail index. We define Z as a return distribution representing the daily returns of an asset based on available data. The distribution function describing Z is denoted F and is not known. Only the extreme values are considered here, as this study focuses on the tail portion of the distribution beyond some threshold θ. Exceedances are defined as $m_\theta = \max(\mathrm{Z}, \theta)$. The resulting probability that Z is greater than θ is denoted p, so probability that it is smaller than or equal to θ is given by 1 – p. The variable p is referred to as tail probability. Based on the above, the conditional distribution F_θ is defined as:

$$F_\theta(x) = \Pr\{Z - \theta \le x | Z > \theta\} = \frac{F(x) - F(\theta)}{1 - F(\theta)}, \theta \le x \qquad \text{Equation 3}$$

Where F_θ is the distribution function of (Z-θ) given that Z exceeds θ. Note that F_θ is not known unless F is known.

Under univariate extreme value theory, following Yamai and Yoshiba (2002b, 2005), F_θ for sufficiently high values of threshold θ, converges to a generalized Pareto distribution $G_{\xi,б}$ (see Embrechts et al., 1997) of the form:

$$G_{\xi,б}(x) = 1 - \left(1 + \xi \cdot \frac{x}{б}\right)^{-1/\xi}, x \ge 0 \qquad \text{Equation 4}$$

Given Equations 3 and 4, and when the value of θ is sufficiently large, the distribution function of the exceedances m_θ can be approximated as follows:

$$\begin{aligned} F_m(x) &\approx \big(1 - F(\theta)\big) G_{\xi,б}(x - \theta) + F(\theta) \\ &= 1 - p\left(1 + \xi \cdot \frac{x - \theta}{б}\right)^{-1/\xi}, x \ge \theta \end{aligned} \qquad \text{Equation 5}$$

Where F_m is hereafter referred to as the distribution of exceedances.

The distribution of exceedances is determined by three parameters: tail probability p, which directly determine the threshold θ; the shape (or scale) parameter б that describes how dispersed the distribution is; and the tail index ξ. Tail index is the parameter with which this study is primarily concerned: it describes how thick the tail of the distribution is. Larger values of ξ indicate thicker tails.

3.2 Tail Risk

For any given shape of the distribution Z, VaR and expected shortfall give the same amount of information about the riskiness of the portfolio since they are simply be scalar multiples of one other. In particular, if the shape of the distribution is normal, both risk measurement are fixed scalar multiples of the standard deviation.

The improvement of expected shortfall over VaR arises when the shape of the distribution changes to have fatter tails. Consider the two return distributions in Figure 3.1 below, assuming that they are plotted to the same scale. The distribution shown in (b) is clearly riskier, since it has a higher probability of extremely high losses while other elements of the distribution to the left of 95% VaR remain the same. While the VaR of these two return distributions is the same, since the value at 95% VaR is the same, ES is different: ES, since it is a probability-weighted average of the top 5% of losses, is increased as a result of the peculiar shape of the tail in (b). So, VaRa = VaRb while ESa < ESb. Since ES increases with increase of risk due to tail shape, it appears to capture the risk associated with the tail shape, while VaR does not. Alternatively, we say that VaR has more tail risk than expected shortfall.

To directly compare the effectiveness of value at risk and expected shortfall at capturing tail risk, we follow the definition of tail risk proposed by Yamai and Yoshiba (2002b, 2005): when comparing the tail indices ξ of two portfolios' return distributions and their respective risk measurements, a risk measure has tail risk if it assigns higher risk to the portfolio with a smaller tail index. Under such conditions, the tested risk measure fails to increase to capture the increase in risk associated with a larger tail. In other words, when $\xi_1 > \xi_2$ but $VaR_1 < VaR_2$ or $ES_1 < ES_2$, that risk measure has tail risk.

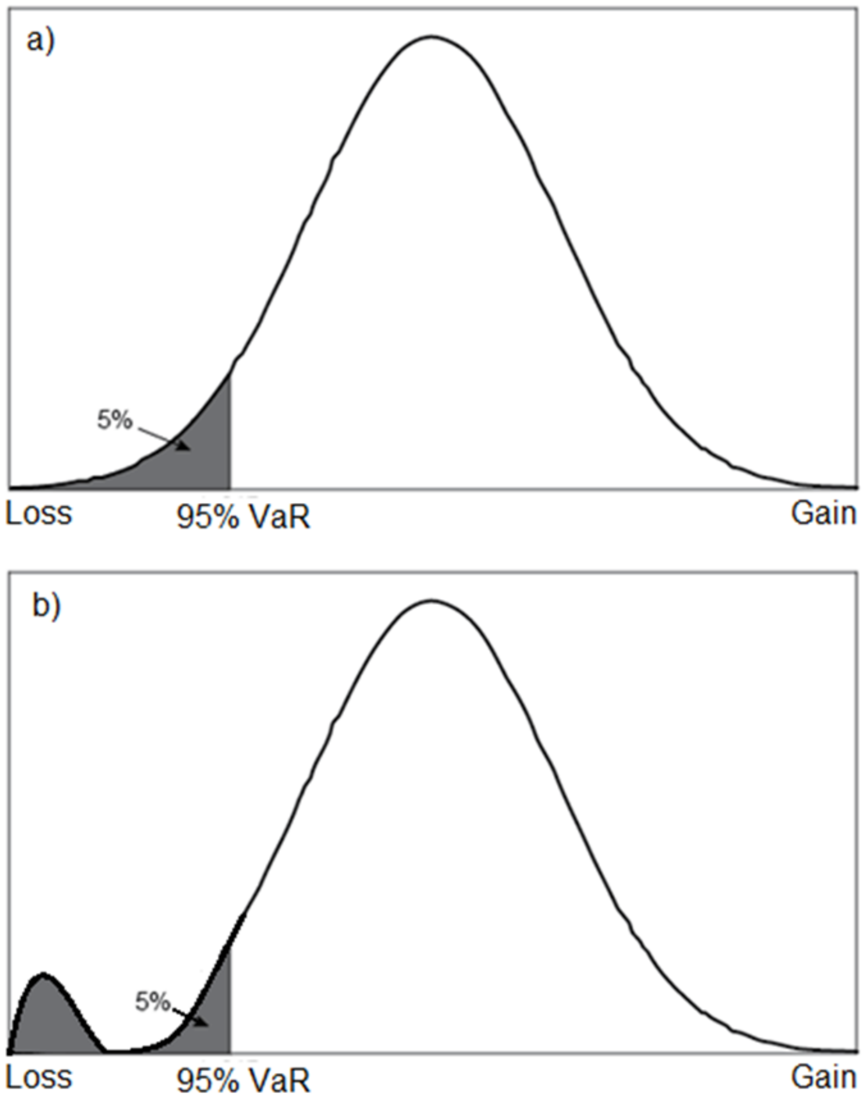

Figure 3-1: Probability density functions of loss distributions that follow (a) a normal distribution and (b) a distribution with extreme tail behavior.

