

INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

8713798

Simaan, Yusif Eed

**PORTFOLIO SELECTION AND CAPITAL ASSET PRICING FOR A CLASS OF
NON-SPHERICAL DISTRIBUTIONS OF ASSETS RETURNS**

City University of New York

Ph.D. 1987

**University
Microfilms
International** 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1987

by

Simaan, Yusif Eed

All Rights Reserved

PORTFOLIO SELECTION AND CAPITAL ASSET PRICING FOR A CLASS
OF NON-SPHERICAL DISTRIBUTIONS OF ASSETS RETURNS

by

YUSIF SIMAAN

A dissertation submitted to the Graduate Faculty in Business
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy. The City University of New York.

1987

© 1987

YUSIF SIMAAN

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Business in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

March 12, 1986
Date

Harry Markowitz
Chair of Examining Committee

March 13, 1986
Date

Alvin Epstein
Executive Officer

Stanley J. Liebowitz
S.A. 13
W. R. R. R. R.
Supervisory Committee

The City University of New York

ABSTRACT

PORTFOLIO SELECTION AND CAPITAL ASSET PRICING FOR A CLASS OF NON-SPHERICAL DISTRIBUTIONS OF ASSETS RETURNS

by

Yusif Simaan

Adviser: Harry Markowitz

The joint distribution that describes the return generating process here is modeled to distinguish between two aspects of portfolio risk by breaking down the variance of the portfolio distribution into two components: Spherical variance and a non-spherical variance. The latter determines skewness and other odd higher cumulants whereas the even cumulants are determined by both components. While any risk averse investor exhibits aversion to the spherical component, he may exhibit either a preference or an aversion to the non-spherical variance depending on his specific utility function. The problem of portfolio selection, for any risk averse investor, is equivalent to a quadratic programming problem that minimizes the spherical variance for a given mean and a non-spherical variance pair. Ranking of portfolio distributions collapses to a ranking of their respective means and risk measures. The mean-variance-skewness

efficient set is derivable as a function of the joint distribution parameters and estimatable from observable security returns. This framework allows an analytical evaluation of the practical optimality of the mean-variance investment strategy for all constant absolute risk averse investors. An optimization premium that reflects the foregone opportunity cost of the mean-variance strategy is derived as a function of the absolute risk aversion measure and the joint distribution parameters. Utilizing monthly returns on ten securities, the derived optimization premium is operationalized empirically. When there is a riskless asset, the optimization premium does not exceed \$0.05 on an invested dollar regardless of the degree of risk aversion. This framework also allows the derivation of a three moment capital asset pricing model in a perfect competitive market.

To my mother: Maryam Odi, and the memory of
my father: Eed Simaan

ACKNOWLEDGEMENT

Throughout the two-year period of working on this dissertation I have drawn help, encouragement and support from many people to whom I feel a need to express my gratitude.

My greatest debt is to Harry Markowitz for his patience in allocating many long hours to discuss this work in its early stage, and for his painstaking reading of the first draft of this dissertation. In addition to the great learning experience that I gained through his supervision of this work, his comments, help and encouragement contributed a great deal to this work and shaped it, indeed, in its final form.

I wish to thank Starvos Thomadakis, Salih Neftci and Gerald Pogue, the other members of the dissertation committee, for their reading of and comments on this dissertation.

I can't escape expressing my gratitude to Starvos Thomadakis who happened to be my teacher in several courses in Financial Theory that contributed a great deal to shape my knowledge in this area. Indeed, this has its effect on this dissertation.

This work also benefitted from many discussions with colleagues and friends. I am particularly thankful for the help of Mohammad Taliba and Frank Scalici.

The help and support of Cheryl Mwaria through the writing period of this dissertation is deeply appreciated. Cheryl edited the text, and her support throughout this period is invaluable.

Table of Contents

Chapter	Page
INTRODUCTION	1
I LITERATURE REVIEW	11
II DISTRIBUTIONS THAT ADMIT MEAN-VARIANCE PORTFOLIO SELECTION	22
2.1 Distributions that imply mean-variance Ranking	24
2.2 Two funds repeating distributions and the admissability of mean-variance portfolio selection	34
III THE EFFICIENT FRONTIER FOR A NON-SPHERICAL JOINT DISTRIBUTION OF ASSETS RETURNS	43
3.1 The distribution of assets returns	47
3.2 The nature of the portfolio problems	56
3.3 The derivation of the efficient set	63
3.4 Estimation of the distribution parameters	76
Appendix 3.A	80
Appendix 3.B	83
IV EVALUATION OF MEAN-VARIANCE PORTFOLIO SELECTION	86
4.1 The joint distribution and the class of utility functions	93
4.2 Derivation of the optimal portfolio and its mean-variance efficient alternative	99
4.3 Empirical demonstration of the optimization premium	117
Appendix 4.A	136
Appendix 4.B	138
Appendix 4.C	140

Chapter	Page
V MEAN-VARIANCE-SKEWNESS CAPITAL ASSETS PRICING . .	147
5.1 Quadratic Characteristic Lines and the CAPM	152
5.2 Mean-variance-skewness CAPM	164
CONCLUSIONS	178
REFERENCES	183

INTRODUCTION

The ability of mean-variance analysis to reach intuitive and insightful results in addition to its simplicity and its practical application made it very popular among academicians and practitioners. The general formulation of the portfolio problem does not lead to such desirable properties. The arbitrary nature of the specification of the return generating process and the investor's utility function do not allow an insightful analysis for the portfolio problems unless the degrees of freedom in their specification are reduced. The mean-variance framework has been successful in providing such a reduction. It reduces the arbitrary nature of the joint distribution of assets returns to an arbitrary specification of their means and covariance structure. Moreover, the ranking of portfolio distributions for an individual investor collapses to the ranking of their mean-variance pairs. Such a reduction in the degrees of freedom of the arbitrary nature of the portfolio problem, indeed, leads to its simplicity. Its intuitive appeal, on the other hand, is due to the direct connection of its constructs to observable assets returns and the ability to estimate the efficient set of portfolios regardless of the utility function. The neat mathematical structure between the risk

characteristics of individual assets returns and those of portfolio returns makes it a successful vehicle for portfolio analysis and other related problems in finance.

The elegant property of the mean-variance framework motivates three questions. First, what are the weakest restrictions on the return generating process and/or the investors preference that warrants its exact validity? Second, what is the opportunity cost of following its investment strategies when they are suboptimal? Third, is it possible to extend such a framework and broaden the domain of its validity while, at the same time, preserving its elegant properties? The smaller the truth domain of each one of the above questions, the greater the relevance of the subsequent question(s). These questions and their answers are the concern and the subject matter of this dissertation.

When the expected utility of the portfolio return is determined by the mean and its variance, mean-variance portfolio selection is perfectly valid. Chamberlain (1983) characterized joint distributions of assets returns that lead to such a property. In his characterization theorems, he ignored the wealth constraint of the portfolio problem. It will be proved that the existence of the wealth constraint makes his restrictions only sufficient, but not necessary, for a mean-variance ranking of random prospects. Namely, the class of distributions that imply mean-variance ranking is

richer than those provided by his restrictions. Moreover, different straightforward proves for the sufficiency of his restrictions for mean-variance ranking will be provided. A distinction between distributions that imply mean-variance ranking of portfolio returns and distributions that admit mean-variance portfolio selection is made. An argument, based on existing results in the literature, is given to demonstrate that mean-variance ranking is not necessary for the admissibility of the mean-variance framework. The class of two funds separating distributions introduced by Ross (1978), to be defined below, admits mean-variance portfolio selection on one hand, while it does not imply mean-variance ranking on the other.

To extend the mean-variance framework a class of non-spherical distributions is introduced to model the return generating process. The return on each security, for this class, is given as a combination of non-spherical random noise, common across securities, and a security specific spherical noise. By its construction, this class exhibits a three funds separation. We prove that the expected utility is determined by the mean, the variance and a non-sphericity parameter of the portfolio return distribution. Moreover, the variance of the portfolio return is decomposable to a spherical variance and a non-spherical variance. It is also proved that any risk averse investor dislikes the

spherical component of the variance. On the other hand, a risk averse investor may either like or dislike the non-spherical variance. Even when he dislikes the non-spherical variance, a risk averse investor exhibits different degrees of dislike to the two components of the variance of the portfolio return. In such a setting we have a two dimensional measure for risk--the two components of the portfolio return variance. The mean and the two dimensional measure for risk provide a three parameter framework for portfolio analysis. The portfolio problem, for any risk averse investor, is reduceable to a mathematical programming problem that maximizes the mean of the portfolio return subject to a given two-dimensional measure for risk. In this framework, the efficient set of portfolios can be derived, for all utility functions, as a function of the distribution parameters. Moreover, the distribution parameters are connected to the first three moments of assets return and are estimatable from observable assets returns. While the mean-variance efficient set provides a menu of efficient portfolios for different levels of expected return, our framework provides a richer one. It provides an efficient portfolio for any pair of mean and skewness parameter. The mean-variance efficient set turns out to be a subset of the efficient set of our class of distributions. Such a framework provides an alternative to the mean-variance portfolio selection. This alternative resembles the simplicity, the intuitive

appeal and the mathematical elegance of the mean-variance framework.

The class of non-spherical distributions, introduced in Chapter III, provides an analytical framework to evaluate the quality of mean-variance investment strategies. To conduct the evaluation, its spherical noise is specified, in Chapter IV, as joint normal while its non-spherical noise is specified as Pearson Type III distribution. Then the expected utility is derived for all preferences that exhibit constant absolute risk aversion. For a given absolute risk aversion, the optimal portfolio is derived. At the same time the portfolio that maximizes the expected utility restricted to the mean-variance efficient set is also derived. Then the two portfolios are compared. To make such a comparison we define an "optimization premium" as the maximum cost that an investor is willing to pay in order to replace his mean-variance efficient portfolio with the optimal one. This optimization premium is derived as a function of the degree of risk aversion and the distribution parameters. Thereafter, the results are illustrated empirically on ten securities with empirical monthly returns. The parameters of the joint distribution are estimated by utilizing a mixture of moments method and the maximum likelihood method.

The analytic evaluation and its empirical demonstration provide interesting insights about the mean-variance

approximation of expected utility maximization. When markets deny a riskless asset, mean-variance strategies are practically optimal up to a certain level of risk aversion. For a measure of absolute risk aversion $A < 15$ the optimization premium is no more than 0.5%--0.5 a penny on an invested dollar. The higher the risk aversion is, the poorer the performance of mean-variance strategies. This reflects the fact that the higher the degree of absolute risk aversion, the more prudent the investor is. Such prudence makes him give more consideration to distributional characteristics beyond the mean and the variance. It is interesting to note that the optimization premium escalates in a linear relationship with the absolute risk aversion. This property reveals a paradoxical nature of a utility function with a high constant absolute risk aversion. For levels of absolute risk aversion $A \geq 300$ the optimization premium exceeds the expected return of the optimal portfolio. This implies that such an investor is willing to pay a certain fee for information that leads to a risky portfolio with an expected return lower than the certain fee that he is ready to pay. Such behavior is, certainly, not rational and therefore, degrees of absolute risk aversion that exhibit such behavior must be ruled out.

When a riskless asset is introduced, the results are remarkably in favor of mean-variance strategies. Regardless of the level of risk aversion the optimization premium

does not exceed 0.05%. In contrast to the previous case, the higher the degree of risk aversion, the lower the optimization premium. Such an antagonistic result to the case when markets deny riskless assets, is due to the fact that the constant absolute risk averse investor tends to what he perceives the safest strategy regardless of the expected return of its alternatives. When there is a no riskless asset, the safest strategy--the strategy that globally minimizes the spherical variance--is not mean-variance efficient. While the geometric distance between the optimal portfolio and its corresponding mean-variance efficient portfolio converges to an upper limit, as the absolute risk aversion tends to be larger, their corresponding optimization premium increases without a limit. But when the riskless asset is introduced, both the optimal strategy and the mean-variance efficient one converge to the riskless asset as the absolute risk aversion increases. Such agreement leads to the opposite trend in the results of the latter case.

The three parameter framework for portfolio selection, in Chapter III, allows a derivation of a mean-variance-skewness capital asset pricing model (CAPM) for a perfect competitive capital market. Kraus and Litzenberger (1976) provided a version of mean-variance-skewness CAPM. They postulated a form of the expected utility that depends on the mean, the variance and the skewness of the portfolio return. Then they took the first order condition of the

individual investor as a description of his equilibrium. They did not prove that such condition is sufficient for maxima and neither did they restrict the form of the expected utility in a way that leads to the sufficiency of the first order condition. Since the skewness of the portfolio return is neither a concave nor a convex function in the portfolio decision variable, the expected utility cannot be postulated as a concave function of the decision variables of the portfolio problem. In fact, according to the way the problem of the individual investor is formulated in Kraus and Litzenberger (1976), the maxima of the expected utility may not exist. Similar to this approach, Kraus and Litzenberger (1983) provided a consumption oriented CAPM. They also provided sufficient conditions on the security returns to exhibit two and three parameter CAPM. It is shown, in Chapter V, that their restrictions, for the two parameter CAPM, are a special case of the restrictions that are necessary and sufficient for a set of assets returns to exhibit two funds separation. It is also proved that their sufficient conditions for a three parameter CAPM are stronger than their sufficient conditions for a two parameter CAPM. In fact, they imply their sufficient conditions for a two parameter CAPM.

On the other hand, the results of Chapter III provide the way for a straightforward derivation of a mean-variance-skewness CAPM. This is done, in Chapter V, for both the

riskless asset case and the case with no riskless asset. When the capital market offers the riskless asset, then the riskfree rate, the expected return on the market portfolio and another efficient portfolio determine the expected return on any individual security. In particular, the third portfolio may be chosen to be neutral for the variance of the market portfolio--a zero beta portfolio. Such a portfolio only contributes to the skewness of the market portfolio. Moreover, this zero beta portfolio may offer a risk premium or a risk discount depending on whether the market investor exhibits dislike or preference for skewness. Kraus and Litzenberger (1976) (1983), and Scott and Horvath (1980) argued the risk averse investor must exhibit preference for skewness. Chapter V includes a counter example which shows that a risk averse investor may not exhibit preference for skewness. According to the derived CAPM, in Chapter V, the risk premium on the market portfolio, the risk premium (or discount) on the zero beta portfolio, the riskfree rate, and a two dimensional measure of systematic risk determine the risk return tradeoff in the capital market. The two dimensional measure of the systematic risk for a given security is its contribution to the variance and its contribution to the skewness of the market portfolio. The analysis also shows that the Sharpe-Linter CAPM can price correctly an individual security under two conditions:

(i) when such a security exhibits the same sensitivity for the variance and the skewness of the market portfolio; or, (ii) when the zero beta portfolio provides neither a risk premium or risk discount. Namely, when the zero beta portfolio earns the riskfree rate. When the riskless asset does not exist, it is proved that the capital market allows creation of a portfolio neutral to both the variance and the skewness of the market portfolio. The expected return on such a portfolio plays the role of the riskfree rate in the CAPM of the previous case.

CHAPTER I
LITERATURE REVIEW

Mean-variance portfolio analysis, introduced by Markowitz (1952), provides on one hand a straightforward normative theory for portfolio selection, and on the other the cornerstone for the first positive theory for capital markets. Its birth, indeed, marked a revolution in financial economic thought. Since its birth, the literature on portfolio theory and capital markets expanded along two different lines, each stimulated the other. One line strove to depict the narrow validity of the mean-variance framework and to propose alternative theories to this framework. [Pratt (1964), Arrow (1965), Borch (1969) and 1974, Levy 1974, Ross (1976), Kraus and Litzenberger (1976), (1983)]. The second line moved in two different directions. One direction moved to relax its dependence on quadratic utility by providing restrictions on the return generating process that warrants its validity [Tobin (1958), Samuelson (1970), Ohlson (1975), Chamberlain (1983), Owen and Ravinowitz 1983]. The other direction moved to defend its normative theory for portfolio selection as a good approximation even for those contingencies outside its exact optimality [Markowitz (1959), Kounig (1959), Tsiang (1972) and (1974), Levy and Markowitz (1979), Pully (1981), and Kroll Levy and Markowitz (1983)].

The first criticism against mean-variance analysis opposed quadratic utility. The fact that quadratic utility does not exhibit decreasing risk aversion as a function of wealth was observed by Pratt (1964). Such an argument does not challenge the validity of the mean-variance framework anymore. This is because restrictions on the return generating process warrant its validity regardless of the utility function. A more challenging argument was provided by Borch (1969). He provided an example which shows that mean-variance ordering is not consistent with expected utility maximization. For a two points distribution of random prospects, he constructed portfolio returns with identical means and identical variances, while one of the prospects stochastically dominated the other. But, since these two prospects have identical means and variances they must provide the same expected utility in a mean-variance framework. The fact that mean-variance portfolio selection can be consistent with expected utility maximization without restricting the investor's preference and asset returns to mean-variance ranking invalidates such a criticism.

There have been several attempts in the literature to replace, or extend, the positive pricing theory implied in a mean-variance framework. While the arbitrage pricing theory of Ross (1976) suggests a completely different approach for assets pricing, Kraus and Litzenberger (1976)

and (1983) attempted to extend the mean-variance framework to a mean-variance-skewness. The latter approach runs into theoretical difficulties which will be discussed in Chapter V. The review of their work is also deferred to Chapter V. Although the validity of the normative theory for portfolio selection has been challenged several times, no alternatives, with its elegance,¹ were proposed either to replace or extend it, especially to a three parameters framework.

Tobin (1958) applied the mean-variance framework to provide a new version of liquidity preference theory under uncertainty. He provided a proof for the validity of mean-variance ordering when probability distribution of assets returns is expressible in two parameters. Such a claim was proved to be false by a counter example provided by Feldstein (1969). He showed that portfolios formed with a riskless asset and a two parameter probability

¹It may occur to the reader that the safety first rule is a challenging alternative to mean-variance portfolio selection. Roy (1952) suggested this rule at the same time that Markowitz (1952) suggested the mean-variance analysis. However, the safety first rule did not have the elegance of the mean-variance framework, it did not prove to be a good vehicle for portfolio analysis and therefore attracted less attention. Pyle and Turnovsky (1970) showed that there is a correspondence between the two rules if all assets are risky. The existence of a riskless asset leads to inconsistency between the two rules.

distribution cannot be ranked, in general, in terms of mean and variance. But Tobin's claim remains valid for the normal distribution.

Samuelson (1970) justified mean variance ordering of random prospects when the probability distribution of wealth is "compact." He used the concept "compact distribution" for distributions with "limiting moments of order higher than the second of smaller magnitude relative to the first two moments of portfolio returns. One case where this arises is when the portfolio revision horizon is short. Ohlson (1975) had shown that a vanishing third moment at a faster rate than the first two is a sufficient condition for a local quadratic approximation.

Another restriction on the distribution of assets' returns, that warrants the optimality of mean variance portfolio selection for any utility, was provided by Chamberlain (1983). He gives a characterization for the joint distribution of assets that implies an expected utility determined by its mean and its variance for any portfolio. When the risk free asset is a feasible asset, the distribution should be a non singular transformation of a random vector with a spherical distribution about the origin. When all assets are risky the restriction on the distribution is weaker. One of the assets could have any distribution, but the other assets, conditional on the first, should be

a non singular transformation of $(n-1)$ random vector spherically distributed about the origin. He proved further that when the expected utility is determined by the mean and the variance of the portfolio, a two funds separability holds. Observe that Chamberlain characterized distributions that imply an expected utility determined by the mean and the variance of the portfolio return. Of course, these distributions admit a mean-variance framework of analysis. However, the determination of the expected utility by the mean and the variance per-se is not a necessary condition for a mean-variance framework. There are other distributions for which the expected utility is not completely determined by the mean and the variance of the portfolio return on one hand, while such distributions admit mean-variance portfolio selection on the other. The class of two funds separating distributions, introduced by Ross 1978, is an example. It will be shown, in Chapter II, that such a class of distributions admits mean variance portfolio selection. The class of distributions introduced by Chamberlain is two funds separating distribution. Chapter II will include details on the class of spherical distributions as well as the class of two separating funds.

Parallel to the research line to extend the domain of mean variance portfolio selection, there was another line that extended the domain of validity of the Sharpe-Linter capital asset pricing model. Owen and Rabinowitz (1983) proved that spherical distributions extend the separation theorem for normal distributions and the results of the capital assets pricing model. The mean covariance framework is generalized to a mean-characteristic matrix framework in which the characteristic matrix of the spherical distribution is a basis for a risk measure. They provide a generalized pricing equation of CAPM.

Most of the defense of mean variance approximation to the expected utility relies on Taylor's expansion of the utility function. Markowitz (1959) provided two methods to approximate the expected utility by a function of the mean and the variance: the first by expanding the utility function around zero return and the second by expanding it around the mean. With empirical distributions and logarithmic utility the first approximation performed reasonably well for rate of return R with $|R| < 0.3$. The second performed markedly better. Pully (1981) used an approximation derived from Taylor's expansion around the mean to approximate the expected utility for logarithm, exponential and power utility functions with specific parameters. He used empirical distributions of rates of

returns on portfolios with short holding periods based on the justification of Samuelson (1970) and Ohlson (1975). Again, the mean variance approximation for the expected utility has been done for specific utility function with specific parameters for some members of the linear risk tolerance family. His report was supportive of mean-variance approximation.

