THE UNIVERSITY OF GEORGIA
TERRY College of Business – Department of Banking and Finance

D I S S E R T A T I O N

Final version v4: November 18, 2001

FINANCIAL RISK MANAGEMENT AND VALUE-AT-RISK: THE IMPACT OF ASSET RETURN-GENERATING MODEL SPECIFICATIONS

MATEJ BLAŠKO

(Under the Direction of JIMMY E. HILLIARD)

ABSTRACT

Value-at-Risk (V@R) is a new, simple and yet informative measure of portfolio risk. This dissertation explores the impact of assumptions about an asset return-generating process on portfolio risk measurement and management within this new framework of enterprise financial risk management. The major contribution of this study is an identification of potential problems and consequences of model misspecification on risk measurement and ultimately on the corporate hedging decision. Chapter One reviews the normative and positive literature on financial risk management. Chapter Two explores the impact of model misspecification if the true (simulated) model returns are from a mixture of normal distributions with feasible parameters. The distributional properties of several assets, such as stocks, currencies and commodities, are examined in Chapter Three. The Maximum Likelihood Method and Method of Moments are used to estimate posited models' parameters and fit empirical distributions. V@R measures calculated from a mixture-of-normals model are then compared to measures from models commonly used by practitioners (e.g., RiskMetrics™), who assume either Gaussian, or some form of an ARCH (EWMA) process. The last chapter empirically explores characteristics of companies using V@R systems with a focus on the benefits and uses of Value-at-Risk systems in financial risk management.

(words: 197)

INDEX WORDS: Financial Risk Management, Value-at-Risk, VaR, Model Risk, Estimation Risk, Model Specification, Return Processes, Mixture of Normal Distributions, Hedging.

FINANCIAL RISK MANAGEMENT AND VALUE-AT-RISK:

THE IMPACT OF ASSET RETURN-GENERATING MODEL SPECIFICATIONS

by

MATEJ BLAŠKO

B.A., Comenius University Bratislava, Slovakia, 1995

M.S., Comenius University Bratislava, Slovakia, 1995

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA, U.S.A.

2001

FINANCIAL RISK MANAGEMENT AND VALUE-AT-RISK:

THE IMPACT OF ASSET RETURN-GENERATING MODEL SPECIFICATIONS

by

MATEJ BLAŠKO

Approved:

Major Professor:	Jimmy E. Hilliard
Committee:	Jeffry M. Netter William D. Lastrapes Qing Zhang Stewart Mayhew

Electronic Version Approved:

Gordhan L. Patel
Dean of the Graduate School
The University of Georgia
December 2001

DEDICATION

I dedicate this dissertation to my parents, Anna and Matej Blaško, for their love and support. Túto dizertačnú prácu venujem mojím rodičom.

ACKNOWLEDGEMENTS

I would like to thank my major professor, Jimmy E. Hilliard, and my dissertation committee members, Jeffry M. Netter, William D. Lastrapes, Qing Zhang, and Stewart Mayhew, for their comments and suggestions that improved this dissertation. I would also like to thank Joseph F. Sinkey, Jr., for his support during my years at the University of Georgia.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

INTRODUCTION 1

CHAPTERS

I FINANCIAL RISK MANAGEMENT AND ASSET RETURN-GENERATING MODELS 4

I.1 Overview 4

I.2 Enterprise Financial Risk Management 5

I.3 Value-at-Risk (V@R) 8

I.4 A Motivating Case Study 10

I.5 The Objective Function of Hedgers 12

I.6 Return Distributions in Finance Theory 14

I.7 Asset Return-Generating Processes and V@R Estimation 15

II ESTIMATION ERRORS AND MODEL SPECIFICATION 21

II.1 Introduction 21

II.2 The Sample Correlation Coefficient 27

II.3 Hedge Ratios, Hedge Volatility, and Hedge Return Distributions 35

II.4 Value-at-Risk and Model Misspecification 49

II.5 Conclusions 54

III ESTIMATION OF DISTRIBUTIONAL PARAMETERS AND V@R METRICS 56

III.1 Introduction 56

III.2 Model Selection and the Goodness-of-Fit Tests 57

III.3 The Mixture-of-Normal-Distributions Model 60

III.4 Data and Results 67

III.5 Conclusions 74

IV FIRM VALUE AND VALUE-AT-RISK 75

IV.1 Introduction 75

IV.2 Risk Management and Financial Institutions 76

IV.3 Sample Selection and Methods 80

IV.4 Empirical Results 81

IV.5 Conclusions 89

V SUMMARY AND CONCLUDING REMARKS 91

V.1 Major Findings 91

BIBLIOGRAPHY 96

INTRODUCTION

Value-at-Risk (V@R) is a new, simple and yet informative measure of portfolio risk. V@R measures the maximum dollar amount that an investor's portfolio can lose over a period of time with some degree of confidence or probability that the actual *(ex-post)* loss will exceed the V@R. The V@R framework has quickly become one of the most popular and important tools of enterprise financial risk management.[1]

This dissertation consists of four chapters. It primarily explores the impact of assumptions about an asset return-generating process on portfolio risk measurement with Value-at-Risk. The major contribution of this study is an identification of potential problems and consequences of model misspecification on risk measurement and ultimately on the corporate hedging decision.[2] In addition, this research investigates the characteristics of financial institutions using V@R systems and provides additional evidence on the benefits (or drawbacks) of financial risk management techniques and corporate hedging strategies. Because financial institutions as well as their regulators are increasingly using V@R risk-measurement tools in setting capital-adequacy requirements, an appropriate model specification is of paramount importance.

This research fits in the literature on corporate hedging and financial risk management following Smith, Stulz (JFQA 1985); Bessembinder (JFQA 1991), Nance, Smith, Smithson (JF 1993), Froot et al. (JF 1993, JFE 1998); Tufano (JF 1996); Mian (JFQA 1996); Minton, Schrand (JFE 1999); Ahn, Boudoukh, Richardson, Whitelaw (JF 1999), and Allayanis and Weston (2001), as well as the literature on asset price processes (Fama, 1965; Merton, 1976a; Cox and Ross, 1976; Bollerslev, 1986; Roll,

[1] Smithson (1996a, b), Bodnar et al. (1995), Jorion (2000)

[2] Tufano (1996) says: "Academics know surprisingly little about corporate risk management practices." Brealey and Myers (2000) quote Arnold Sametz (JF, 1964) who commented: "we know very little about how the great nonroutine financial decisions are made." The question, *"What Risks Should a Firm Take?"* is listed among their *Top 10 Unsolved Problems in Finance*. Because less risk is not always better, the question remains what is the *appropriate* (italics in BM2000) level of risk and how to set up a sensible risk management strategy.

1988), and estimation and model risk (Green and Figlewski, JF 1999; Britten-Jones, JF 1999; and Longin and Solnik, JF 2001, among others). This normative and positive literature is reviewed in CHAPTER ONE.

CHAPTER TWO investigates the potential adverse consequences of return-generating model selection (misspecification) on estimated portfolio risk measured within the framework of Value-at-Risk. The standard V@R model is a natural extension of the classic portfolio theory of Markowitz (1952, 1959), which assumes that asset returns are normally distributed.[3] However, as we know from the research of Fama (1965), Mandelbrot (1966) and more recently Richardson and Smith (1993), and Peiró (1999) among others, asset-return distributions exhibit significant non-normality. Several models have been proposed in order to account for this non-normality of unconditional time-series return distributions. Merton (1976a), Cox and Ross (1976), and Bates (1991) propose models of so-called *jump-diffusion (JD) asset return-generating processes*. Quandt and Ramsey (1978) and Roll (1988) propose a *mixture of normal distributions* model for use in economic and finance modeling. Mixture of normals can be viewed as a discrete-time alternative to *a continuous-time jump-diffusion process*. These models allow for substantial skewness and leptokurtosis of return distributions and are well suited for fitting empirical distributions.

The implications of assumptions about the return-generating model are explored. Since the parameters of a posited process are known by definition, it is possible to calculate the exact V@R metrics of returns generated by that process. True V@R can be compared to an estimate obtained by assuming (as an applied researcher would) that the returns are generated by some other process, such as *Geometric-Brownian Motion (GBM),* an *Autoregressive Conditional Heteroscedasticity (ARCH)* or its version, an *Equally-Weighted Moving Average (EWMA)* process. We hypothesize that the model misspecification leads to biased (distorted) estimates of V@R risk-measures.

CHAPTER THREE empirically examines return distributions of various assets, such as equities, currencies and commodities. To fit empirical distributions, the *mixture-of-normals model* parameters are estimated using the *Method of Moments* of Pearson (1894)

[3] This assumption about a specific stochastic process – Geometric Brownian Motion – restricts the type of admissible preferences (see, e.g., Bick, 1990; He and Leland, 1993; Lo, 1999).

and the *Maximum Likelihood Method*. The chapter focuses on the comparison of the V@R metrics calculated from 1) the mixture-of-normals model, and 2) models employed by RiskMetrics™, which assume either GBM, or a form of an ARCH – EWMA process.

CHAPTER FOUR empirically explores the economic value and benefits of utilizing Value-at-Risk systems in enterprise financial risk management. The Modigliani and Miller (1958) theorem and classic portfolio theory of Markowitz (1952, 1959) suggest that firms should not engage in risk management, since shareholders can costlessly diversify idiosyncratic risks on their own. However, the existence of market frictions provides reasons for corporate-level financial risk management and explains how hedging increases shareholder wealth (see Froot, Scharfstein, and Stein, 1993; Smith and Stulz, 1985; DeMarzo and Duffie, 1995; among others). Therefore, the corporate decision to hedge should have a measurable effect on some firm-value characteristics. We investigate the use of V@R systems by financial institutions and identify benefits (or drawbacks) arising from the use of these financial risk management techniques. Also, recent Basel Committee (1997) regulations allow institutions to use their own risk management systems in setting regulatory capital requirements. This study provides implications for regulatory policy.

Chapter I

FINANCIAL RISK MANAGEMENT AND ASSET RETURN-GENERATING MODELS

I.1 Overview

CHAPTER ONE reviews the existing literature on enterprise financial risk management, asset return-generating processes, and establishes the main research question. The goal and major contribution is an investigation of potential adverse consequences of return-generating model selection on estimated portfolio risk. The Value-at-Risk metric calculated from models assuming leptokurtic returns is compared to the V@R metrics calculated from models with constant and Markov variance processes (GBM, ARCH/EWMA), which assume that discrete-time returns (over a certain period, e.g., one day) are normally distributed. Users of V@R risk management systems and data vendors such as RiskMetrics™ now commonly employ GARCH models. We hypothesize that certain model specifications significantly under- or over-estimate measures of risk (V@R) if prices follow a process that generates leptokurtic returns. The second section of this chapter discusses a short case describing one of the risk management problems facing investors, traders and underwriters.

This research provides additional insight into the portfolio selection problem. If an asset is added to a portfolio, then under certain conditions, the risk of a (hedged) position as measured by V@R may increase despite a decrease in the portfolio's variance. We investigate the conditions that must be satisfied in order to decrease an underwriter's position risk. In general, the model risk and parameter estimation errors are shown to have a substantial impact on the hedge return distributions.

I.2 Enterprise Financial Risk Management

Firms are exposed to various types of risks. In general, these can be classified into *business risks* (product market, technological innovations), *strategic risks* (shifts in the economy or political environment), and *financial risks* (market price risk, credit risk, liquidity risk, operational risk, and legal risk).[4] Enterprise financial risk management can be described as a process by which firms identify, measure, manage and control for various financial risks.[5] Financial risk in general is defined as a measurable possibility of a decrease in an asset's value. However, one of the difficulties in analyzing corporate risk management systems is that the approaches to risk measurement are quite different from the standard measures used in asset pricing models.

Corporations spend enormous resources on risk management systems. According to surveys by Smithson (1996a, 1996b), most firms apply modern financial techniques to manage some of their exposures to financial risks. A survey conducted for a group of financial institutions known as the Group of Thirty (1993, Global Derivatives Project) reports that over 80% of the surveyed corporations considered derivatives important in controlling risk. Moreover, 43% of dealers reported use of some form of V@R and 37% of them indicated plans to adopt V@R-based risk management systems. The 1995 Wharton/CIBC Wood Gundy Survey of derivatives usage reports that 41% of U.S. non-financial firms use derivatives, 48% of end-users use stress testing or scenario analysis and 35% use Value-at-Risk.[6] A survey by the New York University Stern School of Business reports that 60% of responding pension funds use Value-at-Risk.[7]

From the perspective of modern finance theory, corporate hedging may seem puzzling, as shareholders may diversify and hedge risks on their own. In a Modigliani-Miller (1958) framework, corporate hedging and other financial decisions cannot create value for a firm unless they affect either the firm's ability to operate its business or its incentives to invest in the future. Violations of assumptions of the MM theorem have led to several testable hypotheses about the ways risk management enhances firm value.

---

[4] Jorion (2000, p.15)
[5] Jorion (2000, p.3)
[6] Smithson (1996b), Dowd (1998), p.19
[7] Linsmeier and Pearson (1996, p.2), as reported in Dowd (1998, p.19)

The assumptions about frictionless markets are likely to be violated in practice. Most investors face transaction costs and may not have equal access to hedging instruments. In addition, investors may have less information about corporate exposures. These problems may affect idiosyncratic risks. However, most investors hold diversified portfolios. Thus, corporate-level hedging is unlikely to reduce the firm's cost of capital significantly.[8] If hedging cannot reduce the discount rate, then it must increase expected cash flows in order to improve firm values. The primary venues for value creation by corporate hedging are therefore linked to decisions changing tax liabilities, changing contracting costs, or changing investment incentives.

The finance literature identifies several benefits of corporate hedging policies:

1. Hedging can reduce the costs of financial distress and bankruptcy (Mayers and Smith, 1982; Smith and Stulz, 1985).
2. Risk management that lowers earnings volatility will be optimal for shareholders if it reduces expected taxes (Smith and Stulz, 1985). The tax hypothesis assumes that firms face a convex tax function. If effective marginal taxes on corporate income are convex functions of a firm's pre-tax value, then a firm's after-tax value is a concave function of its pretax value. Hedging that reduces the variability of pre-tax values, reduces the expected value of future tax liabilities.
3. Firms employing risk management systems may select a better capital structure (see Harris and Raviv, 1991, for a review). Increased leverage may decrease the agency costs of managerial discretion (Stulz, 1990). Hedging may reduce over-investment problems as debt serves as an effective agency cost control device (Jensen, 1985). As Leland (1995) notes, the benefits of risk management may be larger when the agency costs are low.
4. Hedging also prevents a shortfall in funds and thus the under-investment problem (Myers, 1977; Bessembinder, 1991), by allowing a firm to avoid costs associated with external financing (Froot at al., 1993, 1998).
5. Without firm-level hedging, it may not be feasible to distinguish between losses associated with market exposures and losses due to a negative-NPV project

[8] Grinblatt, Titman (1998, p.714)

selection. Hedging can therefore improve executive compensation contracts and performance evaluation (DeMarzo and Duffie, 1995).

6. An active risk management program can improve internal markets and a management's divisional policy. A central headquarters' ability to allocate capital to internal business units will improve if their profit volatility is low (Grinblatt, Titman, 1998, p. 723).
7. Many industries are regulated and some financial institutions have to conform to risk-based capital requirements imposed by regulators. A simple regulatory process is characterized in Myers (1972) and discussed in Mayers and Smith (1982). Regulators set or allow prices that reflect expected costs plus a normal rate of return. If regulated firms insure or hedge (mostly) idiosyncratic risks, they are able to pass the cost on to consumers. This is not possible for firms in a non-regulated industry where prices are set in a competitive marketplace, independent of whether the firm insures. Reduction in risk may also be a less costly alternative to raising additional capital after a negative project-return shock (Froot and Stein, 1998).

The literature also identifies cases when corporate hedging may lead to a decrease in firm value, in addition to the obvious costs of hedging. Tufano (1998) discusses a potential cost of corporate risk management strategies that are based on cash flow hedging. On one hand, these strategies may allow firms to avoid the deadweight costs of external financing. However, in the presence of agency conflicts between managers and shareholders, hedging strategies can be used to reduce shareholder wealth since they remove the valuable discipline of external financing markets. For an overview of the literature on corporate hedging, see Nance, Smith, Smithson (1993); and Smith (1995). Grinblatt and Titman (1998, Chapter 20) also provide a general survey of financial risk management and corporate strategy.

Empirical evidence on corporate hedging practices is somewhat consistent with theoretical predictions. Nance, Smith and Smithson (1993) and Géczy, Minton and Schrand (1997) document that firms with greater growth opportunities are more likely to hedge. In particular, firms with relatively higher R&D expenditures and higher M/B ratios are more likely to use derivatives. Minton and Schrand (1999) show that higher cash flow volatility is associated with lower average levels of capital expenditures and R&D. This suggests that firms do not use external capital markets to fully cover cash flow shortfalls, but rather to postpone or permanently forgo investment. Nance, Smith and Smithson (1993) find weak evidence that firms with more leveraged capital structures hedge more. These results may, however, reflect the tendency of firms facing high costs of distress to have low leverage ratios. Géczy et al. (1997) report the absence of an association between the leverage and the use of derivatives.

A study by Tufano (1996) documents that managerial incentives have an important effect on the risk management practices in the gold mining industry. The results of Mian's (1996) investigation of 3,022 firms do not provide support for the corporate hedging decision models. Mian's evidence is inconsistent with financial distress cost models and is mixed with respect to contracting cost, capital market imperfections and tax-based models.

Though the evidence gives reasons for corporate hedging and documents that firms' hedging policies are consistent with predictions, there is very little evidence that risk management increases firm value. The first study that *directly* examines the relationship between firm value and hedging is Allayanis and Weston (RFS, 2001).[9] They examine the impact of the use of foreign currency derivatives on firm value in a sample of 720 U.S. non-financial firms between 1990-1995. Their results confirm that markets put a premium of about 5% of firm value on firms that hedge their exposure to exchange rates.

I.3 Value-at-Risk (V@R)

The V@R statistic provides a description of an asset's (or portfolio's) risk-exposure. At the core of the V@R estimation is a forecast of the probability distribution of possible asset values (or equivalently, asset returns) over a specified time horizon,

[9] Nance, Smith and Smithson (1993) do not provide any *direct* evidence that hedging increases the firm value. Their results suggest only that corporate choice to hedge is *consistent* with the theoretical predictions.

usually a few days. In general, the term Value-at-Risk (V@R) can be used in different ways, depending on a particular context:

A. V@R is most commonly referring to a maximum dollar amount that an investor's portfolio can lose over a period of time with some degree of confidence, or probability, that the actual loss will exceed this amount. One possible specification is :

$$V@R_{(N,\,x\%)} = V_{p,t0} * [\, exp(\pi) - 1 \,] , \qquad 1.3.1$$

$$where: \quad \int_{-\infty}^{\pi} \xi(N, z)dz = x\% \ ,$$

and $V_{p,t0}$ is the portfolio's value at time t_0 (today), π is a continuously compounded N-day return (log price difference) subject to the above condition, and $\xi(N, z)$ is the probability distribution function of the asset's continuously compounded N-day returns. Alternative definitions are possible (e.g., simple return or price-distribution based). For example, if returns were expressed as simple returns, the simple return would replace the bracketed term in 1.3.1.

In other words, the V@R metric represents *an absolute dollar loss* (or return) corresponding to a prespecified percentile (usually 5% or less) of the left-hand tail of the asset's price (or return) distribution. For example, managers can quantify risk in a following manner: "We don't expect our portfolio to lose more than 7% (or $7M given our portfolio's value of $100M) in more than 1 out of the following 50 weeks." For example, assuming normality of returns, the specific value of V@R metric can be estimated as follows:

$$V@R\ (N,\ x\%) = a_{x\%}\ \sigma\ N^{1/2} \ , \qquad 1.3.2$$

where x% denotes the confidence level, or the probability that the actual (ex-post) loss will exceed estimated V@R; $a_{x\%}$ corresponds to the x-th percentile of cumulative standard normal distribution Φ, so that $\Phi(a_{x\%}) = x\%$; σ^2 is the asset's daily return variance, and N is the number of days for the V@R metric estimation purposes.

B. The term V@R often means numerical or statistical *procedures* used to calculate V@R metrics and is referred to as V@R estimation procedure. We can also refer to a *V@R methodology* that includes a set of procedures and assumptions.

C. The corporate utilization of V@R figures in financial risk management can form a basis for a distinctive *V@R approach to risk management (VAR-RM).* However, as discussed by Basak and Shapiro (2001), this approach may have some negative consequences. Managers may simply decrease the V@R metric at the expense of the magnitude of losses when losses occur. This practice may increase volatility in down markets.

Jorion (2001) and Dowd (1998) provide excellent surveys of V@R approaches to enterprise risk management, methodologies, estimation procedures, case studies and common problems.

I.4 A Motivating Case Study

On May 14, 1997, Goldman Sachs competed for the biggest block trade in history.[10] The Kuwaiti Investment Office (KIO) was selling a single block of British Petroleum shares worth $2 billion. Goldman Sachs had originally acted as one of the underwriters for the British government when it initially sold the state-owned oil company. In the wake of the 1987 market crash, the British government repurchased the shares from the underwriters and later resold them to KIO. In 1997, KIO held 9.3% of the oil company and was seeking to sell 3%. For Goldman Sachs, the firm that invented block trading, the operation "was a simple matter."[11] At the time of the offering, shares of BP stood at 744 pence. Partners of Goldman Sachs decided that clients would be

[10] Endlich (2000)
[11] Ibid.

interested in the shares at 716 pence and decided to bid 710.5 pence per share on all 170 million shares. The bid included a profit spread of 75 basis points or about $15 million.[12]

There are several risks inherent in similar trades: the price of shares can plummet on the information of a large block sale, lack of liquidity, and most importantly idiosyncratic and market risks of a fall in prices. The BP trade was worth 40% of Goldman Sachs equity, and a substantial loss on this trade was a real possibility. As former co-head of the Goldman's equities division explained: "At the end of the day, it is about risk tolerance . . ." Goldman Sachs needed to take into account the tradeoff between risk and trading profits. From the risk management standpoint, the firm was considering the worst-case scenarios and the potential for a downside. One of these scenarios was a stock market crash in the time between the purchase from KIO and the clients' trade confirmations. The firm estimated that it could lose about $100 million, with potential losses of $50 million being more realistic. The partners felt that "it would be worth losing $25 million . . . in exchange for the elevated profile."

The general strategy of Goldman Sachs is not to hedge block trades. Even in the case of the $2 billion sale of BP, they concluded: "There was no good simple hedge available." A hedge in a form of the sale of a related instrument to reduce the market risk was not found suitable. One person involved in the deal said: "The danger is that the placement fails and the markets rally and you increase your risk exposure."[13]

The most important point of this section is the (lack of) hedging strategy implemented. The interesting question relates to the choice of an optimal strategy. It may be possible, that under some conditions (such as estimation and model uncertainty), the optimal strategy is not to hedge. There is a number of imperfectly diversified investors that are exposed to large idiosyncratic risks. The question remains, how to manage the exposure and how to quantify the risks inherent in large undiversified holdings. We explore several issues addressing these questions in sections that follow.

---

[12] The Wall Street Journal, May 16, 1997, "Big BP Share Sale Yields Easy $15 Million For Goldman Sachs", via Dow Jones Online News.
[13] Endlich (2000).

I.5 The Objective Function of Hedgers

The obvious answer to the hedging issue of Goldman Sachs is to suggest a short sale of another asset positively correlated with BP. The likely assets with these characteristics are shares of another oil company, an equity index, or an oil-related commodity. It may seem natural that *any* such trade would *decrease* the risk of Goldman Sachs' position.

From a basic result in investments we know that investors can decrease the variance of their portfolios by adding another less-than-perfectly correlated asset. In other words, subject to certain restrictions, the objective function may be to minimize portfolio variance:

$$\sigma_P^{\,2} = w'\,\Omega\, w\,, \qquad 1.6.1$$

where $\sigma_P^{\,2}$ is the portfolio variance, w is a column vector of weights of the portfolio's component assets, subject to $\Sigma_i\, w_i = 1$, and Ω is a variance-covariance matrix of asset returns. For a two-asset portfolio this relationship becomes:

$$\sigma_P^{\,2} = w_1^{\,2}\,\sigma_1^{\,2} + (1-w_1)^2\,\sigma_2^{\,2} + 2\,w_1\,(1-w_1)\,cov(r_1, r_2)\,. \qquad 1.6.2$$

However, in the case of Goldman Sachs, the objective function was different. The portfolio selection decision rests on diversification given some total value of an investment. Investors would replace a portion of one asset with another one, rather than adding another contract to their original portfolio. In a block trading or underwriting, it is not possible to rebalance the portfolio. Thus, the underwriter's objective function is to minimize risk given some fixed amount of securities underwritten.

Moreover, even for investors diversifying a given endowment, the objective function expressed by minimization of $\sigma_P^{\,2}$ may lead to a suboptimal decision. The *mean-variance* framework of Markowitz (1952, 1959) rests on the assumption of joint normality of asset returns and investors' quadratic utility functions. If asset returns exhibit jumps or substantially deviate form normality, then portfolio selection based on the *mean-variance* framework may lead to suboptimal decisions if investors maximize their expected utility of future consumption, *E[U]*.[14]

The issue of risk measurement is not confined to underwriters. As Stulz (1999) points out, the mainstream approach to capital budgeting focuses exclusively on the special case where the firm or investment-specific risks do not affect the contribution of a project to the value of the firm. There are numerous theoretically sound explanations for why firms do and should hedge certain risks.[15] If returns on firm projects are not normally distributed, then certain large investments may increase, rather than decrease risk and thus influence firm value (Stulz, 1999). This is due to the fact that as the value of a firm falls, it becomes unable to take advantage of valuable opportunities it could take if it did not face equity capital constrains. Finance theory also suggests that firms in financial distress find it difficult to enter into contracts requiring financial commitments on the part of the counterparties.[16] It follows that a firm with a nontrivial probability of financial distress may not be able to invest in projects that it would find valuable if its probability of distress were zero. Each new project may thus impose costs, and firms have to take these costs into account by quantifying a project's risk. Firms also need to understand how a new project impacts the total risk of the firm. Because the risk that is costly is the risk associated with large losses, the appropriate measures of risk are lower-tail measures of risk such as Value-at-Risk or Cash-flow-at-Risk rather than volatility per se (Stulz, 1999, p.10).

One of the simplest and most intuitive solutions to the risk measurement problem of underwriters and investors/firms can be thus examined within the Value-at-Risk framework. The risk measurement problem is addressed later in this chapter and examines the potential problems that arise if asset returns exhibit jumps (or equivalently, return distributions are skewed and leptokurtic) and traders act as if (conditional or unconditional) returns are normally distributed.

---

[14] See SECTION FIVE for more details.

[15] See SECTION TWO of CHAPTER ONE for a brief review of literature.

[16] This relates to the underinvestment problem of Myers (1977), and the agency theory of Jensen and Meckling (1976), and Jensen (1985), among others.

I.6 Return Distributions in Finance Theory

The role of asset return distributions in finance theory has to be addressed. The investor's portfolio selection or the firm's project acceptance decision may depend on the characteristics of the project return distribution. Many theoretical models and empirical applications in finance assume that log-price differences are normally distributed.[17] However, studies of Mandelbrot, 1963; Fama, 1966; Hagerman, 1978; Richardson and Smith, 1993; Harrison, 1998; Campbell, Lo, and MacKinlay, 1998, among others, document that assets' price returns exhibit significant non-normality. Mandelbrot (1997, p.253) in his case against the lognormal distribution of prices, points out several drawbacks, "each of which suffices to make the lognormal dangerous to use in scientific research." Peiró (1999) documents skewness in some (least capitalized) markets, though his results utilizing the mixture-of-normals, Student's *t* distribution, and distribution-free methods cannot rule out symmetry in many stock-index and forex return-series he examines.

He and Leland (1993) investigate sufficient and necessary conditions that must be satisfied by a general equilibrium asset price process. The existence of a certain class of (power) utility functions implies a skewness preference that is positively valued by investors. Preference for positive skewness may explain the low diversification of many investors' portfolios (Simkowitz and Beedles, 1978; Conine and Tamarkin, 1981) and "the incorporation of skewness into the investor's portfolio decision can cause a major change in the construction of the optimal portfolio" (Chunhachinda *et al.,* 1997).

These studies suggest that better estimation and identification of the assets' return distributions and measurement of risk may improve decisions of traders, underwriters, and other investors in general. V@R suggests itself as a promising venue to address these issues.

---

[17] The normality assumption, or Gaussian hypothesis, is due to Bachelier (1900) and Osborne (1959) and is based on the central limit theorem. If the transaction price changes are independent and identically distributed random variables with a finite variance, then the central limit theorem would lead us to believe that returns over intervals such as a day, or a week will be normally distributed. The documented autocorrelation of returns may lessen the speed of convergence to the normality. Some more complex processes may even lead to the failure of the central limit theorem (see Feller, 1959).

Campbell, Lo and MacKinlay (1997) provide evidence on return moments. Sample estimates of skewness for daily U.S. stock returns tend to be negative for stock indexes but close to zero or positive for individual stocks. Sample estimates of excess kurtosis of daily U.S. stock returns are large and positive for stocks as well as indexes.[18] Early studies of returns modeled this excess kurtosis within a class of stable distributions. Paul Lévy (1924) was among the first to investigate stable distributions.[19] In order to recognize fat-tails of return distributions, applied financial economists explicitly model time-varying second moments. The most important of these studies is the Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Engle (1982). For an in-depth overview of ARCH modeling in finance, theory and evidence, see Bollerslev, Chou, and Kroner (1992), and Gouriéroux (1997).

I.7 Asset Return-Generating Processes and V@R Estimation

The methodologies of V@R estimation employed by RiskMetrics™ and other V@R users mostly assume conditional or unconditional normality of asset returns. However, the correctness of this assumption for large portfolios rests on the special Lindberg-Feller version of the Central Limit Theorem,[20] which posits that sums of variables with *any* distributional properties will converge in distribution to the normal distribution, given that no one random variable (asset) dominates. In practice, cross-correlations and higher moment dependencies can induce severe departures from normality even for large portfolios (e.g., S&P500 index), not to mention distributions of single asset or undiversified portfolio returns.

Distributional assumptions may severely distort the V@R metrics in some cases. The issue is best illustrated by recalling Chebychev's inequality, which states that for any

[18] See Campbell, Lo, MacKinlay (1997, Ch.1) for an introduction to Prices, Returns, and Compounding, and early studies of returns and return modeling.
[19] Normal or Gaussian distribution is a special case of a class of stable distributions. These distributions are stable under addition, e.g., a sum of stable random variables is also a stable random variable. The class is also termed *stable Pareto-Lévy, stable Paretian,* or *L-stable.*
[20] See Greene (1996). The general specification requires independence of random Variables. Extensions without the independence assumption require other restrictions (Feller, 1968, 1971).

random variable X with finite variance, the probability of falling outside a specified interval is:

$$P(|X - \mu| > r\sigma) \leq 1 / r^2 , \qquad 1.7.1$$

where μ, and σ^2 are the known mean and variance of random variable X. The V@R methodology is attempting to estimate the value of $r\sigma$, given some confidence level or probability mass of extreme tails. If the distribution is symmetrical, then the following holds for the left tail:

$$P\{ (X-\mu) < -r\sigma \} \leq ½ r^2 . \qquad 1.7.2$$

The standard V@R methodologies assume normality of returns and would report V@R at the 99% confidence level as:

$$V@R_{Normal} = \alpha_{(99\%)} \sigma = 2.3 \sigma . \qquad 1.7.3$$

However, the inequality *1.7.2* says that an extreme case of distribution would require $r_{(99\%)}$ = 7.07. Thus institutions estimating V@R assuming normality of returns can, at an extreme, understate the true loss potential by a factor of 3.03 (= 7.07/2.3). This correction factor for distributional misspecification has justified the rulings of Basel Committee requiring banks to have equity capital equal to 3 times their V@R.[21]

Moreover, V@R estimates will critically depend on the estimated variance of asset returns. If asset returns exhibit jumps and the estimation procedure uses a short history of returns, the estimated variance may deviate substantially from the true variance and potentially provide misleading V@R risk measures. See the subsection on jump-diffusion processes for a formal statement. What follows is a brief description of the most commonly used models of return-generating processes.

