RISK MANAGEMENT

A Practical Guide

RISKMETRICS GROUP
WWW.RISKMETRICS.COM

Risk Management: A Practical Guide

First Edition (August 1999)

Foreword

The RiskMetrics Group is best known for its leadership in the development of transparent risk estimation methodologies and easy-to-use software tools. The founders of the Group have long recognized, however, that there is much more to risk management that just risk measurement. Indeed, perhaps too much public focus has been placed on the sophistication and apparent precision of risk estimation models, and not enough on the more important managerial and judgmental elements of a strong risk management framework. These include the clarity of risk policies, the strength of internal control, the degree of management discipline, the level of internal risk transparency, and ultimately, the experience and market knowledge of risk management professionals. No technical document, however complete and rigorous, can impart that experience and knowledge.

To contribute to a better understanding of these broader elements of risk management, and in response to frequent client inquiries, the experienced professionals of the RiskMetrics Group have developed this practitioner's oriented guide to Risk Management. While the details of the subject matter can, at times, be technical and complex, the essence of the guide is helping practitioners to get the right information on the right issues to the right people at the right time. Not with a view to producing a single right answer, but with a confidence that the right questions will then be asked, leading to the best informed, experienced judgments.

As with all other elements of risk management, the state-of-the-art guide will itself need to be open to continuous improvement, as new techniques are developed in response to ongoing innovations in markets and risk products. This first edition of the guide focuses on market risk analysis and reporting, while also touching upon closely related issues of counterpart risk reporting and external risk disclosures.

Stephen G. Thieke
Chairman
RiskMetrics Group

Acknowledgments

We would like to thank the many individuals who have helped to create and shape this document.

First, we would like to extend our gratitude to reviewers who have provided valuable feedback on the document: Steve Thieke, David Mengle, Joe Gentile, Shailesh Shah, Robert Young, Chris Leonard, Dorothee Fuhrmann, and Nanaaya Twum Danso from J.P. Morgan, Ken Kecevar from Morgan Stanley Dean Witter, Danilo Guaetolli from Banco Safra, Peter Zangari from Goldman Sachs, Helena Weinberg from Trema, Vincent Orrico from GARP, and David Perregaux from CBL Credit Corporation.

Thanks also to our consulting editor Paige Finger, and to Tracy Penwell and Ronald Hudson for especially thorough reviews and suggestions.

Thanks to everyone at RMG for feedback, ideas, and contributions, in particular Kris Wulteputte, Alvin Lee, Jongwoo Kim, Allan Malz, Gregg Berman, Ken Akoundi, Purvi Tailor, Jorge Mina, John Matero, Chris Finger, Pete Benson, and Nathan Shewmaker.

Special thanks to Ethan Berman of RMG, who conceived of the idea to write this Guide, and to our tireless editor Tatiana Kolubayev.

Author
Alan J. Laubsch
alan.laubsch@riskmetrics.com

Contributors
Andrew Ulmer
andrew.ulmer@riskmetrics.com

Table of contents

Introduction **xi**
Why we wrote this *Guide* xi
Who should read the *Guide* xi
General structure and overview xi

Part I Risk Methodology and Analysis

Chapter 1. Introduction to risk analysis **3**
1.1 History of Value-at-Risk 3
1.2 VaR, relative VaR, marginal VaR, and incremental VaR 4
1.3 Overview of risk methodologies 8
1.4 Confidence level scaling factors 11
1.5 Time scaling of volatility 12
1.6 Components of market risk 13
1.7 Basic dimensions of market risk 15
1.8 Summary 20

Chapter 2. Stress testing **21**
2.1 Why stress test 21
2.2 Two central questions for stress testing 22
2.3 How to use stress tests 23
2.4 What makes a good stress test 24
2.5 Forecasting time frame 26
2.6 How often to stress test 26
2.7 Steps for stress testing 26
2.8 Creating stress scenarios 27
2.9 Summary of stress tests 36

Chapter 3. Backtesting **39**
3.1 Why backtest 39
3.2 Backtesting VaR vs. actual P&L 39
3.3 Accounting for non-position taking income 41
3.4 Backtesting VaR vs. hypothetical trading outcomes 41
3.5 Interpreting backtesting results 42
3.6 Other factors to consider in analyzing backtests 43
3.7 External disclosures of backtests 44
3.8 Backtesting summary 44

Part II Risk Management and Reporting

Chapter 4. Practical problems risk managers face **49**
4.1 Risk reporting 49
4.2 How to use risk reports 50
4.3 What type of information is required 50

4.4	What risk solutions to choose	51
4.5	Summary of issues facing risk managers	53
Chapter 5.	**Generating a risk report**	**55**
5.1	What makes a good risk report	55
5.2	What are the major types of risk reports	58
5.3	How to organize a risk report	60
5.4	Time dimensions in risk reporting	60
5.5	Global bank case study	61
5.6	Leveraged fund case study	68
5.7	Investment manager case study	71
5.8	Corporate case study	74
5.9	Summary of risk reporting issues	79
Chapter 6.	**External risk disclosures**	**81**
6.1	Introduction	81
6.2	Emerging global standards for public disclosures	81
6.3	Voluntary risk disclosure for non-financial corporations	86
6.4	SEC disclosure requirements for derivatives	88
6.5	Summary	89
Chapter 7.	**Using risk information**	**91**
7.1	Linking risk and return	91
7.2	Risk and performance	91
7.3	Risk and capital	93
7.4	Summary	95
Chapter 8.	**Market data for risk reporting**	**97**
8.1	Type and quantity of market data	97
8.2	Deriving volatilities and correlations from raw historical data	98
8.3	Use of historical versus implied volatilities	99
8.4	Exponential weighting of time series	100
8.5	Log price change of GBP/DEM and 95% VaR estimates	100
8.6	What is good market data	100
8.7	The task of the risk data analyst	101
8.8	Where to get market risk data	102
8.9	Summary	102
Chapter 9.	**Position data for risk mapping**	**105**
9.1	The data collection process	105
9.2	What type of position information is required	106
9.3	Principles of cashflow mapping for interest rate risk	107
9.4	Mapping commodities	108
9.5	Mapping equities	108
9.6	Choosing a methodology	109
9.7	Summary	110
Chapter 10.	**Evaluating a risk software vendor**	**111**
10.1	How to choose a risk solution	111
10.2	Summary	113
10.3	Conclusion	114

Appendices

Appendix A. Risk-based limits **117**
Appendix B. Credit exposure of market-driven instruments **119**
Appendix C. The independent risk oversight function **125**

Glossary of terms **127**
Resources **137**
Index **139**

Introduction

Why we wrote this *Guide*

This *Guide* evolved from common client questions about market risk management. While there is a significant amount of quantitative research and high-level risk management literature, there are few practical resources at the risk manager's disposal. The *Guide* addresses the basic issues risk managers face when implementing a market risk measurement process.

We are publishing the *Guide* in our continuing effort to promote improvements in the discipline of risk management and to help our clients develop better risk reporting processes.

The *Guide* focuses on practical issues that arise in the process of risk analysis and reporting. The three major phases of risk reporting consist of (a) compiling position and market data in a timely manner, (b) applying appropriate methodologies to calculate risk (including stress testing), and (c) summarizing portfolio risks in a concise and complete manner. We also include advice on model backtesting.

The world is moving from reporting risk as a single number to viewing a broader ***Picture of Risk***. We emphasize the importance of applying several methodologies to explore different dimensions of risk.

Who should read the *Guide*

The *Guide* is geared for risk monitors and analysts who are responsible for implementing a risk reporting process. After studying the *Guide*, readers should know (a) how to measure risk, (b) why it's important to measure risk—the "so what" or "value-added" of it, and (c) how to present and communicate risk information to management and other interested constituents.

As an introduction, we review basic principles of market risk measurement without relying heavily on statistics and formulas. Quantitatively oriented readers are encouraged to learn the details of risk methodologies in the *RiskMetrics Technical Document*, which are cited throughout the *Guide*.

Our main focus is on the practical issues of communicating about risk. The application of rules and procedures for risk control, while important, are not part of the *Guide*. Through case studies, we consider the risk reporting needs of several types of financial institutions (banks, hedge funds, and asset managers) and corporations. We show what type of input data is needed and give advice on designing effective risk reports. Throughout the *Guide*, we give practical illustrations of how these concepts and data are implemented in real solutions.

For novice risk managers who wish to get a broad overview of corporate risk management, we recommend our on-line *Managing Risk* course, which features live RiskMetrics tutorial support.

General structure and overview

This document is organized in two sections.

Part I addresses risk methodology and analysis, and consists of Chapters 1 through 3:

Chapter 1 introduces the **Value-at-Risk (VaR)** framework for measuring market risk and highlights key input parameters and assumptions. We review the three major methodologies for estimating VaR: **parametric**, **Monte Carlo simulation**, and **historical simulation.**

Chapter 2 gives an overview of *stress testing*, which is a crucial discipline in risk measurement. We emphasize characteristics of effective stress tests and introduce several approaches for creating stress scenarios, including historical and predictive scenario generation.

Chapter 3 addresses *backtesting* of risk models: why it's important, how to do it, when to do it, and what to look for.

Part II addresses risk management and reporting and consists of Chapters 4 through 10:

Chapter 4 defines common problems risk managers face when implementing market risk reporting processes:

- How to produce relevant market risk reports
- How to use risk information
- How to obtain appropriate data
- How to evaluate software for analyzing and reporting risk
- Whether to build or buy risk solutions

In the following chapters, we address these issues and propose solutions.

Chapter 5 describes best practices for *risk reporting*. We make suggestions for designing risk reports (i.e., format, content, and organization), and show sample case study reports representing four types of companies: banks, hedge funds, asset managers, and traditional corporations.

Chapter 6 addresses *external reporting*. We discuss emerging global standards for public risk disclosures of financial and non-financial companies and show examples of actual risk disclosures from leading institutions. We specifically review BIS disclosure recommendations and SEC disclosure requirements.

Chapter 7 discusses how to use risk information to link risk with performance evaluation and capital. We introduce the Sharpe ratio for measuring return on risk of realized revenues. Then, we discuss BIS regulatory market risk capital requirements, and introduce the topic of economic capital allocation.

Chapter 8 focuses on *market data* needed for calculating risk. We define what constitutes good market data and review best practices for data analysts. We review the process of transforming raw historical rates into volatility and correlation forecasts and discuss the use of implied volatility forecasts.

Chapter 9 reviews the *position data* collection process, and the type of information required from position management systems. To simplify the data management process, we introduce the concept of cashflow mapping for fixed income, FX and commodity instruments and show several approaches for treating equities.

Chapter 10 gives advice on choosing a *risk software* vendor. We emphasize the importance of defining risk management objectives and needs up front and propose key evaluation criteria for risk solutions.

In the **Appendices**, we discuss risk based limits, credit exposure of market driven instruments, and the responsibilities of the independent corporate risk management function. We also

provide a **glossary** of risk terminology and a list of resources that includes risk associations and suggested reading.

Throughout this *Guide* we use the following typographic conventions: **boldfaced** terms are defined in the **Glossary**; underlined text indicates a web site hyperlink in the on-line version of the *Guide*.

All reports and graphs were generated using the RiskMetrics RiskManager application.

Part I

Risk Methodology and Analysis

Chapter 1. Introduction to risk analysis

1.1 History of Value-at-Risk

VaR was pioneered by major U.S. banks in the ’80s, as the derivative markets developed. The birth of derivatives represented a new challenge for risk management because traditional measures of exposure were clearly inadequate. For example, two derivative contracts with the same **notional** value could have very different risks. With VaR, banks had developed a general measure of economic loss that could equate risk across products and aggregate risk on a portfolio basis.

Another important stimulus to the development of VaR was the move toward mark-to-market, both for cash instruments and derivatives. Prior to that, the emphasis was on net interest income, where the common risk measure was repricing gap. As trading increased, duration analysis took over, but duration's inadequacies led to the adoption of VaR.[1]

Definition of VaR

VaR is defined as the predicted worst-case loss at a specific confidence level (e.g., 95%) over a certain period of time (e.g., 1 day). For example, every afternoon, J.P. Morgan takes a snapshot of its global trading positions to estimate its Daily-Earnings-at-Risk (DEaR), which is a VaR measure that Morgan defines as the 95% confidence worst-case loss over the next 24 hours due to adverse market movements.

VaR works on multiple levels

The elegance of the VaR solution is that it works on multiple levels, from the position-specific micro level to the portfolio-based macro level. VaR has become a common language for communication about aggregate risk taking, both within an organization and outside (e.g., with analysts, regulators, rating agencies, and shareholders).

Virtually all major financial institutions have adopted VaR as a cornerstone of day-to-day risk measurement. Below is an excerpt describing Chase Manhattan’s use of VaR.

> *Statistical models of risk measurement, such as VAR, allow an objective, independent assessment of how much risk is actually being taken. Chase’s historic simulation methodology permits consistent and comparable measurement of risk across instruments and portfolios, irrespective of the level of aggregation. Historical simulation also makes it easy to examine the VAR for any desired segment of the total portfolio and to examine that segment’s contribution to total risk. The VAR calculations are performed for all material trading portfolios and market risk-related asset/liability management (“ALM”) portfolios. Results are reported at various levels of detail by business unit and in the aggregate.*
>
> – Chase 1998 *Annual Report*

The application of VaR analysis and reporting now extends to non-financial corporations, which has resulted in the adoption of related “at-risk” measures, such as **Earnings-at-Risk** (**EaR**), **Earnings-Per-Share-at-Risk** (**EPSaR**), and **Cash-Flow-at-Risk** (**CFaR**).

[1] For a more detailed discussion, see the *RiskMetrics Technical Document*, Chapter 2, “Historical perspective of VaR,” p. 21.

These measures take into account special considerations of the corporate environment, such as the use of accrual vs. mark-to-market accounting and hedge accounting for qualifying transactions. Furthermore, non-financial corporations focus on longer-term impact of risk on cash flows and earnings (quarterly or even annual) in the budgeting and planning process. The application of risk measurement in the corporate environment is fully discussed in the *CorporateMetrics Technical Document*.

VaRBaR

Risk measurement in the corporate environment

In the corporate environment, uncertainty in future earnings and cash flow is caused not only by uncertainty in a company's underlying business (e.g., sales volumes), but also by a number of other risks, including market risk. Market risk can arise from a number of factors, including foreign exchange exposures, interest rate exposures, commodity price-sensitive revenues or expenses, pension liabilities, and stock option plans. CorporateMetrics, which focuses on market risk, provides a framework centered on the key financial results that corporations monitor.

To address a company's need to quantify the impact of market risk on earnings and cash flow, CorporateMetrics defines the following measures of volatility:

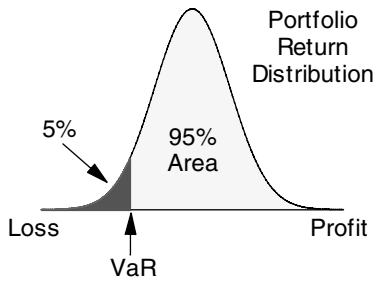
Earnings-at-Risk (EaR). *The maximum shortfall of earnings, relative to a specified target, that could be experienced due to the impact of market risk on a specified set of exposures, for a specified reporting period and confidence level. Since earnings are also usually reported on a per share of equity basis, many companies may prefer to use an* ***Earnings-per-Share-at-Risk (EPSaR)*** *measure.*

Cash-Flow-at-Risk (CFaR). *The maximum shortfall of net cash generated, relative to a specified target, that could be experienced due to the impact of market risk on a specified set of exposures, for a specified reporting period and confidence level.*

Source: CorporateMetrics Technical Document

1.2 VaR, relative VaR, marginal VaR, and incremental VaR

Assuming 95% confidence and a 1-day horizon, a VaR of $11 million means that, on average, only 1 day in 20 would you expect to lose more than $11 million due to market movements.