Three other approaches to assessing the mean-variance framework of analysis did not involve Taylor's approximation. Levy and Markowitz (1979) provide a non local approximation to the expected utility. They fit a quadratic through three different points on the utility function. They suggested a class of approximation by a family of quadratics determined by the following three points: $(E - k\sigma; U(E - k\sigma))$, $(E, U(E))$, and $(E + k\sigma, U(E + k\sigma))$. In fact k determines the locality of the approximation. They approximated members of linear risk tolerance utility function with specific parameters for empirical distributions. For most of the cases the approximation did very well. The smaller k is, the better is the approximation. The approximation for portfolios performed better than for individual securities. In the same paper they proved that the class of approximations that they provide has the same local risk aversion as the approximated initial utility.

Aivazian, Callen, Krinsky, and Kawn (ACKK) used a generalized Box-Cox functional to approximate any utility function by mean and variance. They used the following form:

$$U(\delta) = \alpha_1 E(\lambda) + \alpha_2 V(\lambda) + 1/2 \alpha_3 E(\lambda)^2 + 1/2 \alpha_4 V(\lambda)^2 + \alpha_5 E(\lambda) V(\lambda)$$

where $U(\delta)$, $E(\lambda)$, $V(\lambda)$ are the Box-Cox transformations

$$U(\delta) = (U^{2\delta} - 1) / U^{2\delta}$$

$$E(\lambda) = (E^\lambda - 1) / \lambda$$

$$V(\lambda) = (V^\lambda - 1) / \lambda$$

E is the mean, V is the variance; with λ and δ taking different values, different functional forms are obtained. For $\delta, \lambda \rightarrow 0$ yields a translog form, $\delta = \lambda = 1/2$ gives the generalized Leontief form, and $\delta = \lambda = 1$ translates to a square root quadratic. They derive budget share equations for risky assets by straight-forward portfolio optimization framework. The parameters of the budget share equations were estimated from data on the household sector's holdings of risky assets and the associated rate of returns. A chi-square test was employed to find the most plausible of the three forms. They reported that the generalized Leontief function fits best. Unlike Levy and Markowitz, they did not compare their approximation to some hypothesized utility function. Instead, they validated the generalized

Leontief approximation by showing that, broadly speaking, it yields signs to their budget share equations consistent with theory and intuition. The estimates for the demands for risky assets have been made from aggregate data, and at best describe "representative consumer" utility function.

Kroll, Levy and Markowitz (1983) tested the validity of the mean variance approximation of the expected utility by comparing portfolios that exactly maximize specific expected utility function to mean variance ones for empirical distributions. This has been done by finding a global optimal portfolio for the utility function, and then comparing it to a portfolio that maximizes the expected utility restricted to the mean variance efficient frontier. The results were highly supportive of mean variance approximation. Leverage up to 50% did not affect the robustness of the results.

When the mean variance approximation has been compared to a real utility function there always has been a problem in measuring the closeness of the mean variance approximation to the approximated utility function. Pully (1981) used the ratio of the approximated utility to the real utility function, $\hat{E}U/EU$, where $\hat{U}$ refers to the approximation. Since two utility functions U and $U+A$ give the same ranking to random prospects for any A , his measure is sensitive to the choice of A . In fact, it is possible to choose

a shift parameter A such that his criterion is close to any number. Levy and Markowitz avoided this deficiency by using the correlation coefficient between $\hat{U}$ and U . Since the correlation coefficient is invariant under any positive linear transformation, it provides a better criterion. Kroll, Levy and Markowitz (1983) compared the expected utility of the global optimum portfolio and the optimal portfolio restricted to the frontier by its relative superiority to the naive portfolio over the mean variance restricted portfolio.

None of the previous research evaluated the performance of the mean-variance investment strategy analytically. Such an analytical approach must specify a class of distributions of assets returns for which mean variance portfolio selection is a suboptimal strategy. Then for a class of utility functions, the optimal portfolio decision must be derived analytically and compared to the best mean-variance choice. Thereafter the opportunity cost of the mean-variance strategy must be derived analytically and operationalized empirically. To evaluate the mean-variance portfolio selection, in such an approach, we must have a normative portfolio theory under which mean-variance portfolio selection is a special case. This is why this dissertation, motivated by the problem of assessing the performance of mean-variance investment strategy, led to

a three parameter framework for portfolio selection in which the mean-variance strategies are a subset of a larger class of efficient strategies. Once we have a three parameter framework of portfolio selection, it is possible to evaluate the performance of mean-variance strategies for constant absolute risk aversion investors. Thereafter a derivation of a mean-variance-skewness capital asset pricing model can be done straightforwardly.

CHAPTER II

DISTRIBUTIONS THAT ADMIT MEAN-VARIANCE PORTFOLIO SELECTION

Mean-variance portfolio selection is consistent with expected utility maximization only for a restricted class of distributions of asset returns or a restricted class of utility functions. The only class of utility functions that admits means-variance portfolio selection is the quadratic. While restrictions on the distributions of asset returns are subject to empirical testability, restrictions on investors' preference are not subject to such empirical validation.¹ This fact, besides the criticism against the quadratic class of utility functions, may explain why the body of financial theories for portfolio selection and capital asset pricing favored restricting the return generating process rather than investors' preference structure.

The objective of this chapter is to synthesize existing results in the literature about restrictions on the return generating process that justify the consistency of mean-variance portfolio selection with expected utility maximization. A distinction between distributions that imply mean-

¹There is statistical frameworks to test the validity of competing hypotheses about the distribution of a random variable. To my knowledge, there is no framework to test the validity of competing hypotheses about the utility function of an individual.

variance ranking of portfolio returns and distributions that admit mean-variance portfolio selection is made. An argument, based on existing results in the literature, will be given to demonstrate that mean-variance ranking is not necessary for the admissibility of the mean-variance framework. Chamberlain (1983) provided restrictions on the joint distribution of asset returns as necessary and sufficient for a mean-variance determined expected utility. According to Chamberlain (Ibid.), when a riskless asset exists the class of spherical distributions, to be defined later, is necessary and sufficient for a mean-variance ranking of portfolio return distributions. When all assets are risky, there is one degree of freedom to determine the distribution of the first asset while the joint distribution of all assets conditional on the first, must be spherical.

In his characterization theorems Chamberlain (Ibid.), ignored the wealth constraint of the portfolio problem. It will be proved that the existence of the wealth constraint makes his restrictions only sufficient, but not necessary, for a mean-variance ranking of random prospects. Namely, the class of distributions that imply mean-variance ranking is richer than those provided by his restrictions. Moreover, different straightforward proofs for the sufficiency of his restrictions for mean-variance ranking will be provided. Through this chapter, concepts, definitions, and

results that provide a background and motivation to the developments of the subsequent chapters are also provided.

2.1 Distributions that imply mean-variance ranking

The expected utility of the portfolio return is determined by the mean and the variance of the portfolio return if, and only if, the distribution of the portfolio return is determined by its mean and variance. Denote by Z the random vector of asset returns, with a mean μ and a variance matrix V , and by X any nonrandom vector. A vector, X , is said to be a portfolio if it satisfies the wealth constraint $X'e = 1$ where $e = (1, \dots, 1)'$. The expected utility is determined by the portfolio mean, $X'\mu$, and the portfolio variance, $X'VX$, if and only if the dependence of the expected utility on the portfolio choice, X , operates through $X'\mu$ and $X'VX$. This definition allows the dependence of the portfolio distribution on distributional parameters invariant (or exogenous) to the portfolio choice. As long as the portfolio distribution is controlled only by its mean and its variance we have a mean-variance determined expected utility and hence mean-variance ranking. Before the characterization theorems of Chamberlain are stated some definitions and relevant results in probability theory are provided.

Definition 2.1: A random vector ϵ , with $E(\epsilon) = 0$, is said to be spherically distributed about the origin if, and only if, for all orthogonal matrices R , $\epsilon \stackrel{d}{=} R\epsilon$ (the notation $Y_1 \stackrel{d}{=} Y_2$ denotes Y_1 is equal in distribution to Y_2).

Since a linear transformation is orthogonal if, and only if, it preserves vector norms, a random vector ϵ is spherically distributed about the origin if, and only if, its characteristic function depends on the norm of its arguments. Namely:

$$\phi_{\epsilon} = E \exp \{it'\epsilon\} = g(t't) \quad (2.1)$$

where

$$i = \sqrt{-1}$$

Definition 2.2: A vector Z is said to be spherically distributed if, and only if, Z is a linear transformation of a random vector spherically distributed about the origin. If $Z = A\epsilon + d$ and $W = A'A$ then:

$$\phi_Z(t) = E \exp \{it'Z\} = \exp \{it'd\} g(t'Wt) \quad (2.2)$$

Kelker (1970) used equation (2.2) to define the class of spherical distributions with a mean d and a characteristic matrix W . The notation $Z \sim S_p(d, W)$ denotes a random vector with characteristic function as (2.2). The covariance

matrice for Z is given as kW where k is a scalar independent of d and W . If the density function of Z exists, it has the following form:

$$f_Z(z) = C_p |W|^{-1/2} h[(Z-d)' W(Z-d)] \quad (2.3)$$

Where C_p is constant. A feature that characterizes the class of spherical distributions is the following: A vector Z is spherically distributed if, and only if, there exist a non-negative random scalar V such that the conditional distribution of Z on V (to be denoted $Z|V$) is normal with zero mean and a variance $V \cdot I$. In fact, the spherical distribution is a mixture of n independent normal variates $N(0, V)$ with any apriori mixing distribution for V .²

Lemma 1: Let Z be any random vector with a characteristic function $\phi_Z(t)$, and let X be any vector, the characteristic function of $Y = Z'X$ is given as:

$$\phi_Y(s) = \phi_Z(SX_1, SX_2, \dots, SX_n) \quad (2.4)$$

Proof:

$$\begin{aligned} \phi_Y(s) &= E \exp \{iX'ZS\} = E \exp \left\{ i \sum_{j=1}^n (X_j S) z_j \right\} \\ &= \phi_Z(X_1 S, \dots, X_n S) \quad \text{Q.E.D.} \end{aligned}$$

²These results are drawn from Kelker (1970). The same reference includes further information about the class of spherical distributions.

Corollary 1.1: If $X'Z_1 \stackrel{d}{=} X'Z_2$ for all X , then $Z_1 \stackrel{d}{=} Z_2$.

Proof: Since $X'Z_1 \stackrel{d}{=} X'Z_2$ for all X we have:

$$\phi_{X'Z_1}(S) = \phi_{Z_1}(x_1 S, \dots, x_n S) =$$

$$\phi_{X'Z_2}(S) = \phi_{Z_2}(x_1 S, \dots, x_n S)$$

for all X and S .

If we let $t_i = sx_i$ then $\phi_{Z_1}(t) = \phi_{Z_2}(t)$ for all t and hence identify in t . This implies that

$$Z_1 \stackrel{d}{=} Z_2 \qquad \text{Q.E.D.}$$

In order for corollary 1.1 to hold, X must be unrestricted. Any restriction on X such as the wealth constraint, $X'e = 1$, makes $\phi_{Z_1}(t) = \phi_{Z_2}(t)$ true only for restricted range of t and does not allow the conclusion in corollary 1.1. The following provides the characterization theorems, the different proofs for the sufficiency of their assumptions to mean-variance ranking, and counter examples that demonstrate the inconsistency of their necessity with the wealth constraint.

Theorem 2.1: (The riskless asset case). The distribution of $r = X'Z + C$ is determined by its mean and variance for

any X if, and only if, there exist a non-singular transformation with a matrix A such that $\epsilon = A(Z - \mu)$ is spherically distributed about the origin.

Proof of sufficiency: Suppose there exist a non-singular matrix A such that $\epsilon = A(Z - \mu)$, where $\epsilon \sim S_p(0, I)$. Then $Z = A^{-1}\epsilon + \mu$. Since $\phi_\epsilon(t) = g(t't)$ the characteristic function of Z , $\phi_Z(t)$, is:

$$\phi_Z(t) = \exp \{it'\mu\} g(t'Wt)$$

WHERE $W = (A^{-1})'A$. The characteristic function of r , $\phi_r(S)$, according to Lemma 1, is:

$$\begin{aligned}\phi_r(S) &= \exp \{iCS\} \phi_Z(X_1S, \dots, x_nS) \\ &= \exp \{i(X'\mu + C)S\} g(X'WXS^2)\end{aligned}$$

But since $W = k \text{Var}(Z) = kV$, $E(r) = X'\mu + C = E_p$, and $\text{Var}(r) = X'VX = \sigma_p^2$, we have:

$$\phi_r(S) = \exp \{i E_p S\} g(k\sigma_p^2 S^2) \quad \text{Q.E.D.}$$

The proof of necessity provided by Chamberlain, on the other hand, assumes that X is unrestricted. Through the proof there is an implicit use of Corollary 1.1. Since any restriction on X invalidates this corollary, the class of spherical distributions is not necessary for the

determination of the portfolio distribution by its mean and variance. As an example, consider the following structure of asset returns:

$$Z_i = Y + \epsilon_i \quad i = 1, 2, \dots, n \quad (2.5)$$

where $\epsilon_i \sim S_p(0, I)$ and Y is a non-spherical random variable. Therefore $Z|_Y \sim S_p(Y \cdot e, I)$. The characteristic function of Z is:

$$\begin{aligned} \phi_Z(t) &= E \exp \{it'Z\} = E_Y(E_{Z|Y} \exp \{it'Z\}) \\ &= E_Y [\exp \{it'eY\} g(t't)] = \phi_Y(t'e)g(t't) \end{aligned}$$

Let X_0 be the position in the riskless asset and let R_F be its return. The return on any portfolio is given by:

$$r = Y'Z + X_0 R_F \quad \text{with} \quad X'e + X_0 = 1$$

By Lemma 1 the characteristic function of r is:

$$\begin{aligned} \phi_r(S) &= \exp \{iX_0 R_F S\} \phi_Y(X'eS) g(X'XS^2) \\ &= \exp \{iX_0 R_F S\} \phi_Y[(1-X_0)S] g(X'XS^2) \end{aligned}$$

It is straightforward to show that:

$$E(r) = X_O R_F + (1-X_O)E(Y) = E$$

$$\text{Var}(r) = (1-X_O)^2 \sigma_Y^2 + X'X = \sigma_P^2$$

This leads to:

$$1 - X_O = \frac{E - R_F}{E(Y) - R_F}$$

$$X'X = \sigma_P^2 - \left(\frac{E - R_F}{E(Y) - R_F} \right)^2 \sigma_Y^2$$

Therefore, the distribution of the portfolio return is determined by its mean and its variance. Observe that if $X'e + X_O$ is unrestricted, the distribution of the portfolio return is determined by X_O , $X'e$ and $X'X$ and hence it is not determined by its mean and its variance.

Theorem 2.2: (The case with no riskless asset). If $E(Z) \neq 0$, then the distribution of $X'Z$ is determined by its mean and variance for every X if, and only if, there is a non-singular matrix T such that:

$$TZ = (m, h_2, \dots, h_n)' = (m, H)' = \epsilon$$

where conditional on m , H is spherically distributed about the origin.

Proof of sufficiency:

$$\begin{aligned}
\phi_{\epsilon}(t_1, \dots, t_n) &= E \exp \{it_1' m + it_2' H\} \\
&= E_m(E_{\epsilon|m} \exp \{it_1' m + it_2' H\}) \\
&= E_m(\exp \{it_1' m\} E_{H|m} \exp \{it_2' H\}) \\
&= E_m(\exp \{it_1' m\} \phi_{H|m}(t_2)) \\
&= E_m[\exp \{it_1' m\} g(t_2' t_2)] = \phi_m(t_1' m) g(t_2' t_2)
\end{aligned}$$

where t_2 is the vector t with the first entry deleted.

Recall that $Z = T^{-1} \epsilon = A \epsilon$ and let a_j denote the j^{th} column of A . Then:

$$\begin{aligned}
\phi_Z(t_1, \dots, t_n) &= E \exp \{it' A \epsilon\} = \phi_{\epsilon}(t' a_1, \dots, t' a_n) \\
&= \phi_m(t' a) g\left[\sum_{j=2}^n (t' a_j a_j' t)\right] \\
&= \phi_m(t' a_1) g\left[t' \left(\sum_{j=2}^n a_j a_j'\right) t\right] = \phi_m(t' a_1) g(t' w t)
\end{aligned}$$

where $w = \sum_{j=2}^n a_j a_j'$. In Appendix 2.A it is proved that,

$$w = \text{Var}(Z) - \sigma_m^2 a_1 a_1' = V - \sigma_m^2 a_1 a_1'$$

If we let $r = X'Z$, $E_P = E(r)$, and $\sigma_P^2 = \text{Var}(X'Z)$ then by lemma 1:

$$\begin{aligned}\phi_r(S) &= \phi_m(X'a_1S) g(X'WXS^2) \\ &= \phi_m\left(\frac{E_P}{E(m)}\right) \cdot g(X'VX - \sigma_m^2 X'a a'X) \\ &= \phi_m\left(\frac{E_P}{E(m)}\right) g\left(\sigma_P^2 - E_P^2 \frac{\sigma_m^2}{E(m)}\right)\end{aligned}$$

Again, the restrictions on the joint distribution provided by theorem 2 are not necessary for a mean-variance ranking when the wealth constraint $X'e$ is imposed. This is because the necessity proof of Chamberlain for this theorem also utilizes implicitly corollary 1.1 which is valid only when X is not restricted. As an example, take the following structure of asset returns:

$$z_i = \mu_1 Y_1 + Y_2 + \epsilon_i \quad (2.6)$$

where $\varepsilon_i \sim S_p(0, W)$ and $Y = (y_1, y_2)$ is a two-dimensional random variable with any distribution. The characteristic function of Z is:

$$\begin{aligned}\phi_Z(t) &= E_Y[E_Z|Y \exp \{it'(\mu Y_1 + Y_2 + \varepsilon)\}] \\ &= E_Y(E_Z|Y \exp \{it'\mu Y_1 + t'e Y_2 + t'\varepsilon\}) \\ &= E_Y[\exp \{it'\mu Y_1 + it'e Y_2\} \phi_\varepsilon(t)] \\ &= \phi_Y(t'\mu S, t'e S) g(t'wS)\end{aligned}$$

and by lemma 2.1 the characteristic function of $r = X'Z$ is:

$$\phi_r(S) = \phi_Y(X'\mu S, S) g(X'WX S^2)$$

It is straightforward to validate that $X'\mu$ and $X'WX$ are determined by $E(X'Z)$ and $\text{var}(X'Z)$ for any portfolio X .³

³The cummulant generating function of r , $\psi_r(S) = \log \phi_r(S)$ is:

$$\psi_r(S) = \psi_Y(X'\mu S, S) + \psi_\varepsilon(X'WKS^2)$$

and:

$$E(r) = \bar{1}^{-1} \psi_r'(\theta) = X'\mu E(y_1) + E(y_2)$$

$$\text{var}(r) = -\psi_r''(0) = (X'\mu)^2 \sigma_{Y_1}^2 + \sigma_{Y_2}^2 + X'WX[-\psi_\varepsilon''(0)]$$

The above two equations determine simultaneously $X'\mu$ and $X'WX$ as a function of $E(r)$ and $\text{var}(r)$.

2.2 Two funds separating distributions and the admissibility of mean-variance portfolio selection

There is a myth among scholars in Financial Economics that a mean-variance ordering of random prospects is necessary for a mean-variance portfolio selection. Here, it will be demonstrated that mean-variance ranking is not necessary for a framework of mean-variance portfolio selection. The class of two funds separating distributions introduced by Ross (1978), to be defined below, admits mean-variance portfolio selection on one hand, while it does not imply mean-variance ranking on the other. This fact is of some theoretical significance of its own for at least two reasons. First, it has been demonstrated by Borch (1969) that mean-variance ranking of random prospects is inconsistent with expected utility maximization.⁴ The fact that mean-variance portfolio selection can be consistent with expected utility maximization without restricting the investor's preference and asset returns to mean-variance ranking invalidates such a criticism.

Second, there is a myth regarding the inconsistency of mean-variance portfolio selection with skewed asset returns. The counter examples of section 2.1 demonstrate the consistency of mean-variance ranking with skewed returns.

⁴For two points distribution of random prospects, Borch constructs two distributions with identical mean-variance pairs. Yet one of the two constructed distributions stochastically dominates the other.

However, the skewness of any portfolio, in these examples, is completely determined as an implicit function of the mean and the variance of its return.⁵ The class of two funds separating distributions, on the other hand, allows more flexibility. Distributions of portfolio returns for this class are not determined by their mean-variance pairs. This allows the dependence of the expected utility on the skewness of the portfolio return on one hand, while skewness itself is not determined by the first two moments on the other. The rest of this section will introduce the separating distributions, some results about the two funds separating distributions, and will include the justification of the above claims.