[21] See CHAPTER FOUR for Basel Committee capital adequacy rules for financial institutions.

Geometric-Brownian Motion (GBM), Gaussian distributions.

Many financial applications, including standard V@R methodologies, assume that asset prices follow the process:

$$dP/P = \mu\, dt + \sigma\, dW\,, \qquad 1.7.4$$

where μ (drift) and σ^2 (variance) are constants and dW is an infinitesimal increment of the standard Gauss-Wiener process, and $\Delta W(\tau) \sim N[0, \Delta\tau]$, where "~" means "is distributed as." This model of asset price behavior is also known as *Geometric Brownian Motion (GBM)*. The resulting distribution of *prices* is lognormal and distribution of *returns* (log-price differences, continuously compounded returns) is Normal (Gaussian). The price process can also be written as:

$$P_\tau = P_0\, exp[\, (\mu - \sigma^2/2)\, \tau + \sigma\, W(\tau)] \quad . \qquad 1.7.5$$

Jump-Diffusion (JD) Asset Return-Generating Process

The JD model is due to Merton (1973), Cox and Ross (1976), and Bates (1991). It can be viewed as a mixture of a diffusion process and a jump process. The specification is intuitively appealing as the asset price dynamics can be written as a combination of two types of changes: (a) 'normal' smooth vibrations in price due to a temporary supply/demand imbalance, changes in capitalization rates, or other marginal changes, and (b) 'abnormal' instantaneous jumps in price due to arrival of new important information about an asset's value. It is assumed that the important information arrives only at discrete points in time.

As Merton (1976) points out, the difference between *GBM* and *JD* models is particularly important because the qualitative characteristics of the two processes are fundamentally different. It is reasonable to expect that there will be *ex-post* "active" periods and "quiet" periods depending on information arrival, though the active and quiet periods are *ex-ante* random.

If $P(\tau)$ denotes asset price at time τ, then the posited price dynamics can be represented as:

$$dP/P = (\mu - \lambda k)\, dt + \sigma\, dW + dq \quad , \qquad 1.7.6$$

where $dq=(Y-1)$ if a Poisson (news) event occurs, and 0 otherwise; μ is the asset's instantaneous expected return; σ^2 is the instantaneous asset variance conditional on a no-information arrival; dW is an increment of a standard Gauss-Wiener process, $\Delta W(\tau) \sim N[0,\Delta\tau]$; $q(\tau)$ is the Poisson process independent of dW; λ is the mean number of arrivals of important new information per unit time; and $k \equiv E(Y-1)$ is the expected percentage change (or return, possibly itself a random variable) in the stock price if the information (Poisson) event occurs.

If μ, λ, k, and σ are constants, then the asset price at time τ can be written as:

$$P(\tau) = P_0 \cdot \exp[(\mu - \sigma^2/2 - \lambda k)\, \tau + \sigma\, W(\tau)] \cdot Y(n) \qquad 1.7.7$$

where $Y(n) = 1$ if n=0, and $Y(n) = \Pi_{(j=1,n)}\, Y_j$ for n>0, where Y_j are independently and identically distributed and n is Poisson distributed with parameter $\lambda\tau$. In the special case when $\{ Y_j \}$ are themselves log-normally distributed, then the $P(\tau)$ will be log-normal with the variance parameter a Poisson-distributed random variable.[22]

Autoregressive Conditional Heteroscedasticity (ARCH) models

A class of autoregressive processes due to Engle (1982) has gained widespread acceptance. Bollerslev (1986) has proposed a generalized version of Engle's ARCH model *(GARCH)*. Bollerslev, Chou, and Kroner (1992), and Gouriéroux (1997) provide an in-depth overview of ARCH modeling in finance theory and evidence. The stochastic process for asset price dynamics can be represented in a manner similar to *GBM,* except that the volatility parameter $\boldsymbol{\sigma}$ is not a constant. One possible parameterization, suggested

---

[22] See Merton (1976) and Press (1967) for this specification.

by Engle (1982), is to express $\boldsymbol{\sigma}$ as a linear function of past squared values of the process. The most general model is *GARCH(p, q)*:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t$$
$$\sigma_t^{\,2} = \beta_0 + \Sigma_{(i=1,\,p)}\, \gamma_i\, (\sigma_{t-j})^2 + \Sigma_{(j=1,\,q)}\;\; \beta_j\, (\varepsilon_{t-j})^2\,, \qquad 1.7.8$$

where $\beta_0 > 0$ and $\alpha_i \geq 0$ are constants and $\varepsilon_{t-i} \sim N[0, \sigma_{t-i}]$ is an innovation on day *(t-i)*. The most popular is Engle's original *ARCH(1)* model, which can be obtained from the above *GARCH(p, q)* specification by restricting *p* and *q*, *GARCH(p=0, q=1)*. The commonly used *GARCH(1,1)* model is like an ARCH(∞) process with decaying weights. Also popular is a restricted *GARCH(1,1)* referred to as an *Equally Weighted Moving Average (EWMA)* process. It is the model of choice for data vendors such as RiskMetrics™. This process can be represented as follows:

$$r_t = \alpha_0 + \varepsilon_t \quad ; \quad \varepsilon_t \sim N[0, \sigma_t]$$
$$\sigma_t^{\,2} = \lambda * \sigma_{t-1}^{\,2} + (1-\lambda) * \varepsilon_{t-1}^{\,2}\quad , \qquad 1.7.9$$

where r_t is an asset return at time t, and λ is a decay factor. One additional simplification is frequently made about the drift parameter: α_0=0. Though this assumption leads to $E[r_t]=0$ and contradicts the risk-return tradeoff principle, it is commonly made for short time periods such as a day or a week. The estimator of changing volatility depends on a parameter λ, ($0<\lambda<1$). RiskMetrics™ selects λ=0.96 in its V@R metric estimation. The presumed strength of the EWMA volatility estimator is its ability to respond fast to more recent observations and potential changes in volatility. From the standpoint of Merton (1976), this feature is also its major weakness.

The class of GARCH processes is very popular in applied financial economics. However, as Merton (1976, p.339) points out, if the true process follows a general jump-diffusion process, then the estimated variance per unit time will be affected by the choice of the length of estimation period: "... if investors believe that the underlying process for the stock does not have jumps, then they may be led to the inference that the parameters of the process are not constant when indeed they are."

Suppose that investors observe the price changes over a fixed time period but with a large number of observations. They may be led to believe that they have a very accurate estimate for each time period's variance. If the true asset return-generating process follows the *JD process 1.7.6*, then conditional on *m* jumps having occurred during the observation period, the returns r_t will be normally distributed with variance:

$$E\{[r_t - E(r_t)]^2\} = \sigma^2 + m\delta^2/T \qquad 1.7.10$$
$$= V^2 + \delta^2 (m-\lambda T)/T \quad .$$

If an investor samples returns at high frequencies, then he will believe that the observed variance is very close to the true value of the variance rate. However, if one time period were a quiet period (*ex-post,* e.g., $m<\lambda T$), then he would conclude that the variance rate on the "perceived" process is not constant. Moreover, there would appear to be an effect called regression-to-the-mean, and variance would 'fluctuate' around the 'long-run' average variance, V^2.

For this reason, it may be necessary to have a long enough past history of asset prices to obtain an estimate of true, unconditional variance per unit time of the process, e.g., $V^2 \equiv \lambda\delta^2 + \sigma^2$. This condition is also violated for ARCH specifications since they put more weight on recent data.

Chapter II

ESTIMATION ERRORS AND MODEL SPECIFICATION

"While we can readily describe the process of sharp reversals in confidence, to date economists have been unable to anticipate it. Nevertheless, if episodic recurrences of ruptured confidence are integral to the way our economy and our financial markets work now and in the future, the implications for risk measurement and risk management are significant.

Probability distributions estimated largely, or exclusively, over cycles that do not include periods of panic will underestimate the likelihood of extreme price movements because they fail to capture a secondary peak associated with extreme negative outcomes. Furthermore, joint distributions estimated over periods that do not include panics will underestimate correlations between asset returns during panics."

Alan Greenspan (2000, page 110)

II.1 Introduction

CHAPTER TWO examines the consequences of various model specifications and estimation errors on correlation coefficient estimates, hedge volatilities, hedge distributions, and ultimately the hedging decision. Finance theory provides numerous models guiding the economic agent's decision making. One of the models is the *minimum variance hedge* technique. In practice, parameters have to be estimated and assumptions about the return-generating model made. While there is a large body of literature on some aspects of hedge ratios and correlation coefficients, the area of estimation errors and model misspecification has not received appropriate attention in the academic literature.

Parameter and asset-return model uncertainties are shown to have the ability to increase rather than decrease the risk of a classic minimum variance hedge (see Hull, 1996). The minimum variance hedge may not be optimal, not only in cases of severe parameter uncertainty, but also for assets whose returns are highly leptokurtic. It is also shown that model mis-specification may lead to substantial differences in Value-at-Risk (V@R) measures.

The parameter estimates necessarily rely on a sample of past returns and an assumed return-generating model. The estimated parameters are subject to what in the literature is commonly called *estimation error* and *model error* (or *model risk*). In financial applications, parameters of the models must first be estimated. Black and Litterman (1992) report that investors have often encountered unreasonable results when they have tried to use quantitative models to help optimize the critical allocation decision. The sampling error in estimates of mean-variance efficient portfolio weights was estimated by Britten-Jones (JF, 1999). Using monthly returns on 11 developed countries' stock indexes during 1977-1996, he shows that the sampling error in estimates of the weights of a global efficient portfolio is large. For example, the results provide no *statistical* support for the hypothesis that there are benefits to global diversification for a U.S. investor (page 666). Similarly, Best and Grauer (RFS, 1991) document that sample efficient portfolios are extremely sensitive to changes in asset means, which are subject to large estimation errors. Clearly, a further examination of the estimation and model errors and their potential impact on the optimality of decision makers' actions should be a positive net-present value project.

II.1.1 Methods and Motivation

There is evidence that volatilities vary over time (French, Schwert, Stambaugh, 1987; Schwert, 1989, 1990). This fact led to an ascent of models and methods that weight the most recent data more heavily. The most common estimation procedures use *GARCH* specification (or its variants such as *EWMA*), or assume that returns are draws from a normal distribution (prices follow a Geometric Brownian Motion). In either case, the distribution of returns over subsequent periods is assumed to be Gaussian.

Because of a presumed time varying nature of many parameters that are not directly observable and must be estimated, many practitioners and data vendors use short estimation periods. For example, RiskMetrics™ uses an *EWMA* specification and usually the 25 most recent observations (Longerstaey and Spencer, 1996). However, even if the evidence strongly suggest that the variance of returns this year is different from the variance twenty years ago, it is not obvious that the best way to model daily returns is to employ a short estimation window and an *EWMA* specification. The EWMA processes and its effective estimation periods are discussed in the next section.

Prices and returns may exhibit jumps due to various news events that are unrelated to the volatility of returns on no-news days. Numerous event studies document that news is rapidly incorporated into asset prices. The distributional assumptions as well as the length of estimation period create a new source of errors due to the potential model misspecification and random character of stock prices. Small sample properties of parameter estimators under various distributional assumptions should be of great interest. For models such as *GARCH (EWMA),* even long estimation periods may yield very noisy forecasts since the *effective* estimation period is rather short due to a greater weighting of the most recent observations. Susmel and Hamilton (1994, page 312) note that the forecasting performance of *GARCH* specifications is rather poor in a sample of the value-weighted portfolio of stocks traded on the NYSE from July 1962 to December 1987. If their specifications were correct and the parameters were known with certainty, then σ_t^2 would be the conditional expectation of squared innovations. Hence, a mean squared error loss function MSE = E $[(\varepsilon_t^2 - \theta_t)^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots]$ would be minimized with respect to θ_t by choosing $\theta_t = \sigma_t^2$. However, they show that the *GARCH* MSE is larger than the MSE under the assumption of constant variance during the period.

It is informative to examine first the estimation errors under the normality assumption when the true returns are indeed draws from a normal distribution. An examination of the impact of a short estimation period will provide answers to two related questions. How large are the standard errors of parameter estimates? Alternatively: How short can our estimation period be and still produce “reasonably” precise estimates?

The minimum variance hedge is always optimal for investors with preferences governed by a quadratic utility function. One remaining question relates to the possibility that the minimum variance hedge may not be optimal if: 1) there is a severe parameter uncertainty; 2) an investor has a strong preference for positive skewness; 3) an investor’s utility function declines rapidly below some threshold, and the investor may be interested in the probability that the asset declines below this threshold (e.g., the confidence level of the V@R metrics).

The risk of a *minimum-variance* hedge is shown to be *not* unambiguously smaller than the risk of an unhedged position. Moreover, different measures of risk – such as portfolio variance and V@R metrics – may lead to contradictory conclusions about the change in risk level. It is shown that if asset returns are normally distributed, only an extreme parameter uncertainty and very low absolute levels of return correlation would lead to an increase in risk level, no matter what risk measure is used. However, the same conclusions cannot be made for leptokurtic asset return series.

The impact of model misspecification on V@R measurement is addressed further. An alternative model to the commonly assumed *GBM* model (an assumption of normality of returns) is the mixture-of-normals model (Roll, 1988; Quandt and Ramsey, 1978). The mixtures model can be viewed as a discrete time alternative to the jump-diffusion models of Merton (1976a, b), Cox and Ross (1976), and Bates (1991).[23] The extent and magnitude of errors are illustrated on several examples. It is shown that the normality assumption may lead to severe under-reporting of V@R metrics. Also, back-testing of V@R methodologies using higher confidence levels may be highly misleading if the true returns are leptokurtic.

Monte Carlo methods have been used extensively in prior research (e.g., Figlewski, 1989). They are well suited for estimation of small sample properties of various estimators, as well as for problems that cannot be solved analytically. This paper follows this line of research and simulation techniques. The *Mathematica*™ software package is used because of its symbolic capabilities and a well-tested random-number generator.

Monte Carlo methods are used to illustrate the impact of short estimation periods on parameter estimation (such as correlations, variances and hedge ratios), which directly impact other variables of interest, such as the volatility of a hedged position. Specifically, I generate between 10,000 and 1,000,000 samples of random vectors.[24] Subsequently, random vectors (generated from the posited "true" distribution – either a

[23] See CHAPTER ONE for a further discussion.
[24] The actual number of repetitions (10,000 to 1,000,000) for each Monte Carlo experiment depends on the length of generated random vector samples, which range from 10 to 250.

normal or a mix-normal) are used to examine the probability densities (PDF) of various parameters such as correlation coefficients, volatilities, and V@R metrics. Distributions of hedge returns are contrasted to the distributions of an unhedged asset returns as well as to distributions obtained under incorrect return-model specifications. The parameter estimation window ranges from 10 to 250 periods (estimation sample size), which corresponds to two weeks to one-year's worth of daily data. The errors in Value-at-Risk metrics, $V@R_{(N,\, x\%)}$, induced by model errors are illustrated for distributions with an increasing magnitude of leptokurtosis and at various confidence levels (90% to 99.9%).

II.1.2 The Effective Estimation Period and the EWMA Processes

A class of *Autoregressive Conditionally Heteroscedastic (ARCH)* return-generating processes has gained widespread attention. CHAPTER ONE, SECTION SEVEN, introduces this class of processes as well as a subset of processes called *EWMA* (*Equally Weighted Moving Average*), also known as an *exponential smoothing* model. The general feature of these models is that they weight the most recent observations more heavily. The models are intuitively appealing because the older returns need not reflect future return properties adequately. This feature is their major strength but also their major weakness, because it effectively shortens the estimation period. The *EWMA* model has been widely applied by numerous Value-at-Risk data vendors, including RiskMetrics™.

This section examines the effect of *EWMA* weighting and establishes the notion of an *effective estimation period*. The results suggest that even very long *actual* estimation window may yield noisy estimates because the *effective* estimation window is relatively short. From the model given by equation *1.7.9*, it is possible to derive a volatility estimator of the *EWMA* process. Table II.1 provides volatility estimators of *EWMA* and *GBM* models; where λ denotes the decay parameter, N is the estimation sample size, $\bar{r}$ denotes the mean sample return, and r_t is a return on day-t.

Table II.1 Volatility Estimators

EWMA – Exponential weighting	GBM – Equally weighted
$\sigma_{t,EWMA} = \sqrt{(1-\lambda)\sum_{i=1}^{\infty} \lambda^{i-1}(r_{t-i} - \bar{r})^2}$	$\sigma_{t,Eq.Weight} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(r_{t-i} - \bar{r})^2}$

Because of a finite estimation period (with N observations), the *EWMA* volatility estimator that corrects for weighting (so that Σw=1) is:

$$\sigma_{t,EWMA} = \sqrt{\sum_{i=1}^{N} \frac{\lambda^{i-1}}{\sum_{j=1}^{N} \lambda^{j-1}} (r_{t-i} - \bar{r})^2} \quad . \qquad 2.1.1$$

The “decay factor” λ and an estimation period N play an important role in *effective* weighting of observations. RiskMetrics™ methodologies choose decay factors 0.97 or 0.96 for their V@R estimation procedures, which means that the most recent observation has a weight of 3% or 4%, respectively. With finite estimation periods this weight proportionately increases. Table II.2 illustrates the scope of the problem by providing the cumulative weights of the most recent observations, depending on various decay factors and estimation period lengths. For example, an *EWMA* specification with a decay factor $\lambda=0.94$ provides volatility estimates whose values are *effectively* (to more than 95%) tied to the 50 most recent observations. For $\lambda=0.97$, the effective estimation period is approximately 100 observations. Thus, even procedures using very long estimation periods yield estimates where 95% of the value is determined by the last 50 (or 100) observations. Volatility estimates based on *EWMA* with a decay factor of 0.90 will effectively depend on the last 25 observations.

Therefore, there are good reasons to examine short estimation periods whether long or short estimation windows are used. The RiskMetrics™ system provides volatility and correlation estimates for a large number of financial rates and prices based on a weighted average of the last 25 trading days’ returns, with the bulk of the weight put on the most recent observations (see Longerstaey and Spencer, 1996; and Figlewski, 1997)

Table II.2 Sum of weights of N observations for EWMA process with a decay factor λ

Decay Factor / Length of the Estimation period	λ = 0.94	λ = 0.95	λ = 0.96	λ = 0.97	λ = 0.98
N = 10	46.14%	40.12%	33.52%	26.26%	18.29%
N = 20	70.99%	64.15%	55.80%	45.62%	33.24%
N = 30	84.37%	78.54%	70.61%	59.90%	45.45%
N = 40	91.58%	87.15%	80.46%	70.43%	55.43%
N = 50	95.47%	92.31%	87.01%	78.19%	63.58%
N = 60	97.56%	95.39%	91.36%	83.92%	70.24%
N = 70	98.68%	97.24%	94.26%	88.14%	75.69%
N = 80	99.29%	98.35%	96.18%	91.26%	80.14%
N = 90	99.62%	99.01%	97.46%	93.55%	83.77%
N = 100	99.79%	99.41%	98.31%	95.24%	86.74%
N = 110	99.89%	99.65%	98.88%	96.49%	89.16%
N = 120	99.94%	99.79%	99.25%	97.41%	91.15%
N = 130	99.97%	99.87%	99.50%	98.09%	92.77%
N = 140	99.98%	99.92%	99.67%	98.59%	94.09%
N = 150	99.99%	99.95%	99.78%	98.96%	95.17%

Note:
Decay factor λ = 0.94 means that the most recent observation has a weight of 6% in the *EWMA* volatility estimator, see the *EWMA* specification given by *1.7.9*.

II.2. The Sample Correlation Coefficient

The correlation coefficient plays an important role in asset pricing as well as in risk management. The benefits of portfolio diversification depend on the asset correlation structure. Similarly, a hedging decision is based on the contemporaneous cross-correlation of asset returns. Value-at-Risk metrics, which can be viewed as an extension of classic portfolio theory, often relies on the knowledge of cross-correlation structure. As important as it is, our knowledge of this structure remains elusive. One reason is that the correlation coefficient is a complex function of asset returns.

Kendall, Stuart, Ord (1987, page 524) provide an overview of the statistical literature on the distribution of the sample correlation coefficient. The finance literature that has previously examined the correlation structure among international stock markets includes Karolyi and Stulz, (JF 1996), Longin and Solnik (JF 2001, 1995), Eun and Resnick (JF 1984), and Ramchand and Susmel (JEF 1998). Campbell and Ammer (1993), Keim and Stambaugh (1986), and Kwan (JFE 1996) investigate the correlation between stocks and bonds.

There seems to be a persuasive evidence of time varying volatilities, covariances and correlations (French, Schwert, Stambaugh, 1987; Schwert, 1989, 1990; King, Sentana, and Wadhwani, 1994; Makridakis and Wheelwright, 1974). However, the power and conclusiveness of these results necessarily rests on the examination of *long* time series. Many recent studies point out that the correlation between assets increases in times of large volatility (e.g., during crises). King et al. (Econometrica 1994, RFS 1990) use monthly returns over 1970-1988 period to investigate conditional correlations. Early studies fail to find a common world risk factor or correlation. For example, Hilliard (JF, 1979) examines the structure of international equity market indices during the world financial crisis caused be OPEC embargo of 1973-1974. He does not find significant comovement in the inter-continental equity index series, though it is possible that world markets integration has since changed the underlying structure.[25]

However, as Longin and Solnik (2001) point out, the estimates of correlations in samples *conditional* on the size of returns are misleading. Even if the correlation coefficient is constant and linear in returns, the correlation coefficient measured in conditional return subsamples is different from the true coefficient. An obvious implication is that conclusions about the changing correlation cannot be based on a simple comparison of estimated correlations conditional on different values of one return variable. The distribution of (conditional) correlation coefficient must be specified. In addition, the asymptotic errors may substantially overstate the small sample standard errors and make the results inconclusive.

It is not clear that utilizing short periods of high-frequency data to obtain estimates is the best alternative. Practitioners and data providers such as RiskMetrics™ have widely adopted ARCH/GARCH models of Engle (1982) and Bollerslev (1986) by

[25] The intra-continental equity index series such as New York – Toronto or Amsterdam – Frankfurt *do* have significant comovement structure.

utilizing short estimation periods such as 25 days (see Longerstaey and Spencer, 1996). These estimates are subject to large estimation errors or *noise fitting*.

This section extends this line of literature and provides some insight into the small sample properties of correlation coefficient under different model specifications and compares them to the theoretical asymptotic results under the assumption of normality.

II.2.1 The Distribution of a Sample Correlation Coefficient

Fisher (1915, 1921) derives the distribution of the sample correlation coefficient for normally distributed samples. The joint probability of n sample values $(x_1, y_1) \ldots (x_n, y_n)$ from a bivariate normal population with parameters $\{\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\}$ is

$$dF = \frac{1}{(2\pi)^n \sigma_1^n \sigma_2^n (1-\rho^2)^{n/2}} \cdot \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\sum \frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \sum \frac{(x-\mu_1)(x-\mu_2)}{\sigma_1 \sigma_2} + \sum \frac{(x-\mu_2)^2}{\sigma_2^2}\right\}\right] dx_1 dy_1 \ldots dx_n dy_n$$

. *2.2.1*

There exists a transformation of this distribution that utilizes the five parameters and the corresponding sample estimates (statistics) of these parameters to express the joint distribution of the sample statistics. It turns out that in normal samples the (bivariate) distribution of means is entirely independent of the (trivariate) distribution of the variances and correlation coefficient, which is a characteristic property of multivariate normality. The distribution of the sample correlation coefficient r is then obtained by integrating with respect to s_1 and s_2. Kendall, Stuart and Ord (1987, page 528) provide the following specification of the distribution function of the correlation estimate r:

$$dF = \frac{(1-\rho^2)^{(n-1)/2}}{\pi \Gamma(n-2)} (1-r^2)^{(n-4)/2} \frac{d^{n-2}}{d(\rho r)^{n-2}} \left\{ \frac{\arccos(-\rho r)}{\sqrt{(1-\rho^2 r^2)}} \right\} dr \quad , \qquad 2.2.2$$

where $1 \geq r \geq -1$. The result is due to Fisher (1915). For any ρ, as $n \rightarrow \infty$, the distribution of r tends very slowly to normality, and the density function is increasingly skew as $|\rho|$

increases. The first moment (mean) and the second central moment (variance) of the sample correlation *r* can be approximated by:

$$\begin{aligned}\mu_1'(r) &= \rho\left\{1-\frac{1-\rho^2}{2n}+O\left(1/n^2\right)\right\}\\ \mu_2(r) &= \frac{(1-\rho^2)^2}{n-1}\left(1+\frac{11\rho^2}{2n}\right)+O(1/n^3)\end{aligned} \quad . \qquad 2.2.3$$

In large samples, the distribution of *r* can be approximated by:

$$r \sim N\left[\rho, \frac{1}{n-1}(1-\rho^2)^2\right] . \qquad 2.2.4$$

However, Kendall, Stuart and Ord (1987, page 338) do not recommend this approximation except for very large samples. Exhibit II.2.1 illustrates the extent of the approximation bias in small samples if tests of significance use asymptotic approximations given by relation *2.2.4*.

Fisher (1921) found a remarkable *transformation of r,* which tends to normality much faster than *r* itself. The Fisher's *z-variable* is defined as:

$$\begin{aligned}r &= \tanh(z),\\ z &= \frac{1}{2}\log\frac{1+r}{1-r}\end{aligned} \quad , \qquad 2.2.5$$

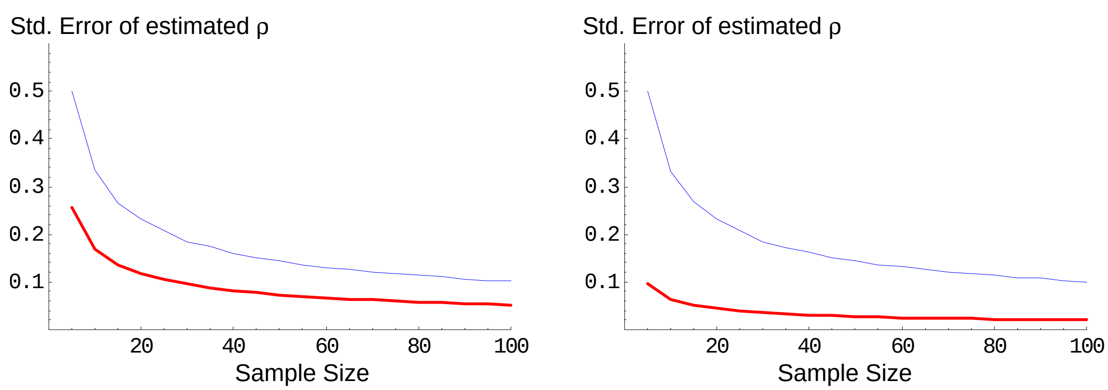
and the recommended test for *r* uses the approximation for *z,* which is asymptotically normally distributed as $z \sim N\,[\,(1/2)\ ln\{(1+\rho)\,/(1-\rho)\},\ (n-3)^{-1}]$. Kendall, Stuart and Ord (1987, page 533) and David (1938) note that this approximation seems satisfactory only for a relatively small ρ and longer time series n>50. However, a simulation reveals that a *z-test* would lead to correct inferences at the usual significance levels (e.g., 5%) even in very small samples (e.g., 15 to 25 observations). Note again, that this approximation holds for bivariate normal random vectors. It is shown later, that the approximation produces spurious results for a sample correlation coefficient of sample return series whose distributions depart substantially from normality.

II.2.2 Correlation: Small Sample Properties vs. Asymptotics; Normal Variates

This section compares the asymptotic and small sample properties of the correlation coefficient if the sample series are from a bivariate *normal* population, as defined by expression *2.2.1*. Exhibit II.2.1 provides two plots of true and asymptotic standard errors of the estimated correlation coefficient as functions of a sample size. Scenarios with two different joint probability density functions with parameters $\rho = 0.7$ and 0.9 are considered (Panels A and B). It is obvious that the asymptotic results given by expression *2.2.4* are not satisfactory. For example, the small sample standard errors exceed the asymptotic by a factor of 2 to 5.

Exhibit II.2.1 Exact and Asymptotic Standard Errors of a Sample Correlation Coefficient

Panel A.: True correlation ρ = 0.7 **Panel B.:** True correlation ρ = 0.9

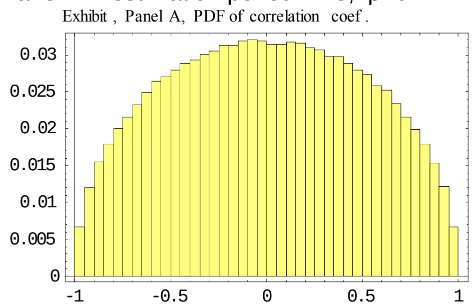


Note: Small sample and asymptotic standard errors of a sample correlation coefficient for random samples from a *bivariate normal* population given by expression *2.2.1*. Correct, small sample standard errors, are depicted in BLUE (thin) curve, and were obtained by a Monte Carlo simulation on 10,000 random samples. The asymptotic standard errors, RED (thick) curve, are calculated from expression *2.2.4*.

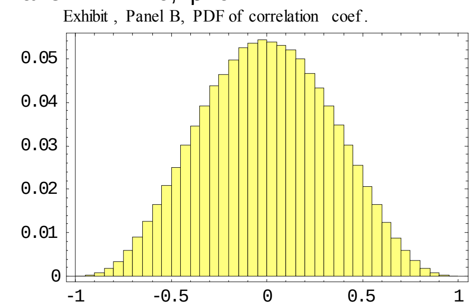
Exhibit II.2.2 depicts PDFs of the sample correlation coefficient if: 1) the returns of assets A and B are normally distributed N[0, 1]; 2) the true correlation of asset returns is $corr(r_A, r_B) = 0$; and 3) the estimation window ranges from 5 days to 250 periods (e.g., the random vector lengths are n=5 in Panel A., n=10 (25, 250) in Panel B. (Panels C., D.).

Exhibit II.2.2 Probability distribution of a sample correlation $\hat{\rho}$, if true ρ=0. *Normal* variates.