This definition of VaR uses a 5% risk level (95% confidence): You would anticipate that losses exceeding the VaR amount would occur 5% of the time (or losses less than the VaR amount would occur 95% of the time).

VaR is a flexible risk measure:

- VaR can be specified for various horizons (generally between 1 day and 1 month) and confidence levels (generally between 90% and 99%).
- VaR can be expressed as a percentage of market value or in absolute currency terms (e.g., USD).

There are three related VaR measures: (a) **relative VaR**, (b) **marginal VaR**, and (c) **incremental VaR**.

Relative VaR

Relative VaR measures the risk of underperformance relative to a pre-defined **benchmark**, such as the S&P 500 Index. It is relevant to many institutional investors, including investment managers and mutual funds, because their performance is often compared to a target benchmark. For example, an Emerging Markets investment manager might have used the J.P. Morgan EMBI+ index as a performance benchmark. If the investment manager's portfolio rose 9% while the EMBI+ index rose 10%, we would say that she underperformed her benchmark by 1%.

Assuming 99% confidence, a 1-month relative VaR of $8 million means that on average, only 1 month in 100 would you expect to underperform your benchmark by more than $8 million due to market movements. Relative VaR is also commonly expressed as a percentage of present value.

An investment manager's risk report might show the following:

Portfolio	VaR*, %	Benchmark†	Relative VaR*, %
U.S. Equities	10	S&P 500 Index	3
Global Equities	11	MS EAFE Index	1
Global Fixed Income	5	JPM GBI+ Index	4
Total Portfolio	8	Custom Global Index‡	3

* 1-month horizon, 99% confidence.

† Refers to the investment manager's target benchmark index used to evaluate relative performance.

‡ Fund managers can have custom benchmarks (e.g., a specific mix of existing benchmarks).

For example, for the U.S. Equities portfolio, the worst-case loss at 99% confidence is equal to 10% of the portfolio's current market value (i.e., 1% probability that losses exceed 10% of market value), whereas the worst-case monthly underperformance, relative to the portfolio's S&P 500 benchmark, is only 3% (i.e., 1% probability of underperforming the benchmark by 3% or more).

This report reveals important differences between *VaR* and *relative VaR*. Global Equities has the highest stand-alone VaR (11%), but considering its benchmark, the smallest relative VaR (1%). On the other hand, the Global Fixed Income portfolio has the smallest stand-alone VaR (5%), but the largest relative VaR (4%).

The 4% relative VaR would be of most concern to a risk monitor, since it deviates most from the benchmark.[2] A large relative VaR may arise when a manager takes positions that do not

[2] Deviation from benchmark could be justified if it is within pre-specified relative VaR limits (and if excess return on risk is acceptably high).

track the benchmark closely. For example, the Global Fixed Income portfolio manager may have underweighted certain countries within the index. As this example shows, portfolio managers using benchmarks should set relative VaR limits (e.g., keep relative VaR below 3%).

Marginal VaR

Marginal VaR measures how much risk a position adds to a portfolio. Specifically, marginal VaR measures how much portfolio VaR would change if the position were removed entirely, (i.e., VaR with position minus VaR without position).[3] Note that marginal VaR can be computed for both absolute VaR and relative VaR.

A marginal risk report might reveal the following:

Position	Market value, MM	VaR, MM	Marginal VaR, MM
Yahoo! Equity	$ 25.1	$ 0.9	$ 0.5
10-year U.S. T-Note	$ 98.2	$ 0.8	$ 0.6

This report implies that, although the Yahoo! stock position has the greater stand-alone VaR (USD 0.9 million), its contribution to portfolio VaR is less than the T-bond's contribution (USD 0.5 million). Often, the largest stand-alone risk positions are not the greatest contributors of risk. This is especially true for hedges, which have a negative marginal VaR.

Marginal VaR is useful for measuring which position (or risk category) is the largest contributor to portfolio risk. It can help answer the question of which position to eliminate **entirely** in order to most effectively reduce risk.

Incremental VaR

Incremental VaR is closely related to marginal VaR. Marginal VaR measures the difference in portfolio risk brought about by removing an entire position, whereas incremental VaR measures the impact of small changes in position weighting. For example, we can estimate incremental VaR by (a) increasing a position weight by 1 dollar and measuring the change in diversified portfolio VaR, and (b) multiplying this change by the position weighting.[4] The sum of all incremental VaRs adds up to the total diversified portfolio VaR. Therefore, incremental VaR may be used to calculate percentage contribution to risk.

One of the most common uses for incremental VaR is to generate reports that rank contribution to risk hedging opportunities. Incremental VaR is useful for identifying best candidates for gradual risk reduction (i.e., where the question is not which position to unwind entirely, but rather which position to partially hedge).

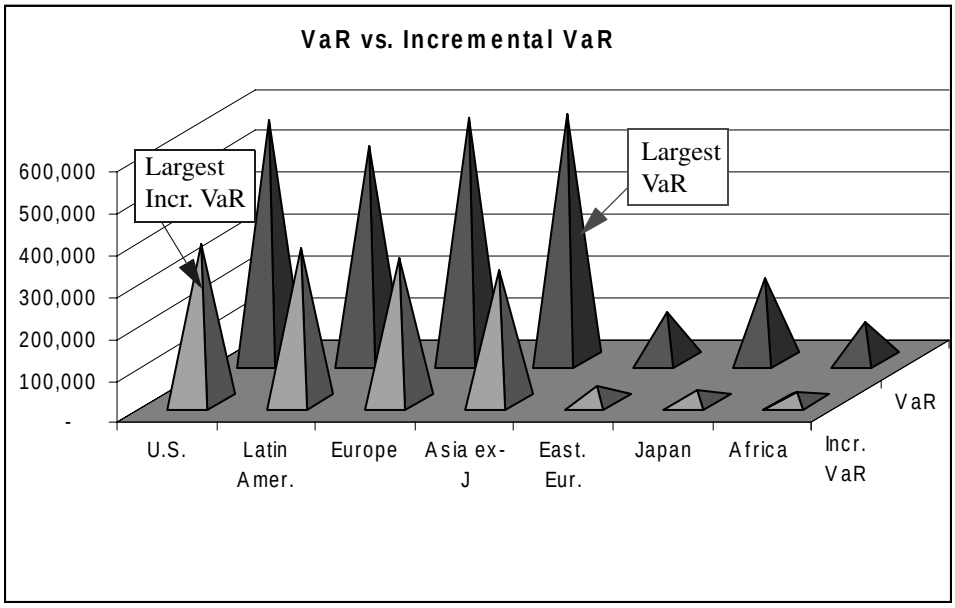
[3] Some practitioners define marginal VaR as incremental VaR.

[4] There is significant confusion in the marketplace about the definition of marginal VaR vs. incremental VaR. Some firms define marginal VaR as the VaR that a position adds incrementally, as measured by the effect on VaR if the position size is increased by a very small amount (this is defined as "incremental VaR" by RMG). Marginal VaR as defined by RMG is the difference in VaR assuming that the position is removed entirely from the portfolio.

The following is a regional risk contribution report, which ranks risk contributors according to their incremental VaR.

Regional Risk Summary **May 1, 1999**

RMG



Risk Contribution Report	Present Value	VaR	Marginal VaR	Incremental VaR	Contribution to Risk
U.S.	71,774,216	574,194	222,075	378,341	25%
Latin America	10,258,887	512,944	220,114	369,626	25%
Europe	64,600,480	581,404	204,358	343,237	23%
Asia ex-Japan	12,693,840	589,734	196,046	317,346	21%
East. Europe	1,948,860	116,932	31,050	40,322	3%
Japan	19,569,450	195,694	48,012	30,068	2%
Africa	4,669,370	93,387	24,423	24,163	2%
Divers. Benefit		*(1,161,186)*			
Aggregate	**185,515,103**	**1,503,103**		**1,503,103**	**100%**

Comments

- Even though Asia ex-Japan is the largest single relative VaR position, it is only the fourth largest contributor to risk, as measured by the percentage of incremental VaR (21%). The best three opportunities for reducing risk through hedges lie in U.S., Latin America, and Europe.
- **VaR** refers to 95% worst-case loss over 1 day due to adverse movements in market rates.
- **Marginal VaR** is the change in aggregate portfolio VaR when the position is taken out of the portfolio.
- **Incremental VaR** is the change in VaR that results when position size is increased by a small amount, multiplied by the portfolio weighting. The sum of all incremental VaRs adds up to total diversified VaR.
- ***Diversification Benefit*** is equal to diversified (aggregate) VaR minus the sum of all individual VaRs. It measures the risk reduction achieved through diversification (i.e., correlation being lower than 1) between risk categories.

1.3 Overview of risk methodologies

Market risk models are designed to measure potential losses due to adverse changes in the prices of financial instruments. There are several approaches to forecasting market risk, and no single method is best for every situation. Over the last decade, Value-at-Risk (VaR) models have been implemented throughout the financial industry and by non-financial corporations, as well. Inspired by modern portfolio theory, VaR models forecast risk by analyzing historical movements of market variables. To calculate VaR, one can choose from three main methods: **parametric**, **historical simulation**, and **Monte Carlo simulation**.[5] Each method has its strengths and weaknesses, and together they give a more comprehensive perspective of risk. Note that we include the **portfolio aggregation** methodology as a subcomponent of historical simulation.[6]

Before comparing these three approaches for calculating VaR, we add a quick note about linear vs. non-linear instruments. A financial instrument is nonlinear if its price changes disproportionately relative to a movement in the underlying asset. The risk of **nonlinear instruments** (e.g., options) is more complex to estimate than the risk of **linear instruments** (e.g., traditional stocks, bonds, swaps, forwards, and futures). To account for the discontinuous payoff of nonlinear instruments like options, risk simulations should use full valuation formulas (e.g., Black-Scholes) rather than first order sensitivities (e.g., **delta**).

The following table describes the three main methodologies for calculating VaR.

Methodology	Description	Applications
Parametric	Estimates VaR with equation that specifies parameters such as volatility, correlation, delta, and **gamma**.	Accurate for traditional assets and linear derivatives, but less accurate for nonlinear derivatives.
Monte Carlo simulation	Estimates VaR by simulating random scenarios and revaluing positions in the portfolio.	Appropriate for all types of instruments, linear and nonlinear.
Historical simulation	Estimates VaR by reliving history; takes actual historical rates and revalues positions for each change in the market.	

Note that Monte Carlo and historical simulations are mechanically identical in that they both revalue instruments, given changes in market rates. The difference lies in how they generate market scenarios. Monte Carlo simulation generates random hypothetical scenarios, while historical simulation takes actual past market movements as scenarios.

From an end-user perspective, the important point to remember is that if you have significant nonlinear exposures in your portfolio, a simulation approach with full position re-pricing will generally be more accurate than a parametric approximation for estimating VaR—however, at the cost of greater complexity.

[5] See the *Managing Risk™ Course*— three methodologies for calculating VaR.

[6] Portfolio aggregation is described by Zangari in the Q2 ’97 *RiskMetrics Monitor* article “A general approach to calculating VaR without volatilities and correlations.”

A summary of the strengths and weaknesses of each methodology is given below:

Methodology	Advantage	Disadvantage
Parametric	• Fast and simple calculation • No need for extensive historical data (only volatility and correlation matrix are required)	• Less accurate for nonlinear portfolios, or for **skewed distributions**
Monte Carlo simulation	• Accurate* for all instruments • Provides a full distribution of potential portfolio values (not just a specific percentile) • Permits use of various distributional assumptions (normal, T-distribution, normal mixture, etc.), and therefore has potential to address the issue of **fat tails** (formally known as "leptokurtosis") • No need for extensive historical data	• Computationally intensive and time-consuming (involves re-valuing the portfolio under each scenario) • Quantifies fat-tailed risk ***only if*** market scenarios are generated from the appropriate distributions
Historical simulation	• Accurate* for all instruments • Provides a full distribution of potential portfolio values (not just a specific percentile) • No need to make distributional assumptions (although parameter fitting may be performed on the resulting distribution) • Faster than Monte Carlo simulation because less scenarios are used	• Requires a significant amount of daily rate history (note, however, that sampling far back may be a problem when data is irrelevant to current conditions, e.g., currencies that have already devalued) • Difficult to scale far into the future (long horizons) • Coarse at high confidence levels (e.g., 99% and beyond) • Somewhat computationally intensive and time-consuming (involves re-valuing the portfolio under each scenario, although far less scenarios are required than for Monte Carlo) • Incorporates tail risk only if historical data set includes tail events

* Accurate if used with complete pricing algorithm.

All three approaches for estimating VaR have something to offer and can be used together to provide a more robust estimate of VaR. For example, a parametric approach may be used for instant risk measurement during a trading day, while a simulation approach may be used to provide a fuller *Picture of Risk* (in particular, nonlinear risks) by the end of the trading day.[7]

Importance of model transparency

At RMG, we emphasize that risk models should not be viewed as black boxes that produce magic numbers. Risk methodologies should be clear, and risk managers should understand the key parameters and fundamental assumptions of each approach. Don't be lulled into a false sense of security through complicated mathematical formulas, even if they were derived by rocket scientists. There is no single correct answer to risk, and risk methodologies are constantly being refined and new approaches invented. This means that risk managers must continually question assumptions, search for new approaches for measuring risk, and keep abreast of the latest research. It's the risk manager—not just numbers— that makes risks transparent.

[7] Distributed computing solutions should eventually enable fast simulation of risks throughout the trading day.

Limitations of VaR

It's important to realize that all three approaches for measuring VaR are limited by a fundamental assumption: that future risk can be predicted from the historical distribution of returns.[8] The parametric approach assumes normally distributed returns, which implies that parametric VaR is only meant to describe "bad" losses on a "normally bad" day. While Monte Carlo simulation offers a way to address the fat-tail problem by allowing a variety of distributional assumptions, volatility and correlation forecasts are still based on statistical fitting of historical returns. While historical simulation performs no statistical fitting, it implicitly assumes that the exact distribution of past returns forecasts future return distributions. This implies that all three approaches are vulnerable to regime shifts, or sudden changes in market behavior. Stress testing should therefore explore potential regime shifts to best complement VaR and create a robust *Picture of Risk.*

Parameters for VaR analysis

Before calculating VaR, we need to specify three parameters: (a) confidence level, (b) forecast horizon, and (c) base currency.

(a) Confidence level

We first choose a confidence level or probability of loss associated with VaR measurement. Confidence levels generally range between 90% and 99%. RiskMetrics assumes 95% confidence as a baseline, but gives users the flexibility to choose other levels. Rather than choose a single parameter, some firms use several confidence levels (e.g., 95% and 99%) and forecast horizons (e.g., 1 day and 1 year).