Definition 2.3: A random variable Y_1 is said to stochastically dominate an alternative Y_2 for a class of utility function, U , if and only if

$$E u(Y_1) > E u(Y_2) \text{ for all } u \text{ belonging to } U$$

⁵Consider the example given in eq. (2.6). The joint distribution of (Y_1, Y_2) may exhibit any skewness. The portfolio return, $y_1 + X'u y_2 + X'e$ will exhibit skewness that depends on X_u and the co-skewness of y_1 and y_2 . Such a skewness may depend only on the mean and the variance of the portfolio return.

Definition 2.4: A vector of asset returns, Z , exhibits K funds separation, if there exist K mutual funds, $X_1, X_2, \dots, X_K$, such that for any utility function u there exists X^* for which

$$\sum_{j=1}^n \lambda_j = 1, \text{ and } X^* = \sum_{j=1}^n \lambda_j X_j \text{ and } Z'X^*$$

stochastically dominates $Z'X$ for any portfolio X .⁶

Ross (Ibid.) introduced the class of separating distributions and provided necessary and sufficient conditions on the distribution of the vector of random return, Z , in order to exhibit K funds separation. According to Ross (Ibid), the following theorem provides sufficient conditions for two funds separation.

Theorem 2.3: A vector of asset returns exhibits two funds separation if there exist two random variables, y and v , a non-random vector b , and two portfolios, X_1 and X_2 , such that:

⁶The above definition permits the choice of $\lambda_1, \dots, \lambda_n$ to depend on the utility function. Ross refers to this definition as seal separation to distinguish from strong separation for which X^* is independent of the utility function.

$$(i) \quad z_i = y + b_i v + \epsilon_i \quad \text{for } i = 1, 2, \dots, n$$

$$\text{with } E(\epsilon_i) = 0$$

(ii) for any t :

$$E[\epsilon_i | y + t \cdot v] = 0$$

$$(iii) \quad \epsilon' X_1 = \epsilon' X_2 = 0 \quad \text{and} \quad X_1' e = X_2' e = 1$$

For the above conditions to be necessary t , in condition (ii), must be restricted properly. The return on any portfolio, X , of the above security returns is given as:

$$X'Z = y + x'b \cdot v + X'\epsilon$$

Observe that the portfolio return is decomposed to three sources of risk: (i) The first source of risk, y , is common across securities and exogenous to the portfolio choice, X . (ii) The second source of risk, v , is also common across securities but its contribution to the portfolio distribution is contingent upon $x'b$ and hence is controllable by the portfolio decision variables. (iii) A security specific noise ϵ_i . The two fund separation is achieved only when the security specific noise satisfies two conditions:

- (a) The security specific noise must be a fair game with respect to any linear combination of the two common sources of noise. (b) There must exist two portfolios, X_1 and

X_2 , where each of them completely diversifies the security specific noise. These two portfolios, in fact, are the separating funds.

Ross (Ibid.) proved that any portfolio is stochastically dominated by a linear combination of X_1 and X_2 . He also proved that when the covariance structure exists any efficient portfolio is mean-variance efficient. That is: it minimizes the variance of the portfolio return among all portfolios with a mean equal to the mean of this portfolio return. As an alternative and different proof for this result consider the following argument: The portfolios X_1 and X_2 determine a line in the portfolio space. This line is determined by two efficient portfolios or two non-equivalent utility functions.⁷ In particular, the optimal portfolio of any two non-equivalent quadratic utilities is on this line. These two portfolios span the line between X_1 and X_2 on one hand and the mean-variance efficient on the other. Therefore, the mean-variance efficient set and the efficient set, determined by X_1 and X_2 , are identical.

One interesting feature of the class of two funds separating distributions is the fact that the return on efficient portfolios have different distributional structure than those

⁷Two non-equivalent utility functions have two different optimal portfolios. Each one of these portfolios is on line determined by X_1 and X_2 . Therefore they span the efficient set.

of inefficient portfolios. The return on any portfolio is given as:

$$X'Z = y + X'b \cdot v + X'\epsilon \quad (2.7)$$

But when X is efficient, $X = \lambda X_1 + (1 - \lambda)X_2$, its return is given as:

$$X'Z = y + (\lambda X_1'b + (1 - \lambda) X_2'b) v = y + X'b v \quad (2.8)$$

The contribution of the security specific risk, $X'\epsilon$, is zero for any efficient portfolio. While the distribution of any efficient portfolio return is controllable through $X'b$, the control parameters (parameters that depend on the decision variable X) of the distribution of inefficient portfolio returns depend on the distribution of $X'\epsilon$ and hence on the joint distribution of the security specific noise ϵ_i . Given the fact that there are $(n-2)$ degrees of freedom to determine such a distribution the ranking of inefficient portfolios does not have to be a mean-variance ranking.⁸ On the other hand, efficient portfolios can be

⁸To create a two funds separating distribution:
 (i) choose two random variables, y and v , with any arbitrary joint distribution; (ii) Choose any joint distribution for $\epsilon_1, \epsilon_2, \dots, \epsilon_{n-2}$ such that each of them is a fair game with respect to any linear combination of y and v ;
 (iii) Let $\epsilon_{n-1} \equiv \epsilon_n \equiv 0$ and $X_1 = (0, 0, \dots, 1, 0)'$ and $X_2 = (0, 0, \dots, 1)$; (iv) choose b as any arbitrary vector.

ranked according to $X'b$ or according to their means as the case in the mean-variance framework. This is because there is one to one mapping between $X'b$ and $E(X'Z)$.⁹

To conclude, the admissibility of the mean-variance framework for portfolio selection is consistent with distributions and utility functions that does not imply mean-variance ranking of random prospects. This fact, invalidates Borch's criticism against the inconsistency of mean-variance portfolio selection with expected utility maximization. Although mean-variance ranking is theoretically inconsistent with expected utility maximization, as Borch (Ibid.) demonstrated, such a criticism does not carry on to mean-variance portfolio selection. Moreover, mean-variance portfolio selection is consistent with a broad domain of skewness structure. The fact that security returns exhibit skewness is not an

Such a choice leads to a two-funds separating distribution. The joint distribution of ϵ can be chosen, in particular, not to admit mean-variance ranking. Moreover, it can be chosen with any skewness structure.

⁹If we apply the expectation operator on equation (2.8) and solve for $X'b$ we have:

$$X'b = \frac{E(X'Z) - E(v)}{E(v)}$$

Since $E(y)$ and $E(v)$ are exogenous, the distribution of the portfolio return of efficient portfolios is completely determined by either $X'b$ or the mean of the portfolio return. The investor preference is completely determined by ranking the means of efficient portfolios. This is precisely what a mean-variance framework implies.

empirical evidence to the inconsistency of mean-variance portfolio selection with observable security returns.

The theoretical validity of mean-variance portfolio selection relies on the hypothesis whether the unavoidable risk is generated by one common source. In fact, it is a hypothesis about the dimensionality of the risk measure. If risk is unidimensional, the variance, or any spread parameter when variance does not exist, can be taken as a risk measure. This implies that all components of variance are disliked by the same risk averse investor at the same rate. One way to think of risk as multidimensional is to break down the variance of portfolio return to different components. Such components may not be disliked at the same rate by the same risk averse investor. Some of these components may not look bad to some risk averse investors. This, in fact, is the approach of this work in the forthcoming chapters.

Appendix 2.A

Since $E(h|m) = 0$ implies $\text{cov}(h_j, m) = 0$ for $j = 2, \dots, n$, we have:

$$\text{Var}(\epsilon) = \begin{pmatrix} \sigma_m^2 & 0 & \dots & 0 \\ 0 & & I & \\ 0 & & & \end{pmatrix} = W$$

and

$$V = \text{Var}(Z) = A'WA = \begin{pmatrix} \sigma_m^2 a_{11} & a_{21} \dots a_{n1} & a_{11} \dots a_{1n} \\ \sigma_m^2 a_{1n} & a_{2n} \dots a_{nn} & a_{n1} \dots a_{nn} \end{pmatrix}$$

this implies that:

$$v_{ij} = \sigma_m^2 a_{1i} a_{1j} + \sum_{k=2}^n a_{ki} a_{kj}$$

Therefore:

$$V = \sigma_m^2 a_1 a_1' + \sum_{k=2}^n a_k a_k' \quad \text{Q.E.D.}$$

CHAPTER III

THE DERIVATION OF THE EFFICIENT SET

The ability to order random prospect according to a "risk measure" affords a great simplification in the portfolio problem and provides results for which investors can develop intuition.¹ Indeed the first breakthrough in portfolio theory, provided by Markowitz (1952), was possible after taking the variance of a random prospect as its single risk measure. The variance of the portfolio return is a proper risk measure, for all risk averse investors, if the joint distribution of asset returns admits two funds separation, and has an existing variance matrix. But for a more general class of joint distributions variance is only one aspect of risk. For example, an increase in the variance of the portfolio distribution that leaves its skewness

¹If Z_1 and Z_2 are two random prospects with identical means, then Z_1 is less risky than Z_2 , for a class of utility functions U , if $E u(Z_1) > E u(Z_2)$ for all u belongs to U with $E u_0(Z_1) > E u_0(Z_2)$ for at least one utility function, u_0 , in U . A risk measure allows a redefinition of the utility function as a function of the risk measure of the random prospects and its mean. The nature of the risk measure and its dimensionality is contingent upon the joint distribution of asset return and the class U . Interesting risk measures are utility free and with few dimensions which are subject to intuitive interpretation. A proper restriction of the joint distribution is the only way to achieve such risk measures.

unchanged can be viewed as worse than the same increase accompanied with an increase in skewness, at least for some risk averse investors.

The return generating process, specified in this chapter, leads to a risk measure of two aspects. Its two aspects correspond to a specific break-down of the portfolio distributions variance to two components. Such a break-down is subject to the control of the investor. While one of the components determines skewness and other higher odd moments of the portfolio distribution, the two components, together, determine the distributions' even moments. The risk aversion of an investor, for the specified joint distribution, is not determined only by specifying his aversion to variance, but by his attitude to its two components. In this setting, variance associated with a high skewness is not as bad as variance associated with no skewness, for some risk averse investors.

While there have been more than one attempt to suggest extensions and alternatives to the positive theory of capital asset pricing, there have been no such attempts in the area of normative portfolio selection. The return generating process, specified in this chapter, leads to a mean-variance-skewness normative theory for portfolio selection. An important feature of this approach is the ability to rank portfolio distribution by the ranking of three control

parameters. Furthermore, the portfolio problem, for any risk averse investor, is reduceable to a mere quadratic programming problem for which the specification of the investor's risk aversion is achieved by determining two of these three control parameters. This leads to a derivation of a mean-variance-skewness efficient set that includes the mean-variance efficient set as a special case. Moreover, the parameters of the joint distribution of asset returns are estimatable from observable asset returns. This leads to a straightforward empirical operationalization of the derived efficient frontier. Such empirical operationalization is of a great value to practitioners. The extension of the results of the normative theory to capital asset pricing is immediate and it is done in Chapter V. The approach of this chapter leads to a three parameter framework for portfolio selection and capital asset pricing that resembles all the desired properties of the mean-variance framework.

The first section of this chapter will introduce the class of distributions employed to describe the return generating process and explores some of its relevant characteristics. In the second section the portfolio problem for this distribution is analyzed and the transformation expected utility maximization problem to a quadratic programming problem is justified. The third section includes the derivation of the efficient set with and without the

riskfree asset. The fourth section discusses the estimation of the distribution parameters from observable asset returns.

3.1 The distribution of assets returns:

The specification of the probabilistic model that generates assets returns is not less difficult than modelling any stochastic process. No matter what parametric distribution we postulate, we can never be certain whether the observable data is truly generated by the postulated model.² A good choice of the probability model has been always a successful compromise between generality and flexibility on the one hand, and mathematical tractability, that drives the analysis far enough to serve its purpose, on the other. The model should be flexible enough to exhibit a stochastic pattern consistent with the observable assets returns. Although a great insight may be reached in analyzing hypothetical unrealistic distributions, the practical application of this analysis is very limited. The mathematical tractability of the model should be sufficient to lead to non-trivial results and allow a reduction of the degrees of freedom of the portfolio problem to a manageable form.

² It may occur to the reader that a Chi-Square test of fitness or the Kolmogorov-Smirnov test allow for testing the truth of the probability model. In fact, the above tests test the consistency of the observable data with the model. But this does not deny the fact that another probability model is the one responsible for generating the data even though the data is consistent with the tested model. The power for such tests are relative to the postulated alternative competing hypothesis. But the specification of the competing hypothesis is a matter of sound judgment of the researcher.

The ability to describe and enumerate the set of efficient portfolios and to operationalize it empirically results in a great reduction in the degrees of freedom of the portfolio problem. Unless it is feasible to estimate the parameters of the distribution with some statistical efficiency, the analysis will be short of providing a normative theory for portfolio selection to practitioners. Nor will it provide testable hypotheses once the normative theory of portfolio selection is extended to determine asset pricing in a given capital market structure. For example, consider the three funds separating distribution described in definition (2.4). The three separating funds x_1, x_2, x_3 that determine the efficient set are themselves parameters of the distribution. Given the generality of its structure it is quite impossible to suggest an estimation for x_1, x_2, x_3 and therefore, the efficient set of portfolios remains a mathematical construct without empirical operationalization. However, the class of spherical distributions has a great mathematical tractability that allows a derivation of an efficient set of portfolios regardless of the utility function. The efficient set depends on the mean and the covariance structure both of which are estimable from the observable returns on assets. The problem of this class is its failure to capture the skewness structure of the assets returns because by definition, assets returns exhibit zero

skewness. But observable assets returns exhibit skewness that cannot be ignored. This motivates an expansion of this class in such a way as to correct for such deficiency on the one hand, and to preserve, as much as possible, the mathematical tractability that allows a comprehensive analysis of the portfolio problem on the other. The distribution described in this section is constructed for such a purpose.

Following the notation of Chapter II, let $z' = (z_1, z_2, \dots, z_n)$ be the vector of assets returns and consider the following structure:

$$z_i = \mu_i + b_i Y + \epsilon_i \quad (3.1)$$

$$i = 1, 2, \dots, n$$

where $\epsilon_i \sim \text{Sp}(0, W)$ (i.e., ϵ_i follow a joint spherical distribution with a zero mean vector and a characteristic $\rightarrow$ matrice W). Y is a random variable that follows non-symmetric distribution. $b' = (b_1, b_2, \dots, b_n)$ and $\mu' = (\mu_1, \mu_2, \dots, \mu_n)$ are vectors of constant parameters. Observe that we introduce Y as a non-spherical noise that underlies all assets. Because Y is skewed and b_i are arbitrary the returns z_i may exhibit any skewness. This gives the distribution a flexibility to capture any skewness of the marginal distribution of assets returns that

spherical distribution fails to capture. The fact that the non-spherical noise is common across assets restricts the coskewness structure of the distribution as will be shown later. Since μ and b are chosen arbitrarily the location and the scale of the distribution of Y can be chosen arbitrarily without any loss of generality.³ For further analysis it will be assumed that $E(Y) = 0$. This allows μ to be the mean vector. Observe that ϵ_i is not a fair game with respect to Y . In fact, the dependence between ϵ_i and Y is not restricted. Before exploring the portfolio problem under this distribution the characteristic function and the moments of the distribution (3.1) will be derived.

Theorem 3.1: The characteristic function of the random vector Z , $\phi_Z(t_1, \dots, t_n)$ is given as:

$$\phi_Z(t) = \exp(it'\mu) \phi_\epsilon(t_1, \dots, t_n) \phi_Y(t'b) \quad (3.2)$$

³ α is a location parameter of the distribution of Y if and only if $f_Y(y-\alpha)$ is independent of α . σ is a scale parameter of the distribution of Y if and only if $f_Y(y|\sigma)$ is independent of σ . In both definitions $f_Y(\cdot)$ is the density function of Y . The choice of the location and the scale of Y is arbitrary because

$$\begin{aligned} z_i &= \mu_i + b_i y + \epsilon_i = \mu_i + \alpha b_i + \sigma b_i \left(\frac{y-\alpha}{\sigma}\right) + \epsilon_i \\ &= \mu'_i + b'_i \left(\frac{y-\alpha}{\sigma}\right) + \epsilon_i \end{aligned}$$

If we choose the non-spherical noise to be $\left(\frac{y-\alpha}{\sigma}\right)$ rather than y , the constant vectors will take the form of μ'_i and b'_i instead of μ_i and b_i .

where $i = \sqrt{-1}$, $\phi_\varepsilon(t)$ and $\phi_Y(s)$ are the characteristic function of the random vector ε and the one-dimensional random variable Y respectively.

Proof: The conditional distribution of Z given $Y = y$ is a spherical with a mean vector $\mu + by$ and a characteristic matrix W . By definition (2.2)

$$\phi_{Z|Y}(t_1, \dots, t_n) = \exp it'(\mu + by) g(t'Wt)$$

where $g(\cdot)$ is any function such that $\phi_{Z|Y}(t)$ is a proper characteristic function. The joint characteristic function of $(Z', Y)' = (z_1, \dots, z_n, y)$ is given as

$$\begin{aligned} \phi_{(Z,Y)}(t_1, t_2, \dots, t_n, s) &= \int \dots \int \exp(it'Z) \cdot \exp(isY) \\ &\quad f(z_1, \dots, z_n, Y) dz_1, \dots, dz_n dy \\ &= \int \dots \int \exp(isY) [\exp(it'Z) (f_{Z|Y}(z_1, \dots, z_n | Y) \cdot \\ &\quad f_Y(Y)) dz_1 \dots dz_n] dy \\ &= \int \exp(isY) \left(\int \dots \int \exp(it'Z) [f_{Z|Y}(z_1, \dots, z_n | Y) \right. \\ &\quad \left. f_Y(Y)] dz_1, \dots, dz_n \right) f_Y(Y) dy \\ &= \int \exp(isY) \phi_{Z|Y}(t_1, \dots, t_n) f_Y(Y) dy \\ &= \int \exp(isY) \exp it'(\mu + bY) g(t'Wt) f_Y(Y) dy \\ &= \exp(it'\mu) g(t'Wt) \int \exp(i(s + t'b)Y) f_Y(Y) dy \end{aligned}$$

$$= \exp(it'\mu)g(t'Wt)\phi_Y((s + t'b)) \quad (3.3)$$

since the marginal characteristic function of Z is given by

$$\phi_Z(t_1, \dots, t_n) = \phi_{Z,Y}(t_1, t_2, \dots, t_n, 0)$$

$$= \exp(it'\mu)g(t'Wt)\phi_Y(t'b)$$

$$= \exp(it'\mu)\phi_\epsilon(t)\phi_Y(t'b) \quad \text{Q.E.D.}$$

The above proof assumes the existence of the density function.

If the density function does not exist, Lebesgue integration may replace the Rieman integration in the proof.

According to (3.2) the joint distribution of Z depends on the mean vector, the joint distribution of ϵ and the vector b . The distribution of the non-spherical noise y determine the nature of the dependence on b . Most important is the fact that the dependence between the non-spherical noise Y and the vector ϵ plays no role in determining the distribution of Z . This is easily seen from (3.3) which gives the joint characteristic function (Z,Y) as a multiplication of the characteristic function of a spherical variable with a mean μ and a characteristic matrix W and the characteristic function of Y at point $(s + it'b)$. This result holds true regardless of the distribution of Y but peculiar to the spherical

property of the distribution of ε . If $\phi_{Z|Y}(t)$ did not take its specific spherical form (3.2) will not be derivable.

It is not possible to derive an explicit form for $f_Z(z_1, \dots, z_n)$ the density function of Z in general as was the case with its characteristic function. Since $Z|Y \sim \text{Sp}(\mu + bY, W)$ the density function of Z conditional on Y is given by:

$$f_{Z|Y}(z_1, \dots, z_n|Y) = C_p |W|^{-\frac{k}{2}} h((Z - \mu - bY)' W (Z - \mu - bY))$$

where $h(\cdot)$ is any non-negative function. The distribution of Z is always given by

$$\begin{aligned} f_Z(z_1, \dots, z_n) &= \int_Y f_{Z,Y}(z_1, \dots, z_n, Y) dY \\ &= \int_Y f_{Z|Y}(z_1, \dots, z_n|Y) f_Y(Y) dY \\ &= C_p |W|^{-\frac{k}{2}} \int_Y h((Z - \mu - bY)' W (Z - \mu - bY)) f_Y(Y) dY \end{aligned} \quad (3.4)$$

The specific form of $f_Z(z_1, \dots, z_n)$ depends on $h(\cdot)$ (i.e., the specific spherical distribution to which ε belongs) and the specific distribution of the non-spherical noise y . Lack of density function for Z is not a great loss.

We could learn a great deal about the distribution of Z from its characteristic function. The next theorem will explore the moment structure of Z .