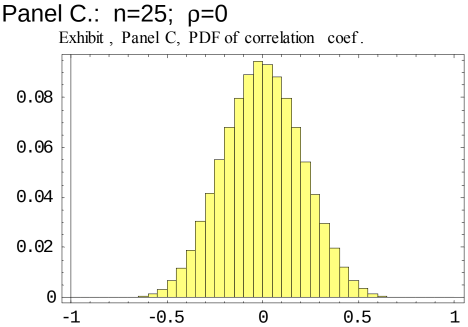
Panel A.: estimation period: n=5; ρ=0



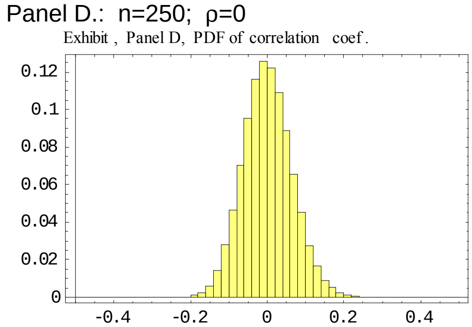
Panel B.: n=10; ρ=0



Panel C.: n=25; ρ=0



Panel D.: n=250; ρ=0



Note: Each bar is 0.05 wide (0.02 wide for Panel D) and represents the probability that a correlation coefficient estimate will be from a respective interval. Some representative probabilities that the correlation coefficient estimate falls into interval *Int.*

Prob ($\hat{\rho}$ ∊ Int):

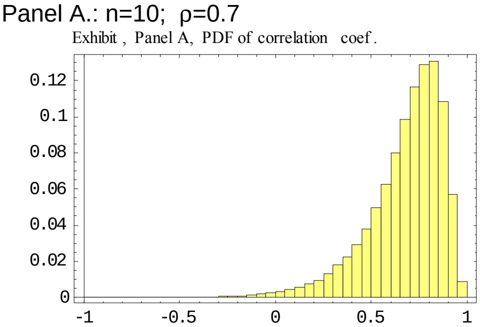
Interval / Size	0.0, 0.2	0.2, 0.4	0.4, 0.6	0.6, 0.8	0.6, 1.0
N=5	12.7%	12.1%	11.0%	9.0%	5.2%
N=10	21.0%	16.4%	9.2%	3.1%	0.3%
N=25	33%	14.5%	2.3%	-	-
N=250	50%	<0.01%	-	-	-

Note: Because of symmetry about zero, the probabilities for negative values are not reported.

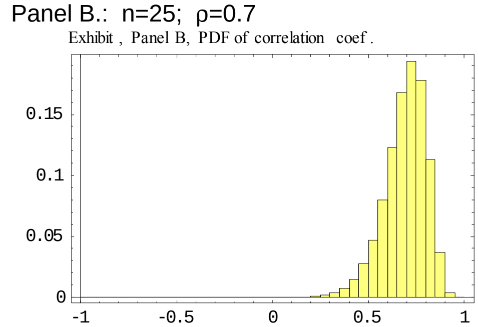
Exhibit II.2.3 shows PDF of the sample correlation coefficient if: 1) the returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0,1]$; 2) the true correlation of asset returns is $corr(r_A, r_B) = 0.7$; and 3) the estimation window is 10 days to 250 periods (e.g., the random vector lengths are n=10 in Panel A., n=25 (100, 250) in Panel B. (Panel C., D.). Sampling reveals a substantial skewness in the distribution of the sample correlation coefficients. However, the skewness seems important (in terms of hypothesis testing and confidence bound deviations) only in small samples or for extremely large correlation coefficients.

Exhibit II.2.3 The Probability Distribution of a Sample Correlation Coefficient $\hat{\rho}$, true ρ = 0.7

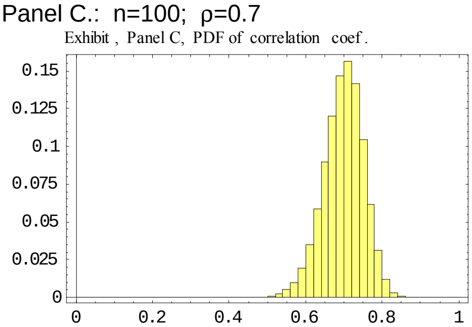
Panel A.: n=10; ρ=0.7



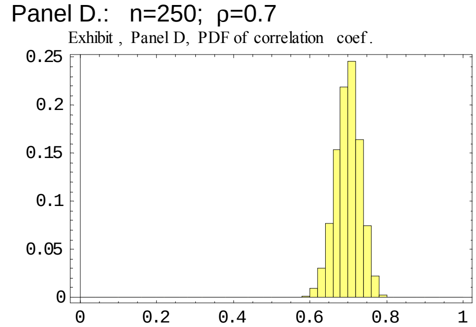
Panel B.: n=25; ρ=0.7



Panel C.: n=100; ρ=0.7



Panel D.: n=250; ρ=0.7



Note: Each bar in Panels A., B. is 0.05 wide (0.02 in Panels C., and D.) and represents the probability that the correlation coefficient estimate will be from the corresponding interval. Number of sampling repetitions: 200,000; 100,000; 30,000; and 10,000 for Panels A., B., C., and D. respectively.
Some representative probabilities that the correlation coefficient estimate falls into interval *Int.*

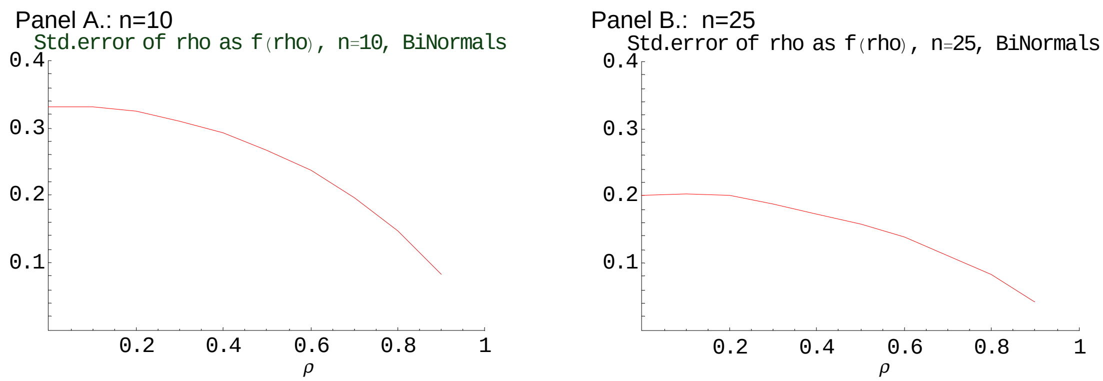
Prob ($\hat{\rho}$ ***∈ Int):***

Int / Size	<0.0	0.0, 0.2	0.2, 0.4	0.4, 0.6	0.6, 0.8	0.8, 1.0
N=10	0.8%	1.9%	6.2%	18%	42.8%	30.3%
N=25	-	<0.1%	1.4%	16.9%	66.1%	15.4%
N=100	-	-	-	3.9%	94.6%	1.5%
N=250	-	-	-	0.3%	99.7%	-

Exhibit II.2.4 shows *standard errors* of the sample correlation coefficient if: 1) asset returns are normally distributed: $r_A \sim r_B \sim N[0,1]$; 2) the true return correlation is $corr(r_A, r_B) = 0.7$; and 3) the estimation sample is 10 (or 25) observation (e.g., the random vector lengths are n=10 in Panel A., and n=25 in Panel B.)

The results given in Exhibits II.2.2 – II.2.4 provide a strong case for the use of small sample results and alternative robust return-generating models in tests of hypotheses. Moreover, the potential estimation errors may lead to errors in some other variables of interest such as optimal hedge ratios. For example, an estimate of a correlation coefficient obtained from a sample of 25 bivariate-normal returns has only a 66.1% chance to take on a value in the (0.6, 0.8) interval, if the true correlation is 0.7.

Exhibit II.2.4 Standard Errors of a Sample Correlation Coefficient $\hat{\rho}$, Bivariate-Normal Returns.



Note: Standard Errors of a sample Correlation Coefficient for random samples from a *bivariate normal* population given by expression *2.2.1*. Estimates obtained by Monte Carlo analysis of 40,000 random samples.

Standard Errors of sample $\hat{\rho}$ as f(ρ, size) **:**

$\hat{\rho}$ / sample size	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
N=10	0.33	0.33	0.32	0.31	0.29	0.27	0.24	0.20	0.15	0.083
N=25	0.20	0.20	0.20	0.19	0.17	0.16	0.14	0.11	0.082	0.043
N=100	0.10	0.099	0.097	0.093	0.083	0.075	0.064	0.050	0.036	0.019

II.2.3 The Sample Correlation Coefficient: Non-Normal Returns

The previous section examined the properties of the sample correlation coefficient if a return sample is from a bivariate normal distribution. However, as reviewed in CHAPTER ONE, financial asset returns exhibit substantial non-normalities. Tests of time-varying correlation and autocorrelation commonly assume normality. The numerical examples below illustrate the scope of potential errors and suggest that great care should be exercised in hypothesis testing (see exhibit II.2.5). One recommendation is to fit non-normal distribution to the sample returns, and investigate the properties of the sample correlation of returns from the fitted distributions.

One alternative to the normality assumption is, as before, to assume that each return series in a sample of two asset returns is generated by a process containing occasional jumps ("news events"), producing a return distribution that is a mixture of two normal distributions (as defined by expressions *3.2.1* and *3.2.2*). Panels A1 and A2 of the Exhibit II.2.5 illustrate the extent of an increase in standard errors of the sample correlation coefficient if asset returns are leptokurtic. Note that the normality case is represented by the point with the leptokurtosis parameter equal to σ_1=1. Panels B1 and B2 of the Exhibit II.2.5 provide the distribution of the sample correlation coefficient for return distributions with parameters $\mu_1 = \mu_2 = 0$, $\sigma_1 = 5$, $\sigma_2 = 1$, the mixing parameter $w = 0.1$ (news frequency), and the cross-correlation parameter $\rho_{12} = 0.7$.[26] The standard error of the sample correlation is $SE_{Mix\text{-}Normal}[\rho] = 0.19$, which exceeds substantially the standard error of the sample correlation if a sample is from a bivariate *normal* distribution ($SE_{Normal}[\rho]=0.11$). Tests of time-varying correlations may provide spurious results and over-reject the null hypothesis of constant correlation if the investigated sample return distributions are leptokurtic.

II.3 Hedge Ratios, Hedge Volatility, and Hedge Return Distributions

The finance literature examining the hedge ratios and optimal hedging with futures is extensive and includes work of Sercu and Wu (2000), Hilliard (1994), Kroner and Sultan (1993), Cecchetti, Cumby, Figlewski (1988), Stulz (1984), Anderson and Danthine (1981), Ward and Fletcher (1971), Stein (1961), Johnson (1960), Telser (1955), among others.[27] However, the academic literature has paid little attention to the impact of estimation risk and model errors.

---

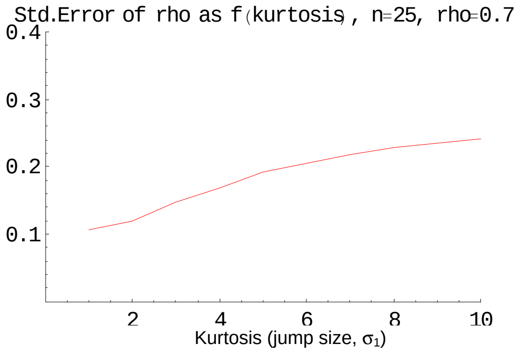
[26] The expected sample correlation is also 0.7. Note that due to the way the correlated vector are simulated, the second return vector is a linear combination of the mix-normal with the posited parameters. E.g., $r_A \sim (0.1)\,N[0, 5] + (0.9)\,N[0, 1]$, and $r_B \sim r_A\,\rho + r_{A,2}\sqrt{(1-\rho^2)}$, where $r_{A,2} \sim r_A$, and $r_A \perp r_{A,2}$.

[27] See also Hull (1997, page 35).

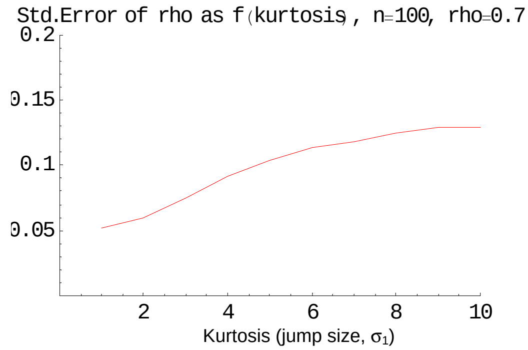
Exhibit II.2.5 Sample Correlation Coefficient Standard Errors and Distributions

PDFs of $\hat{\rho}$ as a function of asset kurtosis; Normal vs. Mix-Normal returns

Panel A1.: Mix-Normal Variates, ρ=0.7, n=25.

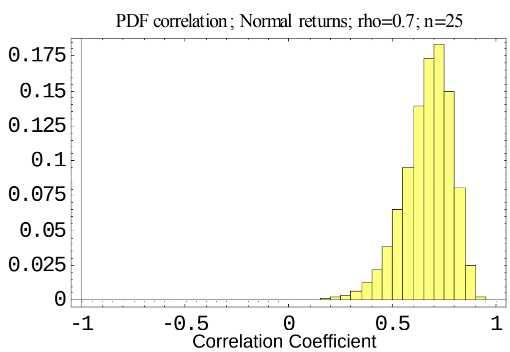


Panel A2.: Mix-Normal Variates, ρ=0.7, n=100.

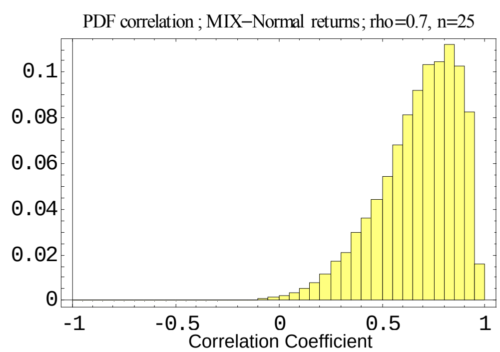


PDFs of a sample correlation coefficient $\hat{\rho}$; Normal vs. Mix-Normal returns

Panel B1.: Normal Variates, ρ=0.7, n=25.



Panel B2.: Mix-Normal Variates, ρ=0.7, n=25.



Notes:
Sample size is n=25 observations, corresponding to about 1 month of daily data, or "effective" estimation period of EWMA processes as discussed in SECTION II.1.2.

Panels A: The sample random vectors r_A and r_B are from a mixture of normal distributions as defined by expressions *3.2.1* and *3.2.2.* Sample vectors are distributed as follows: $r_A \sim (0.1)\, N[0, s_1] + (0.9)\, N[0, 1]$, and $r_B \sim r_A\, \rho + r_{A,2} \sqrt{(1-\rho^2)}$, where $r_{A,2} \sim r_A$, and $r_A \perp r_{A,2}$. $Corr[r_A, r_B] = \rho = 0.7$.

Panel B1: The sample is from a bivariate *normal* population defined by expression *2.2.1* with parameters: $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, $\rho_{12} = 0.7$.

Panel B2: The sample random vectors r_A and r_B are from a mixture of normal distributions as defined by expressions *3.2.1* and *3.2.2.* $r_A \sim (0.1)\, N[0, 5] + (0.9)\, N[0, 1]$. $r_B \sim r_{A,1}\, \rho + r_{A,2} \sqrt{(1-\rho^2)}$. $r_{A,1} \perp r_{A,2}$. $Corr[r_A, r_B] = 0.7$.

This section introduces the hedging problem in general. Portfolio managers often hedge their underlying positions.[28] One common approach is to hedge with futures on other assets. The optimal hedge ratio is usually derived under assumptions of quadratic utility function and uses the minimum variance criterion.[29]

The hedge ratio "h" is the ratio of the value of the position taken in futures contracts to the value of the hedged asset.[30] The *optimal* hedge ratio will depend on the volatilities of the hedged asset and futures, as well as their correlation. The objective is to minimize the variance of the hedged position.

Let's assume that a long position in an asset *"A"* is hedged with an asset *"B"* (e.g., short position in a futures contract). Then the change in the portfolio value is:

$$\Delta\Pi = \Delta S_A - h\,(\Delta S_B) \quad , \qquad 2.3.1$$

where h is the hedge ratio, and ΔS_A and ΔS_B are changes in the spot prices of assets A and B over time period Δt. The variance of the change in value of the hedged position is:

$$V \equiv Variance(\Delta\Pi) = \sigma_A^2 + h^2\sigma_B^2 - 2h\sigma_A\sigma_B\rho \quad , \qquad 2.3.2$$

where $\rho = Corr(\Delta S_A, \Delta S_B)$; σ_A is the standard deviation of ΔS_A; σ_B is the standard deviation and ΔS_B. The minimum variance hedge is then determined by the hedge ratio parameter *"h"* that minimizes variance *"V"*.

$$\partial_h V = 2h\sigma_B^2 - 2\sigma_A\sigma_B = 0 \qquad 2.3.3$$

---

[28] The underwriters of securities and block-traders may serve as an example. CHAPTER ONE introduced a case of Goldman Sachs block-trading.

[29] The classic *mean-variance* portfolio theory of Markowitz (1952, 1959) and *CAPM* of Sharpe, Lintner, Mossin, Black are also based on the assumption that investors care only about the mean and variance of their portfolios, thus restricting the class of acceptable utility functions. The restriction implicitly constrains the allowable set of the return generating processes (for further discussion see Bick, 1990; He and Leland, 1993; Lo, 1999).

[30] The term "hedging" refers to actions designed to decrease or completely eliminate risk. These actions usually involve taking positions in futures or options that tend to offset increases or decreases in prices of the hedged asset.

Solving equation *2.3.3* for *h* gives the optimal hedge ratio:

$$h = \frac{\sigma_A}{\sigma_B}\rho \qquad 2.3.4$$

However, because of the random character of stock prices, one really obtains only an estimate of the optimal hedge ratio:

$$\hat{h} = \frac{\hat{\sigma}_A}{\hat{\sigma}_B}\hat{\rho} \qquad 2.3.5$$

and thus the true variance of the hedged position is:

$$V_{true} \equiv Variance(\Delta\Pi) = \sigma_A^2 + \left(\frac{\hat{\sigma}_A}{\hat{\sigma}_B}\hat{\rho}\right)^2 \sigma_B^2 - 2\left(\frac{\hat{\sigma}_A}{\hat{\sigma}_B}\hat{\rho}\right)\sigma_A\sigma_B\rho \qquad 2.3.6$$

The hedger could be led to believe that the position has a variance:

$$V_{Position,\min} = \sigma_A^2(1-\rho^2), \qquad 2.3.7$$

which is only the lower bound variance, and is attainable only if the true parameters are known.

The question of interest relates to the variance of hedged position returns and how it differs from the variance of an unhedged portfolio and the minimum variance in expression *2.3.7*. The next section illustrates magnitude of the estimation and model misspecification errors.

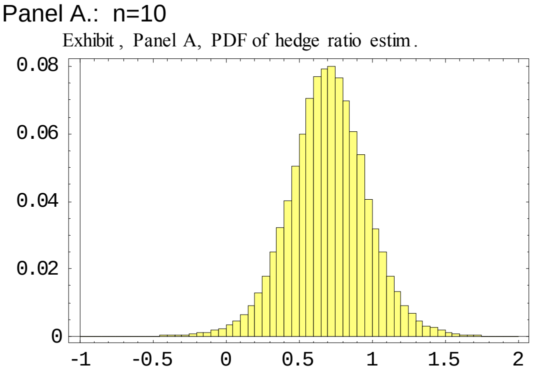
II.3.1 Bivariate Normal Returns

Finance theory in general as well as the minimum variance hedge rule provides unambiguous advice to the decision makers. They need to take certain actions suggested by the theory in order to maximize their objective function. However, the models' parameters create another source of uncertainty. This section examines the small sample properties of the volatility of the hedged portfolio when the true return distribution is normal.

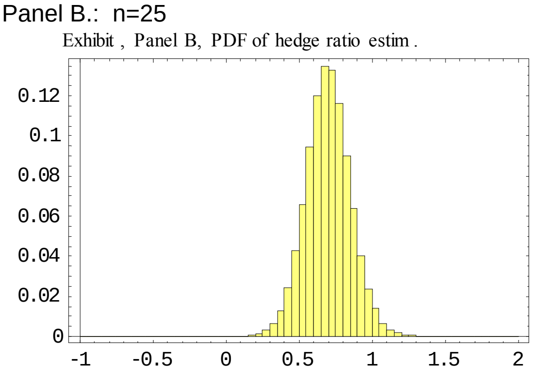
Exhibit II.3.1 shows the PDF of the hedge ratio if: 1) the returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0,1]$; 2) the true correlation of asset returns is $corr(r_A, r_B) = 0.7$, and thus the correct hedge ratio $h = 0.7$; and 3) the estimation period is 10, 25, and 100 days or periods (Panels A, B and C).[31] The uncertainty about the true minimum variance hedge ratio leads to a higher *average* hedge volatility. The deviations do not seem to be trivial.

Exhibit II.3.1
The Probability Distribution of an optimal Hedge Ratio estimate; E[ρ] = 0.7, the optimal *h*=0.7.

Panel A.: n=10



Panel B.: n=25



Note: Each bar in Panel A., and Panel B. is 0.05 wide (0.02 wide in Panel C.) and represents the probability that the hedge ratio estimate will be from the corresponding interval. Number of repetitions used in sampling: 100,000 (Panels A., and B.) and 40,000 (Panel C.). Some representative probabilities that the estimated hedge ratio falls into interval *Int,* given the estimation period N=10 (or 25, 100).

Prob ($\hat{h} \in$ Int):

Int	<0.2	0.2, 0.4	0.4, 0.6	0.6, 0.8	0.8, 1.0	1.0, 1.2	>1.2
N=10	3.2%	8.7%	22.2%	31.7%	22.2%	8.7%	3.2%
N=25	0.1%	2.4%	22.4%	50.4%	22.4%	2.4%	0.1%
N=100	-	-	8.2%	83.7%	8.2%	-	-

[31] E.g., the simulated bivariate normal random samples contain 10, 25 and 100 returns.

The most important question, of course, relates to the distribution of volatilities of the hedged position given imperfect parameter estimates that are due to the random character of prices and possible model errors. Exhibit II.3.2 shows the PDF of standard deviations of the hedge portfolio under the following <u>assumptions:</u>

<u>Panel A</u>: 1) returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0, \sigma=1]$
2) $corr(r_A, r_B) = 0.5$
3) the estimation sample n=10.

<u>Panel B</u>: 1) returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0, \sigma=1]$
2) $corr(r_A, r_B) = 0.5$
3) the estimation sample n=25.

<u>Panel C</u>: 1) returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0, \sigma=1]$
2) $corr(r_A, r_B) = 0.7$
3) the estimation sample n=10.

<u>Panel D</u>: 1) returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0, \sigma=1]$
2) $corr(r_A, r_B) = 0.7$
3) the estimation sample n=25.

<u>Panel E</u>: 1) returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0, \sigma=1]$
2) $corr(r_A, r_B) = 0.9$
3) the estimation sample n=10.

<u>Panel F</u>: 1) returns of assets A and B are normally distributed: $r_A \sim r_B \sim N[0, \sigma=1]$
2) $corr(r_A, r_B) = 0.9$
3) the estimation sample n=25.

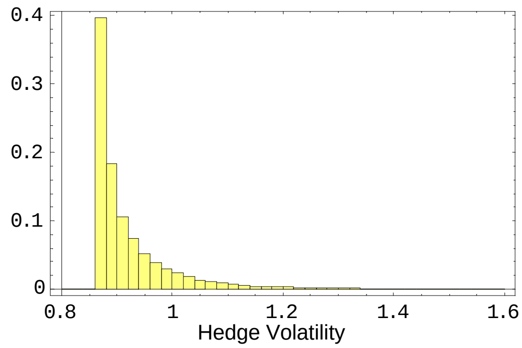
Note that the lower bound of the hedge volatility is given by equation *2.3.7*. For the examples in Exhibit II.3.2 with parameters $\sigma_A = 1$, and $\rho=0.7$ (Panels A.-D.), and $\rho=0.9$ (Panels E.-F.) this translates into the optimal (minimum) hedge volatility equal to:

$$StdDev_{\min,Portfolio,\rho=0.5} = \sigma_A\sqrt{1-\rho^2} = 1*\sqrt{1-0.5^2} = 0.866 \ ;$$

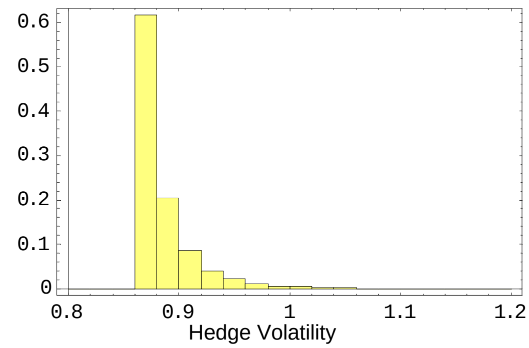
$$StdDev_{\min,Portfolio,\rho=0.7} = 0.714 \ ; \ StdDev_{\min,Portfolio,\rho=0.9} = 0.436 \ .$$

Exhibit II.3.2 PDF of Hedge Volatilities if ρ = 0.5; 0.7, or 0.9; Bivariate-Normal returns.

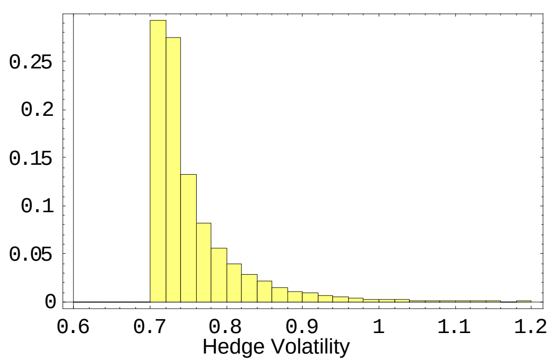
Panel A.:
n=10, ρ=0.5, σ_A=1, σ_F=1, E[σ_P]=0.92



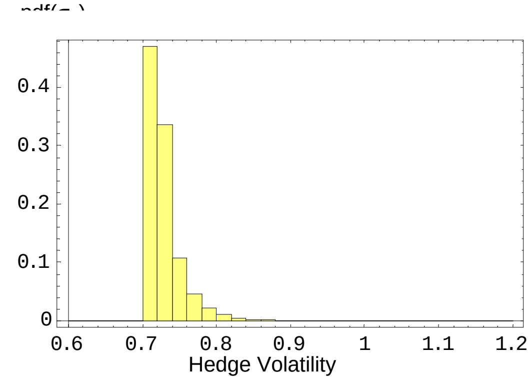
Panel B.:
n=25, ρ=0.5, σ_A=1, σ_F=1, E[σ_P]=0.88



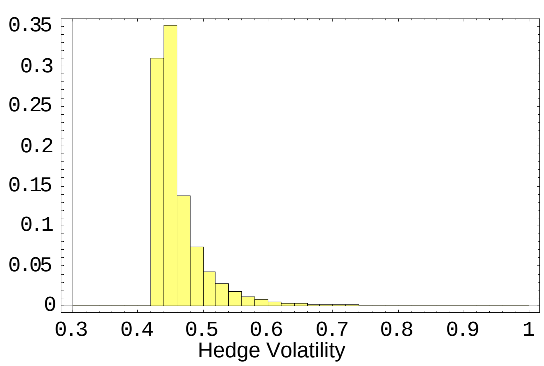
Panel C.:
n=10, ρ=0.7, σ_A=1, σ_F=1, E[σ_P]=0.76



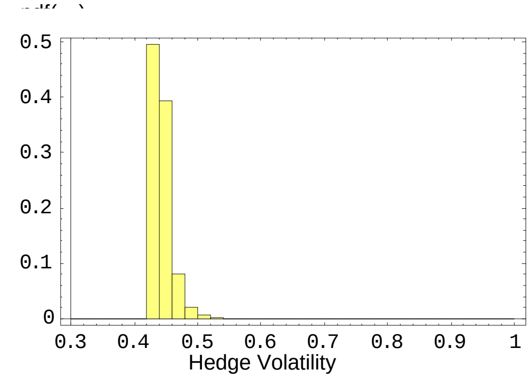
Panel D.:
n=25, ρ=0.7, σ_A=1, σ_F=1, E[σ_P]=0.73



Panel E.:
n=10, ρ=0.9, σ_A=1, σ_F=1, E[σ_P]=0.46



Panel F.:
n=25, ρ=0.9, σ_A=1, σ_F=1, E[σ_P]=0.45



Note: Each bar in Panels A.- F., is 0.02 wide and represents a probability that the volatility (Standard Deviation) of a hedged position will be from the corresponding interval. Number of repetitions used in sampling: 100,000 (all Panels). Some representative probabilities that the estimated volatility is from interval *Int,* given an estimation period N=10 (or 25, 100).

Exhibit II.3.2 continued

Probability ($\sigma_{hedged\ portfolio} \in Int$):

$\sigma_{hedged\ portfolio}$ Interval	< 0.5	0.5, 0.6	0.6, 0.7	0.7, 0.8	0.8, 0.9	0.9, 1.0	> 1.0	Min	Mean
Panel A: σ_A=1, σ_F=1, ρ=0.5, N=10	-	-	-	-	58.1%	30.2%	11.7%	0.866	0.92
Panel B: σ_A=1, σ_F=1, ρ=0.5, N=25	-	-	-	-	82.2%	16.8%	<0.1%	0.866	0.88
Panel C: σ_A=1, σ_F=1, ρ=0.7, N=10	-	-	-	83.6%	11.7%	3.0%	<0.1%	0.714	0.76
Panel D: σ_A=1, σ_F=1, ρ=0.7, N=25	-	-	-	97.9%	2.0%	<0.1%	0	0.714	0.73
Panel E: σ_A=1, σ_F=1, ρ=0.9, N=10	87.4%	10.6%	1.5%	<0.1%	0	0	0	0.436	0.46
Panel F: σ_A=1, σ_F=1, ρ=0.9, N=25	98.9%	1.0%	<0.1%	0	0	0	0	0.436	0.45

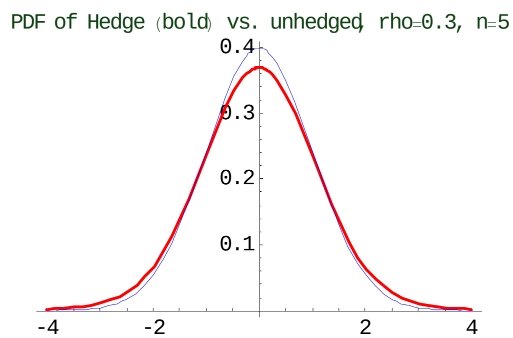
Note: The above probabilities are estimates themselves, but their relative errors are quite small. For example, for the probability of about 0.1%: the relative standard error of this value is about 10% of the value – e.g., the 95% confidence bound is 0.08% to 0.12% (based on a binomial distribution and its normal approximation). The relative standard error is about 3% of the probability value if the estimated interval probability is about 1% (e.g., the confidence bounds are 0.94% –1.06%. The relative standard error is 0.3% if the estimated interval probability is about 50% (e.g., the confidence bounds are 49.4% to 50.6%).

As can be seen from Exhibit II.2.3, taking into account the parameter uncertainty does not lead to an extreme departure of the hedge volatilities from the minimum variance hedge volatility. Thus under the assumption of normality, the hedge volatility *is* generally smaller than volatility of an unhedged asset even for relatively short estimation periods and small correlation coefficient.