How to choose a confidence level

There is nothing magical about confidence levels. In choosing confidence levels for market risk, companies should consider worst-case loss amounts that are large enough to be material, but that occur frequently enough to be observable. For example, with a 95% confidence level, losses should exceed VaR about once a month (or once in 20 trading days), giving this risk statistic a visceral meaning. Risk takers are thus encouraged to compare their daily P&Ls against their VaR and consider return on risk.

Some maintain that using a higher level of confidence, such as 99.9%, would be more conservative. One might also reason, however, that a higher confidence level can lead to a false sense of security. A 99.9% VaR will not be understood as well or taken as seriously by risk takers and managers because losses will rarely exceed that level (we expect a loss of that magnitude to occur about once in four years). Furthermore, due to fat-tailed market returns, a high confidence level VaR is difficult to model and verify statistically. VaR models tend to lose accuracy after the 95% mark and certainly beyond 99%. Note, however, that when using VaR for measuring credit risk and capital, we should apply a 99% or higher confidence level VaR because we are concerned with low-probability, event-driven risks (i.e., **tail risk**).

We can't rely on models to do all the "thinking" for us. Beyond a certain confidence level, rigorous stress testing becomes more important than statistical analysis. The choice of 95% confidence level at J.P. Morgan goes back to former CEO Dennis Weatherstone, who reputedly said, "VaR gets me to 95% confidence. I pay my risk managers good salaries to look after the remaining 5%."

[8] Some argue that using implied volatilities bases risk prediction on market expectations as opposed to past market movements. For a portfolio view of risk, however, historical correlations of market returns must still be applied, as it is nearly impossible to get such information from option prices.

(b) Forecast horizon

Generally, active financial institutions (e.g., banks, hedge funds) consistently use a 1-day forecast horizon for VaR analysis of all market risk positions. For banks, it simply doesn't make sense to project market risks much further because trading positions can change dynamically from one day to the next. On the other hand, investment managers often use a 1-month forecast window, while corporations may apply quarterly or even annual projections of risk.

VaRBaR

Applying longer horizon for illiquid assets

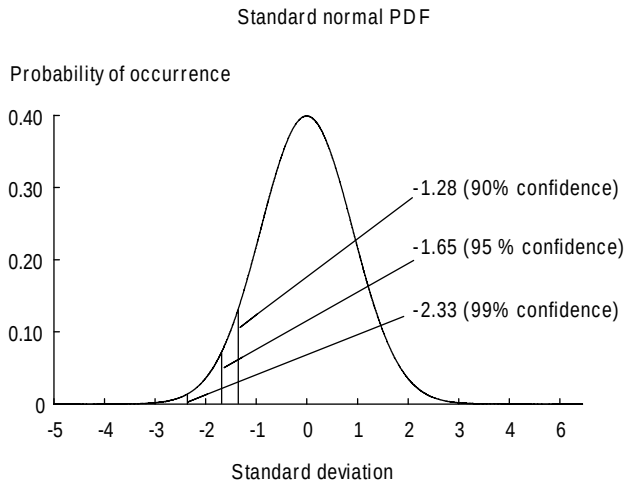
Instead of applying a single horizon, some firms use different forecast horizons across asset classes to account for liquidity risk. One might argue that the unwind period for an illiquid emerging markets asset is much longer than for a G7 Government bond, and that one should therefore use a longer horizon (e.g., 1-week) for emerging markets. However, a better solution is to treat market risk and liquidity risk as separate issues. Currently, the topic of liquidity risk is a hot research topic, and new quantitative methodologies are being developed. Simply using a longer time horizon for illiquid assets is not sufficient, and confuses liquidity risk with market risk. Having a standard horizon for VaR across asset classes facilitates the risk communication process and allows "apples-to-apples" comparison for market risk across asset classes.

(c) Base currency

The base currency for calculating VaR is typically the currency of equity capital and reporting currency of a company. For example, Bank of America would use USD to calculate and report its worldwide risks, while the United Bank of Switzerland would use Swiss francs.

1.4 Confidence level scaling factors

Standard deviations can be used to estimate lower-tail probabilities of loss when the parametric approach to measuring risk is used. Lower-tail probability of loss refers to the chance of loss exceeding a specified amount.



Because returns tend to cluster around the mean, larger standard deviation moves have a lower probability of occurring. To arrive at the tail probability of loss levels and implied VaR confidence levels, we use standard deviations (confidence level scaling factors). This chart shows three confidence level scaling factors and their associated tail probability of loss levels.

Assuming normality, we can easily convert one confidence level (C.L.) to another. For example, we can take J.P. Morgan's 95% confidence level VaR and translate it to the BIS standard of 99% confidence level through a simple multiplication, as shown in the following table.

Reference VaR	C.L.	C.L. scaling factor
JPM VaR	95%	1.65
BIS VaR	99%	2.33
Converting JPM to BIS VaR	95% to 99%	JPM VaR × 2.33/1.65

Views of the U.S. Federal Reserve

For example, in 1998 J.P. Morgan reported that its maximum VaR was USD 55 million, which scales to a VaR of approximately USD 78 million at 99% confidence, assuming normality. Note, however, that regulators, in particular, the U.S. Federal Reserve are increasingly discouraging this simple conversion, because the assumption of normally distributed P&Ls is often an oversimplification (especially when portfolios contain non-linear positions).

1.5 Time scaling of volatility

We know that risk increases with time: the longer we hold a position, the greater the potential loss. But unlike expected returns, volatility does not increase linearly with time. Long-horizon forecasting is complicated due to trending, **autocorrelation**, **mean reversion** of market returns, and the interrelationship of many macroeconomic factors. **Autocorrelation** refers to correlation between successive-days' returns, and **mean reversion** is the tendency for time series to revert to a long-term average (this is observed especially for interest rates). For research on long-horizon forecasting, see the *LongRun Technical Document*.

Square root of time scaling

You may need to time scale VaR estimates, for example when converting a daily VaR to a 10-day horizon regulatory VaR standard. A commonly used method is the **square root of time scaling,** which roughly extrapolates 1-day volatilities as well as 1-day VaR to longer horizons. The method assumes that daily price moves are independent of each other, and that there is no mean reversion, trending, or autocorrelation in markets. Note that we use the number of trading days, as opposed to actual days to scale volatility (5 trading days per week, and 21 days per month).

For example,

- Weekly volatility $= \text{daily volatility} \times \sqrt{5}$
  $= \text{daily volatility} \times 2.24$
- Monthly VaR $= \text{1-day VaR} \times \sqrt{21}$
  $= \text{1-day VaR} \times 4.58$

Views of the U.S. Federal Reserve

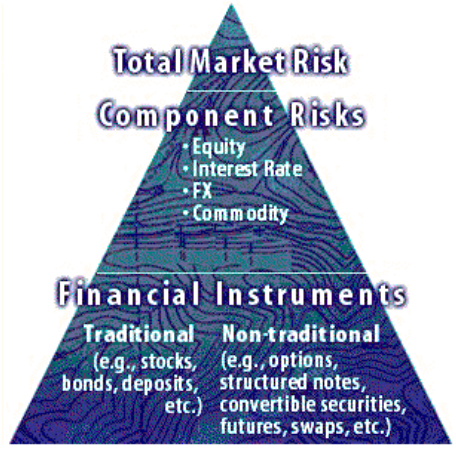
This simple time scaling approach can be useful for converting 1-day management VaR figures to 10-day BIS regulatory VaR standards. Note, however, that the U.S. Federal Reserve is not supportive of this approach, which has prompted some institutions to adopt more accurate methodologies.[9]

[9] For example, the Fed's discouragement of square root of time scaling has prompted J.P. Morgan to calculate 10-day horizon regulatory VaR by basing volatility forecasts on overlapping 10-days of price changes for a 2-year history.

1.6 Components of market risk

Definition of market risk

The BIS defines market risk as "the risk that the value of on- or off-balance-sheet positions will be adversely affected by movements in equity and interest rate markets, currency exchange rates and commodity prices." The main components of market risk are therefore equity, interest rate,[10] FX, and commodity risk.



Source: Managing Risk course

At the top of the pyramid, we have *total market risk*, which is the aggregation of all component risks.

In the middle of the pyramid, we see how financial instruments are driven by the underlying *component risks*.

At the lowest level, market risk arises from fluctuating prices of *financial instruments*.

Residual risks

In addition to market risk, the price of financial instruments may be influenced by the following **residual risks**: **spread risk**, **basis risk**, **specific risk**, and **volatility risk**.[11]

Spread risk is the potential loss due to changes in spreads between two instruments. For example, there is a credit spread risk between corporate and government bonds.

Basis risk is the potential loss due to pricing differences between equivalent instruments, such as futures, bonds and swaps.[12] Hedged portfolios are often exposed to basis risk.

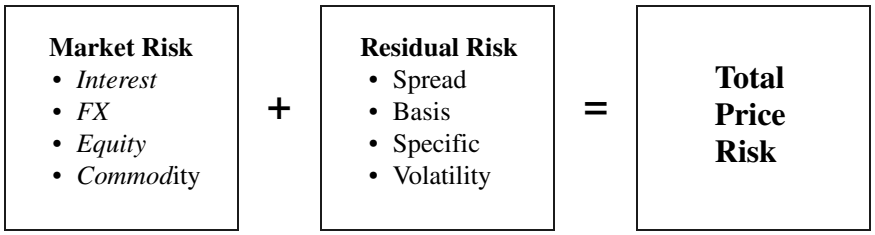
Specific risk refers to issuer specific risk, e.g., the risk of holding Yahoo! stock vs. an S&P 500 futures contract. How to best manage specific risk is a topic of debate. Note that according to the Capital Asset Pricing Model (**CAPM**), specific risk is entirely diversifiable.

Volatility risk is defined as potential loss due to fluctuations in implied option volatilities and is often referred to as "vega risk." Short option positions generally lose money when volatility spikes upward.

[10] Curve risk is a subcomponent of interest rate risk and captures exposure to changes in the shape of a yield curve by calculating interest rate risk for various time buckets.
[11] Some consider credit spread risk and volatility risk to be market risk.
[12] Some consider basis risk as a subcomponent of spread risk.

Total risk

To determine the total price risk of financial instruments, we aggregate market risk with residual risk:



Diversification

Risk is not additive. Total risk is less than the sum of its parts because of diversification between different assets and risk components (i.e., correlation would never be 1). For example, if a USD based investor holds a JPY denominated bond, she is exposed to rising Japanese interest rates and devaluation of JPY relative to USD. Clearly, her total risk is not just the interest rate and FX risk added together, because the likelihood that interest and FX rates both move out of her favor at the same time is less than 100%. This effect is described as **diversification benefit**. Note that we expect high diversification benefit between market and residual risk, due to low correlation.

Definition

Diversification benefit is defined as *total risk* minus the sum of all individual risk components.

Example

The concept of market risk, residual risk and diversification benefit is illustrated in the following risk report:

Portfolio VaR analysis, (All numbers in USD 000s)	**Market risk**				**Residual risk**			***Divers. Benefit***	**Total risk**
	IR	FX	Equity	Cmdty.	Specific	Spread	Vol.		
Aggregate Portfolio	19	94	0	123	400	12	72	***–295***	425
Diversification Benefit	*–198*	*–151*	*–240*	*0*	*–26*	*–2*	*–61*	–	***–266***
Long DAJ Call	2	74	120	–	310	–	50	*–220*	336
Short DAX Call	2	74	120	–	16	–	25	*–114*	123
Long DAJ Callable Bond	111	95	–	–	88	11	20	*-234*	91
Short 5-yr Euro Swap	100	2	–		12	3	–	*–105*	12
Long option on WTI future	2	–	–	123	–	–	38	*–34*	129

Comments

1. Notice how the Short DAX Call hedges the market risk of the long DAJ (DaimlerChrysler) call, but does not affect firm-specific risk ($310). Note that the *Diversification Benefit* reflects the hedge effect between different instruments and risk types.
2. The DAJ Callable Bond is mostly IR hedged with the short Euro Swap, but FX and specific risk is not hedged.
3. The WTI futures position is unhedged.

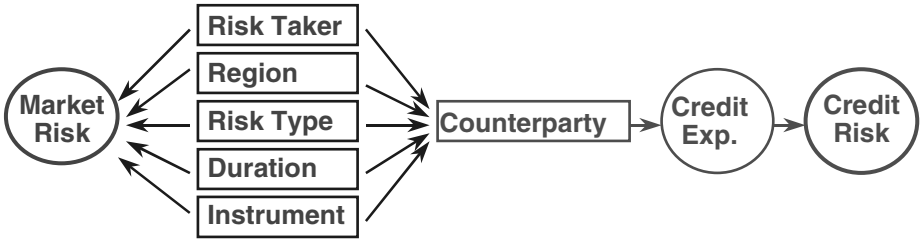
1.7 Basic dimensions of market risk

Risk can be analyzed in many dimensions. Typically, we quantify risk concentrations by:

Dimension	Example
Risk taker	Business unit, desk or portfolio
Risk type	Equity, interest rate, FX, and commodity
Country or region	Europe, Americas, Asia Pacific
Maturity or duration	One week, 1 month, 3 months, 6 months
Instrument type or instrument	Options, forwards, futures, cash
Counterparty	Citibank, Japanese Banks, Thai Corporations

The diagram below illustrates the interrelationship of these risk dimensions. Companies analyze market risk by risk taker, region, risk type, duration (for interest rate risk), and instrument. Fluctuations in market rates can also give rise to counterparty credit exposure and credit risk. Counterparty trading limits should be in place to limit credit exposure due to market driven instruments, such as swaps and forwards. The management of credit exposure for market driven instruments is discussed further in Appendix B.

Key risk dimensions giving rise to market and credit exposures

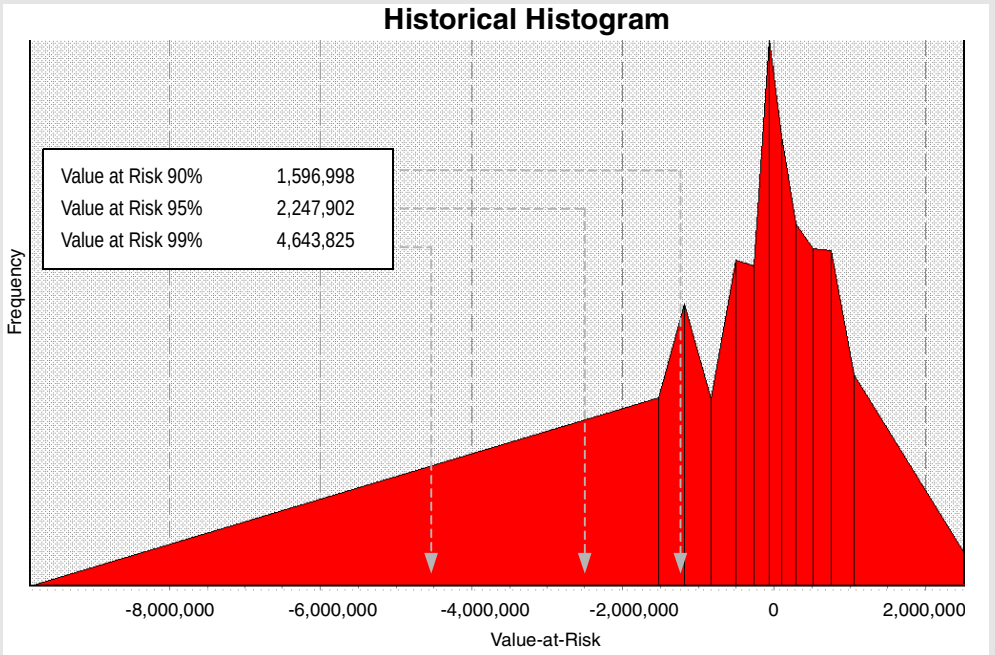


The sample reports in the next pages illustrate different ways to analyze market risk.