Theorem 3.2: The first three moments of the vector Z are given by:

$$E(Z) = \mu \quad (3.5)$$

$$\text{Var}(Z) = W + bb' \sigma_Y^2 = V \quad (3.6)$$

$$\begin{aligned} s_{jkl} &= E(z_j - \mu_j)(z_k - \mu_k)(z_l - \mu_l) \\ &= b_j b_k b_l EY^3 \end{aligned} \quad (3.7)$$

Proof: The cumulant generating function of Z is given as:

$$\Psi_Z(t) = \log \phi_Z(t) = it'\mu + \Psi_\epsilon(t) + \Psi_Y(t'b)$$

where $\Psi_\epsilon(t)$ and $\Psi_Y(t)$ are the cumulant generating function of ϵ and Y respectively.

$$\begin{aligned} E(Z) &= i^{-1} \frac{d}{dt} \Psi_Z(t) \Big|_{t=0} \\ &= i^{-1} [i\mu + \Psi'_\epsilon(0) + ib\Psi'_Y(0)] = \mu \end{aligned}$$

because $\Psi'_\epsilon(0) = E(\epsilon) = \Psi'_Y(0) = E(Y) = 0$ ⁴

$$\begin{aligned}\text{Var}(Z) &= -\frac{d^2\psi_Z(t)}{dt(dt)'} \bigg|_{t=0} = \frac{-d^2\psi_\epsilon(t)}{dt(dt)'} \bigg|_{t=0} + bb'(-\psi''_Y(0)) \\ &= W + bb'\sigma_Y^2\end{aligned}$$

and

$$\begin{aligned}s_{jkl} &= -i^3 \frac{d^3\psi_Z(t)}{dt_j dt_k dt_l} \bigg|_{t=0} = -i^3 \frac{d^3\psi_\epsilon(t)}{dt_j dt_k dt_l} \bigg|_{t=0} + b_j b_k b_l i^3 \psi_Y'''(0) \\ &= b_j b_k b_l E(Y^3)\end{aligned}$$

The first term is zero because ϵ is spherically distributed.

If we denote by $c_Z(n)$ the n^{th} cumulant of Z and $c_\epsilon(n)$ and $c_Y(n)$ the respective cumulants of ϵ and Y we always have

$$c_{Z_j}(n) = c_{\epsilon_j}(n) + b_j^n \cdot c_Y(n) \quad (3.8)$$

A similar result holds for the cross cumulants. Observe how the cross cumulants between ϵ_j and Y play no role in determining the cumulants of Z and hence the moments of Z . One of the interesting properties of this

⁴When $E(Y) \neq 0$ then $E(Z) = \mu + bE(Y)$

distribution is the fact that its co-skewness is constant and equal to 1. If we define a coefficient of co-skewness as an extension to the definition of the correlation coefficient, then it is equal to:

$$\begin{aligned} \text{COSS}(z_i, z_j, z_k) &= \frac{E(z_i - \mu_i)(z_j - \mu_j)(z_k - \mu_k)}{[E(z_i - \mu_i)^3 E(z_j - \mu_j)^3 E(z_k - \mu_k)^3]^{1/3}} \\ &= \frac{b_i b_j b_k EY^3}{(b_i^3 EY^3 b_j^3 EY^3 b_k^3 EY^3)^{1/3}} = 1 \end{aligned} \quad (3.9)$$

(3.9) holds regardless of the distribution of Y as long as ϵ is spherically distributed. In fact (3.9) is a byproduct of the assumption that the skewness structure of assets returns is generated by one common source: the non-spherical noise Y .⁵

3.2 The nature of the portfolio problem:

The choice of (3.1) to model the stochastic process that describes assets returns greatly simplifies the portfolio problem for two reasons. First, the linearity of the

⁵If we define

$$\text{CORR}(z_1, z_2, \dots, z_k) = E \prod_{i=1}^k (z_i - \mu_i) / \left(\prod_{i=1}^k E(z_i - \mu_i)^k \right)^{1/k}.$$

It is straightforward to show that $\text{CORR}(z_1, \dots, z_k) = 1$ for odd $k \geq 3$ because $\text{CORR}(\epsilon_1, \epsilon_2, \dots, \epsilon_k) = 0$ for all odd k . Again this is a byproduct of the fact that Y is responsible for generating all odd moments of the vector Z .

non-spherical noise Y and ϵ_i leads to a three funds separating distribution when ϵ_i follows a two funds separating distribution. Second, the assumption that ϵ_i follows a spherical distribution (a two funds separating distribution) preserves the mathematical elegance of mean-variance analysis. As in the mean-variance framework of analysis, the portfolio problem for any utility function is reduceable to a quadratic programming problem that depends on distribution parameters estimatable from observable returns on assets. The expected utility for any portfolio X depends on three parameters, the mean and the variance of the portfolio return and a parameter that measures the non-sphericity of the portfolio distribution. Controlling the distribution of portfolio return is limited to the control of these three parameters. Moreover, the variance of the portfolio return is broken into a spherical variance and non-spherical variance. It will be shown that the expected utility is decreasing with the spherical variance. This fact allows the reduction of the portfolio problem to a mere quadratic problem. The separating funds are explicitly derivable and estimatable from observable assets returns, and therefore, the efficient set of portfolios is derivable explicitly as a function of parameters estimatable from the observable data. Theorem (3.3) to Theorem (3.6) establish the possibility to transform the general portfolio problem under (3.1) to a quadratic programming problem.

Theorem 3.3: The expected utility of the rate of return on any portfolio X , $Eu(Z'X)$ is determined by $X'\mu$, $X'b$ and $X'WX$ for any utility function.

Proof: $Eu(X'Z)$ is determined by $X'\mu$, $X'b$, $X'WX$ if and only if the distribution of $X'Z$ is determined by $X'\mu$, $X'b$, $X'WX$ for any portfolio X . Since the characteristic function of any random variable completely determines its distribution, it is sufficient to prove that the characteristic function of $X'z$ is determined by $X'\mu$, $X'b$ and $X'WX$.

$$\begin{aligned}\phi_{X'Z}(s) &= E \exp(is(X'Z)) = E \exp(is \sum_{j=1}^n x_j z_j) \\ &= E \exp(i \sum_{j=1}^n (sx_j) z_j) = \phi_Z(x_1 s, x_2 s, \dots, x_n s)\end{aligned}$$

But by (3.2) $\phi_Z(t) = \exp(it'\mu) \phi_\epsilon(t) \phi_Y(t'b)$. Therefore:

$$\phi_{X'Z}(s) = \exp(iX'\mu s) \phi_\epsilon(x_1 s, \dots, x_n s) \phi(X'bs)$$

But by definition (2.2) we have $\phi_\epsilon(t) = g(t'Wt)$. Therefore

$$\begin{aligned}\phi_{X'Z}(s) &= \exp(iX'\mu s) g(s^2 X'WX) \phi_Y(X'bs) \\ &= h(X'\mu, X'WX, X'b) \quad \text{Q.E.D.} \quad (3.10)\end{aligned}$$

Corollary 3.3.1: The expected utility of any portfolio return is determined by its mean $X'\mu$, its variance, $X'VX$, ($V = \text{Var}(Z)$) and its non-sphericity parameter $X'b$.

Proof: By (3.6) of Theorem (3.2)

$$V = W + bb' + \sigma_Y^2$$

and

$$X'VX = X'WX + X'bb'X\sigma_Y^2 = X'WX + (X'b)^2\sigma_Y^2$$

therefore $X'WX = X'VX - (X'b)^2\sigma_Y^2$ ⁶

A substitution of $X'WX$ in (3.10) completes the proof. Q.E.D.

Note that the variance of the portfolio return is decomposable to two parts, a spherical variance $X'WX$ and non-spherical variance $(X'b)^2\sigma_Y^2$. Since the expected utility depends on $(X'b)$, the specific breakdown of the variance of the portfolio return is very relevant for the portfolio choice. In the conventional mean-variance framework both variance components are considered, without

⁶Observe that the distribution of the portfolio return, and hence its expected utility, may depend on the parameters of the distribution of the non-spherical noise Y . But it will still be determined by $X'\mu$, $X'VX$ and $X'b$ for all X since the control of the portfolio distribution from the investor side ends by controlling $X'\mu$, $X'VX$, and $X'b$, whereas the parameters of the distribution of the non-spherical noise are exogenous.

discrimination, as undesirable. But in this framework only the non-spherical variance $X'WX$ is undesirable, as will be proved later in this section. Within this framework the investor taste is specified by deciding on his expected return on the one hand, and deciding on the non-spherical feature of his portfolio distribution on the other. The latter decision determines the specific breakdown of the variance to spherical and non-spherical components. It is worth noting that such a breakdown exists regardless of the covariance structure between the non-spherical noise η and the spherical noise ϵ_i that generates assets returns. This is again due to the sphericity of ϵ_i that allows the proof of Theorem (3.1) and (3.2) thereafter.

The fact that the expected utility is determined by $X'\mu$, $X'WX$, and $X'b$ implies that these three control parameters are sufficient for the ranking of portfolio distributions. If we prove that the expected utility is decreasing with $X'WX$ for risk averse investors, then the portfolio problem will be transformed to a quadratic programming problem. To prove this result we need Theorem (3.4) given by H. Levy (1969).

Theorem 3.4: Let F_1, F_2 be two distribution functions--
 $F_1(t) = P_r(T_i \leq t)$ -- with means μ_1, μ_2 respectively
 such that for some $t_0 < \infty$ $F_1 \leq F_2$ (and $F_1 < F_2$ for
 some $t < t_0$) and $F_1 > F_2$ for $t \geq t_0$. Then F_1

stochastically dominates F_2 for all concave and increasing $u(t)$ if and only if $\mu_1 = \mu_2$.

The proof of this theorem will be omitted. It is theorem 3 of H. Levy (1969).

Theorem 3.5: For a given $X'\mu = E$ and $X'b = B$ the expected utility $Eu(X'Z)$ is decreasing with the spherical component of the variance of the portfolio return, $X'WX$, for all concave u .

Proof:

$$\begin{aligned} Eu(X'Z) &= Eu(X'\mu + X'by + X'\epsilon) = Eu(E + BY + X'\epsilon) \\ &= E_Y(E_\epsilon|_Y u(E + BY + X'\epsilon)) \end{aligned}$$

First it will be shown that $E_\epsilon|_Y u(E + BY + X'\epsilon)$ is decreasing in $X'WX$. Let $x'_1\epsilon$, and $x'_2\epsilon$ be returns on two portfolios such that $\sigma_1^2 = x'_1 W x_1 < x'_2 W x_2 = \sigma_2^2$.

Utilizing Theorem (3.4) it will be shown that $E_\epsilon|_Y u(x'_1 Z) > E_\epsilon|_Y u(x'_2 Z)$. Since $Z|Y \sim Sp(\mu + By, W)$ the conditional distribution of the portfolio return $X'Z$ on Y is a spherical, that is $x'_1 Z|Y \sim Sp(E + By, x'_1 W x_1)$. Since the distribution of $X'Z|Y$ is determined by $X'\mu/y$ and $X'WX$ for any portfolio X the variable $T =$

$\frac{X'Z - X'\mu/y}{\sqrt{X'WX}} \sim Sp(0,1)$ regardless of X . Denote by F the distribution function of T and F_1 , the distribution function of $x'_1 Z|Y$. Then

$$F_i(t) = F\left(\frac{t - E - BY}{\sigma_i}\right) \quad \text{for } i = 1, 2$$

For $t < t_0 = E + BY$

$$F_1(t) - F_2(t) = F\left(\frac{t - E - BY}{\sigma_1}\right) - F\left(\frac{t - E - BY}{\sigma_2}\right) \leq 0$$

Because $\frac{t - E - BY}{\sigma_1} < \frac{t - E - BY}{\sigma_2}$ and F is monotonically decreasing.

For $t > t_0 = E + BY$

$$F_1(t) - F_2(t) = F\left(\frac{t - E - BY}{\sigma_1}\right) - F\left(\frac{t - E - BY}{\sigma_2}\right) > 0$$

Because $\frac{t - E - BY}{\sigma_1} > \frac{t - E - BY}{\sigma_2}$ and F is monotonically

decreasing. By Theorem (3.4) $x_1'Z|Y$ stochastically dominates $x_2'Z|Y$ for all concave u . This implies that for $x_1'Wx_1 < x_2'Wx_2$

$$E_{\varepsilon|Y} u(E + BY + x_1'\varepsilon) \geq E_{\varepsilon|Y} u(E + BY + x_2'\varepsilon)$$

Therefore $E_{\varepsilon|Y} u(E + BY, X'\varepsilon) = h(E + by, X'WX)$ is decreasing in $X'WX$. But

$$Eu(X'Z) = E_Y(E_\varepsilon|_Y u(X'Z)) = \int h(E + BY, X'WX) dF(Y)$$

But one of the fundamental properties of Lebesgue integral is:

$$f(x) > 0 \text{ a.e. implies } \int h(wdF) > 0$$

This completes the proof.

Q.E.D

3.3 The derivation of the efficient set:

The theorems of the previous section allow a derivation of the efficient set of portfolios for all concave utility functions. It has been shown that the expected utility depends on three control parameters $X'\mu$, $X'b$ and $X'WX$. Moreover, the expected utility is decreasing in one of the control parameters, namely $X'WX$. Therefore, for fixed values of the two control parameters $X'\mu$, $X'b$, a portfolio X will maximize the expected utility, if and only if, it minimizes $X'WX$. Let X^* be the optimal portfolio for a utility function u and let $b'X^* = B(X^*)$ and $\mu'X^* = E(X^*)$. Then $X^{*'}WX^* < X'WX$ for all

$X'\mu = E(X^*)$ and $X'b = B(X^*)$. If we allow u to vary over the set of all utility functions, then X^* will vary over the utility functions as well as $E(X^*)$ and $B(X^*)$. Therefore, the set of solutions to the following parametric quadratic programming problem

$$\begin{array}{ll} \underset{X}{\text{minimize}} & \frac{1}{2} X'WX \\ \text{subject to:} & X'\mu = E \\ & X'e = 1 \\ & X'b = B \end{array}$$

includes the optimal portfolio for any concave utility function. Since we have no restrictions on short sales, the solution to the above parametric quadratic problem can be derived explicitly as a function of the control parameters E , and B and the assets returns distribution parameters by a conventional Lagrangian technique. First the efficient set will be derived when all assets are risky. Second a risk-free asset will be added to the other assets.

3.3.1 The case with no risk free asset: The lagrangian and the first order condition are:

$$L = \frac{1}{2}X'WX + \lambda_1(E - X'\mu) + \lambda_2(1 - X'e) + \lambda_3(B - X'b)$$

$$\frac{dL}{dX} = WX - \lambda_1 \mu - \lambda_2 e - \lambda_3 b = 0 \quad (3.11)$$

If W is invertable:

$$X = \lambda_1 W^{-1} \mu + \lambda_2 W^{-1} e + \lambda_3 W^{-1} b \quad (3.12)$$

Since W is a covariance matrix and assumed non-singular, it is a positive definite and $X'WX$ is a convex function in X . Therefore X given by (3.12) is a global minimum. The above result is summarized by the following theorem:

Theorem 3.6: The efficient set of portfolios for distribution (3.1) is spanned by the following three funds.

$$\begin{aligned} \alpha_1 &= (W^{-1} \mu / e' W^{-1} \mu) ; \quad \alpha_2 = (W^{-1} e / e' W^{-1} e) ; \\ \alpha_3 &= (W^{-1} b / e' W^{-1} b) . \end{aligned}$$

The separation property itself is expected by the structure of the joint distribution (3.1). It is not surprising since the spherical distribution of ϵ is a two funds separating distribution and the assets return according to (3.1) are modelled as a linear combination of spherical variable and a common non-spherical noise. The three funds separation is a byproduct of the linearity of the common

non-spherical noise with the spherical one. The importance of Theorem (3.6) comes from the explicit specification of the separating funds as a function of the assets returns distribution rather than being parameters themselves as was the case of the general three funds separating distribution given by definition (2.4). Once we have estimates for μ , b , and W a menu of efficient portfolios is provided as the case of mean-variance framework. The only difference is that this menu is richer because it offers for any pair (E, B) an optimal portfolio rather than an optimal portfolio for any E as in a conventional mean-variance setting. Although B represents the non-sphericity of the distribution, we can think of it as a proxy for the skewness of the portfolio return. Therefore the efficient set described in (3.12) provides an optimal portfolio for different means and skewness measures. In this case the investor preference can be defined by complete ordering of the set of the cartesian product of E and B ⁷ only. It is easier and more intuitive for an investor

⁷This is because all portfolios with parameters E and B , except one portfolio, are not efficient. This only efficient portfolio is the one that minimizes $X'WX$ subject to these parameters. Once the investor determines E and B his optimal portfolio is determined. Therefore, there is no loss of generality if such an investor ranks only the (E, B) pairs, and then chooses the pair that minimizes his expected utility.

to specify a utility function over E and B (i.e., his expected utility) than to specify its original utility function that ranks the set of distributions of portfolio returns.

The efficient set of portfolios described in Theorem (3.6) includes the mean-variance efficient set as a special case. Corollary (3.2) gives an alternative description to the efficient set as a combination of a mean-variance efficient portfolio and a non-mean-variance efficient fund.

Corollary 3.2: Any efficient portfolio is given as a combination of two mean-variance efficient funds $a_1 = (V^{-1}\mu/e'V^{-1}\mu)$, $a_2 = (V^{-1}e/e'V^{-1}e)$ and a mean-variance inefficient fund $a_3 = (V^{-1}b/e'V^{-1}b)$.

Proof: From (2.9) in section (2.4) we know that any mean-variance efficient portfolio is given as a combination of a_1 and a_2 . On the other hand, any efficient portfolio satisfies (3.11.a) written below:

$$WX - \lambda_1\mu - \lambda_2e - \lambda_3b = 0 \quad (3.11.a)$$

(3.11) is equivalent to

develop his preference as a function of E , and B . No matter what is his utility function it will be a function of E , B , $X'WX$ and decreasing with $X'WX$.

$$(W + bb'\sigma_Y^2) - bb'\sigma_Y^2 X - \lambda_1 \mu - \lambda_2 e - \lambda_3 b = 0$$

by (3.6) $W + bb'\sigma_Y^2 = V$. Therefore the first order condition (3.11) is given by

$$VX - \lambda_1 \mu - \lambda_2 e - b(\lambda_3 + b'X\sigma_Y^2) = 0 \quad (3.11.b)$$

or

$$VX = \lambda_1 \mu + \lambda_2 e + (\lambda_3 + B\sigma_Y^2)b$$

and

$$X = \lambda_1 V^{-1} \mu + \lambda_2 V^{-1} e + \lambda_3' V^{-1} b \quad \text{where } \lambda_3' = \lambda_3 + B\sigma_Y^2.$$

Therefore

$$X = \gamma_1 a_1 + \gamma_2 a_2 + \gamma_3 a_3 \quad \text{Q.E.D.} \quad (3.13)$$

The efficient set of portfolios is determined by μ , b , and W according to (3.12) or by μ , b and V according to (3.13). In each case the parameters of the non-spherical noise Y do not appear in either (3.12) or (3.13) and thus they do not influence the efficient set.⁸

⁸This does not rule out the dependence of the weights γ_1 , γ_2 , γ_3 on the parameters of Y . Since the expected utility is a function of the parameters of Y , the maximization of the expected utility may lead to weights that depend on the parameters of Y . CARA utility function in Chapter IV with the distribution assumed there provides an example.

The description of the efficient set in (3.13) has some advantages over the description in (3.12). First, equation (3.13) describes the efficient set in terms of the covariance matrix of the observable assets returns Z rather than the covariance matrix of the unobservable random vector ϵ . Second, it describes any efficient portfolio as a portfolio of a mean-variance efficient fund determined by a_1 and a_2 , and a mean-variance inefficient fund, namely, a_3 . This makes it clear that the mean-variance efficient set is a subset of the efficient set of portfolios.