4: Calculation of VaR and Expected Shortfall

There are many ways to determine VaR and expected shortfall of a portfolio. Here, the historical method based on empirical data is used. This method was chosen because it is the most direct representation of the real market environment, with the least assumptions made to artificially affect the data. As a trade-off, the VaR and expected shortfall values calculated using real-world data are noisy, with less-predictable behavior.

4.1 Method for VaR

A fixed time interval for all indices in both normal and crisis regimes was selected, and daily ticker values were converted to daily log returns. For X% VaR, we simply identify the 1 – X percentile of the data and convert it back into percentage daily return form.

4.2 Method for Expected Shortfall

Because the data are discrete, we use the conditional definition to calculate expected shortfall. According to this method, expected shortfall is the average of all values beyond the threshold of VaR. For X% ES, we take the arithmetic mean of all data points beyond the X% VaR data point.

4.3 Choice of Confidence Level

A time interval of approximately 700 trading days was chosen due to concerns with accurately representing the financial crisis (a detailed explanation can be found in Section 5). The limited size of this data set makes extremely high confidence levels unreliable. Therefore, we choose to focus on 95% confidence level for both VaR and expected shortfall, consistent with common industry practice.

5: Data Selection and Processing

The data used in the analysis were daily log-returns data representing two distinct periods of market operation: a “normal regime” and a “crisis regime.” Index values from Bloomberg were converted to log-returns for these periods, subjected to ARMA/GARCH to isolate the stochastic component of returns, and the tail index ξ is determined for both the log-returns and for the ARMA/GARCH stochastic component.

5.1 Data Retrieval

The data analyzed come from the daily values of thirteen internationally-prominent capital-weighted market indices. Daily index values were retrieved from Bloomberg for the time period from December 13th, 1999 to June 19th, 2014. This time period was used because it straddles the 2007-2008 financial crisis by a wide margin on either side.

The analysis of indices was preferred over individual securities in order to ensure that only beta-risk was captured in the study. The indices used in the study were selected to satisfy two parameters. First, they are capital-weighted. This ensures that any return data on the indices are reflective of returns on capital invested in their respective markets. Second, they are geographically diverse but significant in size. Indices were selected subjectively for this parameter until a sufficiently large group of indices was generated.

The Bloomberg ticker symbols for the selected indices are: CAC (CAC-40 from Paris), CCMP (NASDAQ composite index), UKX (FTSE-100 from London), HSI (Hang Seng from Hong Kong), RTY (Russell 2000 small caps), SPX (S&P 500 large caps), AS51 (S&P/ASX 200 from Australia), SPTSX (S&P/TSX 500 from Canada), SENSEX (BSE from India), IBEX (IBEX-35 from Spain), MEXBOL (Bolsa IPC from Mexico), FMBKLCI (KLCI from Malaysia) and TWSE (TAIEX from Taiwan).

5.2 Trading Day Inconsistencies

Due to differences in trading days between the exchanges, a technique similar to that described by Yamai and Yoshiba (2002b) was applied, where all trading days that exist in any data set were added to all remaining data sets where they were non-trading days. The index value on this non-trading day was then set to the same value as on the previous trading day. The change has no significant effect on the outcome of the analysis since this extra data point has a daily return of zero, which has no effect on the tails of the distribution except for a minor impact on where they begin.

5.3 Determination of Normal and Crisis Regime Dates

Given the base data, it is necessary to specify the starting and ending dates of the "normal regime" and the "crisis regime" to be analyzed; the time between the starting and ending dates is referred to as the "period length." A trade-off exists in the determination of period length, since a shorter period length (under 500 days) offers the best direct representation of the 2007-2008 financial crisis, yet the parameter estimations for ξ, VaR and particularly expected shortfall (see Section 2.2) are all superior for larger data sets. The same period length was used for the normal regime for consistency.

To optimize date selection, MATLAB was used. First, the complete data set was converted to log-returns. Next, it was assumed that, under the normal regime, average daily variance on the log-returns should be low, and that under the crisis regime, average daily downside variance of log-returns should be high. Based on these criteria, for a given period length, all possible combinations of start and end date were analyzed to determine the optimal start and end dates within the period.

The period length of the crisis was determined manually in an iterative process in order to best match the onset of turmoil in the financial markets. Various period lengths were tested in order to obtain the highest possible maximum downside variance, satisfying the above criteria, while maximizing the number of data points used and subjectively starting just before the obvious downturn in markets. The results of the period selection process can be found in Table 5.1

Table 5-1 Start and End Dates for Regimes

Regime	**Start Date**	**End Date**	**Selection Criterion**
Normal	23-Jun-04	19-Feb-07	Minimum average daily variance
Crisis	26-Jun-07	09-Mar-10	Maximum average daily downside variance

5.4 Generalized Pareto Distribution Fit of Stochastic Component

In order to isolate the stochastic component of the daily returns, the data for the two selected regimes were run through an ARMA/GARCH process that ran as follows:

- Convert index values to log-returns
- Iteratively apply ARMA fit until a minimum acceptable fit is found – if no autocorrelation is detected, no ARMA fit is performed
- Apply GARCH(1,1) if variances are correlated; do not apply GARCH if variances are uncorrelated

The result of this automated process is summarized in Table 5.2 for the crisis regime data and in Table 5.3 for the normal regime data. The epsilon values left over from the ARMA/GARCH process – that is, the daily ARMA innovations divided by the squared daily GARCH sigma values – were accepted as the stochastic component of daily log returns.

Finally, MATLAB was used to perform a generalized Pareto distribution fit to the data. This was done twice for each regime. First, it was applied to the log-returns data, consistent with the methodology employed by Yamai and Yoshiba (2005). Second, it was applied to the ARMA/GARCH epsilon values, in order to see if the pure stochastic component demonstrated any significant relationship.

The MATLAB function applies maximum likelihood estimation to a specified tail portion of the data, returning the parameters ξ and σ for a generalized Pareto distribution fit of the data. There is some ambiguity in where the tails of the data ought to begin (that is to say, the quantile cut-off for the start of the generalized Pareto distribution fit). Yamai and Yoshiba (2002b) address this by considering a fit using tails ranging from 10% of the data (90% tails) to 1% of the data (99% tails). The same data are generated for analysis in this study; however, due to the shorter period lengths being examined, difficulties arise in the size of the confidence intervals of the fits.

Table 5-2 Summary of ARMA/GARCH orders used to fit crisis regime data.