Example 1 One of the most common ways to dissect portfolio risk is by region and risk type, as demonstrated below.

VaR: Region vs. Risk Type **April 28, 1999**

RMG



95%, 1-day VaR	Aggregate	*Diversification Benefit*	FX Risk	Interest Rate Risk	Equity Risk
Aggregate	$ 737,190	***–106,562***	178,841	373,118	291,794
Divers. Benefit	***–103,008***		*–2,616*	*–46,981*	*–33,371*
North America	145,729	*–24,054*		97,865	71,919
Europe	165,606	*–120,027*	28,474	82,803	65,191
Japan	104,942	*–121,467*	52,232	52,471	38,334
Emerging Asia	22,727	*–27,185*	12,667	11,364	7,346
Emerging Europe	163,682	*–112,285*	26,552	81,841	61,698
Latin America	187,512	*–171,010*	61,532	93,756	70,676

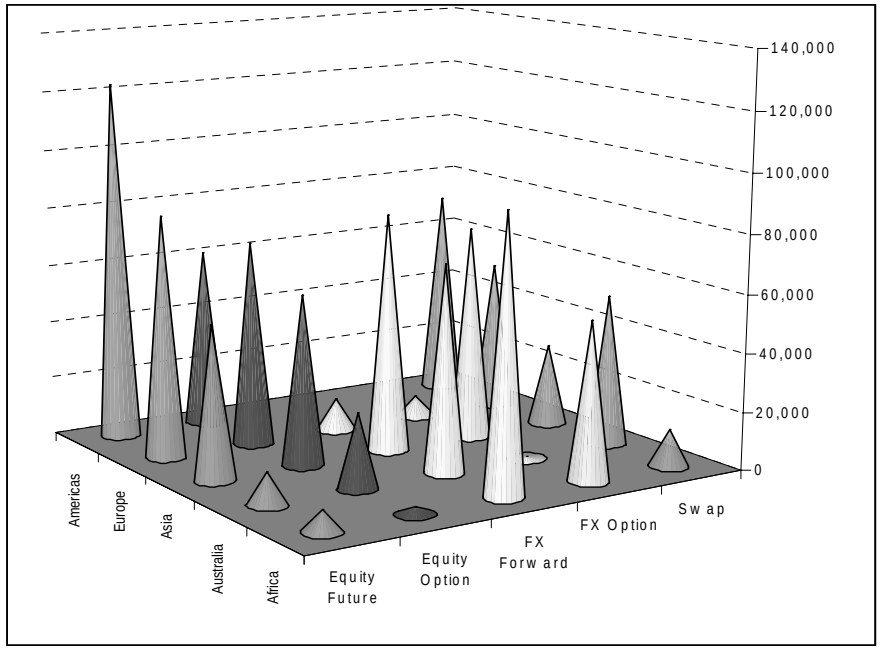
Comments

- Greatest aggregate VaR to Latin America (187,512), followed by Europe (165,606) and Emerging Europe (163,682).
- U.S. shows greatest interest rate risk (97,865) and equity risk (71,919), but no FX risk since base currency is USD.
- Only very small risk positions in Emerging Asia (22,727).

Example 2 Risk can also be analyzed by instrument type, as illustrated below.

VaR by Region and Instrument Type **January 17, 1999**

RMG



99%, 1-day VaR	Aggregate	*Divers. Benefit*	Equity Future	Equity Option	FX Forward	FX Option	Swap
Aggregate	$ 515,085	*–338,234*	220,908	215,598	186,641	101,261	128,910
Divers. Benefit	***–224,643***		*–55,980*	*–3,391*	*–70,996*	*–35,384*	*–88,040*
Americas	186,230	*–70,291*	122,904	61,698			71,919
Europe	145,508	*–78,339*	82,803	70,676	11,364	7,346	51,658
Asia	165,928	*–129,068*	52,471	58,427	81,841	73,783	28,474
Australia	110,681	*–50,129*	11,364	25,569	70,676	969	52,232
Africa	131,381	*–39,554*	7,346	2,619	93,756	54,547	12,667

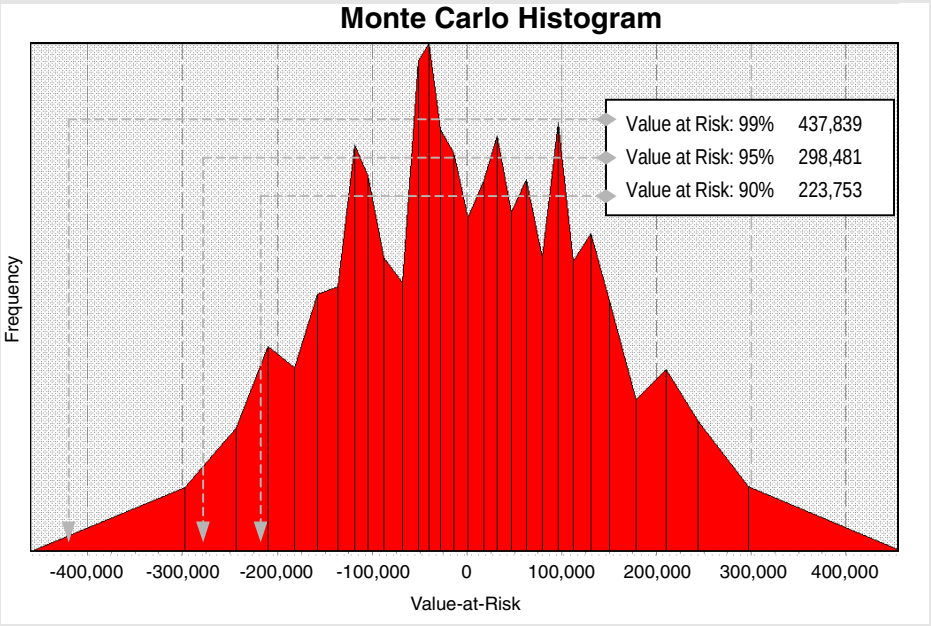
Comments

- The largest instrument risk is to Equity Futures (220,908) and Equity Options (215,598).
- Geographically, risk is well diversified, with the highest activity in the Americas (186,230) and Asia (165,928).

Example 3 Interest rate risk is often presented by maturity or duration bucket to analyze **curve risk**. Curve risk refers to potential losses due to changes in the shape of the yield curve. Below we see a Monte Carlo histogram of returns and a summary of risk vs. instrument type. In this example of a detailed desk level report, swaps and futures are used to hedge the risk of a bond portfolio.

VaR by Maturity **April 28, 1999**

RMG



95%, 1-day VaR	Aggregate	*Divers. Benefit*	Bonds	Swaps	Futures
Aggregate	298,481	***–1,427,188***	914,011	532,429	279,228
Divers. Benefit	*–101,825*		*–783,232*	*–152,032*	*–119,171*
1 to 3 months	12,025	*–200,091*	111,112	101,004	
3 to 6 month	10,957	*–188,618*	103,142	92,136	4,297
6 to 9 months	3,309	*–243,073*	118,810	2,211	125,361
9 to 12 months	137,209	*–119,441*	132,774		123,876
1 to 2 years	67,381	*–26,908*	94,288		
2 to 3 years	34,231	*–289,704*	323,934		
3 to 5 years	2,699	*–332,003*	169,264	165,438	
5 to 10 years	29,643	*–152,508*	152,344	80,867	
10 to 20 years	3,256	*–373,109*	188,111	188,254	
20 to 30 years	99,596	*–454,344*	354,524	54,551	144,865

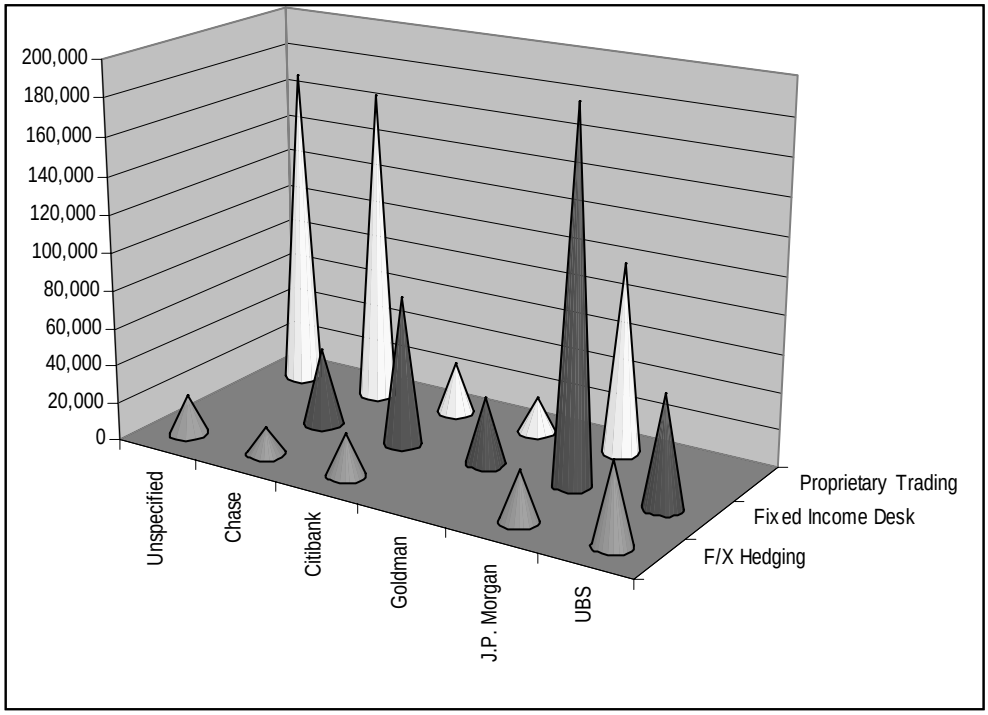
Comments

- In this portfolio, swaps and futures are used to hedge bond positions. For example, see the 1–3 month bucket where bonds and swaps are almost perfectly offsetting (111,112 bond VaR reduced to 12,025).
- Notice the large diversification between instrument types, which indicates hedging activity (–1,427,188 total diversification between instrument types, compared to net aggregate risk of 298,481).

Example 4 VaR can also be analyzed by counterparty to assess potential credit exposure due to market driven instruments. The concept of credit exposure due to market driven instruments such as swaps, forwards and options is explained in more detail in Appendix B.

VaR by Counterparty **June 11, 1999**

RMG



95% 1-week VaR	Total	F/X Hedging	Fixed Income Desk	Proprietary Trading
Total	**476,207**	**64,394**	**341,693**	**294,593**
Unspecified	**170,377**	20,565		170,673
Chase	**166,035**	14,572	40,584	166,026
Citibank	**67,731**	21,837	78,822	27,969
Goldman	**38,943**		35,201	18,578
J.P. Morgan	**208,386**	26,493	192,787	100,558
UBS	**82,094**	43,254	58,647	

Comments

- Largest counterparty exposure is to J.P. Morgan, due to large trades by Fixed Income and Proprietary Trading. There is a 5% chance that our exposure to J.P. Morgan will exceed $208,386 over 1 week.
- Fixed Income Desk is generating the largest total credit exposure ($341,693), followed by Proprietary Trading ($294,593).

1.8 Summary

VaR is a general statistical measure of risk that is used to equate risk across products and aggregate risk on a portfolio basis, from the corporate level down to the individual trading desk. VaR is defined as the predicted worst-case loss at a specific confidence level over a certain period of time.

There are three major methodologies for calculating VaR, each with unique characteristics. Parametric VaR is simple and quick to calculate, but is inaccurate for non-linear positions. The two simulation methodologies, historical and Monte Carlo, capture non-linear risks and give a full distribution of potential outcomes, but require more computational power.

Before calculating VaR, three parameters must be specified: (a) confidence level, (b) forecast horizon, and (c) base currency.

Square root of time scaling of VaR may be applied to roughly extrapolate VaR to horizons longer than 1 day, such as 10 days or 1 month. Square root of time scaling assumes a random diffusion process with no autocorrelation, trending, or mean reversion. Be aware that some regulators are discouraging the use of this simplistic approach. The *LongRun Technical Document* discusses more accurate long-horizon risk forecasting, from 2 months to 2 years.

Financial instruments are subject to both market and residual risks. The four basic components of market risk are interest rate, equity, commodity, and foreign exchange risk. Residual risk includes spread, basis, specific and volatility risk.

Risks can be reported in many dimensions, including risk taker, risk type, region, instrument, and counterparty. In addition, interest rate risk is often analyzed by maturity or duration band.

Chapter 2. Stress testing

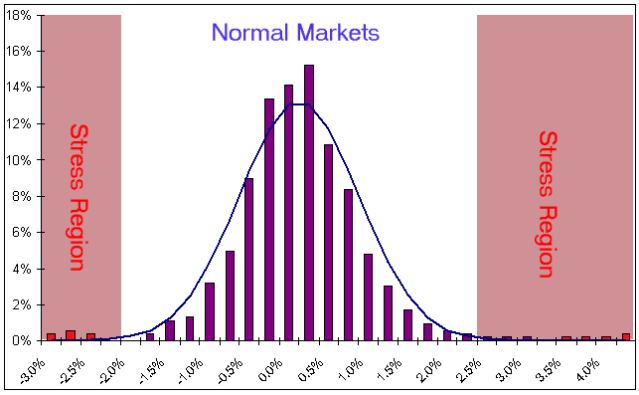
2.1 Why stress test

Stress tests are designed to estimate potential economic losses in abnormal markets. Historical analysis of markets shows that returns have "fat tails," where extreme market moves (i.e., beyond 99% confidence) occur far more frequently than a normal distribution would suggest. Although the discipline of risk management has improved considerably, classical events like natural disasters, wars, and political coups still lie beyond statistical forecasting.

Therefore, regular stress testing is increasingly viewed as indispensable by risk managers and regulators. Stress tests should enhance transparency by exploring a range of potential low-probability events when VaR bands are dramatically exceeded. Stress testing combined with VaR gives a more comprehensive *Picture of Risk*. This sentiment is echoed throughout the risk management community. For example, an excerpt from Chase's 1998 *Annual Repor*t states:

> *Chase's two principal risk measurement tools are VAR and stress testing. VAR measures market risk in an everyday market environment, while stress testing measures market risk in an abnormal market environment…. This dual approach is designed to ensure a risk profile that is diverse, disciplined and flexible enough to capture revenue-generating opportunities during times of normal market moves, but that is also prepared for periods of market turmoil.*
>
> *Source*: Chase Manhattan Corporation

This histogram of the AUD/USD exchange rate illustrates where the focus of stress testing should be. Stress tests should "inspect the tails" of the return distribution. Stress tests could therefore be viewed as a complement to VaR: VaR is for normal markets, and stress tests are for abnormal markets. Together, they paint a broader *Picture of Risk*.