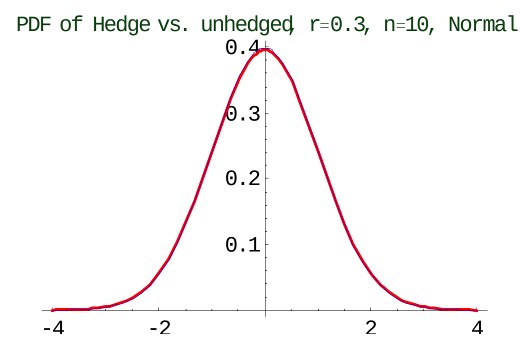
Exhibit II.3.3 compares the distributions of unhedged asset returns to the hedge returns under parameter uncertainty and normality of returns. Results reveal that it would take extreme parameter uncertainty (estimation period of less than 10 days) or extremely low levels of asset correlation ($\rho < 0.3$) to make the hedge volatility larger than the volatility of an unhedged position. The hedge return distribution has several noteworthy features. First, it is leptokurtic, because the distribution of hedge ratios is highly skewed to the right. Second, it is a mixture of normals, because hedge returns are normal for each value of the hedge ratio, e.g., if $r_A \sim r_F \sim N[\mu, \sigma]$, then from $r_H = r_A + h\, r_F$ follows that r_H|(hedge ratio *h*) ~ N[m, s]. Even though the distribution is leptokurtic, the deviations from normality are minor and tend to disappear with longer estimation periods.

Exhibit II.3.3 PDF of Hedge Returns under parameter uncertainty. Normal Returns.

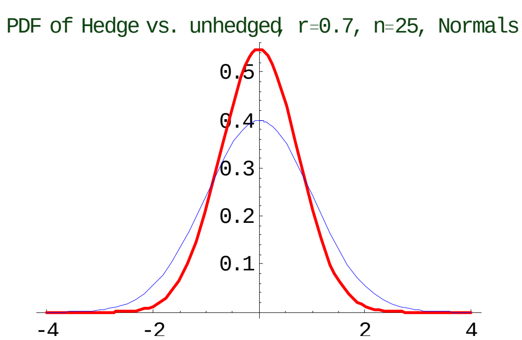
Panel A.:
n=5, ρ=0.3, σ_A=1, σ_F=1; σ_H=1.16, k_H=3.6



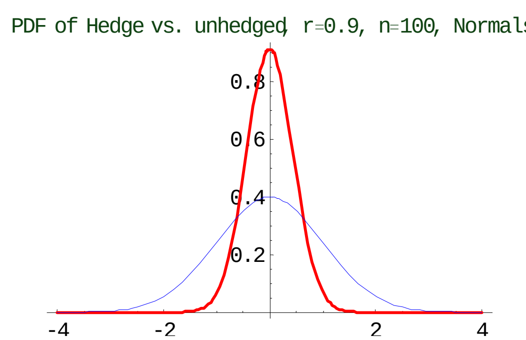
Panel B.:
n=10, ρ=0.3, σ_A=1, σ_F=1; σ_H= 1.02, k_H=0.15



Panel C.:
n=25, ρ=0.7, σ_A=1, σ_F=1; σ_H=0.73, k_H=0.01



Panel D.:
n=100, ρ=0.9, σ_A=1, σ_F=1; σ_H= 0.43, k_H<0.01



Note:

RED (bold) curve is a PDF of the hedge portfolio returns that incorporates parameter uncertainty.
BLUE (thin) curve is a PDF of unhedged asset returns. σ_H denotes the standard deviation of the hedge portfolio that takes into account the estimation error (the distribution of hedge ratio parameter), given the stated estimation and distributional parameters. k_H denotes the excess kurtosis of the hedge portfolio given the stated estimation parameters.

II.3.2 Non-Normal Asset Returns

This section investigates the impact of non-normal returns on the volatility of the hedge. Assume that returns are generate by the following *Mixture of Normals* process:

$$r_A = \begin{cases} \sigma_{A,1}\,\varepsilon_{A,1} & \text{in state } X \text{ with probability "}w\text{"} \\ \sigma_{A,2}\,\varepsilon_{A,2} & \text{in state } Y \text{ with probability "}1\text{-}w\text{"} \end{cases}$$

and $\qquad$ *2.3.8*

$$r_F = \begin{cases} \sigma_{F,1}\,\varepsilon_{F,1} & \text{in state } X \text{ with probability "}w\text{"} \\ \sigma_{F,2}\,\varepsilon_{F,2} & \text{in state } Y \text{ with probability "}1\text{-}w\text{"} \end{cases},$$

where $\varepsilon \sim N[0,1]$; $\mathrm{corr}[\, r_A|X, r_F|X] = \rho$; and $\mathrm{corr}[\, r_A|Y, r_F|Y] = \rho$. This model has been applied by Roll (1988) and Quandt and Ramsey (1978) and is also discussed in CHAPTER THREE, SECTION TWO. To generate a sample of correlated vectors for the simulation purposes, the Cholesky decomposition is commonly used.[32]

The return specification 2.3.8 can be viewed as a switching regime model, where returns are generated either from the distribution X (e.g., "no news" state of the world) or from the distribution Y (e.g., a state of firm-or-economy-wide news events). The model allows for substantial skewness and leptokurtosis in returns.

Exhibit II.3.4 contrasts the distributions of hedge volatilities for normal and mix-normal return distributions under the following assumptions:

Panel A1: (1) returns of assets A and B are normal: $r_A \sim r_B \sim N[0, \sigma=3.3]$

(2) $\mathrm{corr}(r_A, r_B) = 0.7$;

(3) estimation sample is n=25.

---

[32] A sample of two independent random vectors $Y=\{Y_1, Y_2\}$ is generated first: $Y_1 \sim N[0, \sigma_1]$, and $Y_2 \sim N[0, \sigma_2]$. In the next step, a new set of vectors $Z=\{Z_1, Z_2\}$ is calculated as a linear combination of Y_1 and Y_2 in order to induce cross-correlation ρ. $Z_1 = Y_1$; $Z_2 = Y_1\,\rho + Y_2\,\sqrt{(1-\rho^2)}$. This rotation ensures that $\mathrm{corr}[Z_1, Z_2] = \rho$.

Panel A2: (1) returns of assets A and B are from a mixture of normal distributions:

$r_{A,mix} \sim \{10\%\ N[0, 10] + 90\%\ N[0, \sigma=1]\}$;

$r_{B,mix} \sim \rho\ r_{A,mix} \oplus \sqrt{(1-\rho^2)}\ r_{A,mix,\ new}$;

(2) $\rho = corr(r_A, r_B) = 0.7$;

(3) the estimation sample n=25.

Panel B1: (1) returns of assets A and B are normal: $r_A \sim r_B \sim N[0, \sigma=3.3]$;

(2) $corr(r_A, r_B) = 0.5$;

(3) the estimation sample n=15.

Panel B2: (1) returns of assets A and B are from a mixture of normal distributions:

$r_{A,mix} \sim \{10\%\ N[0, 10] + 90\%\ N[0, \sigma=1]\}$;

$r_{B,mix} \sim \rho\ r_{A,mix} \oplus \sqrt{(1-\rho^2)}\ r_{A,mix,\ new}$;

(2) $corr(r_A, r_B) = 0.5$;

(3) the estimation sample n=15.

Panel C1: (1) returns of assets A and B are normal: $r_A \sim r_B \sim N[0, \sigma=1.84]$;

(2) $corr(r_A, r_B) = 0.5$; (3) estimation sample n=15.

Panel C2: (1) returns of assets A and B are from a mixture of normal distributions:

$r_{A,mix} \sim \{10\%\ N[0, 5] + 90\%\ N[0, \sigma=1]\}$;

$r_{B,mix} \sim \rho\ r_{A,mix} \oplus \sqrt{(1-\rho^2)}\ r_{A,mix,\ new}$;

(2) $\rho = corr(r_A, r_B) = 0.5$;

(3) the estimation sample n=15.

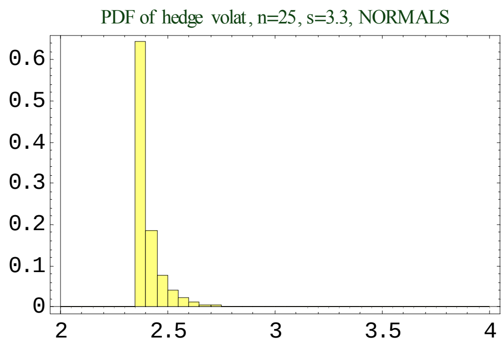
Note that comparisons of hedge volatility distributions under normality vs. non-normality are consistent – e.g., the individual asset volatilities are the same under both normal and non-normal scenarios.

For example, the true asset volatility is: $\sigma = [(0.1)\ 100 + (0.9)\ 1]^{1/2} = 3.3$ and true cross-correlation $\rho = corr(r_A, r_B) = 0.7$ for Panels A1 and A2 of this exhibit. It is apparent that asset return leptokurtosis increases the ex-ante dispersion of the proposed hedge volatilities and the model misspecification may lead to a non-negligible increase rather than decrease in the hedged position volatility.

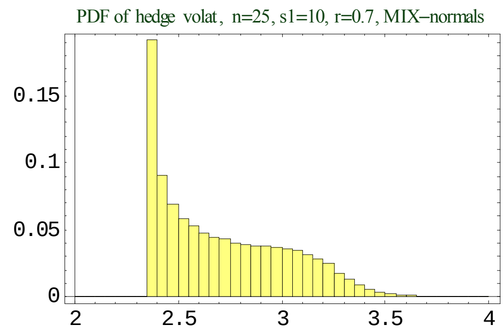
Exhibit II.3.4

Probability Distributions of Hedge Volatilities; Normal versus Mixture-of-Normal Returns

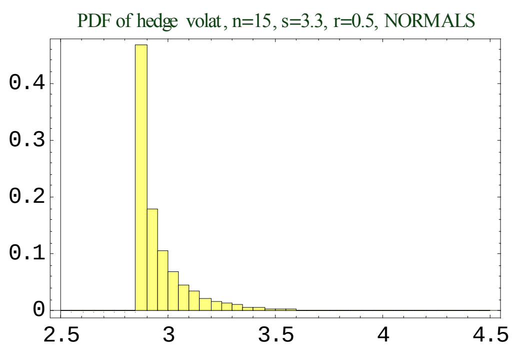
Panel A1:
Normals, σ=3.3, ρ=0.7, n=25



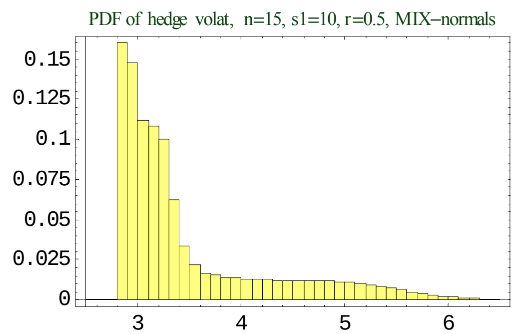
Panel A2:
MixNormals, w=10%, σ_1=10, σ_2=1, ρ= 0.7, n=25



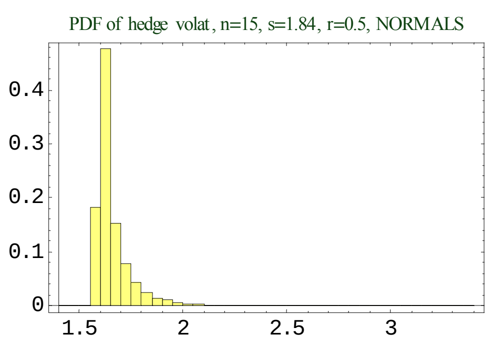
Panel B1:
Normals, σ=3.3, ρ=0.5, n=15



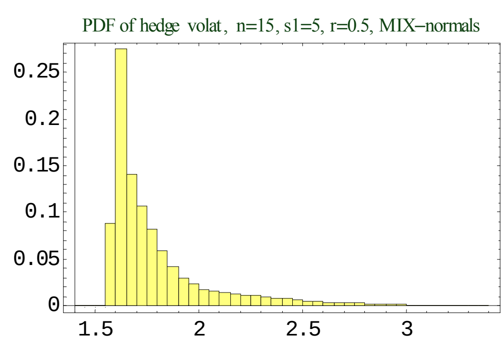
Panel B2:
MixNormals, w=10%, σ_1=10, σ_2=1, ρ= 0.5, n=15



Panel C1:
Normals, σ=1.84, ρ=0.5, n=15



Panel C2:
MixNormals, w=10%, σ_1=5, σ_2=1, ρ= 0.5, n=15



Note:
See the text for a description of parameters. Each bar in Panel A, and Panel B is 0.02 units wide and represents the probability that a realization of hedge volatility falls within some corresponding interval. Number of sampling repetitions is 100,000 for both panels. Some representative probabilities are given below:

Exhibit II.3.4 continued:

Prob ($\sigma_{hedge\ portfolio} \in$ Int):

Interval	< 2.4	2.4–2.7	2.7–3.0	3.0–3.3	> 3.3	$\sigma_{A,B}$	$\sigma_{H,min}$	$\sigma_{H,avg}$
Panel A1: Normal	64.3%	34.6%	1.0%	0.1%	-	3.30	2.36	2.41
Panel A2: MixN	19.3%	36.2%	23.2%	17.3%	4.0%	3.30	2.36	2.72

Prob ($\sigma_{hedge\ portfolio} \in$ Int):

Interval	< 3.0	3.0–3.3	3.3–3.6	3.6–3.9	> 3.9	$\sigma_{A,B}$	$\sigma_{H,min}$	$\sigma_{H,avg}$
Panel B1: Normal	74.7%	20.4%	3.6%	0.9%	0.3%	3.30	2.86	2.97
Panel B2: MixN	31.0%	31.8%	11.5%	4.7%	19.6%	3.30	2.86	3.55

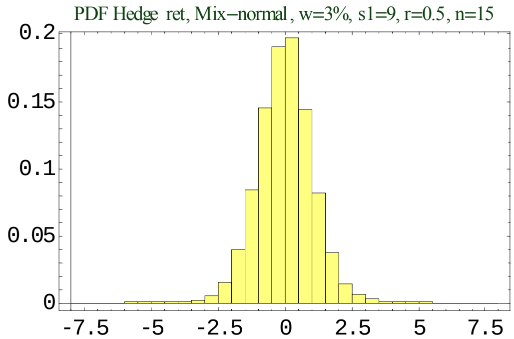
Prob ($\sigma_{hedge\ portfolio} \in$ Int):

Interval	< 1.7	1.7–1.8	1.8–1.9	1.9–2.5	> 2.5	$\sigma_{A,B}$	$\sigma_{H,min}$	$\sigma_{H,avg}$
Panel C1: Normal	80.7%	12.1%	4.2%	2.8%	<0.1%	1.84	1.60	1.66
Panel C2: MixN	50.6%	18.8%	9.9%	16.7%	3.8%	1.84	1.60	1.80

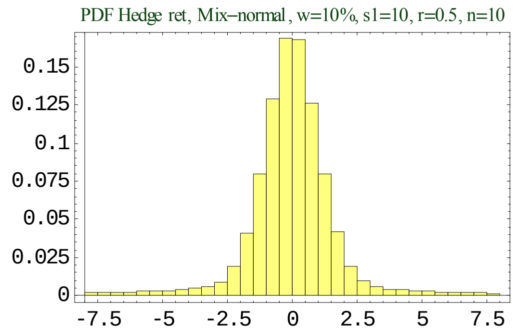
While the hedge volatilities are the usual concern and an objective function of a minimum variance hedge, the ultimate interest rests with the distribution of the resulting hedge returns. While the hedge returns are well described by the second moment if asset returns are normal (apart from the case of an extremely small sample), the higher moments play an important role if asset returns exhibit leptokurtosis. Exhibit II.3.5 provides examples of hedge return distributions under parameter uncertainty and non-normal asset returns. It is shown that the parameter and model uncertainty may lead to an increase in the volatility of the position if hedged, even for relatively larger estimation windows (Panel C). However, the sampling reveals that an alternative measure of risk, V@R, may be smaller for the hedged position relative to an unhedged one. This is an interesting consequence of the central limit theorem – e.g., the portfolio returns will be more normal than the individual asset returns, thus leading to a decrease in V@R. One other result is noteworthy. The two measures of risk, variance versus V@R, may provide contradictory indications about the change in risk of a hedged position. While the variance of normally distributed positions may decrease if hedged with leptokurtic assets, the V@R metrics at some confidence level may actually increase. Panel D provides an example of a normally distributed asset hedged by a leptokurtic one. Even though the estimation window is relatively large (n=20), both variance and V@R(1%) of the hedged position increase substantially.

Exhibit II.3.5 PDF of Hedge Returns under Parameter Uncertainty. Non-Normal Returns.

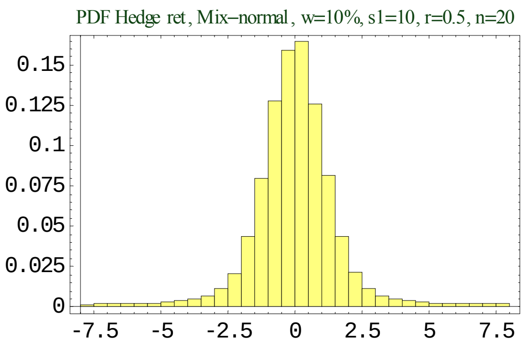
Panel A.: MixNormals
$\sigma_A=\sigma_F=1.84$, ($w=3\%$, $\sigma_1=9$); $\rho=0.5$; $n=15$
$\sigma_{Hedge}=1.85$, $k_H=45.5$



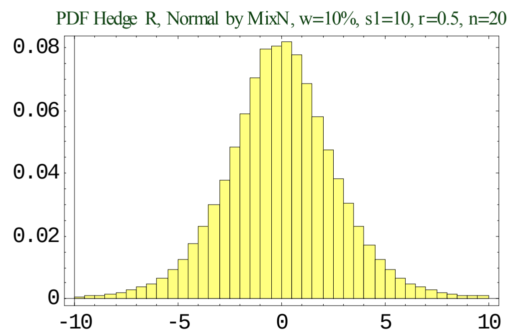
Panel B.: MixNormals
$\sigma_A=\sigma_F=3.3$, ($w=10\%$, $\sigma_1=10$); $\rho=0.5$; $n=10$
$\sigma_{Hedge}=3.75$, $k_H=32.4$



Panel C.: MixNormals
$\sigma_A=\sigma_F=3.3$, ($w=10\%$, $\sigma_1=10$); $\rho=0.5$; $n=20$
$\sigma_{Hedge}=3.5$, $k_H=24$



Panel D.: Normal returns hedged by MixNorm
$\sigma_A=\sigma_F=3.3$, ($w=10\%$, $\sigma_1=10$); $\rho=0.5$; $n=20$
$\sigma_{Hedge}=3.5$, $k_H=15.9$



NOTE:
σ_{Hedge} denotes the standard deviation of the hedge portfolio that takes into account the estimation error (the distribution of hedge ratio parameter), given the stated estimation and distributional parameters.
k_H denotes the excess kurtosis of the hedge portfolio given the stated estimation parameters.

II.4 Value-at-Risk and Model Misspecification

This section illustrates the magnitude of the potential errors posed by the distributional assumptions. One famous example of underestimation of the potential for model error is the case of LTCM hedge fund, which faced near bankruptcy in September 1998 (see Jorion, 2000; Stulz, 2000).

Assume that the true asset returns follow a jump diffusion process, or its discrete time alternative – known as the *mixture-of-normals model*. Roll (1988) has analyzed some aspects of the mixture model and provides early evidence about the complex structure of returns. Quandt and Ramsey (1978) suggest the following specification:

$$r_A = \begin{cases} \sigma_{A,1}\,\varepsilon_{A,1} & \text{in state } X \text{ with probability "}w\text{"} \\ \sigma_{A,2}\,\varepsilon_{A,2} & \text{in state } Y \text{ with probability "}1\text{-}w\text{"} \end{cases}, \qquad 2.4.1$$

where $\varepsilon_{A,1} \sim \varepsilon_{A,2} \sim N[0,1]$. The returns observed conditional on state "X" can be thought of as returns observed during periods with "no asset-specific news". However, returns occasionally exhibit jumps caused by unexpected news events (state Y). Reasonable values for *w* are: $w \in (\,0.1\%,\ 10\%\,)$. This model can be more compactly written as:

$$r_A \sim w\, N[0, \sigma_{A,1}] + (1\text{-}w)\, N[0, \sigma_{A,2}]\ , \qquad 2.4.2$$

where "+" results in a "mixture" or "superposition" of two distribution functions.

Value-at-Risk (V@R) is often calculated at different confidence levels – the most common being 99%.[33] The usual criticism is that V@R neglects to take into account the magnitude of returns. However, this criticism is misplaced, since the V@R metrics is *not* the *absolute* maximum one can lose, but rather a point on a continuous distribution of possible returns. The correct use of V@R techniques involves calculation of V@R measures at various confidence levels. Alternatively, it should provide a complete description of the return distribution.

---

[33] The other commonly used confidence levels are: 97.5% and 95%. Confidence level of 99% means that the probability of returns exceeding $V@R_{99\%\ \text{confidence level}}$ is 1%. In other words, portfolio under consideration will lose more than the specified V@R amount only once in 100 periods.

Exhibit II.4.1 illustrates the potential for understating V@R metrics if prices follow a jump-diffusion process (or an alternative mixture-of-normals process), but the risk-manager acts as if returns are normally distributed. Panels A and B compare the PDF of a posited true distribution: $r_{A,mix} \sim \{95\%\ N[0, 1] + 5\%\ N[0, \sigma=5]\}$,[34] with the PDF of a normal distribution with the same volatility: $r_{A,Gauss} \sim N[0, \sigma=1.48]$. If the above returns are daily returns, and the volatility (standard deviation) is expressed in percentage terms, the unconditional annual volatility of the asset's returns is:

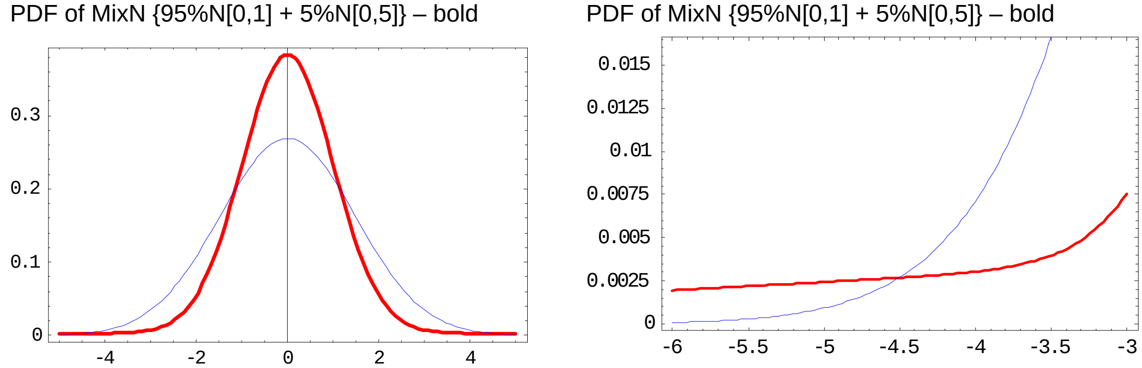
$$\sigma_A = \sqrt{250(0.95 \times 1^2 + 0.05 \times 5^2)} = 23.45,$$

and approximates (in percentage terms) the usual estimates of annual volatility of stocks and stock indexes.

Exhibit II.4.1 Value-at-Risk and model misspecification

Panel A.
PDF of N[0, 1.48] – thin;
PDF of MixN {95%N[0,1] + 5%N[0,5]} – bold

Panel B.
PDF of N[0, 1.48] – thin;
PDF of MixN {95%N[0,1] + 5%N[0,5]} – bold



Note: Panels A and B show the PDF of a posited return distribution (mixture of normals) and the PDF of a Gaussian distribution with the same volatility (standard deviation). The table below shows some lower quantiles of two distributions:
1) bold line: $r_{A,mix} \sim \{95\%\ N[0, 1] + 5\%\ N[0, \sigma=5]\}$, and 2) thin line: $r_{A,Gauss} \sim N[0, \sigma=1.48]$

V@R *Lower Quantile*	0.1%	0.5%	1.0%	2.5%	5.0%	10%
$\Phi^{-1}(r_{A,mix})$	-10.27	-6.41	-4.21	-2.35	-1.83	-1.37
$\Phi^{-1}(r_{A,normal})$	-4.58	-3.82	-3.45	-2.91	-2.44	-1.90

[34] This corresponds to the "no-news" returns occurring 95% of the time, and the "news/jump" returns occurring 5% of the time (e.g., on 12 days during one year).

Exhibit II.4.2 shows the probability distribution of estimated V@R values under the following conditions:

(1) the true return distribution is $r_{A,mix} \sim \{95\%\ N[0, 1] + 5\%\ N[0, \sigma=5]\}$, and

(2) V@R metrics is based on a 99% quantile of a normal distribution $N[0, \sigma]$, where "σ" is estimated using returns observed during the 100 most recent days. This means that a good estimate of 1-day V@R appears to be 2.32 σ. The average V@R that would be obtained under normality assumption and the correct volatility estimate (σ=1.48) is:

Expected [V@R (99%, normality assumption)] = (2.32)(1.48) = 3.43

However, the true Value-at-Risk metric based on the posited return distribution is $V@R_{99\%} = 4.21$. Thus, the true V@R will be 23% [=4.21/3.43 - 1] larger than the mean estimated V@R based on an erroneous model assumption. Even the usual confidence bounds for the average V@R estimate (given 100 returns) is below the true V@R. Specifically, if the return distribution is normal, then confidence bounds for a volatility estimate can be based on the following approximation:

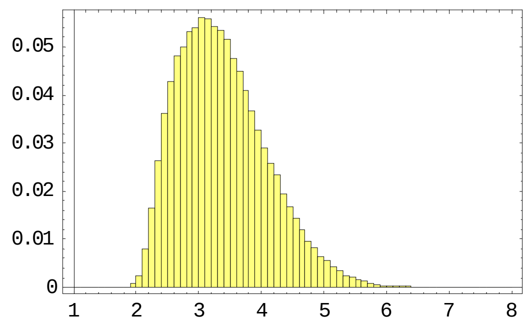
$$\hat{\sigma}^2 \approx N\left[\sigma^2, \frac{2\sigma^4}{n-1}\right] , \qquad \textit{2.4.3}$$

where *"n"* denotes the length of an estimation period. The relative standard error of the volatility and V@R estimates (under normality) is about 7% of the V@R value (given 100 returns). Therefore, the 99% confidence bound for the *E*[V@R] is (2.7, 4.2). In sum, one could "confidently" underestimate the true V@R in 61% of all cases.

Exhibit II.4.3 illustrates the magnitude of the model error for two other experiments. Panels A1 and A2 compare the quantiles of (1) a mixture of normals distribution with returns $r_{A,mix} \sim \{99\%\ N[0, 1] + 1\%\ N[0, \sigma=5]\}$ – thick line; with the quantiles of (2) a normal distribution with the same volatility parameter: $r_{A,Gauss} \sim N[0, \sigma=1.24]$ – thin line.

Panels B1 and B2 compare the quantiles of 1) mixture of normals distribution where : $r_{A,mix} \sim \{97\%\ N[0, 1] + 3\%\ N[0, \sigma=5]\}$ – thick line, with quantiles of 2) normal distribution with the same volatility: $r_{A,Gauss} \sim N[0, \sigma=1.84]$ – thin line. Under perfect knowledge of volatility estimates, the above specifications turn out the 0.1% quantiles (an event once in 4 years) and thus corresponding V@R measures that are 86% and 189% greater for the posited true distributions than V@R measures estimated under erroneous normality assumption.

Exhibit II.4.2 **Probability Distribution of V@R estimate;** (α=1%, n=100, Normal returns)



Note: The Probability Distribution of the Value-at-Risk estimate at the 99% confidence level. V@R metric is calculated using the normality assumption and 100 observations from the assumed leptokurtic return distribution: $r_{A,mix} \sim \{95\%\ N[0, 1] \oplus 5\%\ N[0, \sigma=5]\}$. The true $V@R_{99\%}$ based on the parameters of the assumed distribution is *4.21*. The cumulative probability P[|V@R| < 3.5] = 61%.

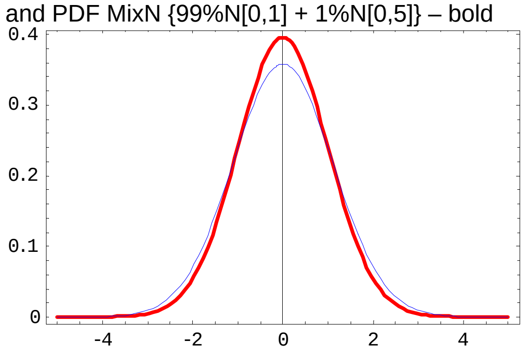
Prob (Estimated $V@R_{99\%} \in$ ***Int):***

Estimated $V@R_{1\%}$ Interval	< 2.5	2.5, 3.0	3.0, 3.5	3.5, 4.0	4.0, 4.5	> 4.5	Average	True
Probability	8.9%	25.0%	27.1%	20.2%	11.4%	7.5%	3.43	4.21

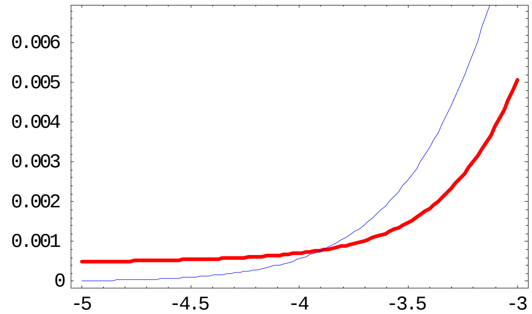
Technical Note: A sample Ω of 5,000,000 returns is generated from the mixture of normals specified above. For each trial, a random subsample of 100 observations is selected from Ω. To estimate V@R, the standard deviation of a subsample is used to calculate a V@R metric as specified by relation 1.3.2. Number of repetitions to obtain PDF: 100,000.

Exhibit II.4.3 Value-at-Risk and model misspecification.