Index	**ARMA(r,m)**		**GARCH(p,q)**	
	r	**m**	**p**	**q**
CAC	2	1	1	1
CCMP	1	0	1	1
UKX	3	1	1	1
HIS	0	0	1	1
RTY	1	0	1	1
SPX	0	1	1	1
AS51	1	0	1	1
SPTSX	3	1	1	1
SENSEX	1	0	1	1
IBEX	3	1	1	1
MEXBOL	0	1	1	1
FMBKLCI	0	1	1	1
TWSE	0	0	1	1

Table 5-3: Summary of ARMA/GARCH orders used to fit normal regime data.

Index	**ARMA(r,m)**		**GARCH(p,q)**	
	r	**m**	**p**	**q**
CAC	1	0	1	1
CCMP	0	0	1	1
UKX	1	0	1	1
HIS	1	0	1	1
RTY	1	0	1	1
SPX	1	0	1	1
AS51	1	0	1	1
SPTSX	0	0	1	1
SENSEX	2	2	1	1
IBEX	1	0	1	1
MEXBOL	1	0	1	1
FMBKLCI	1	0	1	1
TWSE	0	0	1	1

5.5 Determining Tail Parameter ξ

As mentioned above, tail parameter values were calculated for tails beginning at 90% (10% tails) out to tails beginning at 99% (1% tails). The specific values of ξ used in the analysis had to be selected from these. Two requirements were applied in this selection process. First, ξ values should be taken from the smallest tail possible in order to best reflect the behaviour of the most extreme returns in the distribution. Second, they should have reasonably tight confidence intervals since the confidence intervals were quite large relative to the size of each calculated ξ.

The size of the confidence intervals calculated by MATLAB increases within a given index as the size of the tail decreases; this logically reflects the decreasing number of data points available for the estimation of progressively more extreme tails. To balance depth into the tail against confidence interval size, analysis of the complete data set was performed to determine which consecutive set of 5 tail sizes had the smallest variance, constrained by the requirement that MATLAB determine the tail parameter estimates to be statistically significant (such estimates' confidence intervals were reported by the software as statistically irrelevant). The most consistent set of ξ estimates overall was found to be tails ranging from 4% to 8% (alternatively, the tails corresponding to cut-offs at 96% to 92%). The average of these 5 tails was then calculated and used to represent a best-estimate of the tail parameter for each index, within each regime.

Note that the tail parameters calculated correspond to tail thickness that should be apparent in the data set beneath the 95% level. As such, 95% VaR should be able to capture at least some of the risk associated with thickness since the thicker tail would have an effect on returns at the 95th percentile.

6: Analysis and Findings

In this section, the analysis methodology is outlined, and findings and important conclusions are presented. Expected shortfall is found to have less tail risk than VaR, and tail risk is found to be higher in the crisis regime than in the normal regime.

6.1 Analysis Methodology

The analysis methodology used here followed the logic outlined by Yamai and Yoshiba (2002b, 2005). Tail risk must exist in any data where an increase in the tail size, measured by an increase in the value if ξ from the generalized Pareto distribution fit, did not result in a corresponding increase in the VaR and/or ES risk measures. This analysis was performed on both the normal and crisis regimes, using log-return data and using the stochastic component of returns (epsilons remaining after an ARMA/GARCH process), and attempts were made to identify any patterns in the strength of identified patterns.

In order to state definitely that tail risk is at least partially captured by a risk parameter (ES or VaR), it must be true that if $\xi 1 < \xi 2$, then ES1 < ES2 (or VaR1 < VaR2). If this relationship does not hold, then the increase in tail risk from $\xi 1$ to $\xi 2$ is not captured by the risk measure and it therefore has tail risk. The more times this rule is violated, the worse a risk measure is at capturing the risk associated with tail shape. This is the method used by Yamai and Yoshiba (2002b, 2005). Such violations are referred to as "tail exceptions" for the purposes of reporting below.

6.2 Log-Return Data

This analysis follows the methodology of Yamai and Yoshiba (2002b, 2005) by comparing log-returns data.

6.2.1 Normal Regime

Under the normal regime, expected shortfall appears to capture tail risk more effectively than value at risk: expected shortfall has 34 tail exceptions, while VaR has 39. Complete data can be found in Table 6.1.

Table 6-1: Value at risk and expected shortfall as a function of average tail (4% - 8%) ξ for normal regime log returns data.

Log-Returns Normal Regime			
Index	**Left Tail ξ**	**95% VaR**	**95% ES**
CCMP	-0.5576	1.4937	1.8394
SPX	-0.3244	1.0178	1.2838
AS51	-0.2621	1.1151	1.5539
SPTSX	-0.1928	1.1368	1.5681
HSI	-0.1581	1.3405	1.8215
RTY	-0.1260	1.6970	2.0742
CAC	-0.0764	1.2066	1.7689
IBEX	-0.0294	0.9828	1.5681
SENSEX	-0.0279	1.9537	2.9611
MEXBOL	-0.0099	1.8282	2.4341
UKX	0.0241	0.9955	1.4579
FMBKLCI	0.1131	0.7626	1.0295
TWSE	0.2215	1.4071	2.0858
	Exceptions	39	34

There is an interesting trend in the data, observable in Figures 6.1 and 6.2: generally, indices with smaller tails are found in more developed economies (CCMP, SPX, AS51, SPTSX, HIS, RTY, CAC), and within this group, both VaR and ES increase very reliably with ξ. In indices with higher tail parameters (SENSEX, MEXBOL, UKX, FMBKLCI, TWSE), which generally are less-developed or emerging markets with the exception of UKX, the relationship between the risk parameter and the tail size is completely unreliable. It is unclear from the data how their emerging-market status affects these indices' ξ, VaR and ES values, and what (if any) causal relationship exists between these factors. The correlation, however, is quite plain. Interestingly, Yamai and Yoshiba (2002b, 2005) observed similar effects on the tail parameters, expected shortfall and value at risk for emerging markets compared to mature markets in the Asian crisis of the late 1990s.