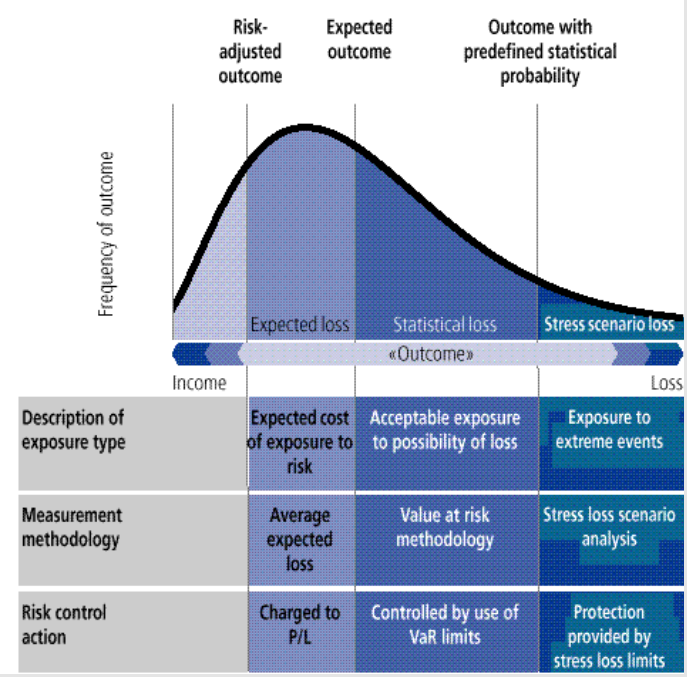


VaRBaR

Stress testing and VaR give a broader *Picture of Risk*

The relationship between stress testing and VaR can be seen in the following diagram, from the UBS Group 1998 *Annual Report*.

Expected, Statistical and Stress Scenario Loss



While UBS uses a value-at-risk measure as the principal measure of its exposure to day-to-day movements in market prices, the experience during the third quarter underlines the fact that these measures are not designed to give an indication of the scale of loss that could occur in the unusual case of extreme market moves. For this reason, UBS supplements its value-at-risk numbers with a system of stress loss simulations in order to monitor its exposure to this type of market shock. These measures seek to assess the scale of loss which UBS might face in the event of large movements in a range of market prices such as equity indices, foreign exchange rates and interest rates. In the light of the events in the third quarter, UBS has revised the range of price changes which it uses to calculate the exposure to stress load and has revised the relevant limit structures.

— UBS Group 1998 *Annual Report*, page 43

2.2 Two central questions for stress testing

Stress tests can be framed around two central questions:

1. **How much** could I lose if a stress scenario occurs, for example the U.S. Equity market crashes?
2. **What event** could cause me to lose more than a pre-defined threshold amount, for example $100 million?

The first question is commonly asked in a top-down approach for stress testing. For example, senior management may ask how much could the firm lose in a major equity market crash.

The second question is best asked at the book or business level. After scenarios are collected from individual risk takers, cross-firm analysis can be done to see if events are diversified or exacerbated. For example, a stress scenario of JPY vs. USD depreciation might be ruled as unimportant due to generally offsetting sensitivities (or no-significant reported sensitivities), while a credit spreads widening scenario could be identified as relevant because many risk-taking units expressed a similar concern. This approach could therefore be viewed as a bottom-

up search for relevant stress scenarios. Relevant stress scenarios should be elevated to the next level of management.

VaRBaR

Searching for vulnerabilities

J.P. Morgan recently introduced a Vulnerabilities Identification (VID) process in which each risk taking unit was asked to (a) qualitatively list what event could cause it to lose more than a specified threshold dollar amount, and (b) assign a probability to each event. The Corporate Risk Management Group (CRMG) polled the entire firm and aggregated the results into a searchable database. CRMG could then conduct cross-firm analysis to see which scenarios were "diversified" away and to identify exacerbating scenarios that many risk taking units were exposed to in common.

Next, J.P. Morgan implemented a web-based VID infrastructure to collect scenarios from risk takers and risk monitors on an ongoing basis. The VID infrastructure reflects hierarchical reporting lines at J.P. Morgan in order to facilitate quick escalation of relevant stress scenarios up the chain of command. For example, with the touch of a button, the head of Fixed Income can log onto the VID system, review stress scenarios submitted by business managers and then escalate relevant stress scenarios to the next organizational level (corporate office). The corporate office therefore receives a filtered collection of critical stress scenarios. Drill-down of stress scenarios, down the reporting line, is also possible. The author of the stress scenario can also be contacted directly for more information.

Efficiently harnessing the firm's collective intelligence, J.P. Morgan's innovative VID process continuously channels relevant stress tests from the bottom up.

2.3 How to use stress tests

The key issue with stress tests is how to create and use them. To be meaningful, stress tests should tie back into the decision making process. Corporate-level stress test results should be discussed in a regular forum by risk monitors, senior management, and risk takers. Just as for VaR limits, companies should have a set of stress loss limits by risk type and risk taking unit. Stress testing should be performed at multiple levels of the micro, macro, and strategic risk pyramid with different frequencies. At a senior management level, stress results should guide the firm's appetite for aggregate risk taking and influence the internal capital allocation process. At the book level, stress tests may trigger discussions on how best to unwind or hedge a position.

VaRBaR

Stress Testing at Chase Manhattan Bank

A descriptive summary of stress testing can be found below in an excerpt from Chase's 1998 *Annual Report.*

Chase's corporate stress tests are built around changes in market rates and prices that result from pre-specified economic scenarios, including both actual historical and hypothetical market events. As with VAR, stress test calculations are performed for all material trading and investment portfolios and market risk-related ALM portfolios.

Stress test scenarios are chosen so they test "conventional wisdom" and focus on risks relevant to the positions taken on Chase's portfolios. A key to the success of stress testing at Chase is continuous review and updating of the stress scenarios. This is a dynamic process that is responsive to changes in positions and economic events, and looks to prior stress tests to identify areas where scenario refinements can be made. Corporate stress tests are performed approximately monthly on randomly selected dates. As of December 31, 1998, Chase's corporate stress tests consisted of seven historical and hypothetical scenarios. These historical scenarios included the 1994 bond market sell-off, the 1994 Mexican Peso crisis and the 1997 Asian markets crisis.

Stress test results are used at all levels of Chase, from the trading desk to the Board of Directors, to monitor and control market risk. Among the controls instituted at Chase are a review of the trading portfolio if potential stress losses exceed Board of Directors-approved advisory limits and the incorporation of stress test exposures into Chase's capital allocation methodology.

2.4 What makes a good stress test

The goal of stress testing is to uncover potential concentrations and make risks more transparent.

Good stress tests should

- be relevant to current positions,
- consider changes in all relevant market rates,
- examine potential regime shifts,
- spur discussion,
- consider market illiquidity, and
- consider the interplay of market and credit risk.

A. Stress tests should be relevant to current positions

Good stress scenarios are designed to stress current positions and probe for portfolio-specific weaknesses. A concentrated portfolio may incur losses from relatively small movements in certain market rates. Therefore, simply stressing portfolios by large movements in generic market rates does not necessarily uncover relevant risks. For example, simulating a simple equity index fall would do little to uncover the risk of a market neutral risk arbitrage book.[1]

In a real world example, Long Term Capital Management (LTCM) had leveraged credit spread tightening positions (i.e., long corporate bond positions were interest rate hedged with short

[1] A market neutral risk arbitrage book would consist of a series of long and short positions, which hedges out market risk but is exposed to firm-specific risk.

Treasurys) in August '98. This portfolio was supposedly "market neutral[2]." A stress test of spread widening (i.e., *flight to safety* phenomenon) would have uncovered the potential for extreme losses.

B. Stress tests should consider changes in all relevant market rates

Stress scenarios should take into account potential changes in a complete set of market rates. A stress scenario in isolation does not reflect reality because market rates don't move in isolation (especially when they are extreme). For example, if we raise the 5-year Euro swap rate by 100 basis points, we need to anticipate potential changes in the rest of the Euro yield curve, other international yield curves, equity markets, and FX rates. Good stress tests represent comprehensive scenarios.

C. Stress tests should examine potential regime shifts

A key question in developing every stress scenario is whether current risk parameters will hold or break down. For example, will observed correlations hold or increase, or could we see a regime shift (i.e., de-coupling of market rates)? For example, during large equity shocks (e.g., '87 crash, '97 and '98 sell-offs), a *flight to safety* often results in a reversal of correlations between stocks and government bonds: as stocks plummet, bonds rise because investors move into safer and more liquid assets.[3] In the market turmoil of September '98, LTCM experienced this problem when credit spreads widened and interest rates fell due to a *flight to safety* [4] (it certainly was not a good time to hedge Corporates with Treasurys). In stress testing, asking the right question (e.g., what could happen), is just as important as providing answers (e.g., what losses would be under those conditions).

D. Stress tests should spur discussion

Stress tests should include some potential rationale for how that adverse scenario could happen, and spur discussion to probe deeper into potential risks. In the case of LTCM, a discussion might have centered on what could happen to cause spread widening and the likelihood of that event. Another discussion might consider how one might best get out of such a concentrated risk position. A good stress test doesn't prevent an event from happening, but it does prepare the risk taker for the possibility and gives the opportunity for taking pre-cautionary measures.

E. Stress tests should consider market illiquidity

Stressed markets are often characterized by significant loss of liquidity. Liquidity can be viewed from two perspectives: the ability to trade positions and the ability to fund positions. Liquidity shocks can be extremely severe in Emerging Markets. For example, Brazilian bond traders reported that bid ask spreads were so wide during the October '97 liquidity crisis that it was unclear whether the local yield curve was upward or downward sloping. When prices in the market place don't exist, it becomes impossible to mark-to-market positions. Furthermore, funding often dries up in these conditions, forcing participants to liquidate positions, which puts even more downward pressure on prices. The inability to fund its concentrated junk bond positions precipitated the demise of Drexel Burnham Lambert in the late eighties. The threat of extreme liquidity risk motivated the recapitalization of LTCM by a consortium of 14 commercial and investment banks in September 1998. In a statement to the U.S. House of Representatives, Chairman Alan Greenspan stated "the consequences of a fire sale triggered by cross-default clauses, should LTCM fail on some of its obligations, risked a severe drying up of market liquidity."[5]

---

2 "Market neutral" refers to a trading style that should be uncorrelated to underlying equity and bond markets.

3 Note that correlation between stocks and bonds is generally highly positive.

4 See "Lessons from LTCM" editorial at www.riskmetrics.com.

5 See http://www.bog.frb.fed.us/BoardDocs/Testimony/1998/19981001.htm.

F. Stress tests should consider the interplay of market and credit risk

Stressed markets often give rise to counterparty credit risk issues that may be much more significant than pure market impacts. For example, a market neutral swap portfolio could result in huge credit exposures if interest rates moved significantly and counterparties defaulted on their contractual obligations. While market rates and credit worthiness are unrelated for small market moves, large market movements could precipitate credit events, and vice versa.

2.5 Forecasting time frame

The forecast horizon for the stress scenario should reflect an institution's typical holding period. Banks, brokers, and hedge funds tend to look at a 1-day to 1-week worst-case forecast, while longer-term investors, like mutual and pension funds, may consider a 1-month to 3-month time frame. Corporations may use up to an annual horizon for strategic scenario analysis.

2.6 How often to stress test

It's important to engage in the discipline of regular stress testing and discussion of stress results. Major financial institutions engage in weekly or even daily stress tests. Special stress tests should be performed when there are unusually large or concentrated risk positions and during abnormal market conditions (i.e., when there is unusually high volatility or when there are impending political or economic events). But stress testing should not be performed so frequently and extensively as to become overwhelming and lose meaning.

2.7 Steps for stress testing

There are three basic steps for stress testing.

Step 1: Generate scenarios
The most challenging aspect of stress testing is generating credible worst-case scenarios that are relevant to portfolio positions. Scenarios should address both the magnitude of movement of individual market variables and the interrelationship of variables (i.e., correlation or causality).

Step 2: Revalue portfolio
Revaluing a portfolio involves marking-to-market all financial instruments under new worst-case market rates. Stress test results are generally changes in present value, not VaR.

Step 3: Summarize results
A summary of results should show expected levels of mark-to-market loss (or gain) for each stress scenario and in which business areas the losses would be concentrated.

In addition to summarizing the effect on the present, a comprehensive analysis could estimate longer-term, indirect effects on a firm's well-being. Such business risk analysis would address how a scenario might affect the level of demand for a business's products and services. For example, during an equity market downturn, one might consider the indirect loss from lower demand for equity underwriting in addition to the direct losses on trading and investment positions.

In considering the anticipated short-term and long-term results of the stress analysis, management can decide whether (and how) the risk profile of the firm should be changed.

Example: Brazilian Company

We illustrate the three steps of stress testing with the following example.

You are a Brazilian consumer products company with a significant amount of unhedged USD-denominated liabilities. You are particularly concerned about the stability of the Brazilian Real (R$), because a devaluation would make USD liabilities prohibitively expensive.

Step 1: Generate scenarios

Your economist presents two potential events:

1. A significant widening of the trade deficit, which puts pressure on the local currency, interest rates, and equity market
2. A narrowing of the trade deficit, which is a positive scenario for local markets

Following are the economist's estimates of the effect of each scenario on the local markets over 1 day:

Stress scenarios	Widening trade deficit	Narrowing trade deficit
R$/US$ exchange rate	up to 20% devaluation	no move
R$ yield curve	interest rates up 50%	interest rates down 25%
BOVESPA equity index	15% fall	10% appreciation

Step 2: Revalue position

The next step involves revaluing the company's financial positions given new market rates. Financial exposures would include all USD and R$ assets and liabilities, as well as equity investments.

P&L (in R$ millions)	Widening trade deficit	Narrowing trade deficit
R$/US$ exchange rate	–20	0
R$ yield curve	+ 5	–2
BOVESPA equity index	– 9	+6
Total	**–24**	**+4**

Step 3: Summarize results

A devaluation scenario could result in a direct financial loss of R$24 million for the company, while a narrowing of the trade deficit could yield a financial gain of R$4 million over 1 day.

Furthermore, management should assess how each scenario might affect underlying business. For example, although a devaluation might hurt domestic sales, it could make exports into other markets more competitive.

Management should then discuss whether it should take action to reduce its risk. The largest potential loss comes from the unhedged USD liabilities—for example, this could be reduced through an FX forward hedge or by purchasing a put option on R$ vs. dollar.

2.8 Creating stress scenarios

There are a variety of approaches to generating stress tests, which we discuss in turn: (a) generating historical scenarios based on days when markets moved violently, (b) introducing market shocks and moving risk factors in isolation by large amounts to gauge sensitivity to each risk

factor, (c) creating anticipatory scenarios in which many market factors are moved in a consistent fashion to approximate real moves of all relevant world markets, and (d) setting up portfolio-specific stress tests, which are based on the weaknesses of the portfolio itself.

Within this framework of scenario generation, we discuss three concepts that are relatively new to stress testing: (1) predictive scenario generation in which a subset of stressed risk factors is used with historical correlations to predict the moves of all other market variables, (2) VaR with stressed volatility and correlations, and (3) portfolio weaknesses as determined by Monte Carlo simulation.

A. Using relevant historical scenarios

A natural approach is to base scenarios on historical periods with extreme market conditions. Some infamous events include: the '87 U.S. stock market crash, the ERM crisis, the Fed rate hike in '94, the '95 Tequila crisis, the '97 Asian crisis, the volatile markets in '98, and the '99 Brazil devaluation. In this approach, data is captured from relevant historical stress periods and a portfolio is valued with historical simulation to measure potential losses.