Panel A1: PDF of N[0, 1.24] – thin;
and PDF MixN {99%N[0,1] + 1%N[0,5]} – bold



Panel A2: PDF N[0, 1.24] – thin;
and PDF MixN {99%N[0,1] + 1%N[0,5]} – bold

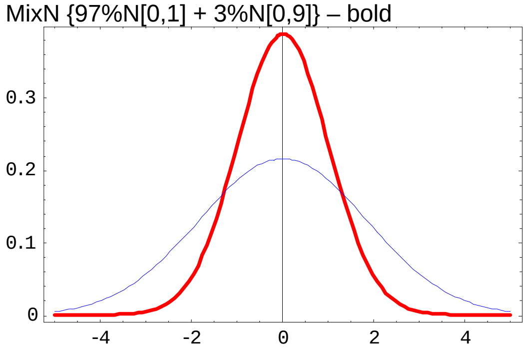


Note: Panels A and B show the PDF of a posited return distribution (mixture of normals) and the PDF of a Gaussian distribution with the same volatility (standard deviation). The table below shows some lower quantiles of two distributions:
(1) bold line: $r_{A,mix}$ ~ {99% N[0, 1] + 1% N[0, σ=5]}, and (2) thin line: $r_{A,Gauss}$ ~ N[0, σ=1.24]

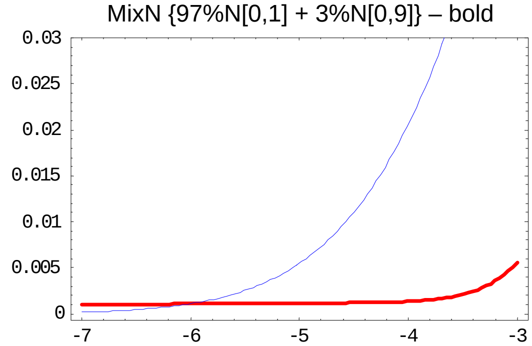
Value at Risk metrics:

Confidence Level (Lower Quantile)	0.1%	0.5%	1.0%	2.5%	5.0%	10%
$\Phi^{-1}(r_{A,mix})$ – thick line	-6.41	-2.85	-2.46	-2.02	-1.68	-1.30
$\Phi^{-1}(r_{A,normal})$ – thin line	-3.44	-2.87	-2.59	-2.18	-1.83	-1.43

Panel B1: PDF of N[0, 1.84] – thin
MixN {97%N[0,1] + 3%N[0,9]} – bold



Panel B2: PDF of N[0, 1.84] – thin
MixN {97%N[0,1] + 3%N[0,9]} – bold



Note: Panels A and B show the PDF of a posited return distribution (mixture of normals) and the PDF of a Gaussian distribution with the same volatility (standard deviation). The table below shows some lower quantiles of two distributions:
(1) thick line: $r_{A,mix}$ ~ {97% N[0, 1] + 3% N[0, σ=9]}, and (2) thin line: $r_{A,Gauss}$ ~ N[0, σ=1.84]

Confidence Level (Lower Quantile)	0.1%	0.5%	1.0%	2.5%	5.0%	10%
$\Phi^{-1}(r_{A,mix})$ – thick line	-16.5	-8.71	-3.91	-2.22	-1.77	-1.34
$\Phi^{-1}(r_{A,normal})$ – thin line	-5.70	-4.75	-4.29	-3.61	-3.03	-2.36

II.5 Conclusions

"The effects of noise on the world, and on our views of the world, are profound. . . . Most generally, noise makes it very difficult to test either practical or academic theories about the way that financial or economic markets work. We are forced to act largely in the dark."

Fisher Black, "Noise," (Journal of Finance, 1986).

The measurement of financial return characteristics and subsequent hedging decisions are subject to large estimation errors and model misspecifications. The various model assumptions and simulations presented here provide an important insight about the impact of the true return dynamics and our misrepresentation of it on the outcome of our hedging decisions.

The risk of a minimum-variance hedge is shown to be *not* unambiguously smaller than the risk of an unhedged position. Moreover, different measures of risk – such as portfolio variance and V@R metrics – may lead to contradictory conclusions about the change in a risk level. It is shown that if asset returns are normally distributed, only an extreme parameter uncertainty and very low absolute levels of return correlation would lead to an increase in a risk level, no matter what risk measure is used. However, the same conclusions cannot be made for a leptokurtic asset return series. If prices are generated by a process resulting in a return distribution that is a mixture-of-normals, the *Gaussian* model assumption may lead to a hedging decision that may greatly increase the risk of an underlying investment position.

Hedging errors and the potential increase in risk is most likely to occur if return distributions exhibit extreme kurtosis and the estimation period is rather short. An investor may be tempted to choose short estimation periods either directly because of a non-stationarity of return series previously reported in finance literature, or indirectly by selecting a model that weights the most recent observations more heavily. It is shown that the *effective* estimation periods of certain ARCH-class models such as *Equally Weighted Moving Average* (EWMA) models can be rather short (25 to 100 days for daily return series) even if very long *actual* data series are used. The short estimation sample introduces extra noise into the parameter estimation and may lead to a hedging decision that increases risk (in terms of variance as well as in terms of V@R) especially if the return distribution is leptokurtic.

The hedging errors can be traced to large standard errors of estimated model parameters such as correlation coefficient. Model misspecification may lead investors to believe that the correlation coefficient of return series is non-stationary. However, existence of jumps in asset return series may lead to the standard errors of sample correlation coefficient estimates that are two to three times larger than under the normality of returns. Subsequently, large errors in the estimates of asset cross-correlations can lead to large errors in hedge ratios. Model and estimation errors may ultimately lead to a return distribution of a hedged position that has undesirable properties, e.g., Value-at-Risk of the hedged position at some confidence level may be substantially larger than if the underlying position were left unhedged. What should an underwriter do to prevent unnecessary model errors? A simple examination of the return properties, alternative models with (preferably) longer *effective* estimation periods should prove highly informative.

It is also shown that V@R may be significantly understated in both economic and statistical sense if a incorrect model is used. The mixture-of-normals model can lead to V@R metrics that dramatically exceed metrics under the normality assumption. For some mixture-of-normals model specifications, the true V@R metric at low (0.1%) confidence levels is a large multiple of V@R metric estimated by (erroneously) assuming that returns are normally distributed. Moreover, back-testing of models at higher confidence levels (e.g., 5% or 10%) may lead to a false information that models used correctly represent the empirical samples. How serious the problem is can be answered only by examining various asset series, which is a subject of CHAPTER THREE.

Chapter III

ESTIMATION OF DISTRIBUTIONAL PARAMETERS AND V@R METRICS

"Any virtue can become a vice if taken to extreme. The models are only approximations to the complex, real world. The practitioner should therefore apply the models only tentatively, assessing their limitations carefully in each application."

Robert Merton
(Comments in 1995, as quoted by *Risk* magazine, October 1998, in "Meriwether Meltdown")

III.1 Introduction

The evidence on asset return distribution is reviewed in CHAPTER ONE, SECTION FOUR. Prior literature documents that asset return distributions depart substantially from normality. The assumption of normality is frequently violated in empirical samples of short-term (daily) returns (Fama, 1965; Campbell, Lo, and MacKinlay, 1997; Peiró, 1999). Most asset return series exhibit leptokurtosis. Moreover, Peiró (1999), employing distribution-free methods, documents that return series in a number of markets (e.g., least capitalized) are not symmetric. Also, Jorion (1989) investigates the existence of discontinuities in the sample path and finds that exchange rates exhibit significant jumps even after allowing for conditional heteroskedasticity in the diffusion process. Therefore, V@R procedures relying on symmetry such as those assuming Gaussian, t-distribution, or *EWMA* based models may be mis-specified resulting in an underestimate of V@R metrics. One can only speculate on how many wrong decisions are made due to the use of an incorrect model.

One of the proposed alternatives is a mixture of normal distributions. Roll (1988) shows that return kurtosis in his sample of stock returns is different when excluding returns on all days surrounding news reports in the financial press, thus "revealing a mixture of returns distribution." He is able to estimate the minimum ratio of news variance to noise (no news) variance and documents that it is substantial. Non-news days also indicate a presence of a distributional mixture, presumably due to traders acting on their private information, but the effect is relatively smaller.

This chapter reviews estimation methodologies of mixture-of-normal-distributions model parameters. The model's parameter estimators are based on Pearson's *Method of Moments* (MM) and *Maximum Likelihood* (ML) surface-search procedure. The ML procedure uses starting values from the MM estimation in order to avoid singularities. Parameter estimates are provided for several assets such as exchange rates, stocks, and precious metals. Resulting mixture-of-normals-ML V@R metrics are compared to the metrics obtained by some alternative models commonly used by academicians and practitioners alike (e.g., *Gaussian, ARCH/EWMA*). The *Pearson-Fisher test of goodness-of-fit* is used to test the agreement of posited models with the empirical samples.

An alternative to the discrete time mixture-of-normals model is the continuous time jump-diffusion model (see equation *1.7.4*) of Merton (1976a), Cox and Ross (1976) and Bates (1991) as discussed in CHAPTER ONE. However, there are substantial difficulties in estimating parameters of a general JD-model in practice. The prior Maximum Likelihood estimation techniques of Beckers (1981) and Ball and Torous (1985) provide estimators in a somewhat simplified setting. Their estimators rely on some parameter restrictions such as the assumption that the jump component is mean zero, which restricts distribution classes to symmetric ones. This would unnecessarily limit the usefulness of this continuous time model in cases where the empirical return distribution is skewed. Also, the Maximum Likelihood estimation with "wrong starting values" may produce volatilities that are negative.[35]

III.2 Model Selection and Goodness-of-Fit Tests

The standard V@R methodologies (e.g., those employed by RiskMetrics™) assume that asset returns are conditionally (*EWMA*) or unconditionally (*GBM*) normally distributed. As noted before, V@R procedures assuming normality may result in a V@R

[35] A similar problem occurs in an ML estimation of parameters for a mixture of normals, but it may be alleviated by a careful selection of starting values. Estimates produced by the method of moments may be used for this purpose.

metric estimate that is substantially different (smaller) from the true metric if returns are draws from a leptokurtic distribution, such as the *mixture-of-normals*. The most important empirical question is the selection of the asset return model that has the ability to fit (or predict) the characteristics of data well (*ex-post* and more importantly *ex-ante*).

While numerous asset return-generating models have been suggested in place of the models assuming unconditional normality of returns, there is a need to test their empirical performance.[36] Susmel and Hamilton (1994, page 312) note that the forecasting performance of *GARCH* specifications is rather poor in a sample of the value-weighted portfolio of stocks traded on the NYSE from July 1962 to December 1987. If their specifications were correct and the parameters were known with certainty, then σ_t^2 would be the conditional expectation of squared innovations. Hence a mean squared error loss function MSE = E $[(\varepsilon_t^2 - \theta_t)^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots]$ would be minimized with respect to θ_t by choosing $\theta_t = \sigma_t^2$. However, they show that the *GARCH* MSE is larger than the MSE under the assumption of constant variance during the period.

While the *MSE* loss function test provides a valuable insight, a goodness-of-fit test is necessary to assess the models' empirical performance, especially in V@R applications. It is possible that even if a particular *GARCH* (*EWMA*) model specification performs well in terms of the *MSE* test relative to the Gaussian model, it is not clear that it is a "good" model. A well-established alternative to judge the performance of suggested models is a *goodness-of-fit* test pioneered by Karl Pearson in 1900, when his paper introducing the X^2 test appeared.[37] Without a statistical test, model selection is often assessed by visual subjective methods. However, as R. A. Fisher (1925) commented: "No eye observation of diagrams, however experienced, is really capable of discriminating whether or not the observations differ from the expectation by more than we would expect from the circumstances of random sampling."[38]

---

[36] As mentioned earlier, the alternative models are usually of the GARCH family of models.
[37] A *goodness-of-fit* procedure is a statistical test of a hypothesis that the sampled population is distributed in a specific way ... for example, that the sampled population is normal (Rayner and Best, 1989).
[38] R. A. Fisher (1925, p. 399) as cited by Rayner and Best (1989, p. 4)

III.2.1 Goodness-of-Fit Test

The Pearson's X^2 smooth test of goodness-of-fit tests a simple null hypothesis that a random sample $\{x_1, \ldots, x_n\}$ of size n comes from a population with a completely specified cumulative distribution function $F(x)$. The sample space is partitioned into m classes, and N_j is defined as the number of observations from the sample that fall into the j-th class. If p_j is the probability of falling into the j-th class when $F(x)$ holds, the Pearson X^2 test statistic is defined as:

$$X_P^2 = \sum_{j=1}^{m} \left(N_j - np_j\right)^2 / \left(np_j\right). \qquad 3.2.1$$

As n increases, the distribution of X^2 tends to be increasingly well approximated by χ^2_{m-1}, the χ^2 distribution with $m-1$ degrees of freedom. The null hypothesis that the sample comes from the specified distribution is rejected if X^2 test statistic is larger than the critical $A(\alpha\%, m-1)$ point of the χ^2 distribution. If this test is to be used, the cells or classes must be constructed. This choice is not innocuous, as different class selection may lead to different conclusions. Rayner and Best (1989, p.24) discuss extensively the class construction and statistical literature on this topic. Since we are primarily interested in how well various models fit empirical data in the "tails", the inferences here are based on a construction of nine classes ($m=9$), one large central class (containing 94% of data) and eight equiprobable (1%-tile) tail classes (four classes for each right and left tail).

If the null distribution depends on a vector $\beta=(\beta_1, \ldots, \beta_q)$ of unknown parameters, then p_i in X^2 must be replaced by an estimate of p_i, say $\hat{p}_i$. If the $\hat{p}_i$ is based on the grouped maximum likelihood estimator, the new statistic is X^2_{PF}, the Pearson-Fisher statistic, which has an asymptotic null χ^2_{m-q-1} distribution. If maximum likelihood estimators are based on the ungrouped observations, the Chernoff-Lehmann X^2_{CL} test statistic is obtained. As noted by Rayner and Best (1989, p.27), it is sufficient to base inferences on the fact that the null distribution of X^2_{CL} is bounded between χ^2_{m-q-1} and χ^2_{m-1}.

One final remark is in order. It is relatively easy to calculate the likelihood of observing a specific sample given some model and its estimated parameters. The ratio of likelihoods of two models may provide some guidance as to the model preference.

However, the *likelihood ratio test* can not be used to test one distributional assumption against another. The test is inappropriate, because the parameter spaces, and hence the likelihood functions of the two cases, are unrelated (Greene, 1993, p. 162).

III.3 The Mixture-of-Normal-Distributions Model

The problem of separating the components of a probability density function, which is a mixture of normal densities, occurs in a wide variety of disciplines.[39] It appears that this problem is one of the oldest estimation problems in the statistical literature, beginning with Pearson (1894), who first suggested the Method of Moments estimator. The simplest possible case of a mixture model is the case in which it is known a priori (or in which the theory suggests) that the number of components is two. In this scenario, a sample of observations $\{x_1, x_2, \ldots, x_n\}$ is given on a random variable x defined as:

$$x \sim N[\mu_1, \sigma_1^2] \quad \text{with probability } \lambda,$$

and 3.3.1

$$x \sim N[\mu_2, \sigma_2^2] \quad \text{with probability } 1-\lambda,$$

where the parameter vector $\{\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2\}$ is unknown. A similar specification can be found in Roll (1988). The density function of the random variable x is:

$$f(x) = \frac{\lambda}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{1-\lambda}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}, \qquad 3.3.2$$

with the corresponding moment generating function (MGF):

$$m(\theta) = \lambda \cdot e^{\mu_1\theta + \frac{1}{2}\sigma_1^2\theta^2} + (1-\lambda) \cdot e^{\mu_2\theta + \frac{1}{2}\sigma_2^2\theta^2}. \qquad 3.3.3$$

The variance (the second central moment) of a mix-normal random variable "*x*" with a density function given by expression *3.3.2* can be obtained from the above MGF

---

[39] Quandt and Ramsey (1978) mention several examples, such as engineering (Young and Coraluppi, 1970), biology (Bhattacharya, 1966), as well as economics (Quandt, 1972; Ramsey, 1975).

by differentiating twice the natural logarithm of $m(\theta)$, or a Cumulant Generating Function CGF = ln m(θ), with respect to θ and letting θ → 0,

$$Variance(x) = \left.\frac{\partial^2 \ln(m(\theta))}{\partial\theta^2}\right|_{\theta\to 0} = \lambda(1-\lambda)(\mu_1-\mu_2)^2 + \lambda\sigma_1^2 + (1-\lambda)\sigma_2^2 \quad 3.3.4$$

The third central moment, or Skewness = E(x - E(x))3, can be expressed as:

$$Skewness(x) = \left.\frac{\partial^3 \ln(m(\theta))}{\partial\theta^3}\right|_{\theta\to 0} = \lambda(1-\lambda)(\mu_1-\mu_2)\left[(1-2\lambda)(\mu_1-\mu_2)^2 + 3(\sigma_1^2-\sigma_2^2)\right] \quad 3.3.5$$

Similarly, for the fourth central moment, or Kurtosis = E[x - E(x)]:

3.3.6

$$Kurtosis(x) = \left.\frac{\partial^4 \ln(m(\theta))}{\partial\theta^4}\right|_{\theta\to 0} =$$

$$= \lambda(1-\lambda)\left[\left((\mu_1-\mu_2)^4 + 6(\mu_1-\mu_2)^2(\sigma_1^2-\sigma_2^2) - 3(\sigma_1^2-\sigma_2^2)^2\right) - 6\lambda\left((\mu_1-\mu_2)^4 + 2(\mu_1-\mu_2)^2(\sigma_1^2-\sigma_2^2)\right) + 6\lambda^2(\mu_1-\mu_2)^4\right]$$

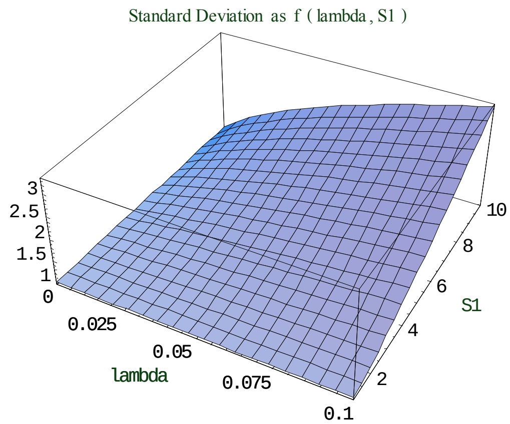
Exhibit III.3.1 shows the *standard deviation* and *coefficient of kurtosis* of the random variable *x* as a function of some distributional parameters. It can be seen that the coefficient of kurtosis can attain fairly large values if the news-related jump-return frequency is low (parameter $\lambda\to 0$) and the jump size is relatively large (parameter $\sigma_1 >> \sigma_2$). High values of coefficient of kurtosis are commonly observed for daily stock returns (see Campbell, Lo, MacKinlay, 1998).

III.3.1 Estimation of Mixture-of-Normals model parameters

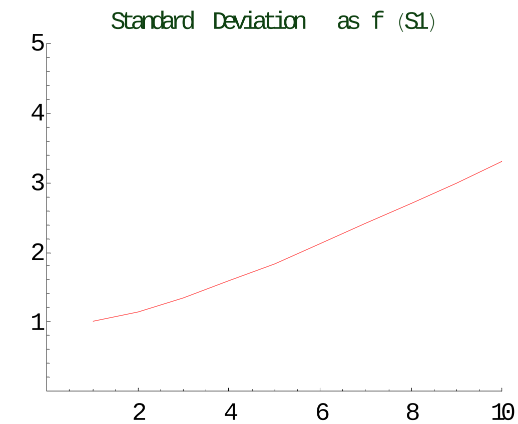
Three principal methods are used for the estimation of the parameter vector $\{ \lambda, \mu_1, \sigma_1, \mu_2, \sigma_2 \}$. The oldest one is the *Method of Moments* suggested by Karl Pearson (1894). Quandt and Ramsey (1978) examine and compare this method to their newly developed method of *Moment Generating Functions.* The last method is based on the *Maximum Likelihood* technique. These methods are discussed below.

Exhibit III.3.1 Standard-Deviation and Coefficient of Kurtosis of a mix-normal random variable.

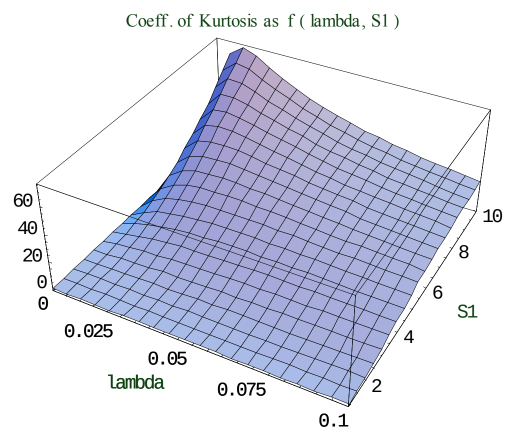
Panel A1: Standard Deviation of ***x***
x ~ MixN { λ N [0, s_1] + (1-λ) N [0, 1] }



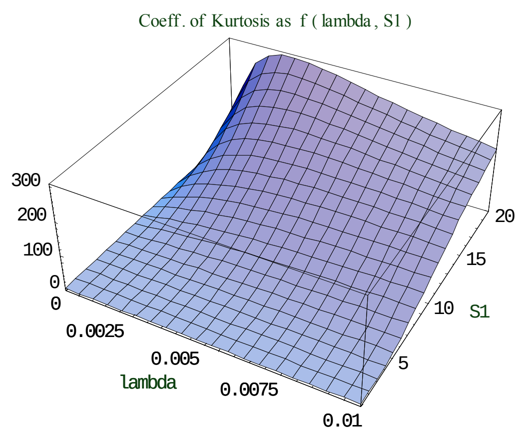
Panel A2: Standard Deviation of ***x***
x ~ MixN { 10% N [0, s_1] + 90% N [0, 1] }



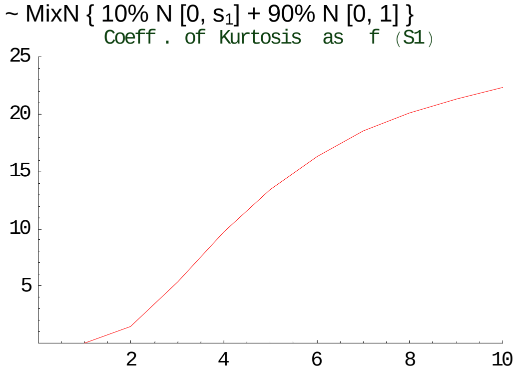
Panel B1: Coefficient of Kurtosis
x ~ MixN { λ N [0, s_1] + (1-λ) N [0, 1] }



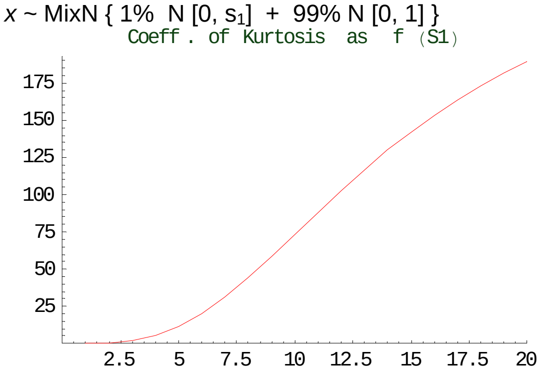
Panel B2: Coefficient of Kurtosis
x ~ MixN { λ N [0, s_1] + (1-λ) N [0, 1] }



Panel B3: Coefficient of Kurtosis
x ~ MixN { 10% N [0, s_1] + 90% N [0, 1] }



Panel B4: Coefficient of Kurtosis
x ~ MixN { 1% N [0, s_1] + 99% N [0, 1] }



1) Method of Moments. This method has been proposed by Pearson (1984) and discussed by Cohen (1967) and Day (1969). The parameters of a density function defined by equation *2.3.2* are estimated by equating the sample mean and second through fifth central moments to the corresponding theoretical moments. This provides five equations in five unknowns that can be solved via a ninth-order polynomial for consistent estimators of the five parameters. Recall that the *k-th* central moment is defined by:

$$m_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k \quad , \; k = 2, 3, 4, 5. \qquad 3.3.7$$

Note that $m_1 = 0$ and m_2 provides an estimate for σ^2. Because $\bar{x}$ converges in probability to μ, the central moments converge in distribution to their true theoretical functions. In general, computing *K* moments and equating them to these functions provides *K* equations that can be solved to provide estimates of the *K* unknown parameters.[40]

The solution requires the negative root of the nonic equation:

$$a_9 z^9 + a_8 z^8 + a_7 z^7 + a_6 z^6 + a_5 z^5 + a_4 z^4 + + a_3 z^3 + a_2 z^2 + a_1 z^1 + a_0 = 0 \; , \qquad 3.3.8$$

where $a_9 = 24$, $a_8 = 0$, $a_7 = 84\, k_4$, $a_6 = 36\,(m_3)^2$, $a_5 = 90\,(k_4)^2 + 72\, k_5\, m_3$,
$a_4 = 444\, k_4\,(m_3)^2 - 18(k_5)^2$, $a_3 = 288\,(m_3)^4 - 108\, m_3\, k_4\, k_5 + 27\,(k_4)^3$,
$a_2 = -[63\,(k_4)^2 + 73\, m_3\, k_5]\,(m_3)^2$, $a_1 = 96\,(m_3)^4\, k_4$, $a_0 = -24\,(m_3)^6$,

where m^i denotes the *i*-th central sample moments and k_j is the *j*-th sample cumulant; e.g., $k_4 = m_4 - 3\,(m_2)^2$, and $k_5 = m_5 - 10\, m_2\, m_3$.

---

[40] Kumar, Nicklin and Paulson (1979) note that the method of moments does not provide estimators which are statistically appealing since the consideration of *only* the first five moments or cumulants may result in a considerable loss of information. Moreover, the higher the order of the moment, the greater the sampling variability. Though any sequence of moments (even fractional, e.g., non-integer) carries information about the parameters in question and can be used to solve for parameters, these procedures are more complicated than solving a nonic equation.

It can be further shown (Cohen, 1967) that if we define the differences:

$$d_1 = \mu_1 - E(x), \text{ and } d_2 = \mu_2 - E(x) \quad , \qquad 3.3.9$$

and if $\hat{z}$ is the negative root which solves equation 3.3.8, and if r is defined by:

$$r = \frac{-8m_3\hat{z}^3 + 3k_5\hat{z}^2 + 6m_3k_4\hat{z} + 2m_3^3}{\hat{z}(2\hat{z}^3 + 3k_4\hat{z} + 4m_3^2)}, \qquad 3.3.10$$

then we obtain as estimates of d_1 and d_2:

$$\hat{d}_1 = \frac{1}{2}\left(r - \sqrt{r^2 - 4\hat{z}}\right), \quad \text{and} \quad \hat{d}_2 = \frac{1}{2}\left(r + \sqrt{r^2 - 4\hat{z}}\right). \qquad 3.3.11$$

From the above results we obtain the estimates of the vector $\{ \lambda, \mu_1, \sigma_1, \mu_2, \sigma_2 \}$:

$$\begin{aligned} \hat{\mu}_1 &= \hat{d}_1 + \bar{x} \ , \\ \hat{\mu}_2 &= \hat{d}_2 + \bar{x} \ , \\ \hat{\sigma}_1^2 &= \frac{1}{3}\hat{d}_1(2r - m_3/\hat{z}) + m_2 - \hat{d}_1^2 \ , \\ \hat{\sigma}_2^2 &= \frac{1}{3}\hat{d}_2(2r - m_3/\hat{z}) + m_2 - \hat{d}_2^2 \ , \\ \hat{\lambda} &= \hat{d}_2/(\hat{d}_2 - \hat{d}_1) \ , \end{aligned} \qquad 3.3.12$$

where $\bar{x}$ is the sample mean, and λ is the mixture coefficient. Though it is known that the sample variances of high-order moment estimates are very large (Kendall, Stuart and Ord, 1998) and therefore affect the variances of parameters, the estimated parameters can be used as starting points in ML estimation which is known to have serious difficulties.[41]

---

[41] See below for further discussion of MLE including a discussion of some common difficulties (see also Quandt and Ramsey, 1978).

2) Method of Moment Generating Function (MGF). This method has been proposed and studied by Quandt and Ramsey (1978), Q&R. They introduce an estimator, which minimizes the sum of squares of differences between the theoretical and sample moment generating functions. They also show the consistency and asymptotic normality of the estimator and compare its finite sample behavior to the *Method of Moments* estimator. The moment generating function is:

$$G(\gamma,\theta_j) = \lambda \cdot e^{\mu_1\theta_j + \frac{1}{2}\sigma_1^2\theta_j^2} + (1-\lambda)\cdot e^{\mu_2\theta_j + \frac{1}{2}\sigma_2^2\theta_j^2}, \qquad 3.3.13$$

where γ is the parameter vector $\{\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2\}$. For any given value of θ_j the quantity $E(e^{\theta x})$ may be estimated by:

$$\bar{y}_n(\theta_j) = \sum_{i=1}^{n} e^{\theta_j x_i} / n, \qquad 3.3.14$$

which converges to $E(e^{\theta x})$ with probability one by the Strong Law of Large Numbers. Since this estimator is, except for sampling error, equal to the moment generating function given by 3.3.13, it is possible to estimate the parameters by minimizing the quantity:

$$S_n(\theta,\gamma) = \sum_{j=1}^{k} [\bar{y}_n(\theta_j) - G(\gamma,\theta_j)]^2 \,. \qquad 3.3.15$$

Q&R recommend selecting the nuisance parameters $\{\theta_j\}_{j=1,\ldots,5}$ so that they are neither close to zero nor too large in order to make $G(\gamma, \theta_j)$ computationally tractable. The MGF estimate γ is obtained by solving five normal equations:

$$S_{n,i}(\theta,\gamma) = \sum_{j=1}^{k} [\bar{y}_n(\theta_j) - G(\gamma,\theta_j)] \frac{\partial G(\gamma,\theta_j)}{\partial \gamma_i} = 0; \quad i=1,2,\ldots,5. \qquad 3.3.16$$

It appears that this method provides more efficient estimates in small samples and when component densities are well separated. However, as Johnson (1978) notes,[42] the MGF method is equivalent to fitting a mixture of two lognormal distributions by using

[42] See his *Comment* that follows the paper by Quandt and Ramsey (1978).

$\{\theta_1, \ldots, \theta_5\}$ moments. Moreover, if λ is close to 0 (as would be expected for distributions of asset returns), the best (ML) estimators of the parameters would be given by the moments, and one would therefore expect the moments estimators to perform relatively better under this assumption. This is also evident from examples provided by Q&R.

3) Method of Maximum Likelihood (ML). Properties of ML estimators are well established (see Green, 1997, page 133). ML estimators are consistent, asymptotically normal, efficient and invariant to re-parameterization. It is possible to estimate the parameters of a normal-mixture distribution defined by *3.3.2* by maximizing the likelihood function:

$$f(X,\gamma) = \prod_{i=1}^{n} f(x_i,\gamma) = L(\gamma \mid X), \qquad 3.3.17$$

where $\boldsymbol{X}$ denotes the sample vector $\{x_i\}_{i=1,\ldots,n}$. Equivalently, it is possible to maximize the log of the likelihood function, which is usually simpler in practice.

The necessary first-order condition for maximizing $ln\ L(\boldsymbol{\gamma}, \boldsymbol{X})$ is:

$$\frac{\partial \ln L(\gamma,\mathbf{X})}{\partial \gamma_i} = 0 \text{, where } \boldsymbol{\gamma} \text{ is the parameter vector } \{\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2\}. \qquad 3.3.18$$

There are several problems with a ML estimation for a mixture of distributions model (see Quandt and Ramsey, 1978). These include problems due to the existence of local maxima, unboundedness of the likelihood function and the potential singularity of the matrix of second partial derivatives of the likelihood function, which is equivalent to a vanishing Jacobian for the set of normal equations derived. The estimation may therefore break down in practice if the components are not well separated (see a *Comment* to Quandt and Ramsey, 1978, by Hosmer). However, it is possible to use the MM estimates as starting values for obtaining the ML estimates, which (naturally) dominate the MM in terms of model probability (likelihood).

III.4 Data and Results

This limited study documents the extent and impact of various model (mis)-specifications in V@R applications. If the estimation procedure correctly predicts the distribution of returns, then we would expect a uniform distribution of the number of *ex-post* returns in individual percentiles predicted by the posited model distribution. For example, given a sample of 10,000 returns, a model and its parameter estimates, we should observe approximately 100 returns (1%) in the first percentile. A significant deviation implies a model misspecification.[43] A goodness-of-fit test, X^2_{PF}, is used to assess the predictive ability of models.