Since the portfolio distribution is determined by $X'\mu = E$, $X'b = B$ and $X'VX = \sigma_p^2$ for any vector X , the efficient set can be described in the three dimensional parameter space determined by E , B and σ_p^2 rather than the n dimensional portfolio space. If we let $\gamma_2 = 1 - \gamma_1 - \gamma_3$ in (3.13) we have

$$(X - a_2) = \gamma_1(a_1 - a_2) + \gamma_3(a_3 - a_2) \quad (3.14)$$

If we multiply (3.14) by μ' and b' respectively and let $E_i = \mu'a_i$ and $B_i = b'a_i$ for $i = 1, 2, 3$ we have the following system of two equations in two unknowns:

Moreover, if the parameters of the distribution are estimated jointly, then the estimates of μ , b and v may depend on the estimates of the parameters of the distribution of Y and therefore the "empirical" efficient set may depend on the estimates of the parameters of Y .

$$(E - E_2) = \lambda_1 (E_1 - E_2) + \lambda_3 (E_3 - E_2) \quad (3.15)$$

$$(B - B_2) = \lambda_1 (B_1 - B_2) + \lambda_3 (B_3 - B_2) \quad (3.16)$$

On the other hand, the variance of the portfolio is given as:

$$\sigma_p^2 = X'VX = \left(\sum_{i=1}^3 \gamma_i a_i \right)' V \left(\sum_{i=1}^3 \gamma_i a_i \right) = \sum_{i=1}^3 \sum_{j=1}^3 (a_i' V a_j) \gamma_i \gamma_j \quad (3.17)$$

By solving (3.15) and (3.16) for γ_1 and γ_2 and substituting the solution in (3.17) it is shown in Appendix (3.A) that

$$\begin{aligned} \sigma_p^2 = & \frac{1}{e'V^{-1}e} + \frac{1}{e'V^{-1}e} \left(\frac{E_1}{E_2} - 1 \right) \left(\frac{E-E_2}{E_1-E_2} \right)^2 \\ & + \frac{1}{e'V^{-1}e} \left(\frac{B_1}{B_2} - 1 \right) \frac{1}{\frac{B_3-B_2}{B_1-B_2} - \frac{E_3-E_2}{E_1-E_2}} \left[\frac{B-B_2}{B_1-B_2} - \frac{E-E_2}{E_1-E_2} \right]^2 \end{aligned} \quad (3.18)$$

The first term of the right hand side of (3.18) is the variance of a_2 , the global minimum variance portfolio. The first two terms together give the variance to the mean-

variance efficient portfolio for a given mean E .⁹ The horizontal distance of any efficient portfolio from the mean-variance efficient frontier in (E, σ_p) plane,

$$|\sigma_p^2 - \sigma_{mVE}^2|^{\frac{1}{2}}, \text{ is linear in } \left| \frac{B-B_2}{B_1-B_2} - \frac{E-E_2}{E_1-E_2} \right|. \text{ The larger}$$

the difference between the two point elasticities of B and E relative to their distance from the global minimum variance portfolio, the larger is the horizontal distance between the efficient portfolio and the mean-variance efficient frontier.

It is possible to describe the efficient set in the three dimensional space E, B and the spherical variance of the portfolio distribution $\sigma_{Sp}^2 = X'WX$. Let $\eta_i = \mu' \alpha_i$ and $\beta_i = b' \alpha_i$ where α_i is given in Theorem (3.6). Replacing E_i, B_i and V by η_i, β_i and W respectively in Appendix 3.A leads to:

$$\begin{aligned} \sigma_{Sp}^2 = X'WX &= \frac{1}{e'W^{-1}e} + \frac{1}{e'W^{-1}e} \left(\frac{\eta_1}{\eta_2} - 1 \right) \left(\frac{E-\eta_2}{\eta_1-\eta_2} \right)^2 \\ &+ \frac{1}{e'W^{-1}e} \left(\frac{\beta_1}{\beta_2} - 1 \right) \frac{\left[\frac{\beta_1-\beta_2}{\beta_1-\beta_2} - \frac{E-\eta_2}{\eta_1-\eta_2} \right]^2}{\frac{\beta_3-\beta_2}{\beta_1-\beta_2} - \frac{\eta_3-\eta_2}{\eta_1-\eta_2}} \end{aligned} \quad (3.19)$$

⁹Any mean-variance efficient portfolio is given by $X =$

Again $\frac{1}{e'W^{-1}e}$ is the variance of the portfolio that globally minimizes the spherical variance. For a given σ_{Sp}^2 and B there are two values of E that satisfy (3.19). However, only the portfolio with the highest mean is efficient. Also for a given σ_{Sp}^2 and E there are two values of B that satisfy (3.19). If the expected utility is monotonically non-decreasing in B , only the one with highest B is efficient. It is also feasible to express σ_{Sp}^2 in terms of E_i, B_i and V since by

$$(3.6) \quad V = W + \sigma_Y^2 b b'.$$

$$\begin{aligned} X'WX &= X'(V - \sigma_Y^2 b b') = X'VX - \sigma_Y^2 X'b b'X \\ &= \sigma_P^2 - \sigma_Y^2 (X'b)^2 = \sigma_P^2 - \sigma_Y^2 B^2 \end{aligned} \quad (3.20)$$

3.3.2 The case with risk free asset: Here we introduce a riskfree asset indexed as asset zero with a fixed rate r_f . Let x_0 be the investment position in this asset. $X = (x_1, \dots, x_n)$ is still an n dimensional vector that denotes

$$\begin{aligned} & \left(\frac{E-E_2}{E_1-E_2} \right) a_1 + a_2 \quad \text{and} \quad X'VX = \left(\left(\frac{E-E_2}{E_1-E_2} \right) a_1' + a_2' \right) V \left(\left(\frac{E-E_2}{E_1-E_2} \right) a_1 + a_2 \right) \\ &= \frac{1}{e'V^{-1}e} + \frac{1}{e'V^{-1}e} \left(\frac{E_1}{E_2} - 1 \right) \left(\frac{E-E_2}{E_1-E_2} \right)^2. \end{aligned}$$

the investment position in the n risky assets. The efficient set is derived by solving the following problem:

$$\min_X \frac{1}{2} X'WX$$

subject to:

$$X'\mu + r_f x_0 = E$$

$$X'e + x_0 = 1$$

$$X'b = B$$

The lagrangian and the first order condition are given as:

$$\begin{aligned} L = & \frac{1}{2} X'WX + \lambda_1 (E - X'\mu - x_0 r_f) + \lambda_2 (1 - X'e - x_0) \\ & + \lambda_3 (B - X'b) \end{aligned}$$

$$\frac{\partial L}{\partial x_0} = -\lambda_1 r_f - \lambda_2 = 0 \quad (3.21)$$

$$\frac{\partial L}{\partial X} = WX - \lambda_1 \mu - \lambda_2 e - \lambda_3 b = 0$$

Replacing W by $V - bb'\sigma_Y^2$ and solving for X and x_0

$$\begin{aligned} X &= \lambda_1 V^{-1} (\mu - r_f e) + (\lambda_3 + b' X \sigma_Y^2) V^{-1} b \\ &= \lambda_1 V^{-1} (\mu - r_f e) + (\lambda_3 + B \sigma_Y^2) V^{-1} b \end{aligned}$$

$$X = \gamma_1 \frac{V^{-1}(\mu - r_f e)}{e' V^{-1}(\mu - r_f e)} + \gamma_3 \frac{V^{-1}b}{e' V^{-1}b}$$

and

$$x_0 = 1 - X'e = 1 - \gamma_1 - \gamma_3$$

Again we have the three funds that span the efficient set, the riskfree asset, $\frac{V^{-1}(\mu - r_f e)}{e' V^{-1}(\mu - r_f e)}$ and a_3 .

The first two are mean-variance efficient and the third fund is mean-variance inefficient. Appendix 3.B includes a derivation of the efficient set, in the three parameter space. It is given as:

$$\begin{aligned} \sigma_p^2 = & \frac{E_1 - r_f}{e' V^{-1}(\mu - r_f e)} \left(\frac{E - r_f}{E_1 - r_f} \right)^2 \\ & + \frac{(E_1 - r_f)}{e' V^{-1}b[(E_1 - r_f)B_3 - (E_3 - r_f)B_1]} [(E_1 - r_f)B - B_1(E - r_f)]^2 \end{aligned} \quad (3.22)$$

Again the first term is the variance of the mean-variance efficient portfolio for a given mean E . Although the mean-variance efficient set is a straight line in (E, σ_p) plane the efficient set is not. Although there is one mean-variance efficient portfolio with expected

return $E = r_f$, there are other portfolios with different values of B that could be theoretically efficient. An extremely risk averse investor faced with spherical assets returns will end up investing all his money in the risk free asset. It is interesting to investigate whether this is also true when he is faced with our distribution. Recall that the spherical variance σ_{Sp}^2 is given by:

$$\sigma_{Sp}^2 = X'WX = X'VX - \sigma_Y^2 X'bb'X = \sigma_p^2 - \sigma_Y^2 B^2 \quad (3.23)$$

An extremely risk averse investor tends to choose a position with $X'WX = 0$ rather than $X'VX = 0$. Inspection in (3.22) and (3.23) reveal that there are different combinations of E and B that result in $\sigma_{Sp}^2 = 0$.

Therefore, an extremely risk averse investor may not necessarily confine himself to the riskfree asset as is the case in the mean-variance framework. If W is singular and V is not and the expected utility is increasing in B , then an extremely risk averse investor will minimize his global spherical variance by choosing a risky portfolio that has the same spherical risk as the riskfree asset but with a positive B . The special case treated in Chapter IV provides an example in which the expected utility is non-increasing in B . This rules out this theoretical possibility and results in a tendency toward the riskfree asset

as the risk aversion increases.

3.4 Estimation of the distribution parameters:

It is desirable to suggest estimates for μ , b and V without further specification of the parametric distribution of the spherical vector ϵ or the non-spherical noise Y . One way of doing this is through equating the sample moments to the distribution moments. According to (3.5), (3.6) and (3.7) we have

$$E(Z) = \mu \quad (3.5)$$

$$COV(Z) = W + bb'\sigma_Y^2 = V \quad (3.6)$$

$$E(z_i - \mu_i)^3 = E(Y^3) b_i^3 \quad (3.7)$$

The sample mean vector is a moment estimate for μ . Such an estimate is unbiased for $E(Z)$ regardless of the distribution of Z . The sample covariance matrix of the observable assets returns is a moment estimate for V and unbiased regardless of the distribution of Z (this holds even when Z does not follow (3.1). If we know the third moment of Y then we have a moment estimate for b_i as

$$\hat{b}_i = \left(\frac{m_{3i}}{EY^3} \right)^{1/3}$$

where m_{3i} is the sample third moment of z_i given as

$$m_{3i} = \frac{N}{(N-1)(N-2)} \sum_{j=1}^N (z_{ij} - \bar{z}_i)^3$$

Again m_{3i} is an unbiased estimate for $E(z_i - \mu_i)^3$ and therefore, for a given value of $E(Y^3)$, $\hat{b}_i$ is an unbiased estimate for b_i . In this approach EY^3 and σ_Y^2 should be specified a priori. However there are restrictions on the specification of $E(Y^3)$ and σ_Y^2 . The estimates b_i , brought by such specification, must lead to a positive definite matrix estimate for the matrix W . If we estimate $\hat{b}_i$ according to (3.23) the estimate for W will be

$$\hat{W} = \hat{V} - \hat{b}\hat{b}'\sigma_Y^2 = \hat{V} - m \cdot m' \left(\frac{\sigma_Y^3}{EY^3} \right)^{2/3} \quad (3.24)$$

where $m_i = (m_{3i})^{1/3}$ and $m' = (m_1, \dots, m_n)$. Recall that $EY^3/\sigma_Y^3 = \alpha_{3Y}$, the coefficient of skewness of Y . Therefore

$$\hat{W} = \hat{V} - m \cdot m' (\alpha_{3Y})^{-2/3}$$

The coefficient of skewness for Y should be large enough to make $\hat{W}$ positive definite for any estimate of V . A choice of a large coefficient of skewness of Y

relative to the highest coefficient of skewness of all assets, typically assures the positive definite estimate for $\hat{W}$.¹⁰

If the distribution of both ϵ and Y are specified then, theoretically, it is possible to derive maximum likelihood estimates for μ , b , V and the parameters of the distribution of Y . The mathematical tractability and the practicality of such an approach depend on the specified distributions. But in general we are able to write the likelihood of the sample. Let $z_1, z_2, \dots, z_N$ be the N independent sample points of the vector of assets return Z . Since $z_i | Y \sim \text{Sp}(\mu + bY, W)$ for $i = 1, 2, \dots, N$

$$f(z_1, \dots, z_N | Y) = C_p | W |^{-\frac{N}{2}} \prod_{i=1}^N h(z_i - \mu - bY)' W (z_i - \mu - bY)$$

The likelihood function is given by integrating $f(z_1, \dots, z_N | Y)$ with respect to the distribution of Y .

$$^{10} \text{For } \hat{W}_i^2 = \hat{v}_i^2 - \frac{m_{3i}}{\alpha_{3Y}}^{2/3} = v_i^2 [1 - (\frac{\alpha_{3i}}{\alpha_{3Y}})^{2/3}] \text{ to}$$

be positive for all assets, it is sufficient that $\alpha_{3Y} > \max\{\alpha_{3i}\}$. But this will not necessarily assure that $\hat{W}$ itself is a positive definite matrix--such a requirement is necessary for W to be a characteristic matrix of the spherical vector ϵ . The empirical experience of this study suggests that α_{3Y} should be at least about 8 times larger than $\max\{\alpha_{3i}\}$ in order for $\hat{W}$ to be a positive definite matrix.

$$\begin{aligned}
 L(\mu, b, W, \theta | Z_1, \dots, Z_N) \\
 = C_P |W|^{-\frac{N}{2}} \int \prod_{i=1}^N h(Z_i - \mu - bY) 'W(Z_i - \mu - bY) dF(Y) \quad (3.25)
 \end{aligned}$$

where θ is a vector of parameters for the distribution of Y . The maximum likelihood estimates are derived by maximizing L over μ, b, W, θ . The practicality of such an approach depends on the function $h(\cdot)$ and the distribution of Y . For the purpose of Chapter IV a mixture of the moments method and the maximum likelihood approach is applied.

APPENDIX 3.A

According to (3.17)

$$\begin{aligned}\sigma_p^2 = & a_1' v a_1 \gamma_1^2 + a_2' v a_2 \gamma_2^2 + a_3' v a_3 \gamma_3^2 + 2a_1' v a_2 \gamma_1 \gamma_2 \\ & + 2a_1' v a_3 \gamma_1 \gamma_3 + 2a_2' v a_3 \gamma_2 \gamma_3\end{aligned}\quad (3.A.1)$$

By the definition of a_1 , a_2 and a_3 we have

$$\begin{aligned}a_1' v a_1 &= \frac{\mu' v \mu}{(e' v^{-1} e)^2} = \frac{1}{e' v^{-1} e} \left(\frac{E_1}{E_2} \right) \\ a_3' v^{-1} a_3 &= \frac{b' v^{-1} b}{(e' v^{-1} b)^2} = \left(\frac{1}{e' v^{-1} e} \right) \left(\frac{B_3}{B_2} \right) \\ a_2' v a_2 &= \frac{1}{e' v^{-1} e} = a_1' v a_2 = a_2' v a_3 \\ a_1' v a_3 &= \frac{1}{e' v^{-1} e} \frac{E_3}{E_2} = \frac{1}{e' v^{-1} e} \left(\frac{B_1}{B_2} \right)\end{aligned}\quad (3.A.2)$$

Therefore

$$\begin{aligned}\sigma_p^2 = & \frac{1}{e' v^{-1} e} \gamma_2^2 + 2\gamma_2(\gamma_1 + \gamma_3) + \left(\frac{E_1}{E_2} \right) \gamma_1^2 + \left(\frac{B_2}{B_2} \right) \gamma_3^2 \\ & + 2\left(\frac{E_3}{E_2} \right) \gamma_1 \gamma_3\end{aligned}$$

Utilizing the identity

$$\gamma_2^2 + 2\gamma_2(\gamma_1 + \gamma_3) = (\gamma_1 + \gamma_2 + \gamma_3)^2 - (\gamma_1 + \gamma_3)^2 = 1 - (\gamma_1 + \gamma_3)^2$$

we have

$$\begin{aligned}\sigma_p^2 &= \frac{1}{e'v^{-1}e} \left[1 + \frac{E_1 - E_2}{E_2} \gamma_1^2 + 2 \frac{E_3 - E_2}{E_2} \gamma_1 \gamma_3 + \frac{B_3 - B_2}{B_2} \gamma_3^2 \right] \\ &= \frac{1}{e'v^{-1}e} \left[1 + \frac{(E_1 - E_2)^2 \gamma_1^2 + 2(E_3 - E_2) \gamma_1 \gamma_3 + (E_3 - E_2)^2 \gamma_3^2 - (E_3 - E_2)^2 \gamma_3^2}{E_2(E_1 - E_2)} \right. \\ &\quad \left. + \left(\frac{B_3 - B_2}{B_2} \right) \gamma_3^2 \right] \\ &= \frac{1}{e'v^{-1}e} \left[1 + \frac{((E_1 - E_2) \gamma_1 + (E_3 - E_2) \gamma_3)^2}{E_2(E_1 - E_2)} \right. \\ &\quad \left. + \left(\frac{B_3 - B_2}{B_2} - \frac{(E_3 - E_2)^2}{E_2(E_1 - E_2)} \right) \gamma_3^2 \right]\end{aligned}$$

Utilizing (3.15) and the fact that $\frac{E_3 - E_2}{E_2} = \frac{B_1 - B_2}{B_2}$,

according to (3.A.2), we have

$$\begin{aligned}\sigma_p^2 &= \frac{1}{e'v^{-1}e} \left[1 + \frac{E_1 - E_2}{E_2} + \frac{(E_1 - E_2)^2}{(E_1 - E_2)^2} \right. \\ &\quad \left. + \frac{B_1 - B_2}{B_2} \left(\frac{B_3 - B_2}{B_1 - B_2} - \frac{E_3 - E_2}{E_1 - E_2} \right) \gamma_3^2 \right] \quad (3.A.4)\end{aligned}$$

Observe that when $\gamma_3=0$ then we get the conventional mean-variance efficient set in E, σ_p^2 plane. If we solve (3.15) and (3.16) for γ_3 we have

$$\gamma_3 = \left(\frac{B-B_2}{B_1-B_2} - \frac{E-E_2}{E_1-E_2} \right) / \left(\frac{B_3-B_2}{B_1-B_2} - \frac{E_3-E_2}{E_1-E_2} \right) \quad (3.A.5)$$

Substituting γ_3 from (3.A.5) into (3.A.4) we have

$$\begin{aligned} \sigma_p^2 = & \frac{1}{e'V^{-1}e} + \frac{1}{e'V^{-1}e} \left(\frac{E_1}{E_2} - 1 \right) \left(\frac{E-E_2}{E_1-E_2} \right)^2 \\ & + \frac{1}{e'V^{-1}e} \left(\frac{B_1}{B_2} - 1 \right) \frac{1}{\frac{B_3-B_2}{B_1-B_2} - \frac{E_3-E_2}{E_1-E_2}} \left[\frac{E-E_2}{E_1-E_2} - \frac{B-B_2}{B_1-B_2} \right]^2 \end{aligned}$$

APPENDIX 3.B

If we denote $a_1 = \frac{V^{-1}(\mu - r_f e)}{e' V^{-1}(\mu - r_f e)}$ and $a_3 = \frac{V^{-1}b}{e' V^{-1}b}$,

then any efficient portfolio is given by

$$X = \gamma_1 a_1 + \gamma_3 a_3 \quad \text{and} \quad x_0 = 1 - \gamma_1 - \gamma_3$$

Multiply by μ' and b' and denote $E_i = \mu' a_i$ and

$B_i = b' a_i$ for $i = 1, 2$ we have

$$X' \mu = \gamma_1 E_1 + \gamma_3 E_3$$

$$X' b = \gamma_1 B_1 + \gamma_3 B_3$$

Recall that $E = X' \mu + x_0 r_f = X' \mu + r_f (1 - \gamma_1 - \gamma_3)$ and

$X' B = B$ therefore we have the following system in γ_1

and γ_3 :

$$(E - r_f) = (E_1 - r_f) \gamma_1 + (E_3 - r_f) \gamma_3 \quad (3.B.1)$$

$$B = B_1 \gamma_1 + B_3 \gamma_3 \quad (3.B.2)$$

The variance of the portfolio is

$$\sigma_p^2 = X' V X = a_1' V a_1 \lambda_1^2 + a_1' V a_3 \lambda_1 \lambda_3 + a_3' V a_3 \lambda_3^2$$

$$\begin{aligned}
 a_1' v a_1 &= \frac{(\mu' - r_f e)' v^{-1} (\mu' - r_f e)}{[e' v^{-1} (\mu - r_f e)]^2} = \frac{(E_1 - r_f)}{e' v^{-1} (\mu - r_f e)} \\
 a_1' v a_3 &= \frac{(\mu - r_f e)' v^{-1} b}{[e' v^{-1} (\mu - r_f e)] [e' v^{-1} b]} = \frac{(E_3 - r_f)}{e' v^{-1} (\mu - r_f e)} \\
 &= \frac{B_1}{e' v^{-1} b} \quad (3.B.3)
 \end{aligned}$$

$$\text{and} \quad a_3' v a_3 = \frac{b' v^{-1} b}{(e' v^{-1} b)^2} = \frac{B_3}{e' v^{-1} b}$$

Therefore

$$\begin{aligned}
 c_p^2 &= \frac{(E_1 - r_f) \lambda_1^2 + (E_3 - r_f) \lambda_1 \lambda_3}{e' v^{-1} (\mu - r_f e)} + \frac{B_3}{e' v^{-1} b} \lambda_3^2 \\
 &= \frac{1}{e' v^{-1} (\mu - r_f e)} \\
 &\quad \left[\frac{(E_1 - r_f)^2 \lambda_1^2 + (E_1 - r_f) (E_3 - r_f) \lambda_1 \lambda_3 + (E_3 - r_f)^2 \lambda_3^2 - (E_3 - r_f)^2 \lambda_3^2}{(E_1 - r_f)} \right] \\
 &\quad + \frac{B_3}{e' v^{-1} b} \lambda_3^2 \\
 &= \frac{1}{e' v^{-1} (\mu - r_f e)} \frac{(E - r_f)^2}{(E_1 - r_f)} + \left(\frac{B_3}{e' v^{-1} b} - \frac{(E_3 - r_f)^2}{e' v^{-1} (\mu - r_f e) (E_1 - r_f)} \right) \lambda_3^2
 \end{aligned}$$

Utilizing (3.B.3)

$$\begin{aligned}\sigma_p^2 &= \frac{(E_1 - r_f)}{e'V^{-1}(\mu - r_f e)} \left(\frac{E - r_f}{E_1 - r_f} \right)^2 + \left(\frac{B_3}{e'V^{-1}b} - \frac{B_1}{e'V^{-1}b} \frac{E_3 - r_f}{E_1 - r_f} \right) \lambda_3^2 \\ &= \frac{E_1 - r_f}{e'V^{-1}(\mu - r_f e)} \left(\frac{E - r_f}{E_1 - r_f} \right)^2 + \frac{E_1 - r_f}{e'V^{-1}b} \left((E_1 - r_f)B_3 - (E_3 - r_f)B_1 \right) \lambda_3^2\end{aligned}$$

but since $\lambda_3 = \frac{(E_1 - r_f)B - (E - r_f)B_1}{(E_1 - r_f)B_3 - (E_3 - r_f)B_1}$

$$\sigma_p^2 = \frac{E_1 - r_f}{e'V^{-1}(\mu - r_f e)} \left(\frac{E - r_f}{E_1 - r_f} \right)^2 + \frac{E_1 - r_f}{e'V^{-1}b} \frac{[(E_1 - r_f)B - (E - r_f)B_1]^2}{(E_1 - r_f)B_3 - (E_3 - r_f)B_1}$$

CHAPTER IV
EVALUATION OF MEAN-VARIANCE
PORTFOLIO SELECTION

The non-spherical class of distributions introduced in the previous chapter together with its results provides a cornerstone for an analytical evaluation of the performance of mean-variance investment strategy. No previous work had challenged analytically the quality of the mean-variance portfolio selection as an approximation of expected utility maximization. Further specification of the joint distribution described in (3.1) together with an adequate choice of a class of utility functions allow a derivation of an explicit term for the expected utility. Once we have an explicit form for the expected utility, one could derive the optimal portfolio and the portfolio that maximizes the expected utility over the mean-variance efficient set only. The latter portfolio is the second best portfolio for an investor who faces the option to maximize his expected utility or the option to choose a mean-variance efficient portfolio. Given the portfolio solution as a function of the risk aversion and the distribution parameters, and given the explicit term for the expected utility in hand, we will

be able to derive a measure of the quality of the mean-variance investment strategy. This measure reflects the utility specific premium achieved by exact optimization over mean-variance portfolio selection. Such an optimization premium tells a practitioner how much it is worth on an invested dollar to carry on the process of exact maximization.