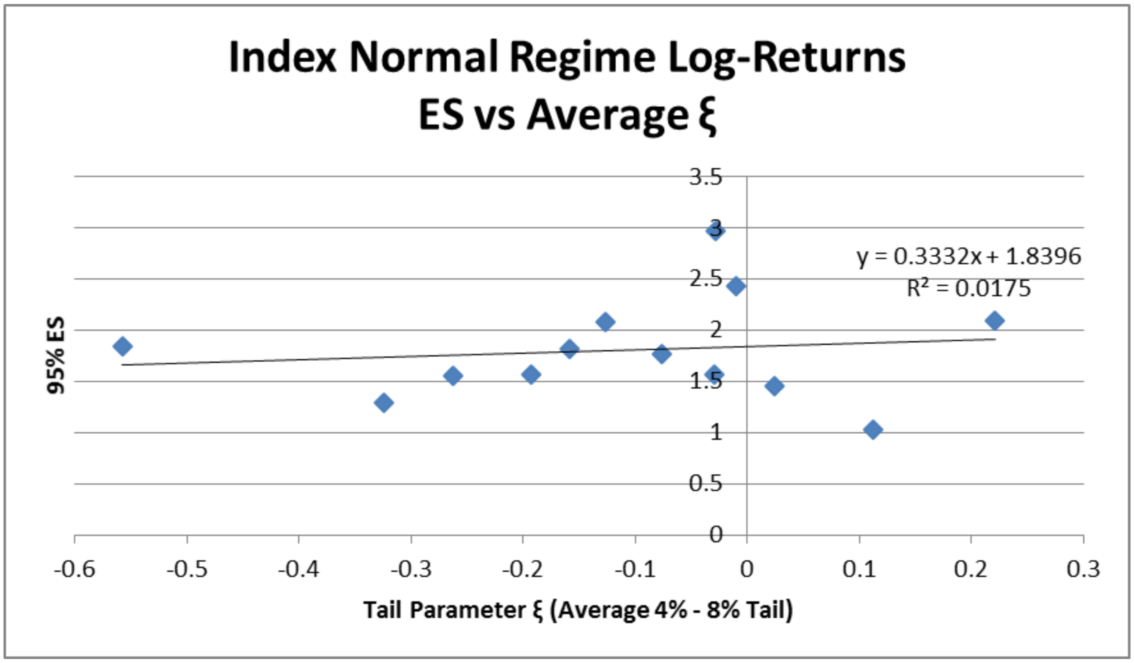

Figure 6-1 Expected shortfall as a function of average tail (4% - 8%) ξ for normal regime log returns data.

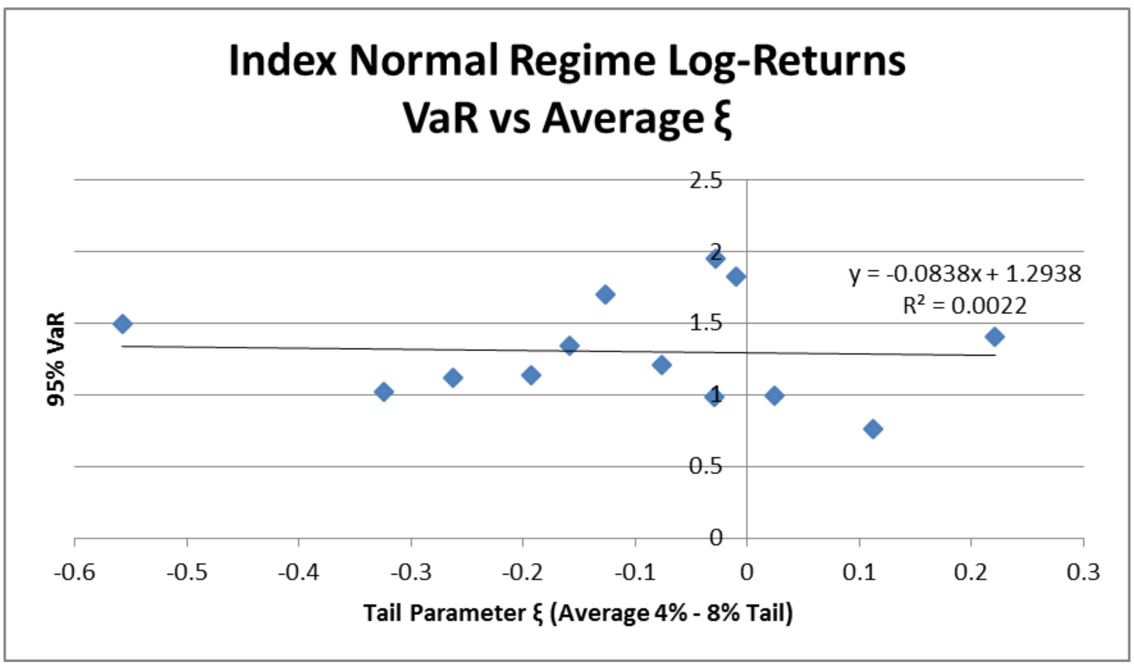

Figure 6-2 Value at risk as a function of average tail (4% - 8%) ξ for normal regime log returns data.

6.2.2 Crisis Regime

In the log-returns data, as expected, tail parameter ξ was generally found to be higher in the crisis regime than it was under the normal regime. This is consistent with the expectation that fatter tails are observed during periods of market stress.

As in the normal regime, value at risk is found to have more tail risk than expected shortfall. ES has 36 tail exceptions, while VaR has 44 in total. For complete results, see Table 6.2.

Table 6-2 Value at risk and expected shortfall as a function of average tail (4% - 8%) ξ for crisis regime log returns data.

Log-Returns Crisis Regime			
Index	**Left Tail ξ**	**95% VaR**	**95% ES**
MEXBOL	-0.2676	2.9212	4.2590
TWSE	-0.1619	3.3476	4.1188
CAC	-0.0526	3.1268	4.5987
IBEX	0.0451	2.8957	4.5145
CCMP	0.0777	3.1373	4.6821
SPX	0.0819	3.0010	4.7703
RTY	0.0842	3.6521	5.4785
SENSEX	0.0973	3.5314	4.9975
UKX	0.1017	2.6252	4.2435
AS51	0.1053	2.5373	3.8510
SPTSX	0.1112	2.9925	4.6411
HIS	0.1134	3.8143	5.3963
FMBKLCI	0.3125	1.5842	2.5639
	Exceptions	44	36

The pattern of smaller tails for mature markets and larger tails for emerging markets was not observed in this data set. The obvious pattern in Figures 6.3 and 6.4 is a very tight cluster of data with ξ = 0.08 to 0.11 where the VaR and ES were quite similar between the different securities (CCMP, SPX, RTY, SENSEX, UKX, AS51, SPTSX, HIS). A logical hypothesis is that this cluster exists due to the increased correlation between markets observed during the financial crisis (Grote, 2012), which should cause index return distributions' shapes (ξ) and extreme losses

(VaR and ES) to converge. Further specific study taking into account correlations between indices (using, for example, vector autoregressive techniques) might reveal interesting trends in the risk-measurement capabilities of ξ, value at risk and expected shortfall as functions of correlations.