The CRSP database is used for stock returns data. Datastream is the primary source of data on commodities and foreign exchange return series. The estimation covers long return series and compares the V@R metrics assuming normal returns and V@R metrics estimated using the estimates of the mixture of normals model parameters.

III.4.1 Exchange Rates, GBP/USD

Exhibit III.1 presents the distribution of 4011 daily returns of the GBP/USD exchange rate over the period January 1986 to May 2001. For comparative purposes, the histogram also includes the *scaled* normal fit (red/bold-dashed curve) and the mixture-of-normals fit using the maximum-likelihood parameter estimates (green/bold-solid curve). The blue/thin curve depicts the mixture model fit using the method-of-moments estimates. The ratio of models' likelihoods reveals that the mixture-of-normals model is enormously more likely than the normal fit. However, as mentioned earlier, the ratio test is inappropriate as a Likelihood Ratio test, since the parameter spaces and likelihood functions are unrelated.

---

[43] The significance of deviation can be tested by the Pearson's test of goodness-of-fit, or in case of two classes by exact binomial distribution. For example, each (1%-wide) percentile should contain about 1% of all observations, subject to a random variation that depends on the sample size. For a sample of 10,000 returns, the 95% confidence interval is 0.008 – 0.012. For a sample of 1,000, the confidence interval becomes 0.004 – 0.016. See Jorion (2001, Chapter six, page 143) for other commonly used model verification methods, e.g., a similar *Kuiper statistic* is used by J.P. Morgan for model verification.

The distribution of a relative number of returns falling in percentiles predicted by the assumed model and parameter estimates is depicted in Exhibit III.1, Panels B and C. If a model is correct, each percentile class should contain approximately 1% (0.01) of all returns. Even though the number of returns in each predicted percentile is itself subject to random variation and a function of the total sample size, it is possible to estimate the 95% confidence level using the binomial distribution. For the sample of 4011 gold returns, the variation should be confined to the 0.007 to 0.013 interval.

The visual inspection of the exhibits confirms that the mixture-of-normals model fits the empirical sample rather well, apart from the rather troublesome frequency of 0% returns, which exceeds the predicted frequency by a factor of 6. More importantly, the frequency of extreme returns (those in the left-and-rightmost percentiles) conforms to the model. The goodness-of-fit confirms that the mixture-of-normals model is a worthy contender. The Pearson-Fisher statistic X^2_{PF} is below the critical value of the χ^2_{m-q-1} distribution and we would (with reservations) accept this model.

On the other hand, the performance of the normal model is rather weak. The frequency of returns in the right-tail's extreme one percentile exceeds the predicted frequency by a factor of 2.2. The goodness-of-fit test confirms that this model should be rejected. It is worth to note that neither of the two other conditional normality models fares any better (see Exhibit III.1, panel C). Both of the *equally-weighted-moving-average* (EWMA) model specifications can be rejected at very high confidence levels. Note that this conditional normality model, similarly to the Gaussian model, severely underestimates the frequency of the most extreme positive or negative observations by a factor of about 2. The EWMA specifications use an estimate of volatility given by *relation 2.1.1*, and a sample of 25 or 100 most recent observations (denoted as EWMA(25) and EWMA(100), respectively.

Several caveats are in order. As Rayner and Best (1989) note, we can probably reject any potential theoretical model by increasing the sample size or by using different X^2_{PF} –test classes. One would reject even the mixture-of-normals model at very high confidence levels if one hundred 1%-wide equiprobable classes are used. Also, the goodness-of-fit test of the mixture-of-normals model is performed using *in-sample* parameter estimates. This "perfect-insight" estimation is not possible in real-world applications. However, a limited robustness check using only a subset of known prior returns to estimate return distribution still yields predictions that are better than predictions of conditional or unconditional normality models.

Given the financial risk management objective of this dissertation, estimates of V@R metrics are of great interest. The table in Exhibit III.1, Panel A, presents estimates of value-at-risk measures given model assumptions. The empirical standard errors of V@R metrics are also provided. The results suggest that distributional assumptions can lead to a substantial underestimation of Value-at-Risk, though not too dramatic for the GBP/USD exchange rate. The mixture-of-normals model yields V@R metrics at probability levels below one percent that are significantly larger than V@R metrics assuming the Gaussian model. The V@R at 0.1% level is about 30% (=2.61/1.98 –1) larger if the normality model is replaced with distributional mixtures model. It is worth noting that tests of V@R procedures and back-testing would not reveal any discrepancies if the probability level (significance level of a firm's reported V@R metric) used is about 5% or larger. Such a back-testing procedure may confirm the validity of the used models and lead to a false sense of security. It seems, that asset crises and the *"six-sigma"* events often reported in popular press are surprising not because the probability theory suggests that we should not observe them in our lifetimes (since we do), but because our models or "sigmas" are incorrect and incapable of capturing the dynamics of asset returns.

III.4.2 Stocks, E-Bay Inc.

Similar to the results for the GBP/USD exchange rate, the mixture-of-normals model performs significantly better than the alternative conditional and unconditional normality models on a sample of 1190 E-Bay's equity returns, April 1996 to December 2000. Exhibit III.2 presents some descriptive statistics, the distribution of returns with models' fits as well as V@R metric for this relatively more successful internet stock. Normality models (Gaussian, EWMA(25) and EWMA(100)) can be rejected at very high confidence levels, since the realized frequencies of extreme returns exceed the models' predictions by a factor of two to three. Given the sample size, each percentile's frequency should be within 0.004 – 0.016 confidence bound.

However, the use of the mixtures model does not lead to larger V@R metrics (not for any practical purposes). This result is due to the relatively low excess kurtosis of the sample data. Naturally, we would expect the deviations to be larger in samples exhibiting large excess kurtosis.

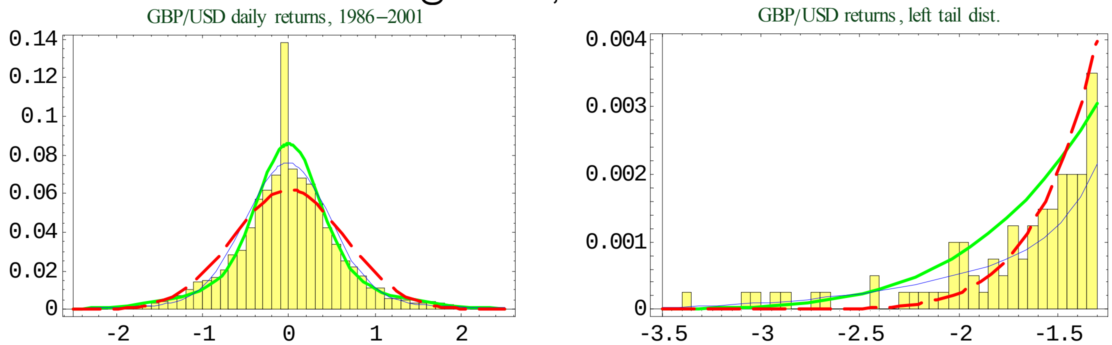
III.4.3 Precious Metals, Gold

Exhibit III.3 presents the distribution of 4540 daily returns of gold over the period January 1984 to May 2001. Similar to the previous assets, the exhibit provides various descriptive statistics. The distribution with scaled model fits and goodness-of-fit tests of various model specifications are also included. The values of the model likelihood function reveal that the mixture-of-normals model is much more likely than the Gaussian model. However, the Pearson-Fisher test rejects even the mixture-of-normals model. The mix-normal model assumption results in V@R estimates at low (but not negligible) probability levels that are significantly and economically meaningfully larger (in absolute value) than V@R metric under the normality assumption. For example, the V@R at 0.1% level is -3.69% for the mixnormal fit versus –2.52% for the normal fit.

Several troubling spots remain. The frequency of zero returns (no price changes) is far larger than could be predicted by any simple model. Also, an examination of the absolute price changes reveals remarkable price discreteness that should not be present for such a highly liquid asset. This price discreteness is even more pronounced for silver (not reported in this study). The frequency of whole dollar changes (and especially exact \$5 dollar changes) is several times the frequency of fractional changes, even though the silver price was between \$200 and \$600. The virtual absence of returns between ±0.4% may be due to this large bid/ask spread phenomenon. The impact of price discreteness on the distributional parameter estimation is still not clear.

Exhibit III.1 GBP/USD exchange rate daily returns, 1986 to 2001

Panel A: The return distribution and V@R metrics, the Normal fit and Mixture-of-Normals fit



Notes: Descriptive statistic of GBP/USD returns (evaluated in 100% terms)**:** Number of obs.: 4011
Mean=0.0025; Standard Deviation=0.642; Skewness=0.219; Excess Kurtosis=3.66.

1) The GREEN curve (bold full, more peaked) corresponds to the Mixture-of-Normals (Max-Likelihood) fit:
$r_{gbp/usd} \sim 0.373\ N[\ \mu_1 = 0.012,\ \sigma_1 = 0.942\ ]\ +\ 0.627\ N[\ \mu_2 = -0.0038,\ \sigma_2 = 0.358\ ]$.

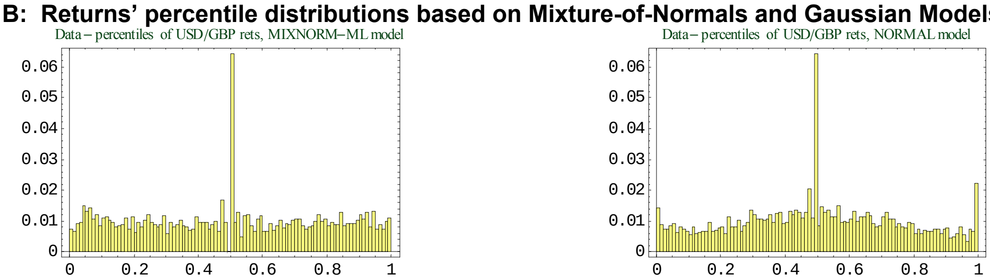
2) The BLUE curve (thin) corresponds to the Mixture-of-Normals (Method-of-Moments) fit:
$r_{gbp/usd} \sim 0.132\ N[\ \mu_1 = 0.113,\ \sigma_1 = 1.25\ ]\ +\ 0.868\ N[\ \mu_2 = -0.015,\ \sigma_2 = 0.483\ ]$.

3) The RED curve (bold dashed) corresponds to the Gaussian fit, e.g., assuming unconditional normality:
$r_{gbp/usd} \sim N[\ \mu = 0.0025,\ \sigma = 0.642]$.

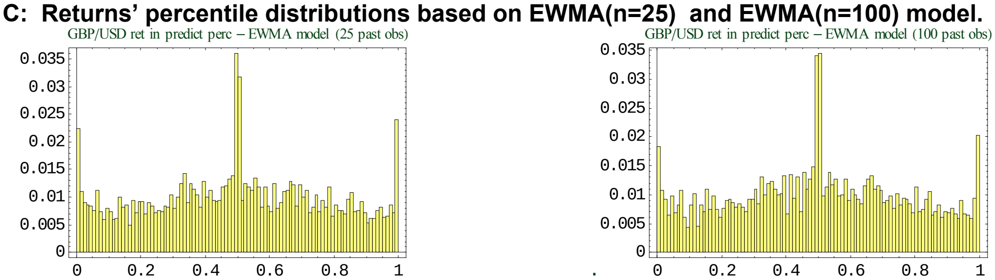
V@R metric *Confid. Level*	0.1%	0.5%	1.0%	2.5%	5.0%	10%
MIXTURE-ML fit : $\Phi^{-1}(\ r_{A,MixN}\ )$ Log-likelihood = –3641	-2.61 stderr=0.13	-2.07 stderr=0.12	-1.80 std.err=0.12	-1.40	-1.04	-0.70
NORMAL fit : $\Phi^{-1}(\ r_{A,normal}\ )$ Log-likelihood = –3916	-1.98 std.err=0.024	-1.65 std.err=0.023	-1.49 std.err=0.020	-1.26	-1.05	-0.82

The ratio of models' likelihoods = $e^{-3641} / e^{-3916} > 10^{119}$.

Panel B: Returns' percentile distributions based on Mixture-of-Normals and Gaussian Models



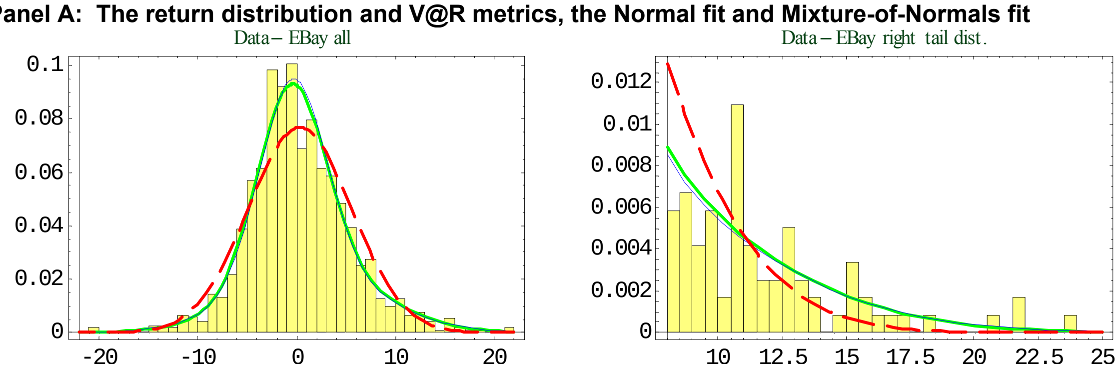
Panel C: Returns' percentile distributions based on EWMA(n=25) and EWMA(n=100) model.



GOODNESS-OF-FIT tests	X^2_{PF} statistic	99% critical value of χ^2_{m-q-1} *dist*	model decision
MIXTURE-ML fit :	12.74	12.83	*accept*
NORMAL fit :	98.06	18.54	reject
EWMA (n=25) fit	149.2	18.54	reject
EWMA (n=100) fit	85.9	18.54	reject

Exhibit III.2 E-Bay Inc. equity daily returns, 1996 to 2000

Panel A: The return distribution and V@R metrics, the Normal fit and Mixture-of-Normals fit



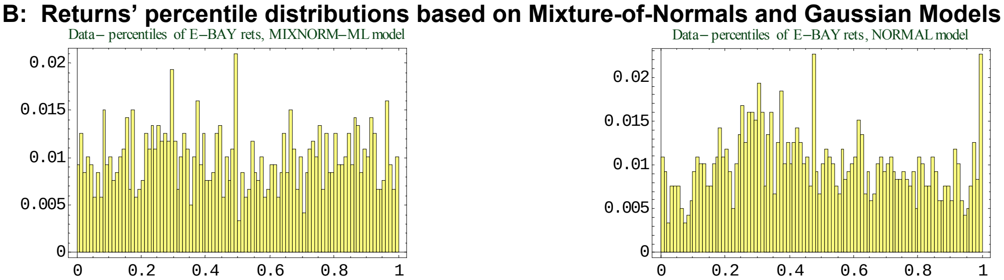
Notes: **Descriptive statistic of E-Bay's returns** (evaluated in 100% terms); Number of obs.: 1190
Mean=0.334; Standard Deviation=5.19; Skewness=0.401; Excess Kurtosis=1.63.

1) The GREEN curve (bold full, more peaked) corresponds to the Mixture-of-Normals (Max-Likelihood) fit:
$r_{E\text{-}Bay} \sim 0.382\ N[\ \mu_1 = 1.71,\ \sigma_1 = 6.98\ ]\ +\ 0.618\ N[\ \mu_2 = -0.511,\ \sigma_2 = 3.41\ ]$.
2) The BLUE curve (thin) corresponds to the Mixture-of-Normals (Method-of-Moments) fit:
$r_{E\text{-}Bay} \sim 0.369\ N[\ \mu_1 = 1.58,\ \sigma_1 = 7.17\ ]\ +\ 0.631\ N[\ \mu_2 = -0.391,\ \sigma_2 = 3.35\ ]$.
3) The RED curve (bold dashed) corresponds to the Gaussian fit, e.g., assuming unconditional normality:
$r_{E\text{-}Bay} \sim N[\ \mu = 0.334,\ \sigma = 5.19\ ]$.

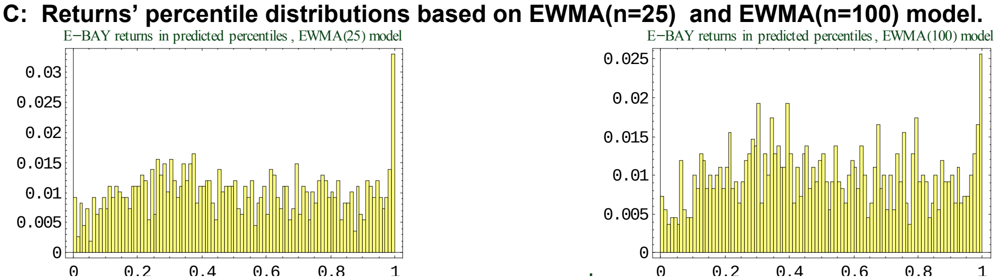
V@R metric *Confid. Level*	0.1%	0.5%	1.0%	2.5%	5.0%	10%
MIXTURE-ML fit : $\Phi^{-1}(\ r_{A,MixN}\ )$ Log-likelihood = –3610	-17.8 stderr=1.26	-13.8 stderr=0.88	-11.9 std.err=0.70	-9.30	-7.40	-5.55
NORMAL fit : $\Phi^{-1}(\ r_{A,normal}\ )$ Log-likelihood = –3651	-15.7 std.err=0.36	-13.0 std.err=0.31	-11.7 std.err=0.30	-9.84	-8.21	-6.32

The ratio of models' likelihoods = $e^{-3610} / e^{-3651} > 10^{17}$.

Panel B: Returns' percentile distributions based on Mixture-of-Normals and Gaussian Models



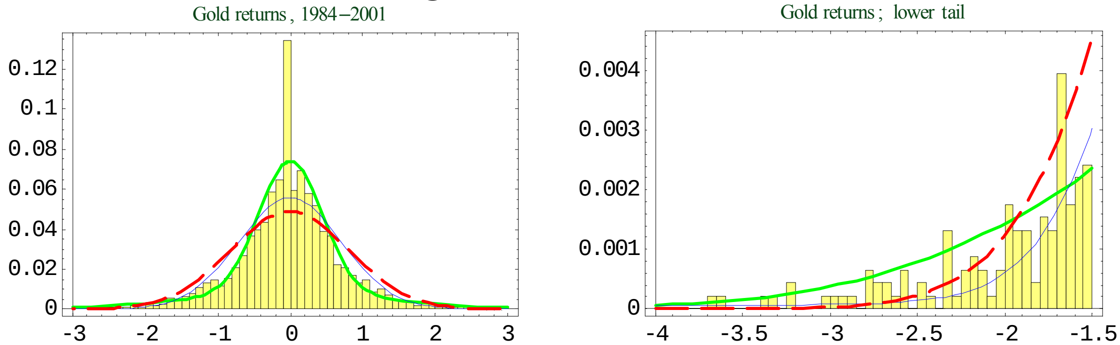
Panel C: Returns' percentile distributions based on EWMA(n=25) and EWMA(n=100) model.



GOODNESS-OF-FIT tests	X^2_{PF} statistic	99% critical value of $\chi^2_{m\text{-}q\text{-}1}$ *distribution*	model decision
MIXTURE-ML fit :	6.76	11.34	*accept*
NORMAL fit :	27.1	16.81	reject
EWMA (n=25) fit	69.5	16.81	reject
EWMA (n=100) fit	42.8	16.81	reject

Exhibit III.3 Gold daily returns, 1984 to 2001

Panel A: The return distribution and V@R metrics, the Normal fit and Mixture-of-Normals fit



Notes: **Descriptive statistic of Gold returns** (evaluated in 100% terms); Number of obs.: 4545
Mean= –0.00454; Standard Deviation=0.815; Skewness=0.0420; Excess Kurtosis=10.31.

1) The GREEN curve (bold full, more peaked) corresponds to the Mixture-of-Normals (Max-Likelihood) fit:
$r_{Gold} \sim 0.283\ N[\ \mu_1 = -0.011,\ \sigma_1 = 1.36\ ]\ +\ 0.717\ N[\ \mu_2 = -0.00064,\ \sigma_2 = 0.435\ ]$.

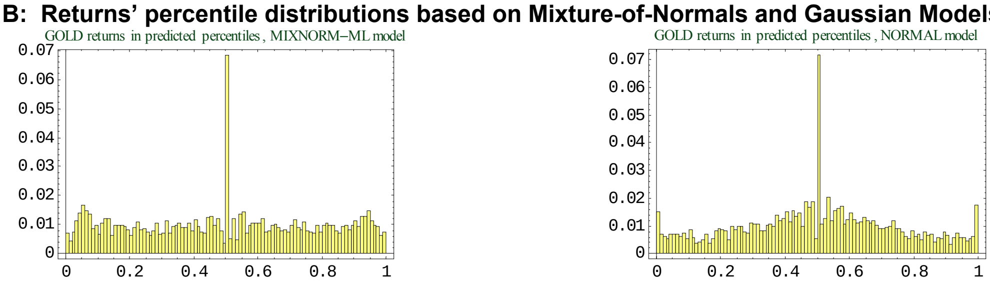
2) The BLUE curve (thin) corresponds to the Mixture-of-Normals (Method-of-Moments) fit:
$r_{Gold} \sim 0.02\ N[\ \mu_1 = 0.044,\ \sigma_1 = 3.12\ ]\ +\ 0.98\ N[\ \mu_2 = -0.0045,\ \sigma_2 = 0.705\ ]$.

3) The RED curve (bold dashed) corresponds to the Gaussian fit, e.g., assuming unconditional normality:
$r_{Gold} \sim N[\ \mu = -0.00454,\ \sigma = 0.815]$.

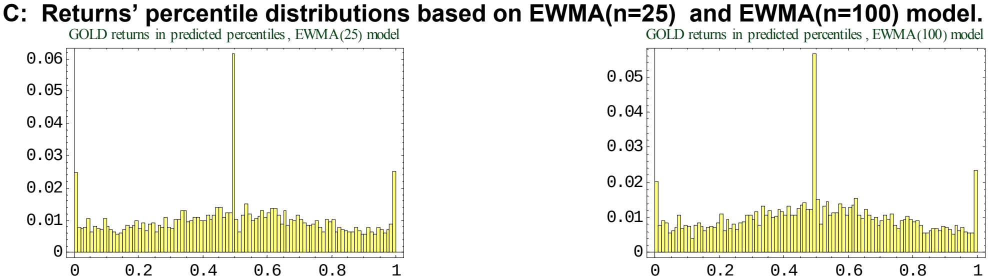
V@R metric *Confid. Level*	0.1%	0.5%	1.0%	2.5%	5.0%	10%
MIXTURE-ML fit : $\Phi^{-1}(\ r_{A,MixN}\ )$ Log-likelihood = –4952	-3.69 stderr=0.24	-2.88 stderr=0.23	-2.48 std.err=0.22	-1.86	-1.30	-0.819
NORMAL fit : $\Phi^{-1}(\ r_{A,normal}\ )$ Log-likelihood = –5518	-2.52 std.err=0.037	-2.10 std.err=0.023	-1.90 std.err=0.022	-1.60	-1.34	-1.05

The ratio of models' likelihoods = $e^{-4952} / e^{-5518} > 10^{245}$.

Panel B: Returns' percentile distributions based on Mixture-of-Normals and Gaussian Models



Panel C: Returns' percentile distributions based on EWMA(n=25) and EWMA(n=100) model.



GOODNESS-OF-FIT tests	X^2_{PF} statistic	99% critical value of $\chi^2_{m\text{-}q\text{-}1}$ *distribution*	model decision
MIXTURE-ML fit :	32.9	11.34	*reject*
NORMAL fit :	85.9	16.81	reject
EWMA (n=25) fit	219.2	16.81	reject
EWMA (n=100) fit	155.1	16.81	reject

III.5 Conclusions

The mixture-of-normals model is shown to fit empirical distributions significantly better than the alternative models that assume conditional (EWMA) or unconditional normality. The Pearson-Fisher *goodness-of-fit* tests forcefully show that daily return dynamics cannot be captured by the normality assumption. The empirical samples examined in this study contain extremely large returns of either sign that exceed the normality models' predicted frequencies by a factor of 2 to 3.

One important caveat is in order. It is very likely that any theoretical model could be rejected by large empirical samples if a goodness-of-fit test with a large number of classes is used. This is true even for the mixture-of-normals model, though it is greatly preferable to the alternative normality models. A mixture-of-normals model is shown to fit empirical samples relatively well and especially well in the "tails".

Several unresolved issues remain. The frequency of zero returns is substantially larger than can be justified by any simple model. Obviously, more complex models are needed to explain this special feature of price dynamics. One possibility is to model the zero returns with Dirac's delta function (a zero frequency "dummy"), commonly used in physics, though this would require an additional (noisy) parameter estimate.

Another potentially troubling feature of short-term returns is the impact of price discreteness. It is apparent in gold and especially silver price series. Price changes of $1 multiples (especially $1 and $5) are much more frequent than any fractional change, even though the silver price oscillated between $200 and $600. Bid-ask spread may distort estimates of model parameters. Distributions relying on estimators obtained using very short (e.g., intraday) returns should be viewed with a healthy dose of skepticism.

The Value-at-Risk metrics at confidence levels below one percent are significantly (and economically substantially) larger if the V@R methodology assumes the mixture-of-normals model rather than the Gaussian model. Return distributions are of great interest to risk managers and financial companies' regulators, as well as option traders. Paraphrasing Robert Merton, we need to be careful in selecting our return dynamics models and assess their limitations, or else be surprised by the next "six sigma" event.

Chapter IV
FIRM VALUE AND VALUE-AT-RISK

IV.1 Introduction

THIS CHAPTER provides an overview of institutional risk management practices and investigates the characteristics of financial institutions using Value-at-Risk approach to risk management (V@R-RM). The main focus is on documenting economic benefits (or value) of the use of V@R methodologies in the risk management of financial institutions.

The extant literature provides numerous valid reasons for corporate management of idiosyncratic risks (Smith and Stulz, 1985; Bessembinder, 1991; Froot, Scharfstein, Stein, 1993, 1998; DeMarzo and Duffie, 1995; among others). It has also been documented that firms generally behave in ways consistent with these theoretical predictions (Nance, Smith and Smithson, 1993; Mian, 1996; Géczy, Minton and Schrand, 1997; Haushalter, 2000). However, we still do not know the answer to the positive question of whether the risk management practices increase firm value. Allayannis and Weston (2001) provide the first evidence along these lines by examining the impact of the firms' use of foreign currency derivatives on equity value. Their sample includes 720 U.S. non-financial firms between 1990-1995. The results confirm that markets put a premium of about 5% of firm value on firms that hedge their exposure to exchange rates.[44]

This chapter complements Allayannis and Weston's study in another dimension. Given the sound theoretical reasons for corporate hedging and risk management, we would expect that markets place a higher value on firms that engage in sophisticated risk

---

[44] Allayanis and Weston (RFS, 2001) are the first to provide *direct* evidence that hedging increases firm value. Other studies such as Nance, Smith and Smithson (1993) only document that corporations behave in ways consistent with the theoretical predictions.

management techniques such as Value-at-Risk.[45] However, a recent analysis of Basak and Shapiro (2001) finds that the V@R risk managers may optimally choose a larger exposure to risky assets than non-risk managers and consequently incur larger losses when losses occur.[46] Their general equilibrium analysis reveals that the presence of V@R managers may amplify the stock-market volatility at times of down markets and attenuate the volatility in up markets. One of the features of their analysis is that although the probability of a loss is fixed, when a large loss occurs, it is larger than when not engaging in V@R risk management. The final effect on the firm value is therefore unknown and a subject for empirical analysis.

The next section provides a brief discussion of institutional details and benefits of using V@R risk management techniques. Section 3 describes data and methodology. Section 4 presents results of univariate and multivariate tests of incremental effect of V@R use, while section 5 describes the results of a probit analysis of the determinants of the Value-at-Risk use by sample firms. The final section offers a summary and conclusions.

IV.2 Risk Management and Financial Institutions

Financial institutions use many different financial risk management techniques.[47] Recent years have witnessed an increase in popularity of the V@R approach to financial risk management.[48] The V@R methods are very attractive and intuitive because the metric is able to describe risks in a form of a single profit/loss number. From a theoretical standpoint, V@R can be viewed as a natural extension of classic portfolio theory of Markowitz (1952, 1959) and subsequent asset pricing models. As described in CHAPTER ONE, V@R is the maximum amount of money the firm can lose over some

---

[45] See THE FIRST CHAPTER for a discussion of theoretical reasons for corporate hedging.
[46] They provide only anecdotal evidence in support of their theory by noting an increased volatility in August 1998, and subsequent problems related to the demise of LTCM.
[47] Other techniques include RAROC (Risk-Adjusted Return On Capital) introduced by Bankers Trust, duration matching of assets and liabilities, cash-flow netting rules, and use of derivatives to manage individual trades.
[48] See Jorion (2001) and Dowd (1998) for an in-depth description of V@R, including estimation procedures, methodologies, history, risk management applications and case studies.

specified period of time with a certain degree of confidence. V@R techniques enable managers to better evaluate risks of certain operations and corporate capital needs. The attractions of Value-at-Risk measures can be summarized as follows: [49]

1. V@R provides managers with a universal (and better?) description of risks, leading to more informed and (probably) superior decisions.[50]
2. Managers are able to establish robust control systems and detect fraud or human error. Using the V@R approach can prevent risk management disasters similar to those of Barings, and Orange County.[51]
3. V@R provides a consistent, integrated treatment of risks across and within institutions, leading to a greater transparency for shareholders and investors of mutual funds, pension funds and trusts.
4. Traders may have incentives to increase risks of their positions if their compensation contracts are imperfect. For example, traders receiving a fixed amount of their trading profits without regard to the risk of their trading positions are effectively given a call option on a portion of their firm's profits. This type of compensation contracts creates a moral hazard problem, as traders have incentives to increase the risk (volatility) of their positions. V@R metrics provides a useful benchmark for improving compensation contracts of traders and managers by taking into account the risks they take.
5. Firms are better able to allocate capital to individual business units and estimate firm-wide capital needs. The approach enables firms to respond to regulations, particularly the capital adequacy regulations that financial institutions face. The costs of regulatory compliance may be lowered by decreasing risks rather than by increasing capital.

---

[49] Jorion (2000, page *xxiii*) and Dowd (1998, page 22) discuss these potential benefits.
[50] V@R metrics and methodologies extend in some respect classic portfolio theory as they describe portfolio risks not only in terms of volatility, but often in terms of the complete probability distribution of portfolio returns So-called *greeks* in derivatives trading, duration, beta, standard deviation, among others, are examples of risk measures used in different settings.
[51] The failures of Orange County and Barings Bank, and losses of Procter&Gamble are generally blamed on lack of risk management systems. See Jorion (2001) among others.

As a result, V@R-RM is being adopted around the world by various market participants including:

1. *Financial Institutions and Asset Managers.* Banks with large trading portfolios have been the pioneers of risk management. Institutions dealing with numerous sources of financial risk and complex instruments are implementing centralized risk management systems. Institutional investors can turn to V@R to report financial risks
2. *Regulators.* The Basel Committee on Banking Supervision, the U.S. Federal Reserve, and the U.S. Securities and Exchange Commission, as well as regulators in the European Union have converged on V@R as a benchmark for risk and regulatory compliance purposes.[52]
3. *Non-Financial Corporations.* As discussed earlier, hedging and risk management can increase shareholder value. For example, multinational corporations face cash-flows in multiple currencies. They can use "Cash-flow at Risk" analysis to predict likelihood of a critical shortfall of funds.