RiskMetrics research has assembled a representative global model portfolio consisting of 60% equities and 40% fixed income to identify pertinent time periods for historical stress tests. Both 1- and 5-day portfolio returns were used to identify extreme loss periods.

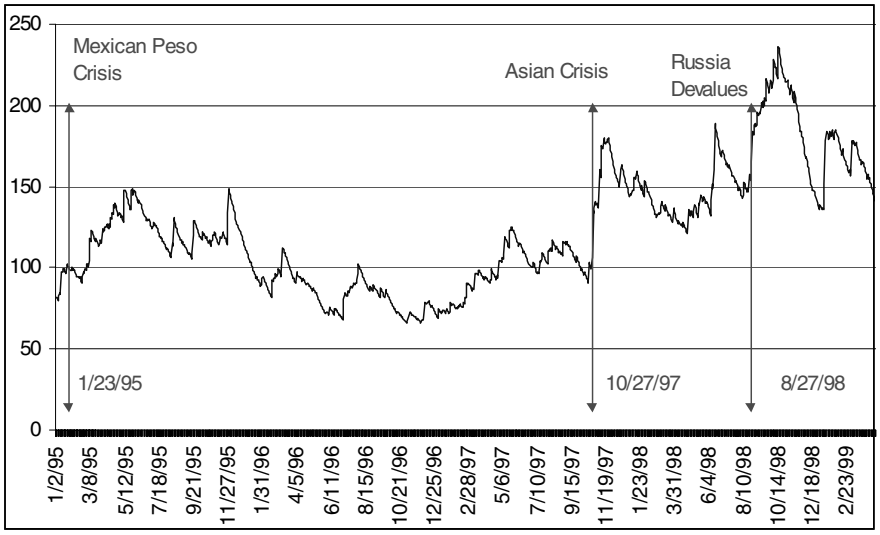
Sample historical scenarios

	Date	1-day return, %	Date	5-day return, %
Black Monday	19-Oct-87	–2.20	20-Oct-87	–5.9
Gulf War	3-Aug-90	–0.90	27-Aug-98	–3.8
Mex Peso Fallout*	23-Jan-95	–1.00	23-Jan-95	–2.7
Asian Crisis	27-Oct-97	–1.90	7-Aug-90	–3.6
Russia devalues	27-Aug-98	–3.80	27-Oct-97	–2.6

* The Mexican peso actually devalued at the end of '94. On 23-Jan-95, the peso lost 6% and several Eastern European markets incurred losses of around 5% to 10%.

Source: A. Ulmer, 1999, unpublished research, RMG

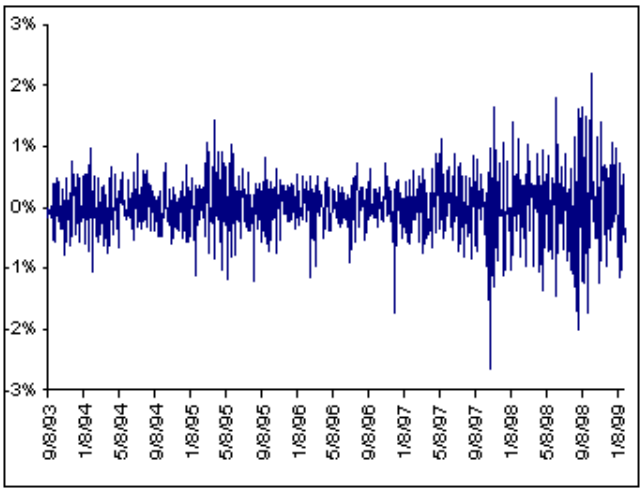
How does the severity of loss depend on global volatility? This figure shows the RiskMetrics Volatility Index (RMVI)[6] for a period which includes three of the historical scenario dates.



Source: DataMetrics

Volatility breeds shocks

Notice that the severity of portfolio losses appears to be related to the level of volatility in the world, as measured by the RMVI. For example, the Mexican peso fallout, which occurred during a period of relative market calm when RMVI was at an average level of approximately 100, only resulted in a 1-day return of –1%, while the Asian Crisis and Russia Devaluation, which occurred while the RMVI was significantly above 100, resulted in more severe portfolio losses of 1.9% and 3.8%. This suggests that it make no sense to run the same static stress tests in all market regimes—more volatile markets require more severe stress tests. To make stress scenarios more responsive to market conditions, the RMVI can be used as a dynamic scaling factor for stress scenarios, as discussed in the next section.



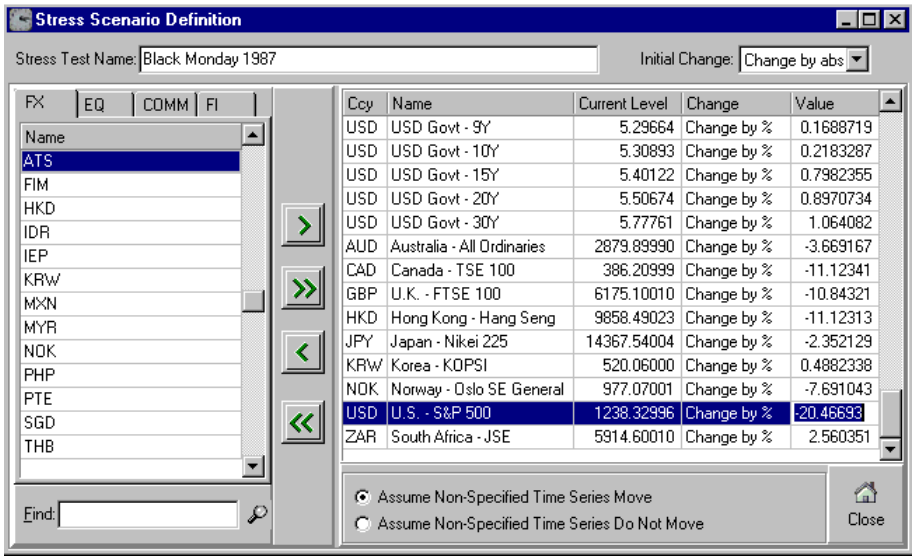
Source: A. Ulmer, 1999, unpublished research, RMG

As further evidence that extreme market moves tend to be larger when market volatility is high, observe the graph of daily returns of our global equity and fixed income test portfolio: large shocks from ’93 to ’96 result in portfolio losses of no more than 1%, while shocks in ’97 and ’98 result in portfolio losses of 2% to 3%. Volatile regimes seem to breed larger shocks.

Historical stress tests and data can be accessed and created in RiskMetrics RiskManager. RM’s stress testing module includes historical stress scenarios tabulated from the perspective of a wide range of base currencies. Because historical data may not exist for all factors (such as many swap rates whose markets were still developing in 1987), RM supplies a method to estimate likely changes in the unknown factors based on present-day correlations. This method is based on the predictive scenario generation, which is described in Section C on page 33, “Anticipatory scenarios.”

In the example below, part of the 1987 Black Monday scenario is presented from an RM screen. Alternatively, one may designate a set of stressed returns based on DataMetrics data by choosing start and end dates for a historical stress test.

[6] The RiskMetrics Volatility Index measures global volatility. The RMVI is composed of equity, fixed income, and foreign exchange markets in 28 countries, as well as three major commodity markets. Observing the daily returns of these 87 markets, we calculate a total volatility across all countries and asset classes and compare it to a historical average. See our web site, http://www.riskmetrics.com and, in particular, the document *RMG Volatility Index: Technical Document,* by Finger, 1998.



B. Applying shocks to market factors or correlations

A second approach to generating stress tests is to shock either market factors or volatilities and correlations by large amounts. This method can provide a good measure of the sensitivity to risk factors and can therefore be useful in identifying trouble spots in a portfolio.

It is straightforward to alter a market rate (e.g., lower the S&P 500 by 10%) or many market rates (e.g., lower all points on the U.S. government yield curve by 50 basis points), but the real challenge is to determine both which market rates to shock and by how much to shock them. In general, these decisions will be based on historical moves, on intuition, and the portfolio itself.

We first illustrate the example of applying shocks to market rates, followed by an example of changing the correlations.

Shocks to market rates example

Below, we show an example of a market shock stress test in which the P/L is considered by both asset class and region. The example shows the effects of both a bull and a bear market on a portfolio.

Geographic region	Interest rates	Equities, %	FX, %
North America	+80bp / –80bp	+/– 8	+/– 10
Europe	+100bp / –100bp	+/– 10	+/– 10
Japan	+50bp / –25bp	+/– 10	+/– 10
Emerging Asia	+250bp / –200bp	+/– 25	+/– 20
Russia & Eastern Europe	+400bp / –300bp	+/– 30	+/– 25
Latin America	+1000bp / –500bp	+/– 35	+/– 20

We can use these scenarios as a basis for stress testing various portfolios. For example, a Global Bank portfolio stress test may look as follows.

Global Bank 1-Day Stress Test ***Tuesday, March 16, 1999***

Global Bank Stress Analysis
(one-day worst-case market moves)

Geographic Region	Interest rates		Equities		FX		Net by Region
	Move, bp	P/L ($mm)	Move, %	P/L ($mm)	Move, %	P/L ($mm)	
North America	80	–5.6	–8	–5.8	–10	–1.5	–10.1
	–80	5.0	8	5.2	10	1.3	12.0
Europe	100	–8.7	–10	–4.6	–10	–1.3	–14.0
	–100	7.0	10	3.7	10	1.1	14.6
Japan	50	4.0	–10	3.0	–10	1.4	9.1
	–25	–3.0	10	–2.5	10	–1.6	–3.5
Emerging Asia	250	–2.0	–25	–3.2	–20	–0.1	–5.0
	–200	1.8	25	2.9	20	0.1	4.9
Russia & Eastern Europe	400	5.0	–30	–3.3	–25	3.1	5.3
	–300	–4.0	30	2.3	25	–2.2	–3.7
Latin America	1000	–12.0	–35	–4.0	–20	5.0	–10.7
	–500	8.0	35	6.0	20	–2.0	13.9
Total	**Up**	**–19.3**	**Down**	**–17.9**	**Down**	**6.3**	**–26.6**
Portfolio	**Down**	**20.8**	**Up**	**17.6**	**Up**	**–3.2**	**37.0**

Stress Test Commentary

- Overall economic sensitivities to stress scenarios are within tolerance limits: daily loss in a global bear market scenario is estimated at $26.6 million, and gain in a global bull market scenario is $37 million.
- Largest asset class exposure is interest rates (–$19.3/20.8), followed closely by equities (–$17.9/17.6).
- Largest regional exposure to a bear market is Europe (–$14.0), followed closely by North America (–$10.1) due largely to corporate bond inventories by N.Y. and London Fixed Income.
- Note net short position in Japan and Russia across asset classes (i.e., losses in a bull market).
- Note also short position in Latin America FX ($5.0) largely through real/USD puts, coupled with long positions in Brady bonds (–$12.0) and equities (–$4.0) by Emerging Markets.
- Note that the individual gains and losses in a row do not, in general, add up to the Net by Region entry (even without non-linear positions), because losses are generally lessened by big moves whereas gains are accented. For example, when the European equities fall by 10% and the FX rate decreases by 10%, the drop in value of a pure equity portfolio when both of these events occur is smaller than 20%, because the equities lose 10% of their value initially, and then lose 10% of their reduced value to FX.

RMVI scaling

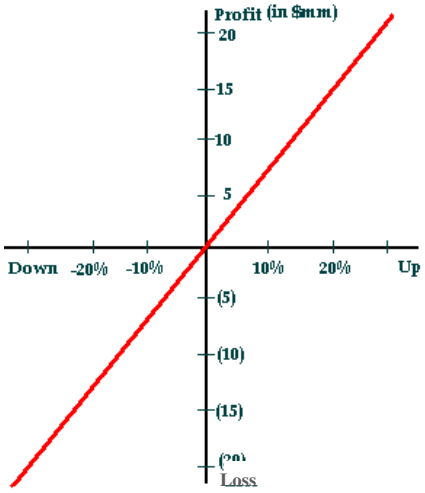
One way to make parameter-shocking better reflect current market conditions is to link the magnitude of extreme market scenarios to the RMVI level. More severe stresses would be applied during volatile markets, and smaller stresses in calm markets. For example base case stress scenarios could be developed for an RMVI level of 100 and scaled linearly to reflect higher or lower volatility (e.g., RMVI level of 150 could reflect stress scenarios of 1.5 times base case).

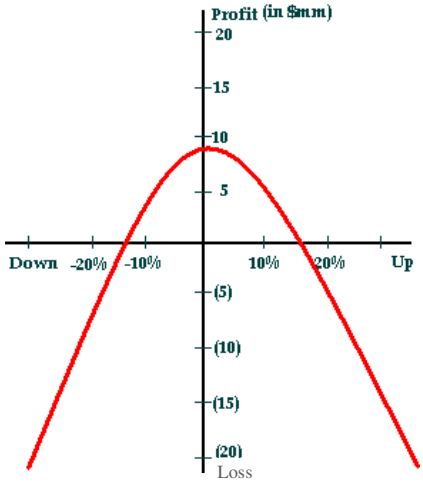
Sliding scales

Another approach to stress testing is to create a sliding scale of P&L vs. key market variables. The Y-axis shows P&L and the X-axis shows change in benchmark market variables: equity, interest rate and commodities indices, and FX rates. Sliding scales are very useful for complex positions with non-linear payoffs long or short option combinations. Sliding scales can be used to test portfolio vulnerabilities: particularly for derivatives portfolios, it is important to understand sensitivities to key market rates and parameters.

For example, a hedge fund may show the following sliding scales for its proprietary positions:

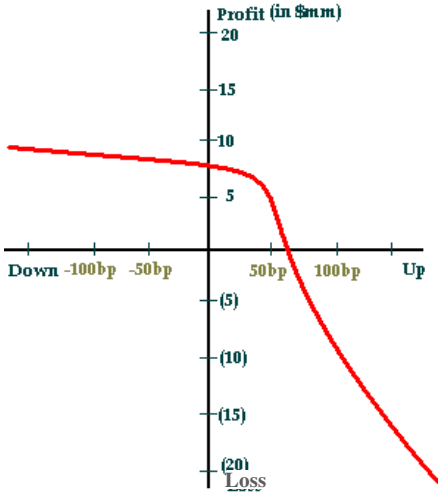
Comments:
This is a classic long asset position, where profits rise and fall linearly with markets.

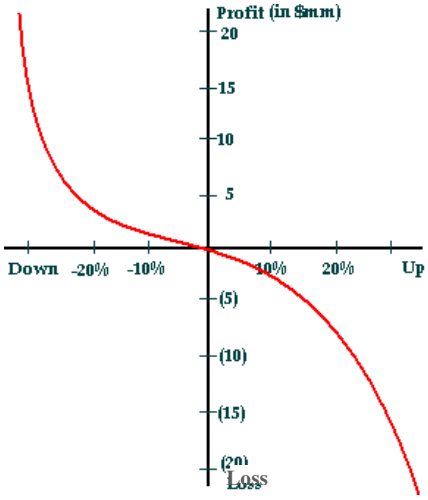




Comments:
This payoff resembles a short strangle (short call and put). The position is profitable if the market stays calm, but loses money if there are large movements either way.

Comments:
This position is sensitive to U.S. rates rising by more than 50 basis points. This could be due to a short cap, or a short put on a bond.