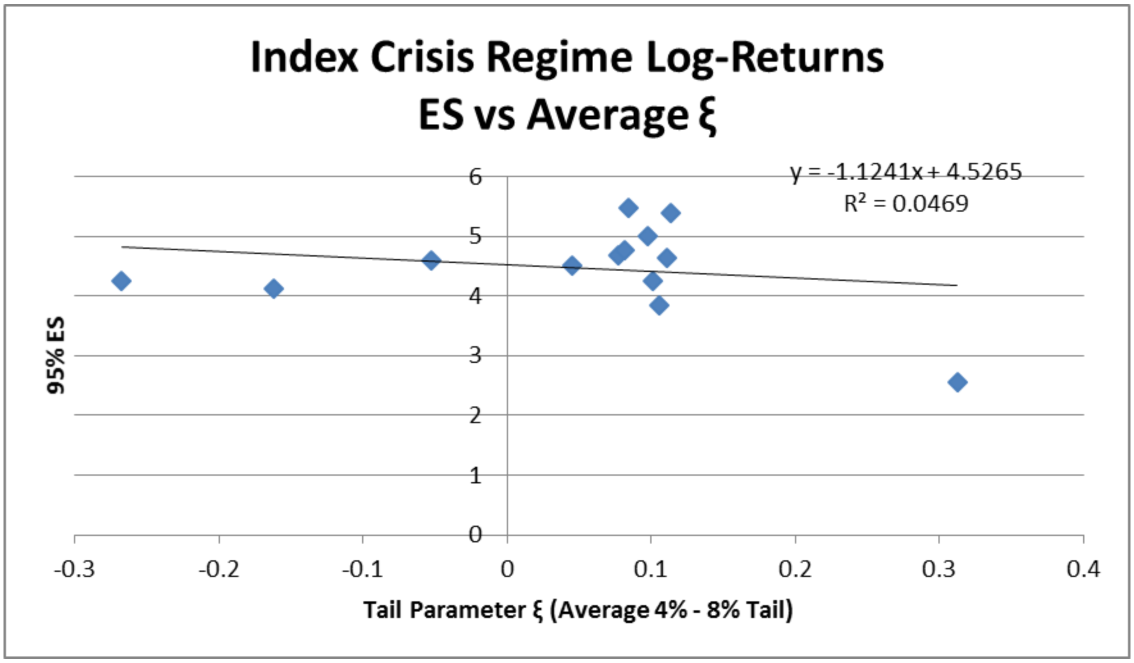

Figure 6-3 Expected shortfall as a function of average tail (4% - 8%) ξ for crisis regime log returns data.

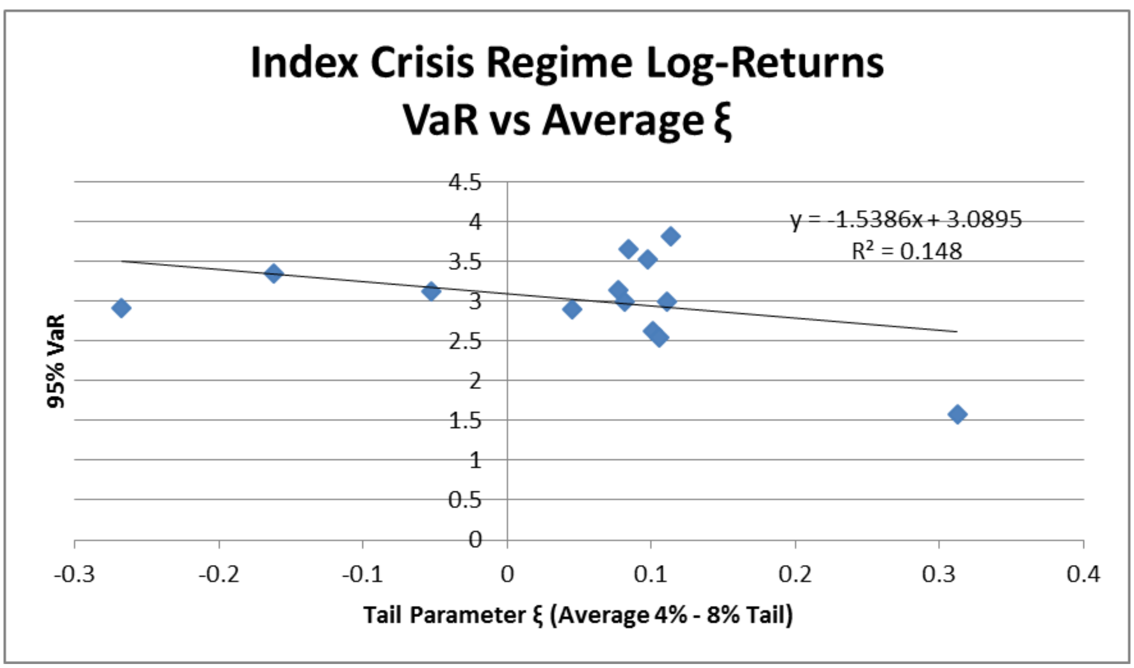

Figure 6-4 Value at risk as a function of average tail (4% - 8%) ξ for crisis regime log returns data.

6.3 ARMA/GARCH Data

This analysis attempts to improve on the log-returns analysis by isolating the stochastic component of returns to attempt to identify any patterns there. This is important to isolate the truly random component of the crisis regime activity, so that the tail parameter ξ is actually a measurement of the stochastic process driving the data and not simply dependent on autocorrelations in the data or its volatility.

6.3.1 Normal Regime

As in the previous data sets, both VaR and ES have tail exceptions, indicating some tail risk in each. Expected shortfall has 31 and value at risk has 37, suggesting again that expected shortfall is better at capturing tail risk than value at risk. Table 6.3 shows the results.

Table 6-3 Value at risk and expected shortfall as a function of average tail (4% - 8%) ξ for normal regime ARMA/GARCH data.

ARMA/GARCH Normal Regime			
Index	**Left Tail ξ**	**95% VaR**	**95% ES**
SPTSX	-0.3215	1.1368	1.5681
CAC	-0.2819	1.2066	1.7689
SPX	-0.2475	1.0178	1.2838
AS51	-0.2363	1.1151	1.5539
CCMP	-0.2341	1.4937	1.8394
MEXBOL	-0.2222	1.8282	2.4341
UKX	-0.2089	0.9955	1.4579
RTY	-0.1710	1.6970	2.0742
HSI	-0.1233	1.3405	1.8215
IBEX	-0.0197	0.9828	1.5681
TWSE	0.0750	1.4071	2.0858
FMBKLCI	0.0887	0.7626	1.0295
SENSEX	0.1136	1.9537	2.9611
	Exceptions	37	31

As illustrated in Figures 6.5 and 6.6, the trend of lower tail risk in developed markets observed in Section 6.2.1 appears to hold true even on the ARMA/GARCH data set, for the most part. Here, data are generally well-distributed with very low tail parameters, but there is a cluster of indices with much higher tail parameters: TWSE, FMBKCLI and SENSEX. This is may be suggestive of a geographical trend in the tail behaviour of the stochastic component of returns, since all of those are Asian indices.

A direct comparison of the exceptions shows that all of the exceptions found in expected shortfall are also exceptions in VaR; as such, there is no individual instance of tail risk in expected shortfall when VaR does not have tail risk. There are, however, six instances of tail risk in VaR that are not observed in expected shortfall. This suggests that, in a normal regime, expected shortfall strictly outperforms VaR. Table 6-4 shows the point-by-point comparison of the exceptions in the two risk measures.