The Goldman Sachs case study presented in CHAPTER ONE, SECTION FIVE provides an example illustrating the risk management problems faced by trading firms in establishing position limits and compensation rules. Larry Becerra was one of the most successful traders of the firm. In 1993, he was reported to have made more than $80 million for the firm that earned a record $2.6 billion in total pretax profits that year. In December 1993, Becerra invested heavily in Italian bonds. After a sharp rise in the Italian bond market, his positions alone made the firm $80 million in a single month. Several signs pointed to the rising overconfidence that seemed to extend to the upper levels of management. One partner reportedly met with a member of the Federal Reserve Board and claimed: "We are too big to fail. A $100 million loss, a $50 million loss, it means nothing. We're too big now, [the Fed] won't let us fail." It seems that hubris had taken hold.[53] Since the firm's proprietary traders felt confident that they could handle a position of almost any

---

[52] Basel Committee, 1995; Securities and Exchange Commission, 1997.
[53] "Hubris" has often been cited as an explanation in the mergers and acquisitions (M&A) literature, (see Roll, JB 1986)

size, and large concentrated bets were the focus, Becerra was allowed to keep a giant position. By February, he lost more than $100 million, just thirty-one days after the gain of $80 million.

IV.2.1 The Hypothesis, Economic Benefits of V@R Approach

Initially, V@R was developed as a practical gauge of financial risk. Recently, it has transformed into a benchmark to compare risk across business units, and even decide on the amount of equity capital necessary to support a trading activity.[54] However, LTCM's failure has been ascribed to its use of Value-at-Risk. V@R's usage may expose current methods used to set capital-adequacy requirements as inadequate[55]. Commercial banks are allowed, under the Basel framework, to use their internal models as the basis for determining the equity capital necessary to support market risk exposures of their trading operations (Basel Committee, 1995, 1997).

This paper investigates a sample of the largest commercial banks (public bank holding companies, BHC) using V@R as their risk management tool to assess economic benefits of the V@R approach. We hypothesize that the 'user' banks have superior valuations (higher P/E ratios, M/B ratios) relative to a benchmark group of 'non-users', controlling for various firm characteristics, as described below.

Consistent with the theoretical predictions, firms that use risk management tools should achieve higher valuations (P/E ratios) because of larger future cash flows, rather than the lower risks they take. From a simple Gordon Growth Model (GGM) follows:

$$L \equiv P/E \approx \frac{1}{r-g} \quad , \qquad 4.2.1$$

where *P/E* denotes the firm's Price/Earnings ratio, *r* is the cost of capital, and *g* is the expected growth rate of earnings (or dividends). Ceteris paribus, a positive change in

[54] Matten (2000)

[55] Jorion (2000) describes the case of Long-Term Capital Management (LTCM) and uses V@R approach to quantify LTCM's (ex-ante) risks to show that principals of LTCM severely underestimated risks of their trading positions.

corporate earnings growth prospects will have a positive impact on it's P/E ratio since $\partial L/\partial g > 0$.

For firms using risk management tools, the higher cash flow growth is based on various theoretical predictions such as lower future tax liabilities, avoidance of financial distress costs, and new capital acquisition transaction costs. Consequently, a better understanding of the risks that firms take allows them to avoid costly states in which they are unable to exercise their valuable growth options.

It would be incorrect to conjecture that firms achieve higher valuations by avoiding risks per se. Since most investors are well diversified, firms *cannot* decrease their cost of capital *r* (and thus increase valuations) by removing idiosyncratic risks. On the other hand, while firms can decrease their cost of capital by removing some systematic risk, in an efficient market this risk is fairly priced and its impact on firm value is inconsequential.

IV.3 Sample Selection and Methods

The sample consists of the100 largest publicly traded bank holding companies (BHC).[56] The time period examined covers years 1993-1998. Table IV.1 provides a basic overview of the sample.

The information on the banks' use of V@R risk management was collected from annual 10-K disclosures, which are available online via the Securities and Exchange Commission's EDGAR database of public firms' disclosure filings. Each report was searched for keywords such as "value-at-risk", "VaR", "earnings-at-risk", and "risk management". From the disclosures reviewed for this study, it is evident that most companies engage in some financial risk management. The extent of disclosure varies widely and indicates that techniques range greatly from a relatively simple interest-rate risk management (GAP and duration matching of assets and liabilities) through scenario

[56] The sample was obtained by selecting 100 largest publicly-traded bank holding companies as of 1998. This sample is held constant over the sample period 1993-1998. Since there were only a few mergers (and no defaults) during this sample period among the largest BHCs, the group of 100 largest BHCs in 1998 is essentially the same one as in 1993.

analysis to more sophisticated methods such as Value-at-Risk (sometimes also referred to as Earnings-at-Risk).

Based on the disclosure in annual reports, each BHC is placed either in a category of users or nonusers of V@R-RM. The test is repeated for all years. The working assumption is that managements voluntarily reporting V@R in their annual reports are more likely to use V@R as an important decision-making tool.

In January 1997, the Securities and Exchange Commission (SEC) issued a ruling requiring public firms to disclose quantitative information about the risk of financial instruments in reports filed with the SEC. This rule applies to all filings for fiscal years after June 15, 1998. Firms are asked to disclose market risks using one of three alternatives: 1) tabular presentation of cash flows; 2) sensitivity/scenario analysis; or 3) Value-at-Risk measures. The sample used throughout this study ends in 1998, as it would be impossible to distinguish between voluntary V@R disclosure reports that are based on a genuine use of V@R for decision-making purposes, and reports designed to satisfy only the regulatory disclosure requirements.

The BHC accounting-level data is obtained from the Sheshunoff database. The data-items used include a) total assets; b) the notional amount of derivatives; c) trading assets; d) total book equity; f) net income. The COMPUSTAT database is used to collect data on the end-of-period earnings, book values of equity, share prices and total shares outstanding. These are used to construct the relative-value measures, such as price-to-earnings (P/E) and market-to-book (M/B) ratios.[57]

IV.4 Empirical Results

Table IV.1 provides a basic overview of the sample. The use of V@R-RM has increased over the years. There are 23 companies reporting V@R-RM in 1998, compared to 2 in 1993 (J.P. Morgan and Chase). There is a wide variation in total assets

[57] The S&P's Compustat database has several files. The market value and book value data is obtained from the industrial file, the end-of-year price and earnings data is obtained from two alternative files: the annual industrial file and dividend/earnings file. There are two alternative P/E measures: the first is based on total annual earnings, the second is based on the last quarter's earnings. Both alternatives return similar results.

among the top 100 BHCs. This is also true about the extent of their derivatives and asset trading activities. The distributions are skewed to the right, as there is tendency of the largest companies to have more trading assets or total derivatives outstanding (usually for trading purposes, but reported separately from other assets held for trading).

Table IV.1 Descriptive Sample Statistics

Panel A.

YEAR	1998	1997	1996	1995	1994	1993
Sample size (no. of BHCs)	100	100	100	100	100	100
Number of BHCs using V@R	23	17	11	6	5	2
Average P/E ratio	20.69	21.16	15.18	14.01	11.00	11.82
Average M/B ratio	2.53	2.25	1.30	0.97	0.77	0.80

Panel B.

YEAR	1998	1997	1996	1995	1994	1993
Total Assets ('000 USD)						
-- Average	45,511,353	32,351,589	27,976,678	23,694,370	21,339,064	18,963,801
-- Median	10,366,580	7,713,245	6,577,847	6,180,686	5,357,955	4,691,900
-- Max	668,641,000	365,521,000	336,099,000	256,853,000	250,489,000	216,574,000
-- Min	2,123,006	948,848	358,752	303,442	728,136	570,433
Total Equity to Total Assets						
-- Average	8.09%	8.24%	8.35%	8.38%	8.03%	8.01%
-- Median	7.84%	8.04%	8.24%	8.18%	7.75%	7.71%
-- Max	13.14%	13.05%	12.70%	14.45%	21.86%	11.80%
-- Min	3.53%	4.07%	4.35%	4.79%	4.85%	4.87%
Trading Assets to Total Assets						
-- Average	1.67%	1.66%	1.60%	1.67%	1.68%	1.51%
-- Median	0.02%	0.01%	0.00%	0.00%	0.00%	0.00%
-- Max	43.63%	42.67%	40.98%	46.05%	48.98%	52.43%
-- Min	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total Derivatives to Total Assets						
-- Average	118.66%	114.13%	91.64%	87.29%	66.11%	56.17%
-- Median	6.28%	6.44%	5.23%	4.26%	3.92%	4.92%
-- Max	3295.37%	2503.15%	2085.85%	1857.29%	1327.99%	1397.77%
-- Min	0.11%	0.00%	0.08%	0.03%	0.01%	0.03%

Note: Total Derivatives denotes the total notional amount of derivates.

IV.4.1 An Analysis of V@R Users and Non-users

Table IV.2 provides a univariate comparison of various characteristics of companies that use V@R-RM techniques and characteristics of companies that don't use these techniques. Panel A reveals that V@R users are significantly larger than non-users (both in economic and statistical sense). This result may be explained by economies of scale, as there may be large fixed costs to establishing an extensive risk management program with large data processing capabilities. Also, as we would expect from the prior discussion, V@R-RM users have larger derivatives positions and their trading assets comprise a larger proportion of their total assets. If V@R metrics provides an efficient monitoring device for managers, we would expect its more extensive use among companies with relatively larger leveraged derivatives positions, for which other risk measures may be hard to interpret or which may not provide sufficient statistic for decision makers.

Panel B contrasts the P/E and M/B ratios of users and non-users. Contrary to our predictions, non-users have higher valuations than users, as their P/E ratios are consistently higher in all years, although only insignificantly so in 2 out of the 6 sample period's years.[58] This result is puzzling. Why would investors put a lower relative value on companies that use V@R risk management systems? This result may be linked to the other firm characteristics, such as size, leverage, and asset composition. The next section provides the results of a multivariate analysis that may shed more light on this puzzle. On the other hand, this result may be consistent with Basak and Shapiro (2001), who show that managers using V@R techniques to constrain their asset selection may incur relatively larger losses if losses exceeding V@R metrics occur. The other alternative measure of firm value, market to book (M/B), is not significantly different for the two sub-samples.

[58] The other commonly used measure of a distribution's centrality is the *median*. However, the use of medians instead of arithmetic averages does not change the observation that the V@R user's P/E and M/B ratios are smaller than the ratios of companies not using V@R-RM methodologies.

Table IV.2 Comparison of V@R-RM users vs. nonusers

Panel A.

	YEAR	1998	1997	1996	1995	1994	1993**
Total Assets (Mean, mill. USD)	**V@R users**	147,263	123,061	132,201	112,557	113,313	141,888
	V@R Non-users	15,118	13,314	16,646	17,835	16,229	16,292
	p-value for difference	0.22%	0.09%	0.5%	3.3%	3.9%	1.8%
Equity to Assets (Mean)	**V@R users**	7.31%	7.11%	7.19%	7.46%	6.67%	7.41%
	V@R Non-users	8.32%	8.48%	8.50%	8.44%	8.10%	8.02%
	p-value for difference	0.85%	0.06%	2.4%	10.3%	5.6%	<0.01%
Trading to Total Assets (Mean)	**V@R users**	6.61%	8.53%	11.88%	12.31%	12.77%	19.34%
	V@R Non-users	0.20%	0.21%	0.27%	0.97%	1.07%	1.13%
	p-value for difference	1.31%	2.19%	3.1%	10.8%	15.4%	35.9%
Derivatives to Assets (Mean)	**V@R users**	478%	603%	699%	776%	557%	985%
	V@R Non-users	11.3%	11.5%	13.9%	41.9%	38.8%	36.0%
	p-value for difference	2.22%	1.4%	1.26%	8.6%	13.9%	11.5%

Panel B.

	YEAR .	1998	1997	1996	1995	1994	1993**
P / E ratio (Mean)	**V@R users**	19.6	18.5	14.7	11.6	8.52	7.66
	V@R Non-users	21.0	21.7	15.2	14.2	11.1	11.9
	p-value for difference	23.3%	0.47%	58%	2.9%	1.1%	1.18%
M / B ratio (Mean)	**V@R users**	2.78	2.39	1.62	0.84	0.67	0.92
	V@R Non-users	2.47	2.23	1.26	0.98	0.77	0.80
	p-value for difference	23.8%	55.3%	20.1%	44.2%	53.5%	86.7%

Notes:
Median values, as an alternative measure of centrality of variable distribution, are very similar to the reported averages and the relative comparisons of the P/E and M/B ratios between the V@R “users” and “non-users” using medians remain the same.
** Though reported, the comparison may not be appropriate for the year 1993, as there were only two V@R users: J.P. Morgan and Chase Manhattan.
“Trading to Total Assets” denotes the ratio of the trading assets to total assets.
“Derivatives to Assets” denotes the ratio of the total notional amount of derivates to total assets.
P-values based on the Satterthwaite Method’s t-test for equality of means of distributions with unequal variances, except where the F-test for equality of variances can’t be rejected.

IV.4.2 Controlling for other firm characteristics

In order to make any conclusions about the impact of corporate risk management policies, it is necessary to control for the effect of other firm characteristics that can have an impact on the firm's value (e.g., on P/E or M/B ratio). It is possible that the V@R premium/discount can be explained by factors previously linked to value premia and factors that theory suggests affect firm value (Fama and French, 1992; Barber and Lyon, 1997; Morck and Yeung, 1991; Lang and Stulz, 1994; Servaes, 1996; Yermack, 1996; Hannan and Hanweck, 1988). The following multivariate tests check to see if the results are robust by controlling for size, leverage, trading activity and market making activity in derivatives. Most of the studies listed above find that P/E and M/B ratios vary depending on the firm investment opportunity set. In this study, it is assumed that the investment opportunity set is the same for all sample firms, as they are all in the same industry (bank-holding-companies) and share many characteristics. The various control variables used in multivariate tests are described below:

1) Size. Various asset-pricing tests reveal that size is related to firm returns (Banz, 1981; Fama and French, 1992; Barber and Lyon, 1997; among others). Though it is unclear whether the size characteristic proxies for a missing risk factor, it is evident that several presumed anomalies may be "explained" by an observed correlation between stock returns and size. Also, due to the economies of scale and potentially substantial fixed start-up costs, relatively larger firms may be more likely to engage in sophisticated risk management practices such as V@R. The size variable used is the natural logarithm of total book value of assets (LNTOTAST).

2) Leverage. A firm's capital structure may affect its value (see Harris and Raviv, 1991, for a survey). This is especially important for financial institutions such as banks that have very high levels of leverage and thus even small differences among these firms may lead to substantial deviations in measures such as P/E ratio (see Barber and Lyon, 1997; Hannan and Hanweck, 1988). The leverage variable used is the ratio of total book equity to total assets (EQTA) expressed in percentage terms.

3) Trading activity, and derivatives market making. Large financial firms are active in extensive asset trading, which includes derivatives. The trading portfolio may be riskier than the other assets. The proportion of assets invested in trading assets (TRTAST) may be related to systematic risks of future bank profits and thus affect the price that investors are willing to pay for bank equity shares. Also, derivatives market making has received lots of negative publicity and lawsuits from large clients whose bets went sour during the sample period.[59] Resulting settlements forced many banks (notably Bankers Trust) to refund derivatives end-users (Procter and Gamble) a portion of their losses. The threat of further ex-post lawsuits from other bank-clients that happen to lose on their derivatives trades may induce markets to revise the price downwards and thus relevant measures of value such as P/E and M/B ratios. The price discount may represent the market's estimate of present value of future liabilities due to derivatives-trade related lawsuits. The extent of derivatives trading may serve as a useful proxy for this effect. Though the market values of derivative portfolios are unknown, the notional amounts of underlying assets should be monotonically related (DERIVTA).

4) Time: Since V@R-RM has gained its widespread acceptance only slowly over the years, and P/E ratios tend to change over time, it is necessary to control for annual fixed effects by using year dummies (YRxx dummy).

Table IV.3 provides results of the least squares regression analyses. The dependent variable is the P/E ratio. The variable of interest is VARDUM (V@R dummy variable).[60] Surprisingly, even after controlling for firm size and other characteristics, the V@R dummy parameter estimate is significantly negative, indicating that markets appear to put a discount on BHCs that employ V@R-RM techniques. The effect is somewhat smaller after including the derivatives variable (DERIVTA). A discount on V@R users is partially a result of the high correlation between the use of V@R and derivatives. This negative relationship between the firm value and the derivatives market making is itself a noteworthy finding.

The results for the M/B ratio as an alternative dependent variable are not reported, because the estimated coefficients of independent variables are mostly not significant.

---

[59] See Jorion (2000) for an overview of derivatives scandals involving large derivatives trading firms such as Bankers Trust and their clients – Procter and Gamble, Gibson Greetings, Orange County.
[60] VARDUM takes on value of 1 if a company uses V@R-RM, and 0 otherwise.

Table IV.3 Multivariate least-squares regression analysis for P/E ratio as the dependent variable

	Dependent Variable is P/E ratio					
	Model 1		**Model 2**		**Model 3**	
Independent Variables	Coefficient estimate	p- value	Coefficient estimate	p- value	Coefficient estimate	p- value
Intercept	23.6	<0.01%	23.2	<0.01%	22.7	<0.01%
VARDUM – V@R dummy 0=non-users, 1= users	-1.57	2.8%	-1.44	5.1%	-1.24	10.2%
LNTOTAST – log of total assets total assets in thousands USD	0.001	99.5%	0.04	81.2%	0.08	65.1%
EQTA – equity to assets expressed in 100% terms	0.084	50.6%	0.055	67.7%	0.041	75.2%
ROE – return on equity expressed in 100% terms	-0.22	<0.01	-0.22	<0.01%	-0.234	<0.01%
TRTAST – trading-to-total assets expressed in 100% terms			-0.029	40.7%		
DERIVTA – derivatives-to-assets expressed in 100% terms					-0.001	15%
YR97dum – year 1997 dummy: 1 if 1997, 0 otherwise	0.62	31%	0.65	29.4%	0.68	27.9%
YR96dum	-5.6	<0.01	-5.6	<0.01%	-5.6	<0.01%
YR95dum	-7.4	<0.01	-7.4	<0.01%	-7.4	<0.01%
YR94dum	-9.9	<0.01	-9.9	<0.01%	-9.9	<0.01%
YR93dum	-9.1	<0.01	-9.0	<0.01%	-9.0	<0.01%
Adjusted R^2 – value	47.9		47.9		48.0	
Regression F – value (Prob>F)	56.6	<0.01	51.0	<0.01%	51.2	<0.01%

IV.4.3 A Logit Analysis of Determinants of the V@R Use

This section estimates a binary response model of V@R risk management system selection. There are common claims that BHCs do (or don't) use V@R because they have large (or negligible) trading or derivatives portfolios. The earlier univariate analysis suggests that portfolio composition may indeed induce companies to select a more sophisticated system. This is consistent with predictions that V@R provides a tool to monitor trades and risk-taking of traders. Table IV.4 provides the results of a logit model for V@R choice.

The results confirm that larger companies as well as those with larger derivatives positions are more likely to use V@R-RM. Importantly, the results indicate that this is also true for firms with higher leverage, though the coefficients are not significant at the usual confidence levels. On the other hand, the impact of trading assets on the decision to implement V@R techniques is inconclusive as the coefficient estimates are insignificant in most specifications. However, a similar logit analysis for year 1998 only gives a significant coefficient for the trading portfolio variable, which is consistent with predictions.[61]

Table IV.4 A Logit Model for a Discrete Dependent Variable VARDUM

	Dependent Variable is a V@R dummy Ordered response 1= users, 0=non-users					
	Model 1		**Model 2**		**Model 3**	
Independent Variables	Coefficient estimate	p- value	Coefficient estimate	p- value	Coefficient estimate	p- value
Intercept	-26.85	<0.01%	-28.35	<0.01%	-25.8	<0.01%
LNTOTAST – log of total assets, in thousands USD	1.60	<0.01%	1.72	<0.01%	1.56	<0.01%
DERIVTA – notional derivatives to assets expressed in 100% terms	0.0011	11.0%			0.0028	6.71%
EQTA – equity to assets expressed in 100% terms	-0.19	27.6%	-0.27	14.4%	-0.23	19.6%
TRTAST – trading-to-total assets expressed in 100% terms			0.016	51.8%	-0.071	18.9%
YR97dum – year 1997 dummy: 1 if 1997, else 0	-0.37	46.6%	-0.38	45.8%	-0.36	47.2%
YR96dum	-1.11	5.55%	-1.10	5.67%	-1.11	5.62%
YR95dum	-2.36	0.18%	-2.20	0.18%	-2.41	0.20%
YR94dum	-2.47	0.15%	-2.40	0.13%	-2.32	0.26%
YR93dum	-3.66	0.02%	-3.51	0.02%	-3.64	0.05%
Likelihood Ratio: Testing global null H_0: coefficients=0	220.7	<0.01%	218.2	<0.01%	222.6	<0.01%

[61] This result may be just a consequence of impending regulation. Effective January 1, 1998, the regulatory agencies began to incorporate market risk into the risk-based capital guidelines. Any bank or bank holding company whose trading activity exceeds either: (1) 10% or more of its total assets, or (2) $1 billion or greater, must measure its exposure to market risk using its own internal value-at-risk model and hold capital in support of that exposure.

IV.5 Conclusions

Despite the enormous popularity of V@R-RM among practitioners, the results of this study do not indicate that sophisticated risk management systems increase corporate value. This conclusion, however, does not mean that the gains do not exist, merely that a positive effect may (though need not) be obscured by other effects. More importantly, contrary to our predictions, the analysis reveals that investors place a negative value on companies using V@R. The relationship between firms' P/E ratios and V@R use remains negative and significant even after numerous robustness checks.[62]

This negative relationship is puzzling as firms using better risk management systems should make more informed decisions and avoid numerous costs such as costs of financial distress, underinvestment costs, and transaction costs related to raising new capital after an adverse return shock. Moreover, an implementation of sophisticated risk management systems should allow firms to establish better monitoring systems and compensation schedules for their traders and managers. As of now, these gains remain elusive.

An alternative explanation of the negative relationship must be then connected with the firms' cost of capital. It is possible that the V@R dummy proxies for an unobserved source of systematic risk, which is not explained by the other explanatory variables. Basak and Shapiro (RFS, 2001) argue that the use of V@R systems may lead managers to mechanically decrease the predetermined V@R metric at some confidence levels, but consequently (and optimally) increase expected losses when losses exceeding the V@R benchmark occur. This strategy may result in a risk profile that is recognized and discounted by investors, but may not show in traditional measures of risk (or their proxies). If V@R use indeed proxies for higher systematic risk and therefore higher cost of capital (ceteris paribus), then P/E ratios would be correspondingly lower. This fact can easily be seen from the simple GGM valuation model represented by relationship *4.2.1*, since $\partial L / \partial r < 0$.

---

[62] The robustness checks included White's heteroskedasticity consistent standard errors, an analysis of dividend payouts, and other partial data analyses.

Also, it is likely that sophisticated bank holding companies using V@R are the same companies dealing with hedge funds, similar to the famed LTCM hedge fund. Contrary to what their names may suggest, many of the hedge funds are not "hedged" (in a sense that they are *not* market neutral) and tend to incur large losses or even go bankrupt at times of market crises. Fung and Hsieh (1997, 1999, 2000) provide empirical evidence on hedge funds and document that these funds employ opportunistic trading strategies on a leveraged basis. Hedge funds are often in a position to exert substantial market impact via their considerable exposures. Also, the risk of their investments may not be captured by standard linear statistical techniques, because of the funds' non-linear dynamic trading strategies (Fung and Hsieh, 1999, p. 232).[63]

BHCs often either extend a large line of credit to hedge funds or finance their investments. Large losses incurred by hedge funds may in turn result in large losses to their creditors. Crises are the most expensive states of the world and least likely to be insured. Though they are rare, crises represent systematic portfolio events and may result in a sizeable discount in market values of firms affected by them. The results of this study may be a manifestation of the above phenomenon, which in the academic literature is also known as the "peso problem."

---

[63] Fung and Hsieh, (1999, p. 313) also report that the standard deviation of hedge funds returns is less than half the standard deviation of the S&P500 index returns. The author is not aware of any systematic study of the hedge funds' bankruptcy rates and returns to the hedge fund creditors during crises. The recent highly publicized examples include bankrupt Long-Term Capital Management hedge fund and Tiger fund. Hedge funds are usually organized as of-shore entities and are not required to disclose their operations or returns. This fact may limit any study attempting to document the behavior of returns to hedge fund investors or banks dealing with them.

Chapter V

SUMMARY AND CONCUDING REMARKS

"The effects of noise on the world, and on our views of the world, are profound. Noise in the sense of a large number of small events is often a causal factor much more powerful than a small number of large events can be. Noise makes trading in financial markets possible, and thus allows us to observe prices for financial assets. . . . Most generally, noise makes it very difficult to test either practical or academic theories about the way that financial or economic markets work. We are forced to act largely in the dark."

Fisher Black, "Noise," (Journal of Finance, 1986).

V.1 Major findings

MODEL MISSPECIFICATION AND ESTIMATION ERRORS are shown to play an important role in financial risk management. Moreover, they seem to be largely unavoidable. Because our understanding of the true asset price processes is still rather vague, theoretical models of finance have to be applied carefully in light of their underlying assumptions. Practitioners need to assess each model's limitations carefully. THIS CHAPTER provides a summary of major findings as well as their discussion and suggestions for future research.

Model misspecification and estimation errors can lead to substantial deviations from optimal hedging policies and erroneous assessment of investment risk. Various model assumptions and simulations presented here provide an important insight about the impact of the true return dynamics and its potential misrepresentation on the outcome of an investor's hedging decisions and measurement of risk.

Investors with different utility functions may find more informative different risk measures (e.g., variance, V@R metrics). The risk of a minimum-variance hedge is shown to be *not* unambiguously smaller than the risk of an unhedged position. Moreover, different measures of risk – such as portfolio variance and V@R metrics – may lead to contradictory conclusions about the portfolio's risk level. It is shown that if asset returns are normally distributed, only an extreme parameter uncertainty and very low absolute levels of return correlation would lead to an increase in a risk level, no matter what risk measure is used. However, the same conclusions cannot be reached for leptokurtic asset return series. If prices are generated by a process resulting in a return distribution that is a mixture-of-normals, a *Gaussian* model assumption and the minimum-variance hedge methodology may lead to a hedging decision that can increase rather than decrease the risk of an underlying investment position.

The hedging errors and the potential increase in risk is most likely to occur if return distributions exhibit extreme kurtosis and the estimation period is rather short. Short estimation periods may be applied either directly due to the non-stationarity of return series previously reported in finance literature, or indirectly by selecting a model that weights more recent observations more heavily. It is shown that the *effective* estimation periods of certain ARCH-class models such as *Equally Weighted Moving Average* (EWMA) models can be rather short (25 to 100 days for daily return series) even if very long *actual* data series are used in parameter estimation. The short estimation sample introduces extra noise into the parameter estimation that may lead to a hedging decision that increases risk (in terms of variance as well as in terms of V@R).

Uncertainty about the optimal hedging decision and the (change in) risk level can be traced to large standard errors of various parameter estimates such as the correlation coefficient. Model misspecification may lead investors to believe that the correlation coefficient or variance of return series is non-stationary (Merton, 1976). Existence of jumps in asset return series leads to the standard errors of sample correlation coefficient estimates that may be several times larger than standard errors of estimates under the normality of returns. Subsequently, large errors in the estimates of asset cross-correlations can lead to large errors in hedge ratios. Model and estimation errors may thus ultimately lead to a return distribution of hedged position that has undesirable properties, e.g., Value-at-Risk of a hedged position may be substantially larger (at some small but not negligible confidence level) than V@R of an unhedged position. What should an underwriter do to prevent unnecessary model risk and ultimately hedging errors? A simple examination of sample return properties, use of alternative models and preferably models with longer *effective* estimation periods should prove to be highly informative and a positive net present value project.

It is also shown, that V@R may be significantly understated in both economic and statistical sense if an incorrect model is used. The mixture-of-normals model can lead to V@R metrics that dramatically exceed metrics under the normality assumption. For some mixture-of-normals model specifications, the true V@R metric at low (0.1%) confidence levels is a multiple of the V@R metric estimated by erroneously assuming that returns are normally distributed. Moreover, back testing of models at higher confidence levels (e.g., 5% or 10%) may lead to a false sense of security that models used are correctly representing the empirical return dynamics. How serious the problem is can be answered only by examining empirical samples.

CHAPTER THREE investigates characteristics of various asset returns. A mixture-of-normals model is shown to fit empirical distributions significantly better than the alternative models that assume conditional (*EWMA*) or unconditional normality (*GBM*). The Pearson-Fisher *goodness-of-fit* tests forcefully show that daily return dynamics cannot be captured by the normality assumption. The empirical samples examined in this study are shown to contain extremely large returns of either sign, which exceed frequencies predicted by normality models by a factor of 2 to 3. One important caveat is in order. It is very likely that any theoretical model could be rejected by large empirical samples if goodness-of-fit test with a large number of classes is used. This is true even for the mixture-of-normals model, though it is greatly preferable to the alternatives. A mixture-of-normals model is shown to fit empirical samples relatively well overall and especially well in the "tails", the major concern of financial risk managers.

Several unresolved issues remain. The frequency of zero returns (no price changes) is substantially larger than can be justified by any simple model. Obviously, more complex models are needed to explain this special feature of price dynamics. Another potentially troubling feature of short-term returns is an impact of price discreteness. Bid-ask spread may distort estimates of model parameters, and distributions relying on estimators obtained using very short (intraday) returns should be viewed with a healthy dose of skepticism.

The Value-at-Risk metrics estimates at confidence levels below one percent are significantly larger (in a statistical and economic sense) if the V@R methodology assumes the mixture-of-normals model rather than the Gaussian model. Return distributions are of great interest to risk managers, financial companies' regulators, as well as option traders. Paraphrasing Robert Merton, we need to be careful in selecting our return dynamics models and assess their limitations, or else be surprised by the next "six sigma" event.

CHAPTER FOUR examines the potential benefits and value of V@R systems in managing financial institutions' investments and portfolio choices. Despite enormous popularity of V@R-RM among practitioners, the results of this study do not indicate that sophisticated risk management systems increase corporate value. This conclusion however does not mean that the gains do not exist, merely that the positive effect may (though need not) be obscured by other effects. More importantly, contrary to our predictions, the analysis reveals that investors place a negative value on companies using V@R. The relationship between firms' P/E ratios and V@R use remains negative and significant even after numerous robustness checks.

This negative relationship is puzzling as firms using better risk management systems should make more informed decisions and avoid numerous costs such as costs of financial distress, underinvestment costs and transaction costs related to raising new capital after an adverse return shock. As of now, these gain remain elusive.

An alternative explanation of the negative relationship must be then connected to the firms' cost of capital. It is possible that the V@R dummy proxies for an unobserved source of systematic risk, which is not explained by the other explanatory variables. Basak and Shapiro (RFS, 2001) argue that the firms' use of V@R systems may lead to an increase in expected losses when losses exceeding the V@R benchmark occur. This strategy may result in a risk profile that is recognized and discounted by investors but does not show in traditional measures of risk. If V@R use indeed proxies for higher systematic risk and therefore higher cost of capital (ceteris paribus), then P/E ratios would be correspondingly lower.