Comments:
This payoff resembles a short out-of-the money call and a long out of the money put. This bearish position makes money if oil futures fall and loses if oil prices rise.

Shocks to volatilities and correlations

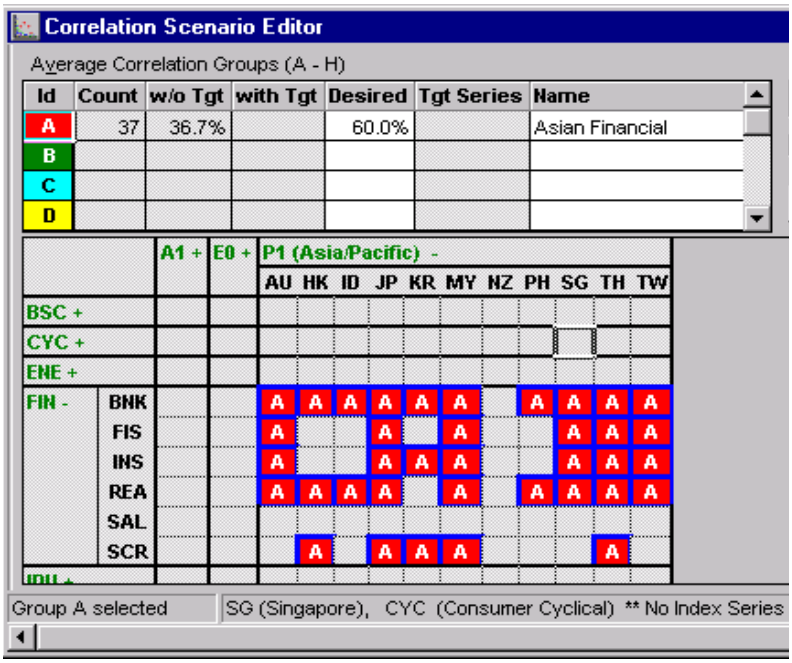
VaR estimates can be stressed by applying shocks to volatilities or correlations. While volatilities can be adjusted up and down like market rates, special care must be taken with correlation matrices because nonsensical correlation structures can often be created from the interrelationship between factors. These correlation structures can, in turn, result in nonsensical VaRs (imaginary).

For example, consider a three-party government. If party A and party B always vote in opposite directions (correlation of –1), it is impossible for party C to be positively correlated with both A and B. RiskMetrics research has published a methodology to adjust correlations in a mathematically consistent fashion. The general idea is to mix in an average correlation term into a prespecified group of assets, and then to adjust all diagonal terms (see the *RiskMetrics Monitor*, Q4 ’97).

Correlation example

This methodology is implemented in CreditManager, where it can be used to change the average correlation among industries and countries. In the example below, we show a stress test where the average correlation between Asian financials is increased from 37.6% to 60%. The boxes marked “A” indicate the factors that were selected to be in this group of increased correlation.

Detail from CreditManager Correlation Scenario Editor



C. Anticipatory scenarios

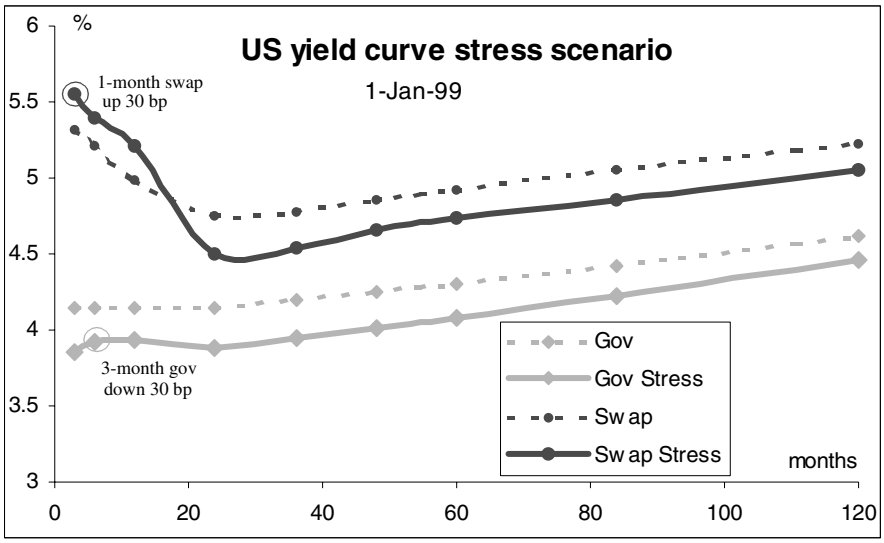
In generating an anticipatory scenario, a risk manager must determine (1) the event of interest (e.g., flight to quality within Asian markets, stock market crash), (2) the severity of the event, (e.g., from once-in-a-year to once-in-a-decade), and (3) the effects of such an event on the global market. For the third determination, it is essential to move all market rates in a consistent fashion, e.g., in a flight to quality, not only will government/corporate spreads widen, but equity prices will fall.

Generating a stress scenario

Below we conduct a stress test of short-term swap spreads widening on January 1, 1999, in order to forecast 1-day portfolio losses using the RM stress module. Note that a swap spread widening scenario is a simulation of the *flight to safety* phenomenon, where we would expect risky assets to fall. We increase 1-month swap rates by 30 basis points, lower 3-month government rates by 30 basis points, and let RM estimate the likely effect on other yield curve points and market rates based on historical correlations from January 1, 1999 to July 1, 1998.

Estimating impact on relevant market variables

Below are the original and stressed government and swap yield curves. Observe how the swap yield curve twists around the 18-month mark, with short-term yields higher and longer-term yields lower.



Forecasted stress effects on other market variables are given by the 1-day returns in the table below:

1-day return, %	Asset class
–7.80	U.S. - S&P 500
–8.76	Spain - IBEX 35
–17.59	Argentina - SE Blue Chip
–6.44	THB
–5.44	Sweden - OMX
–5.23	Mexico - IPC
–5.22	Singapore Dollar
–4.23	Germany - DAX
0.42	Japan - Nikei 225
–5.00	France - CAC 40
–3.79	JPY
–0.95	GBP
0.22	FRF
0.22	DEM

This stress scenario is largely consistent with a *flight to safety* phenomenon.

The impact on equity markets is bearish: U.S. equity markets plummet by 7.8%, and Europe and Latin America also suffer large losses.

In FX, Thai baht, JPY, and Singapore dollar depreciate significantly (vs. USD). FRF and DEM move by the same amount due to Euro peg, while GBP depreciates by 0.95%.

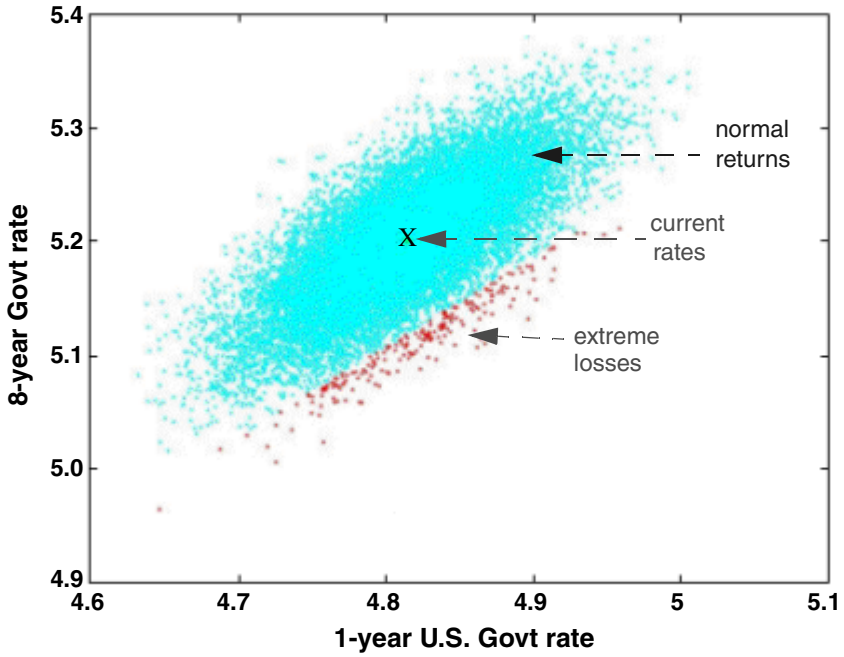
The next step in the stress test would be to re-value the portfolio under the above scenario, and to analyze sensitivities and summarize the results.

D. Applying portfolio-specific stress tests

Another stress testing approach searches for stress scenarios by analyzing the vulnerabilities of the specific portfolio in question. One way to discern the vulnerabilities is by conducting a historical or Monte Carlo simulation on a portfolio, and searching for all scenarios that could cause a loss exceeding a defined threshold amount. Instead of specifying the scenario and calculating potential losses as in the three approaches described previously, we specify what constitutes a severe loss and search for scenarios.

To illustrate this approach, consider a long 1-year note and a short 8-year note position. We run a Monte Carlo simulation and highlight the 5th percentile worst-case losses. We can see that losses are worst when long-term rates fall relative to short-term rates—that is, we are exposed to a tilt in the yield curve. In more complex portfolios, the tail events may cluster into a number of groups. These weak spots should be considered in stress tests.

Monte Carlo simulation result of 1-year vs. 8-year rates



VaRBaR

External disclosures of stress tests

To provide investors with greater risk transparency, companies may provide stress scenarios and sensitivities.

1. Citicorp interest rate stress test

Citicorp Earnings-at-Risk (impact on pre-tax earnings)

In Millions of Dollars at December 31, 1998	Assuming a U.S. Dollar Rate Move of Two Standard Deviations Increase	Assuming a U.S. Dollar Rate Move of Two Standard Deviations Decrease	Assuming a Non-U.S. Dollar Rate Move of [1] Two Standard Deviations[2] Increase	Assuming a Non-U.S. Dollar Rate Move of [1] Two Standard Deviations[2] Decrease
Overnight to three months	$ (85)	$ 87	$ (23)	$ 23
Four to six months	(34)	38	(30)	30
Seven to twelve months	(29)	31	(40)	40
Total overnight to twelve months	**(148)**	**156**	**(93)**	**93**
Year two	(28)	22	(51)	51
Year three	12	(22)	17	(16)
Year four	54	(64)	22	(21)
Year five	119	(152)	24	(23)
Effect of discounting	(29)	39	(26)	26
Total	**$ (20)**	**$ (21)**	**$ (107)**	**$ 110**

(1) Primarily results from Earnings-at-Risk in Thai baht, Singapore dollar and Hong Kong dollar.
(2) Total assumes a standard deviation increase or decrease for every currency, not taking into account any covariance between currencies.

In its 1998 *Annual Report*, Citicorp discloses a simple interest rate stress test, which consists of perturbing interest rates by 2 standard deviations. Although this simple analysis is not comprehensive, shareholders and analysts get a rough perspective of Citicorp's Net Interest Earnings (NIE) sensitivity to domestic and foreign interest rates.

2. UBS Group stress scenarios

Country	Foreign exchange	Interest rates	Equity
	Price	Libor/Govt.	Price
Europe	+/– 10%	+/– 100 bps	+/– 15%
North America	+/– 5%	+/– 120 bps	+/– 15%
Japan	+/– 15%	+/– 100 bps	+/– 25%
Emerging markets	+/– 40%	+500/– 300 bps	+/– 40%

(+) = Market appreciation. (–) = Market depreciation.

UBS Group reveals a sample of its stress scenarios in its 1998 *Annual Report*. However, loss levels are not indicated.

2.9 Summary of stress tests

Most models make assumptions that don't hold up in abnormal markets. Stress tests are therefore essential for a comprehensive *Picture of Risk* and should be an integral component of the risk analysis and communication.

As with VaR analysis, stress testing must be done at different levels of the organization. The organizational hierarchy for stress testing is even more important than for VaR reporting. At the

desk level, traders are interested in stressing individual positions and specific risk factors. On a corporate level, senior management is concerned about macro stress scenarios that could pose a threat to firmwide operations.

The process of generating and discussing stress scenarios is a collective exercise in risk analysis. Stress tests are an opportunity to consider scenarios that most view as unlikely, but are possible. Make stress tests workable, realistic, and timely. Rather than stress everything, focus on relevant position-specific stresses. It is important for stress tests to tie back to the decision making process: stress results should guide corporate risk appetite decisions, impact limits, and be a judgmental factor in capital allocation.

Stress testing can be viewed from two perspectives: what would be the potential losses if certain events occurred, or what stress events could lead to losses of a certain magnitude? There are four major approaches for generating stress scenarios. The first uses historical scenarios and the second shocks market rates to examine portfolio sensitivities and concentrations. The third approach considers hypothetical future scenarios, based on current market conditions. The fourth approach searches for stress scenarios by analyzing portfolio vulnerabilities.

Chapter 3. Backtesting

3.1 Why backtest

Models are designed to reflect reality. **Backtests** compare realized trading results with model-generated risk measures, both to evaluate a new model and to reassess the accuracy of existing models. Although no single methodology for backtesting has been established, banks using internal VaR models for market risk capital requirements must backtest their models on a regular basis. The BIS imposes a penalty on institutions whose VaR models perform poorly. Banks generally backtest risk models on a monthly or quarterly basis to verify accuracy. In these tests, they observe whether trading results fall within pre-specified **confidence bands** as predicted by VaR models.[1] If the model performs poorly, they probe further to find the cause (e.g., check integrity of position and market data, model parameters, methodology).

Risk measurement can always be improved. The pragmatic question is whether the improvement in performance is worth the investment. Backtesting can help in this cost benefit analysis.

Example

For example, due to data constraints a dealer might be forced to use the HKD swaps yield curve to approximate the risk of HKD government bonds. Hypothetical backtests of VaR vs. P&L could show the difference between using a swaps or government curve. If VaR using actual government yields is not noticeably better, it may not be worth upgrading the data model.[2]

To mirror the three levels of internal VaR reporting, backtesting of risk models should be performed from the corporate to the desk level. The most important backtest is that of total diversified trading revenues at the corporate level—it shows how well the total aggregation of VaR estimates works. Corporate level backtesting is also necessary when using VaR to comply with **BIS market risk capital requirements**. The BIS outlines backtesting best practices in its January 1996 publication *Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements.*[3]

Some key issues to consider for backtesting include the following:

- What P&L measures to use?
- What to compare VaR against? Zero P&L or expected P&L?
- How to treat market-making businesses vs. pure proprietary trading businesses

3.2 Backtesting VaR vs. actual P&L

Financial control should keep a database of daily VaR and mark-to-market trading revenues for all trading desks and business units. Trading revenues should be defined as the change in mark-to-market in positions, plus any trading-related interest income or other revenue. The most straightforward way to backtest is to plot daily P&L against predicted VaR and to monitor the number of **excessions**, or departures, from the confidence bands. Hypothetical results, or "No-action P&Ls" may also be used for backtesting. No-action P&L assumes that we keep today's positions until the forecast horizon, at which point they are re-valued. Preferably, 90 days or

---

1 See the *RiskMetrics Technical Document*, pp. 219–223.

2 Note that positions reflecting typical trading strategies should be used for backtesting; a portfolio where HKD swaps are hedged with HKD government bonds (e.g., long and short positions) is different from a long swaps and bond portfolio. In the hedged portfolio, we have basis risk between swaps and government bonds, which can only be captured if you have both swaps and government curves. A backtest would show this.