Table 6-4 Point-by-point comparison of exceptions as found in VaR and/or ES under the normal regime. Text indicates what, if any, risk measures have exceptions in each comparison.

Normal Regime Exception Comparison												
	AS51	CAC	CCMP	FMBKLCI	HSI	IBEX	MEXBOL	RTY	SENSEX	SPTSX	SPX	TWSE
CAC	Both											
CCMP	None	None										
FMBKLCI	Both	Both	Both									
HSI	None	None	Both	Both								
IBEX	**VaR Only**	Both	Both	Both	Both							
MEXBOL	None	None	None	Both	Both	Both						
RTY	None	None	None	Both	Both	Both	Both					
SENSEX	None	None	None	None	None	None	None	None				
SPTSX	Both	None	None	Both	None	Both	None	None	None			
SPX	None	Both	None	Both	None	**VaR Only**	None	None	None	Both		
TWSE	None	None	**VaR Only**	Both	None	None	Both	**VaR Only**	None	None	None	
UKX	Both	Both	Both	Both	None	**VaR Only**	Both	None	None	Both	**VaR Only**	None

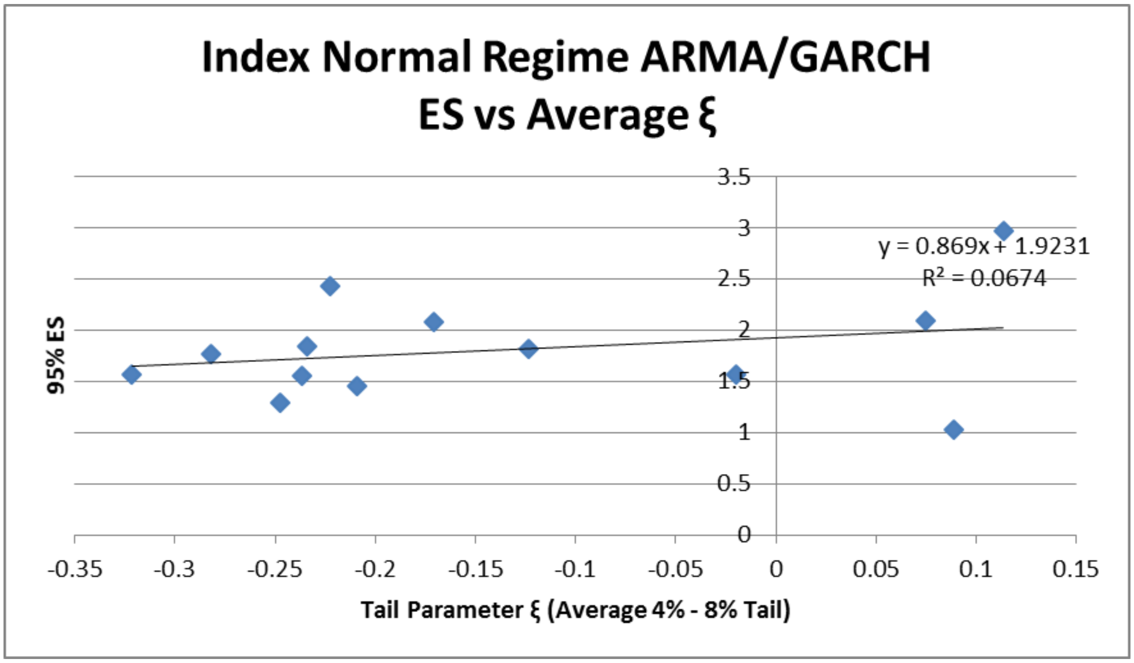

Figure 6-5 Expected shortfall as a function of average tail (4% - 8%) ξ for normal regime ARMA/GARCH data.

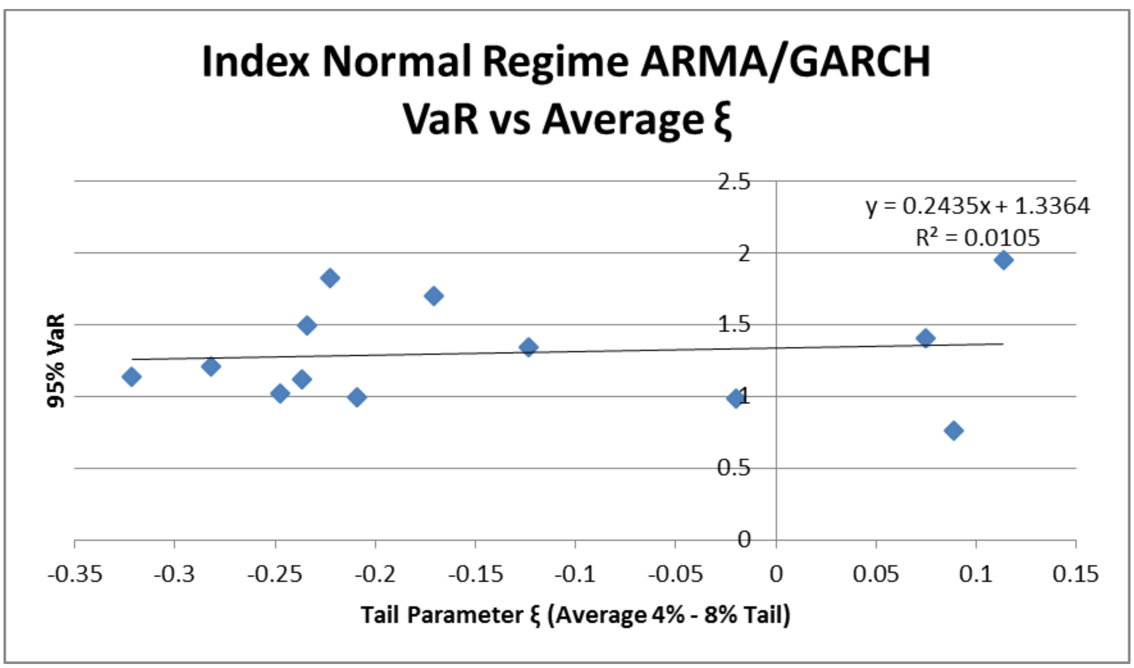

Figure 6-6 Value at risk as a function of average tail (4% - 8%) ξ for normal regime ARMA/GARCH data.

6.3.2 Crisis Regime

The ARMA/GARCH crisis regime data suggest that neither expected shortfall nor value at risk is able to compensate effectively for the risk associated with bigger tails, as indicated by the fact that both ES and VaR have huge numbers of tail exceptions – far larger than in any other data set. ES has 49 tail exceptions while VaR has 51, suggesting slightly worse capture of tail risk by VaR – however, both risk measures clearly are unable to capture differences in tail risk under crisis conditions. Results for this analysis can be found in Table 6.5.