Several hypotheses may be put forward. Sophisticated banks that use V@R-RM deal with numerous hedge funds. These hedge funds are not "hedged" (in a sense that they are *not* market neutral) and may incur large losses or even go bankrupt at times of market crises. Fung and Hsieh (1997, 1999, 2000) provide empirical evidence on hedge funds and document that these funds employ opportunistic trading strategies on a leveraged basis. Also, the risk of their investments may not be captured by standard linear statistical techniques, because of the funds' non-linear dynamic trading strategies (Fung and Hsieh, 1999, p. 232).

BHCs often either extend a large line of credit to hedge funds or finance their investments. Large losses incurred by hedge funds may in turn result in large losses to their creditors. Crises are the most expensive states of the world and least likely to be insured. Though they are rare, crises represent systematic-risk portfolio events and may result in a sizeable discount in market values of firms affected by them. The results of this study may be a manifestation of this "peso problem" phenomenon.

Financial risk management is a rather new and still evolving discipline. The effects of noise are profound. While we cannot eliminate estimation and model errors, we should be able to control their impact on our hedging and risk management decisions. Future research should address the tradeoffs between simple models and complex models with noisy parameters. The impact of noise and model uncertainty may force us to act largely in the dark. Paraphrasing Fischer Black (1986, p. 530), if my conclusions are not accepted, I will blame it on noise. Congratulations, you have reached the end, what follows is just noise.

BIBLIOGRA PHY

Books :

Brealey, Richard A., and Stewart C. Myers, 2000, *Principles of Corporate Finance*, Irwin/McGraw Hill.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press.

Dowd, Kevin, 1998, *Beyond Value at Risk : The New Science of Risk Management*, John Wiley Publ.

Endlich, Lisa, 2000, *Goldman Sachs - The Culture of Success*, Touchstone publ., New York, NY 10020.

Feller, W., 1968, 1971, *An Introduction to the Theory of Probability and its Applications*, Volume 1 (3rd edition) and Volume 2 (2nd edition), Wiley, New York.

Gouriéroux, Christian, 1997, *ARCH Models and Financial Applications*, Springer Series in Statistics, Springer-Verlag, New York.

Greene, William H., 1996, *Econometric Analysis*, Prentice Hall.

Grinblatt, Mark, and Sheridan Titman, 1998, *Financial Markets and Corporate Strategy*, Irwin/McGraw-Hill Companies, Inc.

Group of Thirty, 1993, *Derivatives: practices and principles*, New York: Group of Thirty.

Hull, John C., 1997, *Options, Futures and Other Derivatives*, 3rd Edition, Prentice Hall.

Hull, John C., 2000, *Options, Futures and Other Derivatives*, 4th Edition, Prentice Hall

Jorion, Philippe, 2000, *Value-at-Risk. The New Benchmark for Managing Financial Risk*, Second Edition, McGraw-Hill, New York.

Kendall, Maurice, and Alan Stuart, J. Keith Ord, 1987, *Kendall's Advanced Theory of Statistics: Distribution Theory*, Vol. 1, Oxford University Press, 5th ed.

Kendall, Maurice, and Alan Stuart, J. Keith Ord, 1998, *Kendall's Advanced Theory of Statistics: Distribution Theory*, Vol. 1, Oxford University Press, 6th ed.

Longerstaey, Jacques, and Martin Spencer, 1996, *RiskMetrics™ – Technical Document*, 4th Edition, Morgan Guaranty/J.P.Morgan/Reuters.

Longerstaey, Jacques, and Martin Spencer, 1997, *RiskMetrics Monitor*, Fourth Quarter 1997, J.P.Morgan/Reuters.

Mandelbrot, Benoit B., 1997, *Fractals and Scaling in Finance, Discontinuity, Concentration, Risk,* Springer-Verlag New York.

Mood, Alexander McFarlane, and Franklin A. Graybill, and Duane C. Boes, 1998, *Introduction to the Theory of Statistics*, The McGraw-Hill Comp.

Neftci, Salih N., 2000, *An Introduction to the Mathematics of Financial Derivatives*, Academic Press

Pearson, Karl, 1894, "Contributions to the Mathematical Theory of Evolution," *Philosophical Transactions of the Royal Society*, A, 185, 71-110.

Rayner, J. C. W., and D. J. Best, 1989, *Smooth Tests of Goodness of Fit*, Oxford University Press

Stulz, René, 2002, *Derivatives, Financial Engineering, and Risk Management,* South-Western College Publishing, forthcoming.

Articles :

Ahn, D.-H., and Jacob Boudoukh, Matthew Richardson, Robert Whitelaw, 1999, "Optimal risk management using options," *The Journal of Finance* 54, 359-375.

Alexander, Carok, and C. Leigh, 1997, "On the covariance matrices used in V@R models," *Journal of Derivatives* 4, 50-62.

Allayannis, George, and James P. Weston, 2001, "The Use of Foreign Currency Derivatives and Firm Market Value," *Review of Financial Studies* 14, 243-276.

Anderson, Ronald W., and Jean-Pierre Danthine, 1981, "Cross Hedging," *Journal of Political Economy* 89, 1182-1196.

Bachelier, L. J. B. A., 1900, ”Theorie de la speculation.”, Paris: Gauthier-Villars. Reprinted in Paul H. Cootner (ed.), *The Random Character of Stock Market Prices*, Cambridge, Mass: MIT Press, 1964, 17-78.

Bakshi, Gurdip, and Charles Cao, Zhiwu Chen, 1997, ”Empirical performance of alternative option pricing models,” *The Journal of Finance* 52, 2003-2049.

Ball, Clifford A., and Walter N. Torous, 1985, ”On jumps in common stock prices and their impact on call option pricing,” *The Journal of Finance* 40, 155-173.

Ball, Clifford A., 1988, ”Estimation bias induced by discrete security prices,” *The Journal of Finance* 43, 841-865.

Barber, Brad M., and John D. Lyon, 1997, “Firm Size, Book-to-Market Ratio, and Security Returns: A Holdout Sample of Financial Firms,” *Journal of Finance* 52, 875-883.

Barone-Adesi, Giovanni, and Giannopoulos Kostas, Les Vosper, 1999, ”V@R without correlations for nonlinear portfolios,” *Journal of Futures Markets* 19, 583-602.

Basle Committee on Banking Supervision, 1995, *An Internal Model-Based Approach to market Risk Capital Requirements*, Basle, Switzerland: BIS.

Basak, Suleyman, and Alexander Shapiro, 2001, “Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices,” *Review of Financial Studies* 14, 371-405.

Bates, David S., 1996, ”Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options,” *The Review of Financial Studies* 9, 69-107.

Bates, David S., 1991, ”The crash of '87: Was it expected? The evidence from options markets,” *The Journal of Finance* 46, 1009-1044.

Battacharya, C.G., 1967, “A Simple Method of Resolution of a Distribution into Gaussian Components,” *Biometrics* 23, 115-135.

Beckers, Stan, 1981, ”A note on estimating the parameters of the diffusion-jump model of stock returns,” *Journal of Financial and Quantitative Analysis* 16, 127-140.

Benninga, Simon, and Zvi Wiener, 1998, ”Value at Risk”, *Mathematica in Education and Research* 7, 1-8.

Bessembinder, H., 1991, “Forward contracts and firm value: Investment incentive and contracting effects,” *Journal of Financial and Quantitative Analysis* 26, 519-532.

Best, Michael J., and Robert Grauer, 1991, “On the sensitivity of mean-variance efficient portfolios to changes in asset means: Some analytical and computational results,” *Review of Financial Studies* 4, 315-342.

Bick, A., 1990, “On Viable Diffusion Price Processes of the Market Portfolio,” *Journal of Finance* 45, 673-689.

Black, Fischer, and Myron Scholes, 1972, “The valuation of option contracts and a test of market efficiency,” *The Journal of Finance* 27, 399-417.

Black, Fischer, and Myron Scholes, 1973, “The pricing of options and corporate liabilities,” *Journal of Political Economy* 81, 637-659.

Black, Fischer, 1986, “Noise,” *Journal of Finance* 41, 529-543.

Black, Fischer, and Robert Litterman, 1992, “Global portfolio optimization,” *Financial Analysts Journal* 48, 28-43.

Bodnar, Gordon M., and Gregory S. Hayt, Richard C. Marston, Charles S. Smithson, 1995, “Wharton Survey of Derivatives Usage by U.S. Non-Financial Firms,” *Financial Management 24.*

Bollerslev, T., 1986, ”Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics* 31, 307-327.

Bollerslev, Tim, and Ray Y. Chou, and Kenneth F. Kroner, 1992, “ARCH modeling in finance,” *Journal of Econometrics* 52, 5-59.

Breeden, Douglas T., and Robert H. Litzenberger, 1978, ”Prices of state-contingent claims implicit in option prices,” *Journal of Business* 51, 621-651.

Britten-Jones, Mark, 1999, “The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights,” *Journal of Finance* 54, 655-671.

Butler, J. S., and Barry Schachter, 1998, ”Estimating Value at Risk with a precision measure by combining kernel estimation with historical simulation,” *Review of Derivatives Research* 1, 371-390.

Campbell, J., and J. Ammer, 1993, “What moves the stock and bond markets? A variance decomposition for long-term asset returns,” *Journal of Finance* 48, 3-37.

Canina, Linda, and Stephen Figlewski, 1993, ”The informational content of implied volatility,” *Review of Financial Studies* 6(3), 659-681.

Cecchetti, Stephen G., and Robert E. Cumby, Stephen Figlewski, 1988, "Estimation of the Optimal Futures Hedge," *Review of Economics and Statistics*, Vol 70, Issue 4, 623-630.

Chappell, David, and Kevin Dowd, 1999, "Confidence intervals for V@R," *Financial Engineering News*, March

Chunhachinda, P., and K. Dandapani, S. Hamid, A.J. Prahash, 1997, "Portfolio selection and skewness: Evidence from international stock markets." *Journal of Banking and Finance* 21, 143-167.

Cohen, A. Clifford, 1967, "Estimation in Mixtures of Two Normal Distributions," *Technometrics* 9, 15-28.

Conine, T.E., and M.J. Tamarkin, 1981, "On diversification given asymmetry in returns." *Journal of Finance* 36, 1143-1155.

Cootner, P.H., 1964, *The Random Character of Stock Market Prices,* Cambridge, MA, M.I.T. Press.

Coval, Joshua D., and Tyler Shumway, 2001, "Expected option returns," *Journal of Finance* 56*,* June 2001

Cox, John C., and Stephen A. Ross, 1976, "The valuation of options for alternative stochastic processes," *Journal of Financial Economics* 3, 145-166.

Culp, Christopher, and Merton Miller, Andrea Neves, 1998, "Value at Risk: uses and abuses," *Journal of Applied Corporate Finance* 10, 26-38.

David, F. N., 1938, "Tables of the Correlation Coefficient," *Cambridge Univ. Press*, London.

Day, N. E., 1969, "Estimating the Components of a Mixture of Normal Distributions," *Biometrica* 56, 463-474.

DeMarzo, Peter, and Darrell Duffie, 1995, "Corporate incentives for hedging and hedge accounting," *Review of Financial Studies* 8, 743-772.

Dennis, Patrick, and Stewart Mayhew, 2000, "Implied volatility smiles: evidence from options on individual equities," working paper

Dowd, Kevin, 1999, "A Value at Risk approach to risk-return analysis," *Journal of Portfolio Management* 25, 60-67.

Duffie, Darrell, and Jun Pan, 1997, ”An overview of Value at Risk,” *Journal of Derivatives* 4, 7-49.

Eberlein, Ernst, and Ulrich Keller, Karsten Prause, 1998, ”New insights into smile, mispricing and Value at Risk: the hyperbolic model,” *Journal of Business* 71, 371-406.

Engle, R., 1982, ”Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK inflation, ” *Econometrica* 50, 987-1008.

Esty, B.C., 1997, “Organizational Form and Risk Taking in the Savings and Loan Industry,” *Journal of Financial Economics* 44, 25-55.

Esty, B.C., Tufano, P., Headley J., 1994, “Banc One Corporation: Asset and Liability Management,” *Journal of Applied Corporate Finance* Vol. 7, No. 3, 33-65.

Eun, Cheol S., and Bruce G. Resnick, “Estimating the Correlation Structure of International Share Prices,” *Journal of Finance* 39, 1311-1324.

Falkenstein, Eric, 1997, ”Value at Risk and derivatives risk,” *Derivatives Quarterly* 4, 42-50.

Fama, Eugene, 1965, ”The behavior of stock-market prices,” *Journal of Business* 38, 34-105.

Fama, Eugene, 1970, ”Efficient capital markets: A review of theory and empirical work, ” *Journal of Finance* 25, 383-417.

Fama, Eugene, 1991, ”Efficient capital markets: II, ” *Journal of Finance* 46, 1575-1618.

Fama, Eugene, and Marshall E. Blume, 1966, ”Filter rules and stock-market trading,” *Journal of Business* 39, 226-241.

Fama, Eugene, and Kenneth French, 1992, ”The cross-section of expected stock returns,” *The Journal of Finance* 47, 3-56.

Figlewski, Stephen, 1989, ”Options Arbitrage in Imperfect Markets,” *The Journal of Finance* 44, 1289-1311.

Figlewski, Stephen, 1997, “Forecasting Volatility”, *Financial Institutions, Markets, and Instruments,* Vol 6., N.1, Jan. 1997, 1-88.

Fischer, R. A., 1915, “Frequency-distribution of the values of the correlation coefficient in samples from an indefinitely large population,” *Biometrica* 10, 507.

Fisher, R. A., 1921, "On the probable error of a coefficient of correlation deduced from a small sample," *Metron* 1.

French, Kenneth R., and G. William Schwert, Robert F. Stambaugh, 1987, "Expected stock returns and volatility," *Journal of Financial Economics* 19, 3-29.

Froot, Kenneth A., David S. Scharfstein, and Jeremy C. Stein, 1993, "Risk management: Coordinating corporate investment and financing policies", *Journal of Finance* 48, 1629-1658.

Froot, Kenneth A., and Jeremy C. Stein, 1998, "Risk management, capital budgeting, and capital structure policy for financial institutions: an integrated approach," *Journal of Financial Economics* 47, 55-82.

Fung, William, and David A. Hsieh, 1997, "Empirical characteristics of dynamic trading strategies: the case of hedge funds," *Review of Financial Studies* 10, 275-302.

Fung, William, and David A. Hsieh, 1999, "A primer on hedge funds," *Journal of Empirical Finance* 6, 309-331.

Fung, William, and David A. Hsieh, 2000, "Measuring the market impact of hedge funds," *Journal of Empirical Finance* 7, 1-36.

Géczy, Christopher, and Bernadette Minton, and Caherine Schrand, 1997, "Why Firms Use Currency Derivatives," *Journal of Finance* 48, pp. 1629-1658.

Godfrey S., and R. Espinosa, 1998, "Value-at-Risk and corporate valuation," *Journal of Applied Corporate Finance* 10, 108-117.

Golub, B. W., and L. M. Tilman, 1997, "Measuring yield curve risk using principal components analysis, Value at Risk, and key rate durations," *Journal of Portfolio Management* 23, 72-84.

Green, T. Clifton, and Stephen Figlewski, 1999, "Market Risk and Model Risk for a Financial Institution Writing Options," *Journal of Finance* 54, 1465-1499.

Greenspan, Alan, 2000, "Technology and Financial Services," *Journal of Financial Services Research* 18, 109-113.

Hagerman, Robert L., 1978, "More evidence on the distribution of security returns", *The Journal of Finance* 33, 1213-1221.

Hamilton, James D., and Raul Susmel, 1994, "Autoregressive conditional heteroskedasticity and changes in regime," *Journal of Econometrics* 64, 307-333.

Hannan, Timothy H., and Gerald A. Hanweck, 1988, “Bank Insolvency Risk and the market for Large Certificates of Deposit,” *Journal of Money, Credit, and Banking* 20, 203-212.

Harris, Milton, and Artur Raviv, 1991, “The Theory of the capital structure,” *Journal of Finance* 46, 297-355.

Harrison, Paul, 1998, ”Similarities in the distribution of stock market price changes between the eighteenth and twentieth centuries, ” *Journal of Business* 71, 55-79.

Haushalter G.D., 2000, “Financing policy, basis risk, and corporate hedging: Evidence from oil and gas producers,” *Journal of Finance* 55, 107-152.

He, Hua, and Hayne Leland, 1993, ”On equilibrium asset price processes,” *Review of Financial Studies* 6, 593-617.

Hilliard, Jimmy E., 1979, “The Relationship Between Equity Indices on World Exchanges,” *Journal of Finance* 34, 103-114,

Hilliard, Jimmy E., 1994, “Finite Horizon Hedge Ratios for American Options: A Minimum Variance Solution,” *Journal of Financial Engineering* 3, 1-17.

Hilliard, Jimmy E., and Jorg Reis, 1999, ”Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot,” *Journal of Financial and Quantitative Analysis* 33, 61-86.

Hogan, A. M. B., Malmquist, D. H., 1999, “Barriers to Depository Uses of Derivatives: An Empirical Analysis,” *Journal of Multinational Financial Management* 9, 419-440.

Hull, John, and Alan White, 1998, ”Incorporating volatility updating into the historical simulation method for value-at-risk,” *Journal of Risk* 1, 5-19.

Hull, John, and Alan White, 1998, ”Value at Risk when daily changes in Market Variables are not normally distributed,” *Journal of Derivatives* 5, 9-19.

Jackson, Patricia, and David Maude, William Perraudin, 1997, ”Bank capital and Value-at-Risk,” *Journal of Derivatives* 4, 73-90.

Jackwerth, Jens Carsten, and Mark Rubinstein, 1996, ”Recovering probability distributions from option prices,” *The Journal of Finance* 51, 1611-631.

Jensen, M.C., and W.H. Meckling, 1976, "Theory of the firm: managerial behavior, agency costs and ownerhip structure," *Journal of Financial Economics* 3, 305-360.

Jensen, M.C., 1986, "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review* 76, 659-665.

Johnson, Leland L., 1960, "The Theory of Hedging and Speculation in Commodity Futures," *Review of Economic Studies* 27, 139-151.

Jones, Charles M., and Gautam Kaul, Marc L. Lipson, 1994, "Information, trading, and volatility," *Journal of Financial Economics* 36, 127-154.

Jones, D., 2000, "Emerging Problems with the Basel Capital Accord: Regulatory Capital Arbitrage and Related Issues," *Journal of Banking and Finance* 24, 35-58.

Jorion, Phillippe, 1985, "International Portfolio Diversification with Estimation Risk," *Journal of Business* 58/3, 259-278.

Jorion, Philippe, 1988, "On Jump Processes in the Foreign Exchange and Stock Markets," *The Review of Financial Studies* 1, 427-445.

Jorion, Philippe, 1995, "Predicting volatility in the foreign exchange market," *The Journal of Finance* 50, 507-528.

Jorion, Philippe, 2000, "Risk Management Lessons from Long-Term Capital Management," *European Financial Management 6*.

Karolyi, G. A., and Rene M. Stulz, 1996, "Why do markets move together? An investigation of U.S.-Japan sotck return comovement," *Journal of Finance* 51, 951-986.

Keim, S. and R. F. Stambaugh, 1986, "Predicting returns in the stock and bond markets," *Journal of Financial Economics* 17, 357-390.

King, M., and S. Wadhwani, 1990, "Transmission of volatility between stock markets," *Review of Financial Studies* 3, 5-33.

King, M., and E. Sentana, S. Wadhwani, 1994, "Volatility and links between national stock markets," *Econometrica* 62, 901-934.

Kroner K. F., and J. Sultan, 1993, "Time-Varying Distributions and Dynamic Hedging with Foreign-currency Futures," *Journal of Financial and Quantitative Analysis* 28/4, 535-551.

Kumar, K. D., and Eh. H. Nicklin, A. S. Paulson, 1979, “Comment on ‘Estimating Mixtures of Normal distributions and Switching Regressions’”, *Journal of the American Statistical Association* 74, 52-55.

Kupiec, Paul, 1995, ”Techniques for verifying the accuracy of risk measurement models,” *Journal of Derivatives* 3, 73-84.

Kupiec, Paul, 1999, ”Risk capital and V@R,” *Journal of Derivatives* 7, 41-52.

Kwan, Simon H., 1996, “Firm-specific information and the correlation between individual stocks and bonds,” *Journal of Financial Economics* 40, 63-80.

Lamoreux, Christopher G., and William D. Lastrapes, 1993, ”Forecasting stock-return variance: Towards an understandign of stochastic implied volatilities,” *Review of Financial Studies* 6, 293-326.

Lang, L., and René Stulz, 1994, “Tobin’s Q, Corporate Diversification and Firm Performance,” *Journal of Political Economy* 102, 1248-1280.

Lawrence, Colin, and Gary Robinson, 1995, ”How safe is RiskMetrics™?,” *Risk* 8, 26-29.

Leland, Hayne E., 1998, “Agency Costs, Risk Management, and Capital Structure,” *Journal of Finance* 53, 1998 AFA Presidential Address.

Lévy, Paul, 1924, “Théorie des Erreurs. La Loi de Gauss et les Lois Exceptionelles,” *Bull. Soc. Math.* 52, 49-85.

Linsmeier, T.J., and N.D. Pearson, 1996, *Risk measurement: an introduction to Value at Risk*, Working Paper, University of Illinois at Urbana-Champaign.

Lo, Andrew W., and A. Craig MacKinlay, 1988. ”Stock market prices do not follow random walks: evidence from a simple specification test, ” *Review of Financial Studies* 1, 41-66.

Lo, Andrew W., 1999, “The Three P’s of Total Risk Management,” *Financial Analyst Journal*, 13-26.

Longin, Francois M., and Bruno Solnik, 1995, “Is the correaltion in international equity returns constant: 1960-1990?”, *Journal of International Money and Finance* 14, 3-26.

Longin, Francois M., 1996, “The asymptotic distribution of extreme stock market returns,” *Journal of Business* 63, 383-408.

Longin, Francois M., 1999, ”From Value at Risk to stress testing: The extreme value approach,” *Journal of Banking and Finance* 66.

Longin, Francois M., and Bruno Solnik, 2001, “Extreme Correlation Of International Equity Markets,” *Journal of Finance* 56, 649-676.

Macbeth, James D., and Larry J. Merville, 1979, ”An empirical examination of the Black-Scholes call option pricing model,” *The Journal of Finance* 34, 1173-1186.

Makridakis, S.G., and S.C. Wheelwright, 1974, “An analysis of the interrelationships among the major world stock exhanges,” *Journal of Business Finance Accounting* 1, 195-216.

Malz, Allan M., 2000, ”Do implied volatilities provide early warning of market stress?” *The RiskMetrics Group*, working paper.

Mandelbrot, Benoit, 1963, ”The variation of certain speculative prices,” *Journal of Business* 36, 394-417.

Mandelbrot, Benoit, 1966, ”Forecasts of future prices, unbiased markets, and 'martingale' models,” *Journal of Business* 36.

Matten, Christopher, 2000, *Managing Bank Capital: Capital Allocation and Performance Measurement*, Wiley John & Sons Inc.

Markowitz, Harry, 1952, “Portfolio Selection,” *Journal of Finance* 7, 77-91.

Markowitz, Harry, 1959, *Portfolio Selection: Efficient Diversification of Investments*, Wiley, New York.

Mayers, David, and Clifford W. Smith, 1982, “On the Corporate Demand for Insurance,” *Journal of Business* 55, 281-296.

Merton, Robert C., 1976a, ”Option pricing when underlying stock returns are discontinuous,” *Journal of Financial Economics* 3, 125-144.

Merton, Robert C., 1976b, ”The impact on option pricing of specification error in the underlying stock price returns,” *The Journal of Finance* 31, 333-350.

Mian, S.L., 1996, “Evidence on corporate hedging policy,” *Journal of Financial and Quantitative Analysis* 31, 419-439.

Minton B.A., and C. Schrand, 1999, “The impact of cash flow volatility on discretionary investment and the costs of debt and equity financing,” *Journal of Financial Economics* 54, 423-460.

Modigliani, Franco, and Merton H. Miller, 1958, "The Cost of Capital, Corporation Finance, and the Theory of Investment," *American Economic Review* 48, pp. 261-297.

Morck, R., and B. Yeung, 1991, "Why Investors Value Multinationality," *Journal of Business* 64, 165-187.

Myers, S.C., 1977, "The determinants of corporate borrowing," *Journal of Financial Economics* 4, 147-175.

Myers, S.C., 1972, "The application of finance theory to public utility rate cases," *Bell Journal of Economics and Management Science* 3, 58-97.

Nance, Deana R., and Clifford W. Smith, Jr., and Charles W. Smithson, 1993, "On the Determinants of Corporate Hedging," *Journal of Finance* 48, 267-284.

Neftci, Salih N., 2000, "Value at Risk calculations, extreme events, and tail estimation," *The Journal of Derivatives* 7/3, spring 2000, 23-37.

Osborne, M.F.M., 1959, "Brownian motion in the stock market," *Operations Research* 7, 145-173, reprinted in Cootner (1964).

Pagan, Adrian, 1996, "The econometrics of financial markets," *Journal of Empirical Finance* 3, 15-102.

Pena, Ignacio, and Gonzalo Rubio, Gregorio Serna, 1999, "Why do we smile? On the determinants of the implied volatility function," *Journal of Banking and Finance* 23, 1151-1179.

Peiró, Amado, 1999, "Skewness in financial returns," *Journal of Banking and Finance* 23, 847-862.

Press, S. James, 1967, "A Compound Events Model for Security Prices," *Journal of Business* 40, 317-335.

Pritsker, Matthew, 1997, "Evaluating Value at Risk methodologies: accuracy versus computational time," *Journal of Financial Services Research* , 201-241.

Quandt, Richard E., and James B. Ramsey, 1978a, "Estimating Mixtures of Normal Distributions and Switching Regressions," *Journal of the American Statistical Association* 73, 730-738, (with discussions by various Authors).

Quandt, Richard E., and James B. Ramsey, 1978b, "Estimating Mixtures of Normal Distributions and Switching Regressions. Rejoinder," *Journal of the American Statistical Association* 73, 751-752.

Quandt, Richard E., 1972, "A New Approach to Estimating Switching Regressions," *Journal of the American Statistical Association* 67, 306-310.

Ramchand, Latha, and Raul Susmel, 1998, "Volatility and cross correlation across major stock markets," *Journal of Empirical Finance* 5, 397-416.

Ramsey, James B., 1975, "Mixtures of Distributions and Maximum Likelihood Estimation of Parameters Contained in Finitely Bounded Compact Spaces," Econometrics Workshop Paper No. 7308, Michigan State University

Richardson, Matthew, and Tom Smith, 1993, "A test for multivariate normality in stock returns," *Journal of Business* 66, 295-321.

Rubinstein, Marc, 1985, "Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978," *The Journal of Finance* 40, 455-480.

Rubinstein, Mark, 1994, "Implied binomial trees," *The Journal of Finance* 49, 771-818.

Roll, Richard, 1986, "The Hubris Hypothesis of Corporate Takeovers," *Journal of Business* 59, 197-216.

Roll, Richard, 1988, "R^2," *Journal of Finance* 43, 541-566.

Samuelson, Paul A, 1965, "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review* 6, 41-49.

Schmalensee, Richard, and Robert R. Trippi, 1978, "Common stock volatility expectations implied by option premia," *The Journal of Finance* 33, 129-147.

Schmidt, Peter, 1982, "An Improved Version of the Quandt-Ramsey MGF Estimator for Mixtures of normal distrubutions and Switching Regressions," *Econometrica* 50, 501-516.

Schrand C., and Haluk Unal, 1998, "Hedging and coordinated risk management: Evidence from thrift conversions," *Journal of Finance* 53, 979-1013.

Schwert, G. William, 1989, "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance* 44, 1115-1153.

Schwert, G. William, 1990, "Stock returns and real activity: a century of evidence," *Journal of Finance* 45, 1237-1258.

Sercu P., and X. P. Wu, 2000, Cross- and delta-hedges: Regression- versus price-based hedge ratios," *Journal of Banking and Finance* 24/5, 735-757.

Servaes, Henry, 1996, "The Value of Diversification During the Conglomerate Merger Wave," *Journal of Finance* 51, 1201-1225.

Simkowitz, M.A., Beedles, W.L., 1978, "Diversification in a three-moment world," *Journal of Financial and Quantitative Analysis* 13, 927-941.

Smith, Clifford W., and René Stulz, 1985, "The determinants of firms' hedging policies," *Journal of Financial and Quantitative Analysis* 31, 419-439.

Smith, Clifford W., 1995, "Corporate Risk Management: Theory and Practice," *Journal of Derivatives*, Summer 1995.

Smith, Clifford W., and Warner, J.B., 1979, "On financial contracting: an analysis of bond covenants," *Journal of Financial Economics* 5, 117-161.

Smithson, Charles, 1996a, Surveys surveyed, *Risk* 9 (No. 3, March 1996), pp. 56-57.

Smithson, Charles, 1996b, Who uses what, *Risk* 9, (No. 5, May 1996), pp. 50-51.

Stein, Jerome L., 1961, "The Simultaneous Determination of Spot and Futures Prices," *American Economic Review* 51, 1012-1025.

Stulz, René M., 1984, "Optimal Hedging Policies," *Journal of Financial and Quantitative Analysis* 19, June 1984, 127-140.

Stulz, René M., 1990, "Managerial discretion and optimal financing policies," *Journal of Financial Economics* 26, 3-28.

Stulz, René M., 1999, "What's Wrong with Modern Capital Budgeting," *Financial Practice and Education* 9, No.2, 7-11.

Stulz, René M., 2000, "Why risk management is not rocket science," *Financial Times (of London),* June 27, 2000.

Tan, W.Y., and W. C. Chang, 1972, "Some Comparisons of the Method of Moments and the Method of Maximum Likelihood in estimating Parameters of a Mixture of Two Normal Densities," *Journal of the American Statistical Association* 67, 702-708.

Telser, Lester G., 1955, "Safety First and Hedging," *Review of Economic Studies* 23, 1-16.

Tufano, Peter, 1996, "Who Manages Risk? An Empirical Examination of Risk Management practices in the Gold Mining Industry," *Journal of Finance* 51, pp. 1097-1137.

Tufano, Peter, 1998, "Agency Costs of Corporate Risk Management," *Financial Management* 27.

Wall, Larry, and J. Pringle, 1989, "Alternative Explanations of Interest Rate Swaps: A Theoretical and Empirical Analysis," *Financial Management* 18, 59-73.

Ward, Ronald W., and Lehman B. Fletcher, 1971, "From Hedging To Pure Speculation: A Micro Model of Optimal Futures and Cash Positions," *Americal Journal of Agricultural Economics* 53, 71-78.

Yermack, D., 1996, "Higher Market Valuation of Companies with a Small Board of Directors," *Journal of Financial Economics* 40, 185-211.

Young, Tzay Y., and G. Coraluppi, 1970, "Stochastic Estimation of a Mixture of Normal Density Functions Using an Information Criterion," *IEEE Transactions on Information Theory*, IT-16, 258-263.

Additional useful bibliography is provided in Jorion (2001) and Dowd (1998).

For an updated list of the published research and working papers on various V@R related topics: check out the Barry Schachter's website: *www.gloriamundi.org*.