A number of factors distinguish this approach from all previous attempts made to evaluate the quality of mean-variance portfolio selection vis-à-vis expected utility maximization. The comparison in all previous works was based on empirical distribution of individual securities or portfolios. Such an approach stops short of any specification of the parametric structure of either the joint assets returns or the individual securities and portfolios compared. But the "art" of portfolio choice is essentially an "art" of constructing, for a given preference, the "best" probability distribution of a linear combination of a given number of random variables which share a joint stochastic nature. The admissibility of a mean-variance framework, regardless of investors preference, is essentially due to a "closure" property that the joint distribution exhibits. Such a property leads to the fact that the distribution of any combination of random variables is controllable¹ by two

¹A controllable parameter is endogenous to the portfolio problem. Its value varies with the portfolio proportion. For

choice parameters that have one-to-one mapping with the mean and the variance of the portfolio distribution. Since the admissibility of mean-variance portfolio selection is a parametric property of the joint distribution, the validity of a mean-variance framework as an approximation of expected utility maximization rests upon the parametric structure of the joint return generating process. Therefore, an examination of the performance of mean-variance portfolio choice in a framework which presumes a joint distribution with a parametric structure that does not admit mean-variance framework provides a greater challenge for the mean-variance investment strategy than an approach that stops short of its specification.

Second, the structure of previous approaches took the following form:² A heuristic mean-variance approximation for the expected utility was suggested. Then, for empirical distributions, the expected utility and its mean-variance approximation were calculated for returns on randomly selected portfolios and individual securities. Thereafter the divergence between the expected utility and its

example, a portfolio X from assets generated by distribution (3.1) has a return $X'Z = X'\mu + X'bY + X'\epsilon$. The distribution of $X'Z$ depends on $X'\mu$, $X'b$, $X'WX$ and the parameters of the distribution of Y . In this case $X'\mu$, $X'b$ and $X'WX$ are endogenous to the portfolio problem whereas the parameters of Y are exogenous.

²Except Kroll, Levy and Markowitz (1983).

approximation was analyzed. Such an approach heavily discriminates against mean-variance portfolio selection. This is because it requires the approximation to hold for randomly chosen portfolios, or individual securities, and hence, everywhere. But what really counts is whether the optimal portfolio choice could be "approximated" with another choice on the mean-variance efficient frontier. Consider the mean standard-deviation plane as the continuum of all possible portfolios. Suppose further that we have the following set of optimal portfolios, for different risk aversions, $(E_1, \sigma_1), \dots, (E_n, \sigma_n)$. Suppose that their corresponding second best portfolios are $(\hat{E}_1, \hat{\sigma}_1), \dots, (\hat{E}_n, \hat{\sigma}_n)$. The mean-variance consists of a "good" approximation if the utility placed on (E_i, σ_i) is "close" to the utility placed on $(\hat{E}_i, \hat{\sigma}_i)$. The divergence of the expected utility from its approximation for randomly selected portfolios does not necessarily imply poor performance. If we denote u and f the expected utility and its approximation respectively, even a divergence of $u(E_i, \sigma_i)$ from $f(E_i, \sigma_i)$ -- or $u(\hat{E}_i, \hat{\sigma}_i)$ from $f(\hat{E}_i, \hat{\sigma}_i)$ -- per-se does not imply poor performance. An approximation may have a poor performance for randomly chosen portfolios, yet a fair performance when it is compared at the points of an expected utility choice vis-à-vis the mean-variance choice. Moreover, the heuristic nature of the approximation does not necessarily

provide the "best" mean-variance approximation. The performance of mean-variance portfolio selection should be evaluated by analyzing the divergence of the expected utility from its "best" mean-variance approximation.

Third, the lack of analytical framework of all previous studies did not allow a systematic investigation of the sensitivity of the quality of mean-variance approximation to the parameters of the portfolio problem: risk aversion, joint distribution parameters, and the existence of a riskless asset. This approach allows such an investigation. It has the advantage of pointing out contingencies under which mean-variance performs poorly or fair. One of the interesting things that this approach points out is that for certain CARA investors the existence of a riskfree asset leads to a remarkably good performance for mean-variance strategies while its non-existence leads to an opposite result, for some investors.

Fourth, in previous studies the divergence of the expected utility from its approximation was measured by constructs that are short of providing a practitioner with an idea about the forgone opportunity cost from the failure to maximize the expected utility. In this work a measure that gives the utility specific premium placed by an investor to replace his mean-variance strategy by an optimal one is derived and operationalized empirically. This

"optimization premium" gives the maximum amount that an investor is willing to pay for information that leads to the optimum portfolio while he has the mean-variance efficient frontier for free. The cost of information to obtain the optimal portfolio is perhaps the most relevant factor for a practitioner to decide whether mean-variance selection fits him or if it does not.

The distribution of the spherical noise in (3.1) is specified as normal in this chapter, and the distribution of the non-spherical noise is specified as Pearson Type III. The class of utility function that admits the derivation of its expected utility in this setting is the CARA class. The analysis together with its empirical illustration reveals an interesting feature of the mean-variance approximation problem. When markets offer a riskless investment opportunity, the maximum fee that a CARA investor is willing to pay, no matter what his risk aversion, is 0.05% (0.05 of a cent on an invested dollar). Such a fee is ten times less than a typical brokerage fee. A high risk averse investor will pay an even lower fee. On the other hand, when a CARA investor is denied a riskless investment opportunity, the results and their trend are different. The premium placed on optimization is higher and after the risk aversion reaches a certain point, mean-variance strategy starts to perform poorly. The nature of our framework of analysis

allows us to reveal a "paradoxical" behavior of the highly risk averse CARA investor. As absolute risk aversion increases, the CARA investor has an increasing tendency to invest in what he views as the "least risky" alternative regardless of the size of the expected return of more risky choices. Since the "minimum risk" portfolio is not on the mean-variance efficient frontier, in contrast to the risk-free case, he exhibits an increasing dislike to mean-variance efficient portfolios that escalate in a positive linear relation with the measure of absolute risk aversion. Perhaps such behavior could be described as "obsession" and "frustration." He has an "obsession" to invest in the "minimum risk" portfolio even when he is lured with a "slightly" more risky prospect that may strike a fortune. Moreover, his "frustration" from the denial to reach this minimum risk portfolio (recall that this portfolio is not a mean-variance efficient one) makes him willing to pay a certain fee to get this portfolio that greatly outweighs his uncertain expected return from the investment on the least risk portfolio. Although the "obsession" is very strange it is difficult to argue against its rationality. But his "frustration" is indeed irrational. The "obsessive" nature of a CARA investor was pointed out by Levy and Markowitz (1977) for an absolute risk aversion equal to 10. They provided an example which demonstrates that such an

investor will prefer a 10% riskfree asset over a gamble that either breaks even or strikes a fortune. What is interesting is the fact that the existence of a riskfree investment prospect makes the mean-variance efficient menu more likable to a CARA investor and its existence per-se is in favor of mean-variance and not the opposite.

This chapter will include three sections. The first section determines the specification of the joint distribution, the class of utility function, and the derivation of the expected utility. The second section provides the derivation of the optimal portfolio and its mean-variance second best when all assets are risky and when a riskless asset is introduced. It also includes the derivation and the justification of the optimization premium as a measure of the forgone opportunity cost of mean-variance portfolio selection. The third section gives the empirical demonstration of the results of the previous section. It includes the description of the data employed, the estimation of the joint distribution parameters, the calculated results and their analysis.

4.1 The joint distribution and the class of utility functions:

To evaluate analytically the quality of mean-variance portfolio selection, for a given class of utility functions, both the portfolio that maximizes the expected utility, the

optimal one, and the portfolio that maximizes the expected utility restricted to the mean-variance efficient frontier have to be derived and compared. This cannot be done to the class (3.1) unless both the parametric distribution of the spherical vector ϵ and the non-spherical noise are specified further. Thereafter, the expected utility has to be derived and maximized, first globally and next over the mean-variance efficient frontier. The feasibility of deriving an expression for the expected utility depends on the specific choice of the assets returns distribution and the mathematical form of the utility function. Again the choice that has been made for both the class of distributions and utility function is a compromise between mathematical tractability, on the one hand, and generality on the other.

It is assumed through this chapter that the spherical random vector ϵ in (3.1) follows a joint normal distribution with zero mean and a covariance matrix W (to be denoted $\epsilon \sim N(0, W)$.) The only justification for such a choice is the mathematical tractability of this class. For the non-spherical noise it is important to specify more than two parameters of non-symmetric distribution. Since the choice of the location and the scale of the distribution of Y is arbitrary, it is desirable to have a third parameter that allows for different levels of skewness for the non-spherical noise. The choice made in this chapter

has been in favor of Pearson Type III distribution given as:

$$f_Y(Y) = \frac{1}{\lambda \Gamma(p)} \left(\frac{Y-\alpha}{\lambda} \right)^{p-1} \exp - \left(\frac{Y-\alpha}{\lambda} \right) \quad (4.1)$$

where $\alpha < Y < \infty$, $-\infty < \alpha < \infty$, $0 < \lambda$, $p < \infty$.

Its first three central moments and its characteristic function are given as:³

$$E(Y) = \alpha + \lambda p \quad (4.2)$$

$$\sigma_Y^2 = p\lambda^2 \quad (4.3)$$

$$m_{3Y} = E(Y - EY)^3 = 2\lambda^3 p \quad (4.4)$$

$$\gamma_1 = m_{3Y}/\sigma_Y^3 = 2/\sqrt{p} \quad (4.5)$$

$$\phi_Y(s) = \exp(ips) (1 - i\lambda s)^{-p} \quad (4.6)$$

α is a location parameter for Y and λ is a scale parameter. Since their choice is arbitrary for distribution (3.1) the scale parameter λ will be chosen as $\lambda = 1$ and the location parameter will be chosen such that $E(Y) = 0$. Such a restriction and (4.2) lead to $\alpha = -p$.

³When $\alpha = 0$ Pearson Type III distribution is identical to the well known Gamma distribution with a scale parameter λ and a shape parameter p . For information about Pearson systems and Gamma distributions see Johnson and Katz, Continuous Distributions (1970), Chapter 1 and Chapter 17 respectively.

Utilizing the form of the characteristic function of the normal distribution, (4.6) and (3.2), the characteristic function of the joint distribution of the vector of assets returns Z is given by:

$$\begin{aligned}\phi_Z(t_1, \dots, t_n) &= \exp(it'\mu) \cdot \exp\left(\frac{1}{2}t'Wt\right) \cdot \exp(-it'bp)(1-it'b)^{-P} \\ &= \exp\{it'(\mu-bp) + \frac{1}{2}t'Wt\}(1-it'b)^{-P} \quad (4.7)\end{aligned}$$

where μ , b are vectors of constants as defined in (3.1) and Y follows (4.1) with mean zero and scale parameter $\lambda = 1$.

The class of utility functions to be considered is the class of Constant Absolute Risk Aversion (CARA) given as:

$$\begin{aligned}u(r) &= -\frac{1}{A} \exp(-Ar) \\ r &> 1, A > 0\end{aligned} \quad (4.8)$$

Where $r = Z'X$ which is one plus the return on a portfolio X (i.e., $X'e = 1$). The mathematical tractability of the CARA is the main reason for its choice. The derivation of the expected utility for this class with distribution (4.7) is straightforward. For other classes of utility functions such as the logarithmic class, $\log(r+A)$ or the Constant Relative Risk Aversion (CRRA)

r^{α} , it is not feasible to derive an explicit expression of either the expected utility or the optimal portfolio when assets returns follow (4.7).⁴

When all assets are risky the portfolio rate of return is $X'Z$. Utilizing (3.10) and (4.7) the characteristic function of $X'Z$ is given as

$$\begin{aligned}\phi_{X'Z}(s) &= \exp(iX'\mu s) \phi_e(X'WXs^2) \phi_Y(X'bs) \\ &= \exp\{i(X'\mu - pX'b)s + \frac{1}{2}X'WXs^2\} (1 - iX'bs)^{-P} \quad (4.9)\end{aligned}$$

and the expected utility of a CARA is given by:

$$\begin{aligned}Eu(X'Z) &= E\left\{-\frac{1}{A} \exp(-AX'Z)\right\} = -\frac{1}{A} \phi_{X'Z}(-Ai^{-1}) \\ &= -\frac{1}{A} \exp\{-A(X'\mu - pX'b) + \frac{A^2}{2} X'WX\} (1 + AX'b)^{-P} \quad (4.10)\end{aligned}$$

If we recall that $\text{Var}(Z) = V = W - \sigma_Y^2 b b' = W - p b b'$ and denote $X'\mu = E$, $X'b = B$ and $X'VX = \sigma^2$

⁴ Even when the assets returns are normally distributed, the expected value of $\log(r+A)$ and r^{α} are indefinite truncated Gaussian integrals. This does not allow an explicit derivation of the optimal portfolio as a function of the distribution and the utility function parameters.

$$\begin{aligned}
Eu(X'Z) &= -\frac{1}{A} \exp\{-A(X'\mu - pX'b) + \frac{A^2}{2} X'VX \\
&\quad - p \frac{A^2}{2} X'b b'X\} (1+AX'b)^{-P} \\
&= -\frac{1}{A} \exp\{-A(E-pB) + \frac{A^2}{2} \sigma^2 \\
&\quad - p \frac{A^2}{2} B^2\} (1+AB)^{-P} \tag{4.11}
\end{aligned}$$

If a riskfree asset is available and if we denote by r_f its rate of return and x_0 the investment in the riskfree asset, then the rate of return on the portfolio $r = X'Z + r_f x_0$ and

$$\begin{aligned}
Eu(r) &= -\frac{1}{A} \exp(-Ar_f x_0) Eu(X'Z) \\
&= -\frac{1}{A} \exp\{-A(X'\mu + r_f x_0 - pX'b) + \frac{A^2}{2} X'VX \\
&\quad - p \frac{A^2}{2} (X'b)^2\} (1+AX'b)^{-P} \tag{4.12} \\
&= -\frac{1}{A} \exp\{-A(E-pB) + \frac{A^2}{2} \sigma^2 - p \frac{A^2}{2} B^2\} (1+AB)^{-P}
\end{aligned}$$

The derivation of the expected utility of a CARA investor for the class (3.1) is always straightforward when the characteristic function of ϵ and the characteristic function of Y take a neat form. For example, the characteristic function of a joint t distribution is an infinite series that does not allow the derivation of (4.10). The same is true if a beta distribution for the non-spherical

noise had been chosen. A choice of an Inverse-Gaussian distribution for the non-spherical noise Y , instead of Pearson Type III, also allows a straightforward derivation for the expected utility of a CARA investor. Since the Inverse-Gaussian is a two parameter family and Pearson Type III allows for three parameters the latter has been preferred. The choice of the joint distribution allows the assets returns to exhibit any skewness levels and does not admit two funds separation. The chosen class of utility functions allows different levels of risk aversion as well. These are, indeed, minimal requirements for an analytical framework to challenge the performance of the mean-variance investment strategy.

4.2 Derivation of the optimal portfolio and its mean-variance efficient alternative:

Throughout this chapter short sales will be allowed with no restrictions. The analysis will be performed when all assets are risky and when an extra riskless asset is introduced to the existing assets. First, the portfolio that maximizes the expected utility will be derived and next, forcing any portfolio to be a combination of two mean-variance efficient funds, the expected utility will be maximized. The latter is the "best" portfolio that the investor can choose if he is offered the menu of

mean-variance efficient portfolios. Once both portfolios are derived as a function of the parameters of the assets distribution and the measure of the risk aversion, an evaluation criterion should be determined. The choice of such a criterion and its justification will be discussed in section 4.2.3.

4.2.1 The case with no riskless asset: Instead of maximizing $Eu(r)$ we may maximize $h = -\log(-Eu(r))$. Since h is monotonically increasing in $Eu(r)$, both functions reach their maximum at the same point. The maximization of $h(X)$ rather than $Eu(X'Z)$ saves unnecessary algebraic work. Utilizing (4.11) we have:

$$\begin{aligned} h(X) &= -\log(-Eu(X'Z)) = \log A + A(X'\mu - pX'b) - \frac{A^2}{2} X'WX \\ &\quad + p \log(1+AX'b) \\ &= A(E-pB) - \frac{A^2}{2} \sigma_{Sp}^2 + p \log(1+AB) \end{aligned} \quad (4.12)$$

It is clear from (4.12) that $h(E, \sigma_{Sp}^2, B)$, and therefore the expected utility, is increasing in the mean of the portfolio return, decreasing in the spherical variance $\sigma_{Sp}^2 = X'WX$, and decreasing in the non-spherity parameter of the portfolio distribution B . This is because

$$\frac{dh(\cdot)}{dB} = -pA \frac{AB}{1+AB} < 0$$

The fact that the expected utility is decreasing in the spherical variance is not surprising and predicted by Theorem 3.5 for all utility functions. The positive contribution of the mean is expected for all increasing utility functions. However, the marginal contribution of B to the expected utility does not have uniformly the same sign across utility functions but depends on the specific nature of the utility function and the particular specification of the assets returns distribution. An increase in B has two effects at the same time: it increases the variance of the portfolio return on the one hand, and it increases its skewness on the other. Since the first is typically undesirable and the second is desirable, the ambiguity of its effect, regardless of the utility function, is clear. With our specific utility function the expected utility is decreasing in B but it is decreasing with $X'WX$ faster.⁵ Whereas we have a one-dimensional measure of risk in the conventional mean-variance framework, the variance of the portfolio return, a two-dimensional measure of risk is allowed within this framework, the spherical variance of the portfolio

⁵ The difference between the negative marginal contribution of the spherical variance and the non-sphericity parameter makes the optimal strategy disagree with the mean-variance efficient strategy. If both have the same marginal contribution to the expected utility then both strategies will agree. When the non-sphericity of the joint distribution exhibits this property across all utility functions, the assets returns distribution belongs to the two funds separating class.

distribution σ_{Sp}^2 and its non-spherity parameter B . Since both σ_{Sp}^2 and B are positively related to σ^2 the expected utility is also decreasing with the portfolio variance. Recall that $X'VX = X'(W-pbb') = X'WX - p(X'b)^2$. Therefore:

$$\begin{aligned} h(X) &= A(X'u - pX'b) - \frac{A^2}{2} X'VX + p(X'b)^2 + p \log(1+AX'b) \\ &= A(E-pb) - \frac{A^2}{2} \sigma^2 + pB^2 + p \log(1+X'b) \end{aligned} \quad (4.14)$$

Observe that for the expression of the expected utility given in (4.14) the expected utility is increasing in B since:

$$\frac{d}{dB} h(\cdot) = Ap \frac{(AB)^2}{1+AB} > 0$$

This reflects the fact that an increase in B that leaves the variance unchanged is desirable.