Table 6-5 Value at risk and expected shortfall as a function of average tail (4% - 8%) ξ for crisis regime ARMA/GARCH data.

ARMA/GARCH Crisis Regime			
Index	**Left Tail ξ**	**95% VaR**	**95% ES**
RTY	-0.3831	3.6521	5.4785
SPTSX	-0.3542	2.9925	4.6411
TWSE	-0.3050	3.3476	4.1188
MEXBOL	-0.3031	2.9212	4.2590
SPX	-0.2145	3.0010	4.7703
SENSEX	-0.1343	3.5314	4.9975
UKX	-0.1126	2.6252	4.2435
HSI	-0.0795	3.8143	5.3963
CCMP	-0.0633	3.1373	4.6821
AS51	-0.0462	2.5373	3.8510
FMBKLCI	0.1441	1.5842	2.5639
CAC	0.1833	3.1268	4.5987
IBEX	0.2610	2.8957	4.5145
	Exceptions	51	49

The dramatic increase in the number of tail exceptions in the purely stochastic component compared to the log-returns (Section 6.2.2) is surprising. The magnitudes of the tail parameters themselves are far smaller on the ARMA/GARCH data. This suggests that much of the shape of the tails in the crisis data can be modelled by data and variance serial correlation, and is not merely random movement.

Furthermore, the dense clump of data points observed in the crisis regime using log-returns (Figures 6.3 and 6.4) is not apparent in the ARMA/GARCH data (Figures 6.7 and 6.8). This is an extremely interesting result, and suggests that the increases in correlation in financial crises can be modelled alternatively as serial correlation of daily returns and variances.

Point-by-point comparison of the tail risk exceptions is less predictable than what is found in the normal regime. In the crisis regime, five instances of tail risk that are not observed in VaR are found in expected shortfall, while seven exceptions from VaR are not found in expected shortfall. This indicates that expected shortfall is better on average, with two fewer overall exceptions. However, the strict improvement of expected shortfall over VaR observed in the normal regime is not replicated in the crisis regime data. In some data points, VaR performs better than expected shortfall and vice versa. Detailed comparisons can be found in Table 6-6.

Table 6-6 Point-by-point comparison of exceptions as found in VaR and/or ES under the crisis regime. Text indicates what, if any, risk measures have exceptions in each comparison.

Crisis Regime Exception Comparison												
	AS51	CAC	CCMP	FMBKLCI	HSI	IBEX	MEXBOL	RTY	SENSEX	SPTSX	SPX	TWSE
CAC	None											
CCMP	Both	Both										
FMBKLCI	Both	None	Both									
HSI	Both	Both	Both	Both								
IBEX	None	Both	Both	None	Both							
MEXBOL	Both	None	None	Both	None	**VaR Only**						
RTY	Both	Both	Both	Both	**ES Only**	Both	Both					
SENSEX	Both	Both	Both	Both	None	Both	None	Both				
SPTSX	Both	**ES Only**	None	Both	None	Both	Both	Both	None			
SPX	Both	**ES Only**	**ES Only**	Both	None	Both	None	Both	None	None		
TWSE	Both	**VaR Only**	**VaR Only**	Both	None	**VaR Only**	**VaR Only**	Both	None	**ES Only**	**VaR Only**	
UKX	Both	None	None	Both	None	None	Both	Both	Both	Both	Both	**VaR Only**

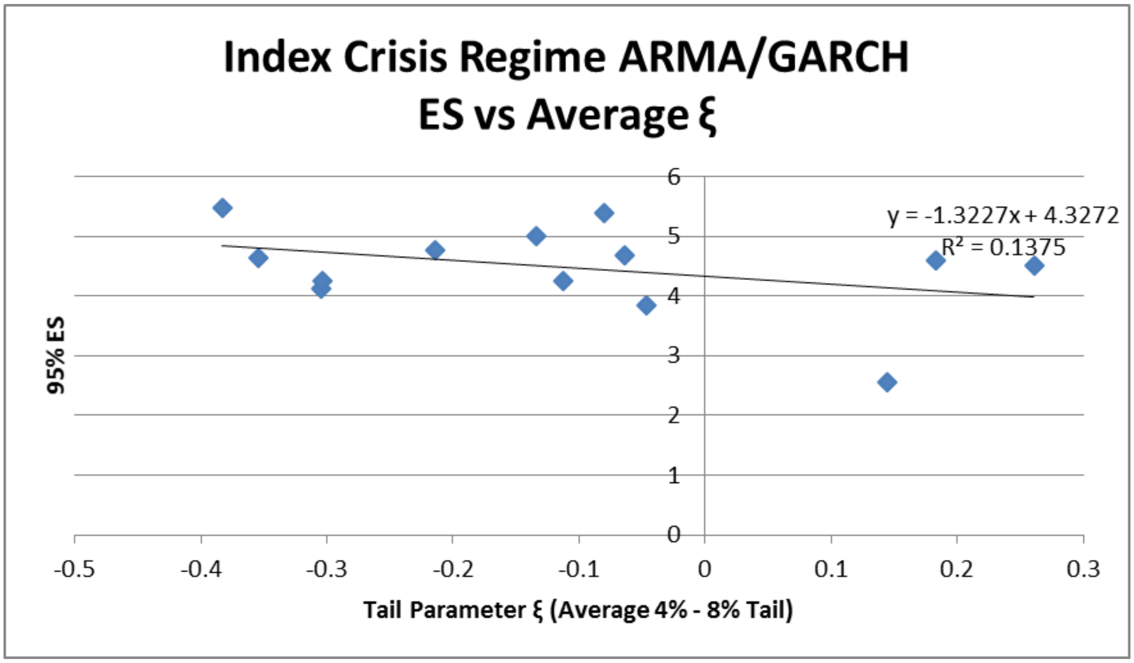

Figure 6-7 Expected shortfall as a function of average tail (4% - 8%) ξ for crisis regime ARMA/GARCH data.