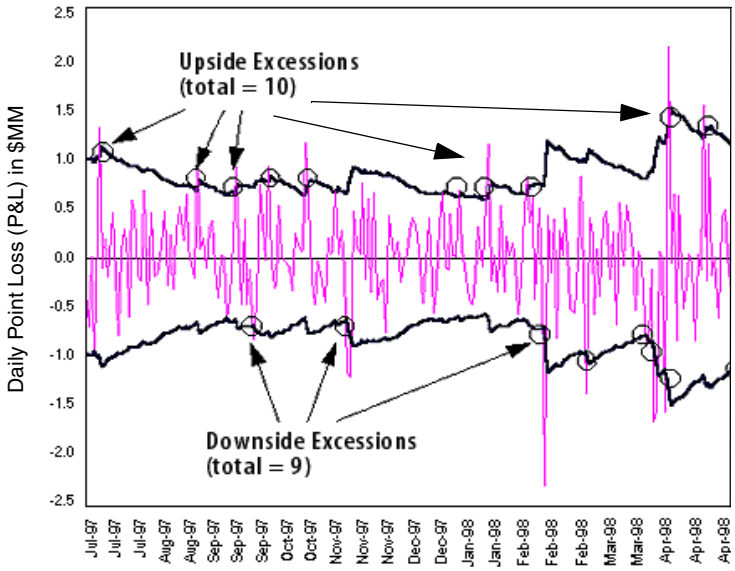
3 Document is available at http://www.bis.org/publ/index.htm.

more of history should be available for backtesting. According to the BIS, national regulators should use the number of excessions over the most recent 12 months of data (or 250 trading days) as the basis for supervisory response. Excessions should be within confidence level expectations: if you have a 1-day 95% confidence VaR, you should expect about 5% downside excessions over time. If actual excessions are significantly different, you can take steps to track down the source of error.

Example

In the example below, we see that the backtest results for Global Bank are reasonably close to expectation: 4% downside excessions instead of 5%.

Daily P&L (–) vs. predicted 95% confidence VaR (▬)



	Upside excessions	Downside excessions
Number	10	9
Percentage	4%	4%

VaRBaR

Confidence levels and backtesting

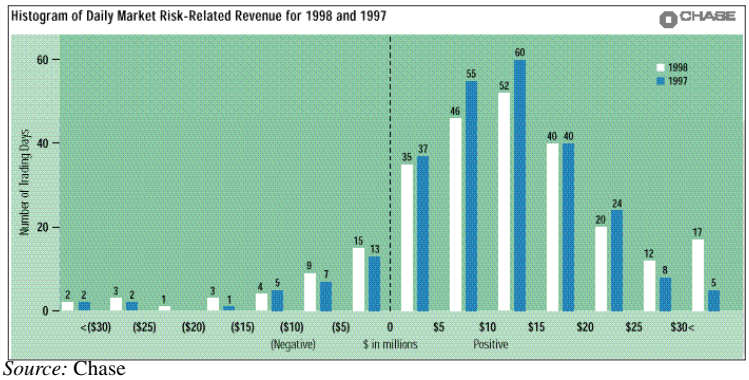
A 95% daily confidence level is practical for backtesting because we should observe roughly one excession a month (one in 20 trading days). A 95% VaR represents a realistic and observable adverse move. A higher confidence level, such as 99%, means that we would expect to observe an excession only once in 100 days, or roughly 2.5 times a year. Verifying higher confidence levels thus requires significantly more data and time. Even if your firm calculates VaR based on a high confidence level, it may make sense to test at other confidence levels as well in order to dynamically verify model assumptions (e.g., test at 90%, 95%, 97.5%, 99%).

An even better test would be to compare the actual distribution of returns against the predicted distribution of returns (i.e., how close is the picture of predicted risk to the actual risk). Such an approach for testing VaR models was proposed by Drachman and Crnkovic of J.P. Morgan. Instead of just checking excessions at a specific confidence level (e.g., 95%), the model tests all confidence levels, compares the distribution of forecasting errors against the uniform distribution, and assigns a "Q-Test" score between 0 and 1, with a lower score being better. For a full description of the methodology, see "Quality Control" in the September 1996 issue of *Risk Magazine*.

3.3 Accounting for non-position-taking income

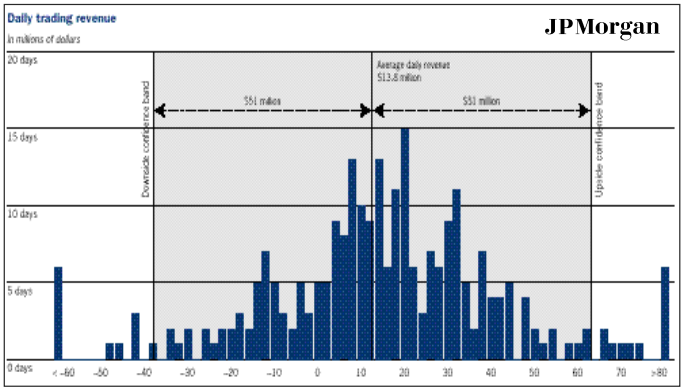
One debate among risk practitioners centers on how to account for non-position-taking income, such as fees. When backtesting VaR models, many financial institutions find that VaR tends to overestimate losses when compared against zero expected revenues, particularly on market-making books. Market-making books derive significant revenues from client flows, where firms can often earn a spread without taking directional views. To account for such non-position-taking franchise revenue, revenues can be backtested against VaR relative to expected revenues, instead of zero.

For example, if 95% confidence VaR on March 24, 1999 is $3 million, and average daily revenues are $0.5 MM, the 95% confidence bounds for that day would be –$2.5 million and +$3.5 million (not –$3 million and +$3million). It makes theoretical sense to take into account expected revenues because risk can be defined as unexpected loss, or deviation from expectation.

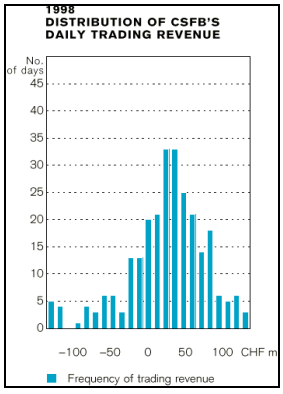


Source: Chase

Non-position-taking factors are significant especially for banks with strong market-making franchises. For example, notice the positively centered revenue distributions of Chase, CSFB, and J.P. Morgan, as published in their respective annual reports.



Source: J.P. Morgan



Source: CSFB

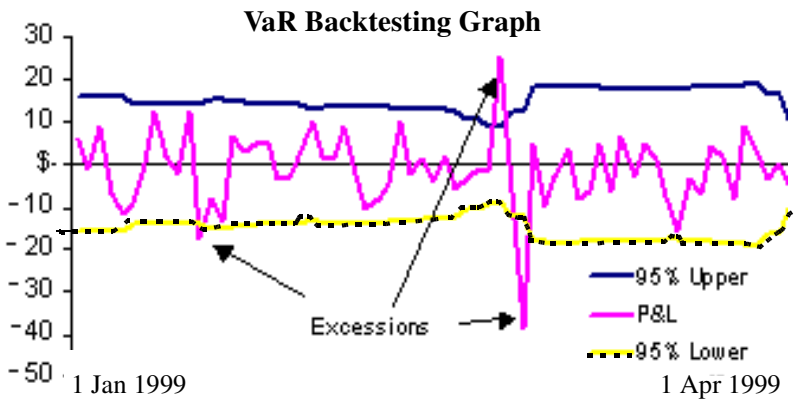
3.4 Backtesting VaR vs. hypothetical trading outcomes

The Basle Committee encourages financial institutions "to develop the capability to perform backtests using both hypothetical and actual trading outcomes." Hypothetical results, or "No-action P&Ls," are particularly useful for longer-horizon VaR estimates. For example, backtesting a 10-day horizon VaR for banks makes more sense with No-action P&Ls than with actual trading results because positions vary greatly on a daily basis.

Backtesting against No-action P&Ls follows the same format as testing against actual trading results.

Example **VaR vs. No-action P&L**

This test of hypothetical trading revenues shows only one upside (1.7%) and two downside excessions (3.3%) in over 60 trading days. Even though this model seems conservative in terms of percentage excessions, the actual excessions are large and clustered.



VaRBaR

Should VaR be conservative or accurate?

Risk managers may be tempted to excuse VaR models if they err on the side of conservatism. However, VaR should be accurate, not conservative. People don't pay attention to overly conservative models because they often are unrealistic. Supposedly conservative models may actually lead to excessive risk taking by giving a false sense of security.

When building models, it can be tempting to include many conservative assumptions. But when conservatism is layered upon conservatism inside the model, the result is unclear. Is it 99% VaR? 99.5% VaR? The important thing is to first get a precise VaR. If we then want to be more conservative, for example to estimate capital requirements, we can apply a transparent multiple to VaR (e.g., the Basle Committee specifies a multiple of 3 times the 99% confidence 10-day VaR as minimum regulatory market risk capital).

3.5 Interpreting backtesting results

How can we tell if the backtesting results are reasonable? Note that even if you sample a perfectly known distribution, where you are absolutely sure about the 5% level, and you take 100 samples, there is a reasonable probability of 4 or 6, or even 3 or 7 exceptions. In backtesting, we have two possible errors. Type 1 errors refer to rejecting a theoretically sound model that performed poorly due to chance. Type 2 errors refer to accepting a flawed model that performed well due to chance.

Given these anomalies, the Basle Committee recommends classifying outcomes in three categories: green, yellow, and red zone. Green indicates high probability of model validity, while red implies high probability of model flaw. Yellow is an ambiguous zone where "a supervisor should encourage a bank to present additional information before taking action."[4]

For a further discussion of this topic, refer to the BIS document.

[4] Document is available at http://www.bis.org/publ/index.htm.

3.6 Other factors to consider in analyzing backtests

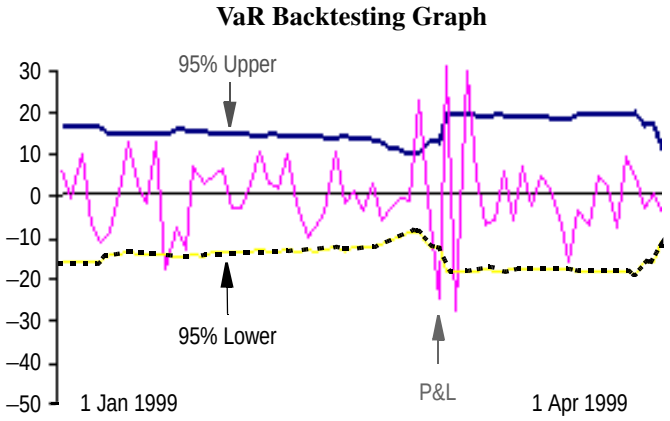
Even if excession percentages are within tolerance, it may make sense to probe further.

A. Look for clustering of excessions

In addition to counting the percentage of VaR band excessions, risk monitors should watch out for excession clustering. For example, even if a quarterly backtest shows exactly 5% upside and 5% downside excessions, it would be a disturbing sign if these excessions were clustered in one narrow time period. Clustering of excessions could imply high autocorrelation in risks, which may manifest itself as a losing streak.

Example **P&L vs. VaR test: clustered excessions**

In this graph of P&L vs. VaR, both upsides and downside excessions are within tolerance: we have three upside and three downside excessions in 65 trading days (i.e., 4.6% excessions each for upside and downside). Notice, however, the clustering of excessions in the beginning of March, when VaR was unresponsive to the increased revenue volatility. This may be due to a missing risk factor (e.g., spread risk) or poor parametrization (e.g., exponential weighting).



Ideally, one would see an even distribution of excessions through high- and low-volatility regimes, which would demonstrate that the VaR model is responsive to a variety of market conditions.

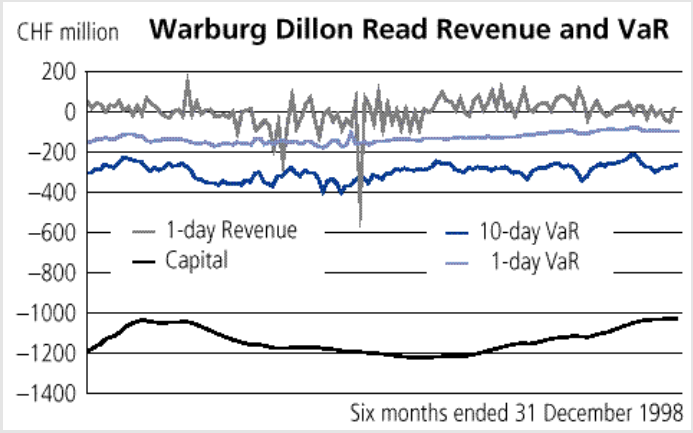
B. Examine magnitude of excessions

Another factor to consider when analyzing backtests is how large are excessions when they do occur. Extreme excessions (e.g., $5MM for $1MM VaR) are a red flag and point to the high probability of event risk. Thorough stress testing of extreme market moves should be performed regularly to estimate event risk. In the longer run, researchers might look at improving distributional assumptions to include event risk.[5]

[5] See the *RiskMetrics Monitor,* Q4 ’96 and Q4 ’97.

P&L vs. VaR test: large excessions

In this backtest of UBS Warburg Dillon Read's published trading revenues vs. VaR, we see two large clustered excessions and one extreme loss during the global market rout from August to October '98. These three excessions during this 6-month observation period total 2.4%, which is larger than expected, given UBS's 99% VaR confidence level. Also noteworthy is the size of the excessions. The graph reveals a 1-day loss of close to CHF 600 million, which is almost half of the CHF 1.2 billion capital allocated for market risk. Such results may prompt UBS to examine its VaR model's responsiveness to extreme market volatility and also continue its focus on stress testing.



3.7 External disclosures of backtests

Several major global banks have led the way in disclosing VaR and backtesting results. Below is one example.

From Chase Manhattan's 1998 *Annual Report*:

The Chase VAR methodology assumes that the relationships among market rates and prices that have been observed over the last year are valid for estimating risk over the next trading day. In addition, Chase's VAR estimate, like all other VAR methodologies, is dependent on quality of available market data. Recognizing these shortcomings, Chase uses diagnostic information to continually evaluate the reasonableness of its VAR model. This information includes the calculation of statistical confidence intervals around the daily VAR estimate and the use of daily "backtesting" of VAR against actual financial results.…

Chase conducts daily VAR "backtesting" for both regulatory compliance with the Basle Committee on Banking Supervision market risk capital rules and for internal evaluation of VAR against trading revenues. During 1998, a daily trading loss exceeded that day's trading VAR on 2 days. This compares to an expected number of approximately 3 days.

Considering the unsettled markets of 1998, Chase believes its VAR model performed at a very high level of accuracy during 1998.

3.8 Backtesting summary

Practitioners use backtesting to verify the accuracy of VaR models and to help analyze the costs and benefits of improving VaR models. Regulators require regular backtesting from banks that

they approve to use internal models for calculating market risk capital requirements. Voluntary disclosure is also becoming prevalent, as evidenced by several leading banks disclosing their trading revenues and VaR backtests.

Although there is no broadly established standard for backtesting, tests generally compare actual or hypothetical trading results against VaR bands to help determine if percentage excessions are within tolerance. Clustering and magnitudes of excessions are also informative and worth examining. Causes of outliers should be analyzed, and attributed to changes in volatility, correlation, or other factors.