The optimal portfolio is derived by:

$$\begin{aligned} \max_X \quad & h(X) \\ \text{subject to} \quad & X'e = 1 \end{aligned}$$

The Lagrangian and the first order condition are given by:

$$L = A(X'\mu - pX'b) - \frac{A^2}{2} X'WX + p \log(1+AX'b) - \lambda(1-X'e)$$

$$\frac{dL}{dX} = A(\mu - pb) - A^2 WX + pA \left(\frac{1}{1+AX'b} \right) + \lambda e = 0$$

divide by A and denote $\gamma = \frac{\lambda}{A}$

$$AWX = \mu + \gamma e - p \left(\frac{AX'b}{1+AX'b} \right) b \quad (4.15.a)$$

If we let $W = V - pbb'$ we have the following alternative form for the first order condition:

$$AVX = \mu + \gamma e + p \left(\frac{(AB)^2}{1+AB} \right) b \quad (4.16.a)$$

Solving the first order condition for X :

$$AX = W^{-1}\mu + \gamma W^{-1}e - p \left(\frac{AX'b}{1+AX'b} \right) W^{-1}b$$

or

$$AX = V^{-1}\mu + \gamma V^{-1}e - p \left(\frac{AX'b}{1+AX'b} \right) V^{-1}b$$

If we denote $\alpha_1 = (W^{-1}\mu / e'W^{-1}\mu)$, $\alpha_2 = (W^{-1}e / e'W^{-1}e)$, $\alpha_3 = (W^{-1}b / e'W^{-1}b)$ and $X'b = B$ the solution to (4.15.a) can be written as:

$$X = \frac{eW^{-1}\mu}{A} \alpha_1 + \frac{(e'W^{-1}e)}{A} \gamma \alpha_2 - p \left(\frac{AB}{1+AB} \right) e'W^{-1}b \alpha_3$$

If we multiply by e' and solve for γ and substitute its value, the solution for (4.15.a) takes the following form:

$$(X - \alpha_2) = \frac{e'W^{-1}\mu}{A} (\alpha_1 - \alpha_2) - p \frac{e'W^{-1}b}{A} \frac{AB}{1+AB} (\alpha_3 - \alpha_2) \quad (4.15.b)$$

If we denote $a_1 = (V^{-1}\mu/e'V^{-1}\mu)$, $a_2 = (V^{-1}e/e'V^{-1}e)$ and $a_3 = (V^{-1}b/e'V^{-1}b)$ and solve for X in (4.16.a) we arrive to a similar form:

$$(X - a_2) = \frac{e'V^{-1}\mu}{A} (a_1 - a_2) + p \left(\frac{e'V^{-1}b}{A} \right) \frac{(AB)^2}{1+AB} (a_3 - a_2) \quad (4.16.b)$$

To determine X , either (4.15.b) or (4.16.b) could be multiplied by b' and solved for B . After the determination of B , X will be given as a function of A and the parameters of the joint distribution of returns only. If we let $B_i = b_i' a_i$ and multiply (4.16.b) by $A \cdot b'$ we have:

$$AB - AB_2 = e'V^{-1}\mu(B_1 - B_2) + pe'V^{-1}b(B_3 - B_2) \left(\frac{(AB)^2}{1+AB} \right) \quad (4.17)$$

(4.17) is a quadratic equation in (AB) with two

solutions. If we denote $k_1 = pe'V^{-1}b_j(B_3-B_2)$ and $k_2(A) = AB_2 + e'V^{-1}\mu(B_1-B_2)$ and $AB = t$ we have the following equation:

$$t^2 = k_2(A) + k_1\left(\frac{t^2}{1+t}\right)$$

or equivalently

$$(1-k_1)t^2 + (1-k^2(A))t - k^2(A) = 0$$

The above equation has at most two solutions. Since the expected utility is defined only for $1 + AB > 0$ any solution with $t \leq -1$ is not relevant. The second order condition is given by:

$$\begin{aligned} H &= \left| \frac{\partial^2 L}{\partial x_i \partial x_j} \right| = \left| -V + pbb' \quad 1 - \frac{1}{(1+AB)^2} \right| \\ &= \left| -W - \frac{p}{(1+AB)^2} bb' \right| < 0 \end{aligned} \quad (4.18)$$

Since W is positive definite H is typically negative. For all the empirical results, to be discussed later in this chapter, there has been only one solution to equation (4.17) for which $AB > -1$. This solution always satisfied the second order condition for maxima given by (4.18).

If we restrict the investor to the mean-variance efficient frontier his portfolio will be a combination of the two mean-variance efficient funds a_1 and a_2 . That is:

$$X = \lambda a_1 + (1-\lambda)a_2 = \lambda(a_1 - a_2) + a_2 \quad (4.19)$$

If we multiply (4.19) by μ' and let $E_i = \mu'a_i$ and solve for λ we have $\lambda = (E - E_2)/(E_1 - E_2)$. If we multiply by b' and solve for λ also we have $\lambda = (B - B_2)/(B_1 - B_2)$. On the mean-variance efficient frontier we have:

$$\lambda = \frac{E - E_2}{E_1 - E_2} = \frac{B - B_2}{B_1 - B_2} \quad (4.20)$$

and since

$$X'VX = (a_1 - a_2)'V(a_1 - a_2)\lambda^2 + (a_1 - a_2)'Va_2\lambda + a_2'Va_2 \quad (4.21)$$

the expected utility, on the efficient frontier, can be written as a function of λ and the distribution parameters, or alternatively as a function of E only or B only. Imposing (4.19) in (4.14) and utilizing (4.20) and (4.21) $h(\cdot)$ can be written as:

$$\begin{aligned}
h(B) = & A(E(B) - pb) - \frac{A^2}{2} ((a_1 - a_2)' V(a_1 - a_2) \lambda^2(B) \\
& + (a_1 - a_2)' V a_2 \lambda(B) + a_2' V a_2) \\
& + p \frac{A^2}{2} B^2 + p \log(1+AB)
\end{aligned} \quad (4.22)$$

The first order condition is given as:

$$\begin{aligned}
A \left(\frac{E_1 - E_2}{B_1 - B_2} \right) - A^2 \left(\frac{(a_1 - a_2)' V(a_1 - a_2)}{(B_1 - B_2)^2} \right) (B - B_2) \\
+ \frac{(a_1 - a_2)' V a_2}{B_1 - B_2} + pA \frac{(AB)^2}{(1+AB)} = 0
\end{aligned} \quad (4.23)$$

Again we have a quadratic equation in $AB = t$. If we denote

$$\begin{aligned}
k_1 &= \frac{p (B_1 - B_2)^2}{(a_1 - a_2)' V(a_1 - a_2)} \\
k_2(A) &= \frac{(E_1 - E_2) (B_1 - B_2)}{(a_1 - a_2)' V(a_1 - a_2)} + \left(\frac{(a_1 - a_2)' V(B_2 a_1 - B_1 a_2)}{(a_1 - a_2)' V(a_1 - a_2)} \right) A
\end{aligned}$$

then we have an equation in $t = AB$ with a similar form as (4.17)

$$t = k_2(A) + k_1 \frac{t^2}{1+t}$$

or alternatively

$$(1-k_1)t^2 + (1-k_2(A))t - k_2(A) = 0$$

The second order condition is

$$-A^2 \frac{(a_1-a_2)'V(a_1-a_2)^2}{(B_1-B_2)^2} + pA^2 \left(1 - \frac{1}{(1+AB)^2} \right) \quad (4.24)$$

Similar to the case of the exact optimization only solutions with $t > -1$ are admissible since the expected utility is defined for $AB > -1$. In the empirical demonstrations, to be discussed in section 4.3, one of the pair solutions has been admissible for all the cases. This solution always satisfied the second order condition for maxima.

4.2.2 The case with riskless asset: Using (4.12), the proper objective function of the investor is:

$$h(X) = \log A + A(X'\mu + r_f x_0 - pX'b) - \frac{A^2}{2} X'VX \\ + p \frac{A^2}{2} (X'b)^2 + p \log(1+AX'b)$$

where $X = (x_1, x_2, \dots, x_n)$ is the investment position in the n risky assets and x_0 is the investment position in the riskless asset with a rate r_f . The Lagrangian and

the first order condition are given as:

$$L = h(X) + \lambda(1 - x_0 - X'e)$$

$$\frac{dL}{dx_0} = Ar_f - \lambda = 0$$

$$\frac{dL}{dX} = A\mu - Apb - A^2VX + pA^2(X'b)b + \frac{Ap}{1+AX'b} - \lambda e = 0$$

Substituting the value of λ from the first equation in the second, and solving for X leads to:

$$X = \frac{V^{-1}(\mu - r_f e)}{A} + \frac{p}{A} \frac{(AX'b)^2}{(1+AX'b)} V^{-1}b \quad (4.25)$$

To determine X , multiply (4.24) by $A \cdot b'$ and denote $X'b = B$

$$AB = b'V^{-1}(\mu - r_f e) + p \frac{(AB)^2}{(1+AB)} b'V^{-1}b \quad (4.26)$$

Equation (4.25) is a quadratic equation in AB . After the determination of AB , the optimal investment in the risky asset, X , will be determined, the investment in the riskfree asset will be $x_0 = 1 - X'e$. Again (4.25) is very similar in its form to (4.17) and (4.23). Only one of the two solutions is greater than -1 and satisfies the second order condition for maxima.

When the investor is forced to choose a mean-variance efficient portfolio he will choose a combination of the risk free asset and $\hat{a}_1 = V^{-1}(\mu - r_f e) / e' V^{-1}(\mu - r_f e)$. The following restriction on X should be imposed:

$$X = \lambda \hat{a}_1$$

If we multiply by μ' and b' respectively we get the following relationship between λ , E , and B

$$\lambda = \frac{E - r_f}{\hat{E}_1 - r_f} = \frac{B}{\hat{B}_1}$$

where $\hat{E}_1 = \mu' \hat{a}_1$ and $\hat{B}_1 = b' \hat{a}_1$. The objective function $h(\cdot)$ can be expressed as a function of either λ , E , or B . As a function of B it is given as:

$$h(B) = \log A + A(E(B) - pB) - \frac{A^2}{2} \hat{a}_1' V \hat{a}_1 \lambda^2(B) + p \frac{A^2}{2} B^2 + p \log(1+AB)$$

$$\frac{d}{dB} h(B) = A \left(\frac{\hat{E}_1 - r_f}{\hat{B}_1} \right) - A^2 \left(\frac{\hat{a}_1' V \hat{a}_1}{\hat{B}^2} \right) B + pA \frac{(AB)^2}{(1+AB)} = 0 \quad (4.27)$$

Similar to the previous cases we have a quadratic equation in (AB) . Once AB is determined B will be determined as a function of the absolute risk aversion A ,

and therefore, the best mean-variance efficient portfolio.

It is interesting to observe how the investment behavior of an investor changes with his measure of risk aversion A . According to (4.15.b), for the case without riskless asset, as A goes to infinity the right hand side of (4.15.b) approaches zero and therefore

$$X_{A \rightarrow \infty} + \alpha_2 = (W^{-1}e / e'W^{-1}e).$$

It is important to note that

α_2 is the portfolio that globally minimizes the spherical variance $X'WX$. An extremely risk averse investor with a CARA utility function will end up choosing a portfolio that minimizes the spherical variance regardless of the mean vector μ or the non-spherity parameter, the vector b , of the assets distribution. The only thing that counts for him is the covariance structure of the spherical component of the assets distribution. Such obsession will lead to an absurd behavior as will be shown later. It is worth noting that there is no reason for α_2 to be a mean-variance efficient portfolio and if it turns out to be mean-variance efficient that will be accidental. Therefore an extremely risk averse investor is most likely to disagree with the mean-variance investment strategy when he is denied a riskless asset. But once he is offered a riskless asset his position in risky assets, according to (4.25) tends to zero and therefore he will choose to invest only in the

riskless asset. Since the mean-variance efficient set includes the riskfree asset, an extremely risk averse investor will agree with the mean-variance efficient strategy. This is clear from both equations (4.26) and (4.27). In both cases, in contrast to the case with all risky assets, AB is fixed. Therefore the non-spherity parameter, B , that a CARA investor chooses for his portfolio distribution is unitary elastic to his degree of risk aversion.⁶ As A tends to infinity B tends to zero. When he is forced to choose a mean-variance efficient portfolio λ tends to zero and hence his investment in risky assets. These two distinctive features of the riskless asset case, the unitary elasticity and the agreement between the optimal portfolio and the mean-variance efficient frontier, lead to antagonistic results for the two cases regarding the performance of mean-variance portfolio selection for CARA investors.

4.2.3 The optimization premium: To evaluate the quality of mean-variance portfolio selection a criterion has to be chosen to compare the exact optimal portfolio with its mean-variance efficient second best. Needless to say, the

⁶Since the solutions to (4.26) and (4.27) are different, the optimal portfolio and the best portfolio on the mean-variance efficient frontier are different for all finite values of A also.

difference between the expected utilities that each portfolio provides is sensitive to arbitrary displacement and scale shifts of the utility function and hence is not a proper criterion.⁷ Comparing the geometric distance between the optimal portfolio and its mean-variance efficient second best is invariant to positive linear transformation because the portfolios themselves are. On the other hand, it fails to reflect the subjective frustration from the failure to achieve the exact optimal portfolio. Indeed, investors with different utility functions may place different utility on the distance between the optimal portfolio and its mean-variance efficient second best.

The invariance of the comparison criterion to an arbitrary positive linear transformation of the utility function and its dependence on the original utility function are necessary conditions for a proper evaluation criterion. But this slightly restricts the possibilities of choice that

⁷This is because the utility function $u(r)$ and any other utility function given as a positive linear transformation $= c_1 u(r) + c_2$ where $c_1 > 0$ (the transformation displaces the index by c_2 and scales it by c_1) leads to the same portfolio choice and therefore they are identical. On the other hand, the distance between the compared portfolio is sensitive to the choice of c_1 and c_2 . The ratio between the expected utilities has been used by Pully (1981) but it has the same deficiency, pointed out by Kroll, Levy and Markowitz (1983). Namely it is sensitive to the displacement parameter c_2 of the positive linear transformation and therefore it does not qualify either.

we are facing. The correlation coefficient between the utility function and its mean-variance approximation is one of the choices used by Levy and Markowitz (1979). The relative advantage of the optimal portfolio and the mean-variance efficient second best over a fixed portfolio, such as the naive portfolio⁸ is a second choice used by Kroll, Levy and Markowitz (1983). The difference in certainty equivalent between the two portfolios is a third choice suggested by Ziemba (1979), Pully (1983). Certainly this does not exhaust the list.

Since a main objective of an evaluation of the mean-variance portfolio selection is to help practitioners identify contingencies under which it is "practically" safe to follow mean-variance investment strategy, it is desirable to develop a criterion that tells what is the premium to optimize, or what is really the foregone opportunity cost from the choice of mean-variance investment strategy as a second alternative to the exact optimization. Consider a consultant that offers the mean-variance efficient frontier to his clients for free, but charges a rate θ from an

⁸The naive portfolio is a portfolio with equal proportions in all existing assets. The criteria that had been used was the ratio of the difference between the expected utility of the optimal portfolio and the naive portfolio to the difference between the expected utility of mean-variance second best and the naive portfolio.

investor that comes with a utility function $u(r)$ that he insists on maximizing. The relevant question for a practitioner is: What is the highest value of θ that the consultant could charge? If the joint distribution of assets returns admits two funds separation, then $\theta = 0$. But with our distribution θ is not necessarily zero. If we denote by X and $\hat{X}$ the optimal portfolio and its mean-variance efficient second best with returns r and $\hat{r}$ respectively, then θ must satisfy:

$$Eu(r-\theta) = Eu(\hat{r}) \quad (4.28)$$

Observe that θ is invariant under any positive linear transformation of the utility function $u(r)$ and a specific of the utility function itself. For CARA utility function, θ can be expressed explicitly as a function of the expected utility of r and $\hat{r}$.

$$\begin{aligned} Eu(\hat{r}) &= -\frac{1}{A} E \exp(-A\hat{r}) = Eu(r-\theta) = -\frac{1}{A} E \exp(-A(r-\theta)) \\ &= -\frac{1}{A} \exp(A\theta) \cdot E \exp(-Ar) \end{aligned}$$

solving for θ

$$\theta = \frac{1}{A} [\log Eu(\hat{r}) - \log Eu(r)] \quad (4.29)$$

$$\Theta = \frac{1}{A} [h(r) - h(\hat{r})]$$

$$\Theta = (E - \hat{E}) - \frac{A}{2} (\sigma_{Sp}^2 - \hat{\sigma}_{Sp}^2) + \frac{B}{A} \log \left(\frac{1 + AB}{1 + A\hat{B}} \right) \quad (4.30)$$

According to (4.30) the optimization premium Θ is higher the greater the difference between the mean of the optimal portfolio and its mean-variance second best mean in a linear fashion. It is also positively related to difference between the non-spherity parameter B . It is worth noting that the optimal portfolio minimizes the spherical variance σ_{Sp}^2 for a given E and B . Therefore, when E and B are close to $\hat{E}$ and $\hat{B}$ respectively it is typical that $(\sigma_{Sp}^2 - \hat{\sigma}_{Sp}^2) < 0$. For the case with no riskless asset, when A tends to be very large the investor will tend to choose a portfolio that globally minimizes the spherical variance, namely $\alpha_2 = (W^{-1}e/e'W^{-1}e)$. In this case, since α_2 is not necessarily mean-variance efficient, $\sigma_{Sp}^2 < \hat{\sigma}_{Sp}^2$ and the optimization premium increases linearly with A without limit. Certainly it may exceed the mean of $Z'\alpha_2$. This demonstrates how absurdly a very risk averse CARA investor could behave. He has such an obsession to get the globally minimum spherical variance portfolio, even if that results in a charge higher than the expected return on his investment. But when a riskless asset exists, as A increases to infinity the investor

will tend to invest only in the riskless asset. Since the mean-variance efficient set offers this asset, the three terms of the right hand side of (4.30) will approach zero and hence the optimization premium θ .

4.3 Empirical demonstration of the optimization premium:

The results of the previous section give the optimal portfolios and their corresponding mean-variance efficient alternatives as a function of the distribution parameters and the risk aversion. Still we cannot give an empirical operationalization of these results unless we can tell something about their numerical size--especially the optimization premium. We have to tell what may be the size of the optimization premium if empirical data has been generated by distribution (4.7) and how it behaves with risk aversion under each contingency. Through this process no claim has been made regarding the true distribution that generates assets return. Whether distribution (4.7) is a true description of real empirical returns is not a relevant question here. The objective of this empirical demonstration is to answer a hypothetical question. Namely, what would have been the size of the optimization premium if observable data were generated from distribution (4.7).

There are two approaches to carry on this illustration. In the first approach we may simulate assets returns from

distribution (4.7) by jointly simulating a vector of random returns and a random variable from (4.6) to describe the non-spherical noise and utilizing (3.1) to create the assets returns. This approach requires an apriori choice of the distribution parameters μ , b and W in a matter consistent with observable assets returns. In the second approach, we take historical returns on a given number of assets and then estimate the distribution parameters by imposing its parametric structure on the return generating process. Since the optimization premium depends on the distribution parameter only, the second approach that extracts them from observable returns has been preferred to the one that required their apriori determination.

For the purpose of the demonstration monthly returns on ten securities, arbitrarily chosen from the CRSP, the period horizon is January 1973 - December 1982 are used. Exhibit (4.1) gives the names of these securities and the first four moments of the marginal distribution of each security return. The mean and the variance are scaled to reflect the monthly rates in annual scale. Such descriptive statistics of security returns, especially skewness coefficients and kurtosis, are typical to security returns. We have positive skewness for all securities that range from zero to more than one.