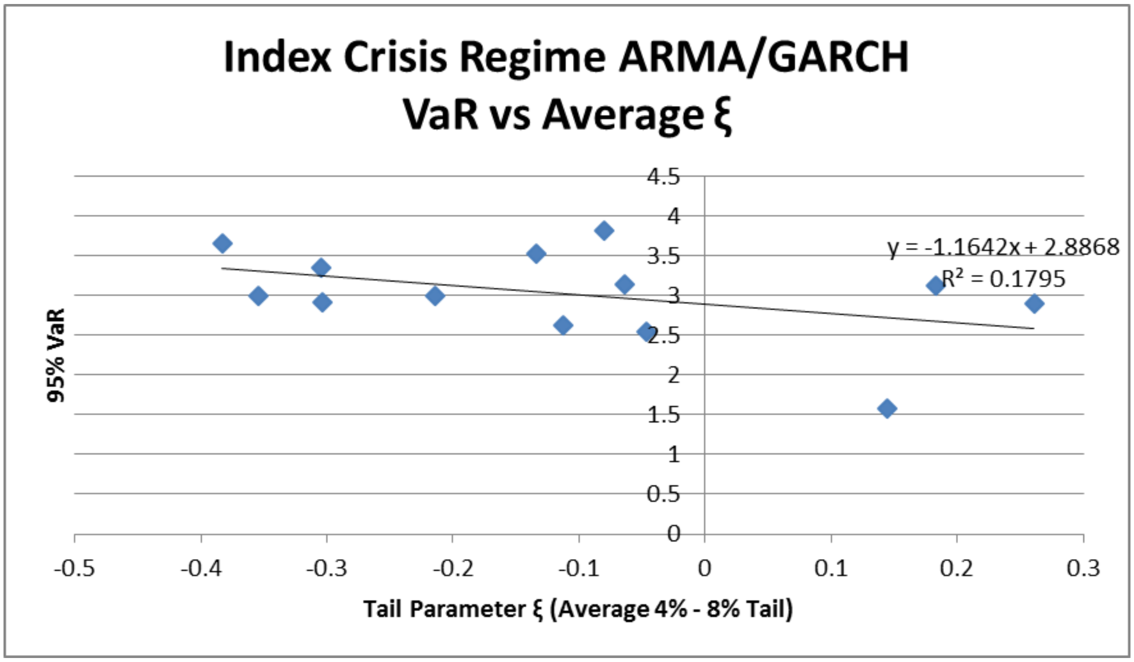

Figure 6-8 Value at risk as a function of average tail (4% - 8%) ξ for crisis regime ARMA/GARCH data.

6.4 Relationships and Trends

The analysis performed allows comparisons along two primary vectors: between the normal and crisis regimes; and between serially correlated and uncorrelated time series (log-returns vs. ARMA/GARCH epsilons). Several core relationships bear emphasis here:

1. Expected shortfall has less tail risk than value at risk. All data confirm that, under any regime and regardless of correction for autocorrelation, expected shortfall has fewer tail exceptions than value at risk.

2. Tail risk increases in the crisis regime compared to the normal regime. Both expected shortfall and value at risk exhibit more tail exceptions in the crisis regime than in the normal regime, suggesting that their abilities to capture tail risk are lower in a crisis regime. Furthermore, the strict improvement of expected shortfall over VaR observed in the normal regime disappears in crisis.

3. Neither expected shortfall nor value at risk fully captures tail risk in any situation. Since tail exceptions were observed in all data sets, both expected shortfall and value at risk have significant tail risk, despite expected shortfall's superior relative performance. This is an important finding because it highlights the fact that neither risk measure is perfect.

4. Adjustment for autocorrelation in the data severely impedes tail risk capture in the crisis regime. Furthermore, the improvement of expected shortfall over value at risk under the crisis regime is extremely small when the autocorrelation-adjusted ARMA/GARCH data are used.

6.5 Other Interesting Results

While Section 6.4 summarizes the results pertaining scope of this study, several other interesting trends in the data may be worthy of future study.

6.5.1 Emerging vs. Mature Markets

The size of the tail parameter ξ appears to be significantly larger in emerging markets compared to in developed markets, without a significant corresponding increase in the risk measures. Although the trend is most obvious in the normal-regime log-returns data, it is also present to some extent in the normal-regime ARMA/GARCH data (see Tables 6.1 and 6.3). It would be interesting to examine whether or not the apparent extra tail risk is priced into emerging markets differently, or whether or not an increase in the number of both mature and emerging markets would validate this apparent trend.

6.5.2 Correlations in Crisis Regimes

This is most pronounced in Figures 6.3 and 6.4, where 9 of the 13 indices all have very tightly-grouped ξ (0.08 to 0.11), ES (3.9 to 5.5) and VaR (2.5 to 3.8) values. This cluster is consistent with the observation that returns converge in periods of stress (Grote, 2012).

The clustering effect appeared to be removed by the application of ARMA/GARCH to the data. However, it would be interesting to perform a multivariate analysis to see how the correlations between the indices vary through the onset of the crisis, and to see if the data would behave significantly differently when subjected to such an analysis. The ability of ARMA/GARCH to effectively model an apparent correlation phenomenon is noteworthy.

Relating back to section 6.5.1, the data not in the cluster in Figures 6.3 and 6.4 appears to be disproportionately made up of emerging markets (MEXBOL, TWSE and FMBKCLI). The inclusion of a broader array of emerging markets would be interesting here – for example, validation of this trend might suggest that the lower correlation observed in the crisis regime may be the result of less integration with developed markets.

7: Summary

This study compared the effectiveness of value at risk and expected shortfall in terms of their ability to capture tail risk in crisis and normal regimes. The data used were daily index values over a selected time period in each regime. Value at risk and expected shortfall were determined empirically from the data. Tail parameters were determined by applying extreme value theory and fitting the data to generalized Pareto distributions using maximum likelihood methods. Tail risk quantification followed the univariate method employed by Yamai and Yoshiba (2002b, 2005).

We find that, while neither risk measure captured tail risk perfectly in any regime, expected shortfall was found to have less tail risk than value at risk. Furthermore, both risk measures were found to be less effective at capturing tail risk in a crisis regime compared to a normal regime. They were particularly ineffective at capturing the tail risk in the stochastic component in a crisis regime, suggesting that unpredictable losses under such conditions cannot be effectively quantified by either risk measure.

The superior performance of expected shortfall indicates that it would be sensible for the financial industry and regulators to adapt risk measurement practices to work with expected shortfall instead of VaR. This will also require adaptation to the limitations of expected shortfall, such as the need for larger data sets. Furthermore, risk managers must be conscious of the fact that both risk measures cannot be relied upon to capture extreme tail behaviour, and expected shortfall must be used in conjunction with stress-testing to ensure that risks under such conditions are anticipated.

Reference List

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

Artzner, P., Delbaen, F., Eber, J.M., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9 (3), 203–228.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin.

Grote, Jim 2012. Going to one: Is diversification passé? Journal of Financial Planning 25(6), 20-25.

Yamai, Y., Yoshiba, T., 2002a. Comparative analyses of expected shortfall and VaR (2): Expected utility maximization and tail risk. Monetary and Economic Studies 20 (2), 95–115.

Yamai, Y., Yoshiba, T., 2002b. Comparative analyses of expected shortfall and VaR (3): Their validity under market stress. Monetary and Economic Studies 20 (3), 181–237.

Yamai, Y., Yoshiba, T., 2005. Value-at-risk versus expected shortfall: A practical perspective. Journal of Banking & Finance 29, 997–1015.