EXHIBIT 4.1

DESCRIPTIVE STATISTICS OF MONTHLY RETURNS ON THE SECURITIES USED
FOR EMPIRICAL DEMONSTRATION (SAMPLE PERIOD IS 1/1973 --- 12/1982)

VARIABLE	MEAN*	STANDARD* DEVIATION	SKEWNESS	KURTOSIS
MCD MC DONALDS CORP	0.114830	1.045776	0.456943	4.997324
OXF OXFORD MFG INC	0.431890	1.404093	0.671820	1.022959
PRIZER CHAS & CO INC	0.121300	0.973932	0.783679	3.569307
STANDARD PRESSED STL CO	0.285980	1.560280	0.426785	-0.161410
SEAGRANG LTD	0.171100	1.091264	0.906994	2.842856
SEARS ROEBUCK & CO	0.077910	0.940444	0.750765	1.198725
STORER BROADCASTING CO	0.280940	1.160455	0.020960	0.124583
SUPERMARKETS GEN COP	0.296720	1.278464	0.798701	2.310712
UNION CARBIDE	0.145240	0.790892	1.090410	1.884392
WFC WELLS FARGO & CO	0.165430	1.103427	0.403002	2.001686

* QUOTED IN ANNUAL DIMENSION

4.3.1 The estimation of the distribution parameters: Equation (3.25) gives the likelihood function of N sample points from distribution (3.1). Utilizing our particular specification of the spherical joint distribution of ϵ as normal and the specification of the non-spherical noise as Pearson Type III with $\lambda = 1$, $\alpha = -p$, the likelihood function of the sample takes the following form:

$$L(\mu, b, W, p) = \frac{\pi^{-\frac{Nn}{2}} |W|^{-\frac{N}{2}}}{\Gamma(p)} \int_{-p}^{\infty} \exp\left\{-\frac{1}{2} \sum_{t=1}^N (z_t - \mu - by)' W (z_t - \mu - by)\right\} (y+p)^{p-1} \exp\{-y+p\} dy \quad (4.31)$$

When we have n assets the number of parameters to be estimated is $\frac{n(n+3)}{2} + 1$. Perhaps the most desirable way to estimate these parameters is the maximum likelihood method. But this is not an easy job to do. An alternative approach as a mixture of moment method and maximum likelihood is used. For a given p , we have a moment estimate for μ by the sample mean vector, for V by the sample covariance matrix and for b according to (3.23) as

$$b_i = \left[\frac{m_{3i}}{EY^3} \right]^{1/3} = \left[\frac{m_{3i}}{2p} \right]^{1/3} \quad (4.32)$$

where m_{31} is the sample third moment around the mean. Therefore we have the following estimate for W :

$$W = V - bb'\sigma_Y^2 = \hat{V} - \hat{b}\hat{b}'p \quad (4.33)$$

If we substitute these estimates in (4.31) we get a likelihood function that depends only on the common non-spherical noise parameter p . Using the sample data about the observable assets returns and the moments estimates as a function of p , the likelihood function can be evaluated for different values of p . The estimate for p will be chosen so that the likelihood function is maximized. We have a lower limit on p as zero. The upper limit for p can be calculated. Since any estimate for the matrix W should lead to a positive definite matrix, the upper limit on p will be the largest value that allows the matrix W to be positive definite. After calculation of the eigen values of W for different values of p , it has been found that with a $p = 0.065$ the lowest of the eigen values becomes very small and becomes negative for $p = 0.066$.⁹ An empirical bound

⁹ A symmetric matrix is positive if, and only if, all its eigen values are positive. According to this theorem the eigen values of W have been calculated for different values of p . The minimum of the eigen values is reported in Exhibit I in Appendix 4.B.

on p has found to be 0.065 with our empirical data.

To calculate the likelihood function, the integral on the right hand side of (4.31) has to be evaluated numerically. Appendix 4.A described the method of integration employed. It is important to note that if we assume a uniform apriori distribution of p on the range $0 \leq p \leq 0.065$, then the posterior distribution of p is the likelihood function of p standardized by dividing the likelihood function by its integral over p . Therefore the posterior probability and the maximum likelihood function are maximized at the same point. Exhibit 4.B.1 in Appendix 4.B gives the posterior distribution for p and the minimum of the eigen values of the matrix W for different values of p . Exhibit 4.B.2 in the same appendix gives the graph of the posterior distribution for p . Both the posterior distribution of p and its likelihood function are maximized for $p = 0.019$. Once we have an estimate for p the estimates for b and W are determined.

Since the non-spherical noise is assumed to be the same across securities returns, its parameter p must be estimated utilizing joint information about assets returns and cannot be estimated from observation on marginal assets returns. Perhaps the above estimation method is not the most statistically efficient but it is sufficient for the purpose of our empirical demonstration for the results of

previous sections. Still it is an open question to suggest statistically efficient estimates for the parameters of the common spherical noise, especially for its mean and variance, independent of the peculiar specification of the distribution of ε or Y . This is important once the results of Chapter III are applied for portfolio selection under the assumption that portfolio returns follow distribution (3.1).

4.3.2 The empirical results: The estimates of the distribution parameter from section (3.1) were used to calculate the optimal portfolios and their mean-variance efficient second best portfolios for different levels of absolute risk aversion. For the case that rules out the existence of the risk free asset, the quadratic equation (4.17) has been solved for different levels of risk aversion. This equation has two solutions. One solution was always less than -1 and hence, not admissible, since the expected utility is not defined for $AB \leq -1$. The second solution satisfied the second order condition for maxima. This is because the estimates for p have been chosen such that the matrices W are positive definite. The solution of (4.17) was combined with equation (4.16.b) to produce the optimal portfolio for different levels of risk aversion in this case. Similarly the mean-variance efficient second

best portfolio was obtained by utilizing equations (4.23) and (4.19) for the same levels of risk aversion. Exhibit 4.2 for different levels of A provides the parameters of the optimal portfolio distribution: namely, the mean E ; the non-spherity parameter B ; and the standard deviation SD . It also provides the corresponding parameters for the corresponding mean-variance efficient second best: namely, the mean $EMVE$; the non-spherity parameter $BMVE$; and the standard deviation $SDMVE$. It also provides the optimization premium, $THETA$, based on (4.30) and the Geometric norm between the two portfolios, namely $\|X - \hat{X}\|^{1/2}$ under the title $NORM$. Exhibit 4.3 shows graphically the optimal portfolios and their corresponding mean-variance efficient ones.¹⁰ Exhibit I in Appendix 4.C gives the portfolio proportions for both the optimal and the mean-variance efficient ones for the different levels of A . For degrees of risk aversion, $A < 2$, the investor will follow highly aggressive short sales strategies that lead to portfolios with an expected return higher than 80%. Since such strategies are unrealistically aggressive they are not reported.

¹⁰ Although all the calculations have been made with unscaled monthly returns, the means, the standard deviations, the optimization premium, and the non-spherity parameter were scaled to annual dimensions by multiplying their respective values by 12.

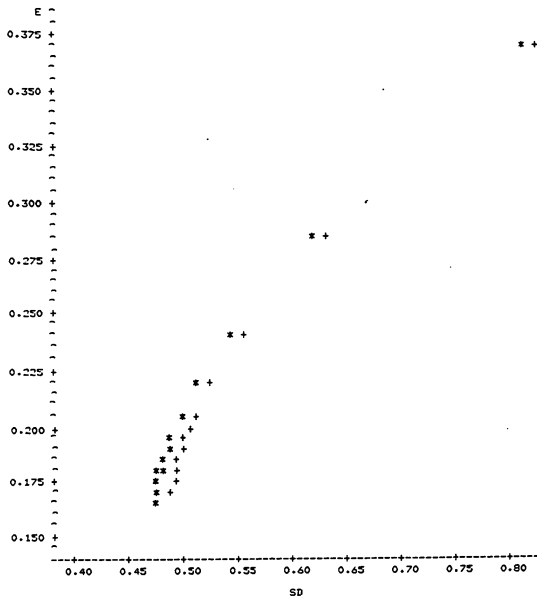
EXHIBIT 4.2

THE PARAMETERS OF THE OPTIMAL PORTFOLIO AND ITS MEAN
 MEAN VARIANCE EFFICIENT ALTERNATIVE, THE OPTIMIZATION
 PREMIUM, θ , AND THEIR GEOMETRIC NORM
 (WHEN ALL ASSETS ARE RISKY)

A	E	B	SD	ENVE	BVE	SDVE	θ	NORM
2	0.807790	3.33918	2.04685	0.805316	3.19213	2.03865	0.00023	0.066033
4	0.480859	3.03918	1.10910	0.478317	2.85734	1.09956	0.00073	0.081811
6	0.372108	2.95104	0.82489	0.369529	2.74596	0.81408	0.00144	0.092372
8	0.317822	2.91168	0.69823	0.315222	2.69036	0.68638	0.002289	0.099766
10	0.285293	2.89030	0.63089	0.282481	2.65705	0.61822	0.00323	0.105206
15	0.241989	2.86536	0.55616	0.239361	2.61269	0.54212	0.00581	0.114058
20	0.220375	2.85485	0.52736	0.217739	2.59056	0.51253	0.00859	0.119366
25	0.207421	2.84933	0.51342	0.204782	2.57729	0.49808	0.01147	0.122902
30	0.198792	2.84601	0.50566	0.196152	2.56846	0.48996	0.01441	0.125625
40	0.186015	2.84233	0.49779	0.185373	2.55742	0.48162	0.02040	0.128786
50	0.181554	2.84039	0.49407	0.178911	2.55081	0.47762	0.02646	0.130922
60	0.177248	2.83921	0.49203	0.174606	2.54640	0.47539	0.03257	0.132400
70	0.174174	2.83841	0.49079	0.171532	2.54325	0.47400	0.03870	0.133483
80	0.171869	2.83789	0.48998	0.169227	2.54089	0.47308	0.04485	0.134311
90	0.170077	2.83748	0.48942	0.167435	2.53906	0.47244	0.05102	0.134965
100	0.168444	2.83717	0.48901	0.166001	2.53759	0.47197	0.05719	0.135494
150	0.164344	2.83631	0.48804	0.161703	2.53319	0.47079	0.08811	0.137114
200	0.162196	2.83592	0.48769	0.159554	2.53099	0.47033	0.11909	0.137943
250	0.160907	2.83570	0.48752	0.158265	2.52967	0.47010	0.15010	0.138447
300	0.160047	2.83556	0.48734	0.157406	2.52879	0.46996	0.18111	0.138785
400	0.158973	2.83540	0.48730	0.156333	2.52769	0.46982	0.24316	0.139211
500	0.158329	2.83530	0.48729	0.155689	2.52703	0.46974	0.30521	0.139469

EXHIBIT 4.3

THE OPTIMAL PORTFOLIOS AND THEIR CORRESPONDING MEAN
VARIANCE EFFICIENT IN THE MEAN STANDARD DEVIATION PLANE
(WHEN ALL ASSETS ARE RISKY)



SYMBOL USED FOR OPTIMAL PORTFOLIS IS '+'
SYMBOL USED FOR MEAN VARIANCE EFFICIENT ONE IS '*'

We learn from Exhibits (4.2) and (4.3) that as risk aversion increases, both the geometric distance and the optimization premium increases. While the geometric distance approaches an upper limit of 0.1395 the optimization premium escalates linearly with the risk aversion. It reaches about 1% for $A = 25$ for portfolios that provide an approximate expected return of 20% and it comes to be larger than the expected return for $A = 300$ for portfolios with an expected return about 16%. Such behavior is indeed absurd. An investor with such risk aversion is willing to pay a consultant a certain fee for advice about a portfolio with a risky return that provides an expected return that is less than the fee. If a CARA utility function is to be used to describe investors' behavior toward risk, its parameter A must be restricted to rule out such irrational behavior.

As pointed out in section 4.2, when the risk aversion is very large the investor will tend to choose the portfolio $\alpha_2 = (W^{-1}e/e'W^{-1}e)$, the global minimum spherical variance portfolio, regardless of the mean vector μ of the joint assets distribution or its non-sphericity parameter b . Such an investor will prefer portfolio α_2 even if he is offered another portfolio with an expected return a hundred times more than the expected value of α_2 if this portfolio has a spherical variance slightly higher than

$\alpha_2'W\alpha_2$. Although the presumed assets returns distribution yields a two-dimensional measure for risk, the spherical variance σ_{Sp}^2 and B , the risk aversion of our investor to σ_{Sp}^2 dominates his risk aversion to B . Extremely large values of A reflect a blind obsession to avoid the spherical variance regardless of the alternatives he is facing.¹¹

Levy and Markowitz (1977) pointed out a strange behavior of this kind even for $A = 10$. If an investor with risk aversion $A = 10$ is faced with a gamble that has a 50-50 chance of breaking even (0 return) or hitting a 100% return, or faced with a 10% riskfree prospect, he will prefer the

¹¹When our investor is faced with normal assets returns, it is straightforward to show that as his risk aversion increases he has an obsession to choose the global minimum variance portfolio, $a_2 = (V^{-1}e/e'V^{-1}e)$. Since α_2 is not a mean-variance efficient portfolio, the disagreement on the portfolio choice increases and approaches

$$\|(\alpha_2 - a_2) + \frac{(a_1 - a_2)'W(a_1 - a_2)}{(a_1 - a_2)'Wa_2}(a_1 - a_2)\|$$

This is because the portfolio

$$\frac{-(a_1 - a_2)'W(a_1 - a_2)}{(a_1 - a_2)'Wa_2}(a_1 - a_2) + a_2$$

minimizes $X'WX$ subject to the restriction that $X = \lambda(a_1 - a_2) + a_2$. Observe that the disagreement increases with $\|\alpha_2 - a_2\|$, with $\|a_1 - a_2\|$, and with the ratio $\text{var}((a_1 - a_2)'X)/\text{cov}((a_1 - a_2)'X, a_2'X)$.

riskfree prospect. He will not be lured to change his mind even if hitting the Jackpot means 500% or 1000%. Certainly such an investor will not go for oil drilling or gold mining regardless of the treasure under the ground, unless he is certain to find the treasure. Again, we have a demonstration of how, for such an investor, what is regarded as risk is avoided and considered inferior to the safest route relative to the risk measure regardless of the good fortune that risk taking may offer under some contingencies.

An examination of Exhibit 4.4 and Exhibit 4.5 reveals a different picture for the riskfree asset case. Both the geometric distance and the optimization premium decline with risk aversion. Observe that for $A = 2$ the geometric distance is 2.4865 (18 times more than 0.1395, the upper level of the geometric distance in the previous case), yet the optimization premium is less than 0.05%. While an investor with $A = 2$ is not willing to pay 0.05% to replace his suboptimal mean-variance efficient portfolio that yields an expected return of 91%, his friend with $A = 300$ is willing to pay a fee of 16%, exactly equal to his expected return, to replace his suboptimal mean-variance efficient portfolio with an optimal one--even though his mean-variance efficient portfolio is 18 times closer to the optimal one. In any case, Exhibit 4.4 and 4.5 demonstrate that a CARA investor will pay a fee less than 0.05% (\$5

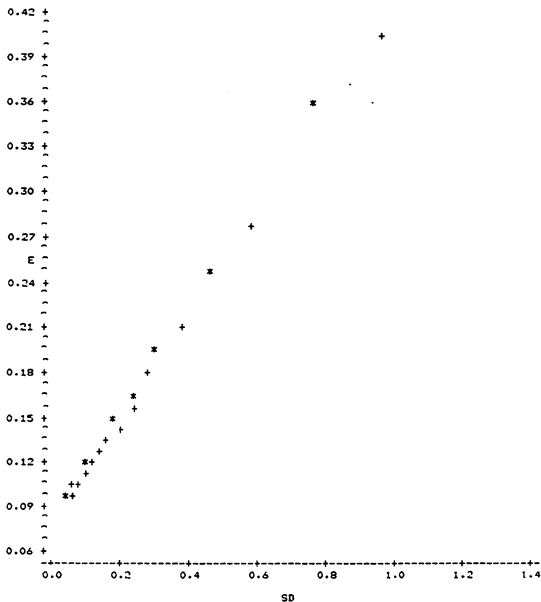
EXHIBIT 4.4

THE PARAMETERS OF THE OPTIMAL PORTFOLIO AND ITS MEAN
 MEAN VARIANCE EFFICIENT ALTERNATIVE, THE OPTIMIZATION
 PREMIUM, θ , AND THEIR GEOMETRIC NORM
 (THE RISK-FREE ASSET CASE)

A	E	B	SD	EMVE	BMVE	SBMVE	THETA	GD
2	1.05178	12.4770	2.89111	0.916789	5.79445	2.30958	0.000495078	2.48653
4	0.56589	6.2385	1.44555	0.498395	2.89723	1.15479	0.000247539	1.24326
6	0.40393	4.1590	0.96370	0.358930	1.93148	0.76986	0.000165026	0.82884
8	0.32395	3.1192	0.72278	0.289197	1.44861	0.57739	0.000123770	0.62163
10	0.27436	2.4954	0.57822	0.247358	1.15889	0.46192	0.000099016	0.49731
15	0.20957	1.6636	0.38548	0.191572	0.77259	0.30794	0.000066010	0.33154
20	0.17718	1.2477	0.28911	0.163679	0.57945	0.23096	0.000049508	0.24865
25	0.15774	0.9982	0.23129	0.146943	0.46356	0.18477	0.000039606	0.19892
30	0.14479	0.8318	0.19274	0.135786	0.38630	0.15337	0.000033005	0.16577
40	0.12859	0.6238	0.14456	0.121839	0.28972	0.11548	0.000024754	0.12433
50	0.11887	0.4991	0.11564	0.113472	0.23178	0.09238	0.000019803	0.09546
60	0.11239	0.4159	0.09637	0.107893	0.19315	0.07689	0.000016503	0.08288
70	0.10727	0.3565	0.08260	0.103908	0.16556	0.06599	0.000014145	0.07104
80	0.10429	0.3119	0.07228	0.100920	0.14486	0.05774	0.000012377	0.06216
90	0.10160	0.2773	0.06425	0.098595	0.12877	0.05132	0.000011002	0.05526
100	0.09944	0.2495	0.05782	0.096736	0.11589	0.04619	0.000009902	0.04973
150	0.09296	0.1664	0.03855	0.091157	0.07726	0.03079	0.000006601	0.03315
200	0.08972	0.1248	0.02891	0.088368	0.05794	0.02310	0.000004951	0.02487
250	0.08777	0.0998	0.02313	0.086694	0.04636	0.01848	0.000003661	0.01989
300	0.08648	0.0832	0.01927	0.085579	0.03863	0.01540	0.000003301	0.01658
400	0.08486	0.0624	0.01446	0.084184	0.02897	0.01155	0.000002475	0.01243
500	0.08389	0.0499	0.01156	0.083347	0.02318	0.00924	0.000001980	0.00995

EXHIBIT 4.5

THE OPTIMAL PORTFOLIOS AND THEIR CORRESPONDING MEAN
VARIANCE EFFICIENT IN THE MEAN STANDARD DEVIATION PLANE
(THE RISKFREE ASSET CASE)



SYMBOL USED FOR OPTIMAL PORTFOLIOS IS '+'
SYMBOL USED FOR MEAN VARIANCE EFFICIENT ONE IS '*'

on \$10,000) for a service that replaces a mean-variance efficient strategy with exact expected utility maximization. Indeed, this describes a remarkable performance for the mean-variance investment strategy.

The divergence of the results between the all risky assets case and riskfree assets case may be the result of two factors. While the non-existence of a riskless asset denies the investor the ability to mix his investment strategy with the least risky portfolio its existence provides this option. Given the obsession of increasingly risk averse investors to approach the least risky option, regardless of anything else, the disagreement between the mean-variance and the optimal strategy diverge in the first case and converge in the second. It is important to realize that the convergence between the two in the latter case is due to the fact that both converge to the riskless asset. Without such a common limit, the higher the degree of risk aversion the greater the tendency to base decisions not only on the mean and the variance, but on other information about the distribution. A risk neutral investor will base his decision on the mean only, disregarding other information about the distribution. Higher risk aversion means more prudence, reflected by examining more and more parameters of the distribution. Indeed, such a tendency is confirmed in the trend reflected in Exhibit 4.2. That is, higher risk

aversion leads to a greater tendency to consider information beyond mean and variance.

We began with a joint distribution that implies a theoretical inconsistency between mean-variance portfolio selection and expected utility maximization. Yet the results exhibit a larger domain for which mean-variance strategies lead to a fairly good approximate choice of expected utility maximization than that of Levy and Markowitz (1979). In Levy and Markowitz (1979) CARA utility started to perform poorly for $A > 5$. Exhibit 4.2 reflects that until $A = 15$ the forgone opportunity cost of mean-variance strategy is about 0.5%. It reaches a penny on the dollar for $A = 25$. When a riskless asset is introduced the optimization is almost 0.05% regardless of the level of risk aversion. This could be attributed to the fact that comparing the expected utility with its approximation for randomly selected portfolios rather than comparing the optimal portfolio to its mean-variance second best discriminates against mean-variance. Moreover, the fact that the approximation employed is heuristic, and not necessarily the best, together with the need to interpret conservatively the correlation coefficient between the expected utility and its approximation, results in a further discrimination. Whereas it is difficult to interpret what a correlation coefficient of 0.90 between the expected utility

and its approximation, the criteria used here is not subject to such difficulty and hence conservative interpretation.

We may conclude that when markets offer a riskless investment opportunity, mean-variance portfolio selection exhibits a very remarkable (for a CARA investor) performance regardless of the degree of risk aversion. The forgone opportunity cost is no more than 0.05 of the penny on an invested dollar. Such a forgone opportunity cost is ten times less than a typical brokerage fee. But when markets deny a riskless investment prospect mean-variance strategies exhibit a fair performance for a limited range of risk aversion. Its opportunity cost reaches 0.5 of a penny on an invested dollar for $A = 15$ and then it increases linearly with the degree of absolute risk aversion. It is worth noting that such a weak performance is due to a strange obsession on the part of a highly risk averse investor to pursue what he perceives to be the safest strategy regardless of the reward of even slightly more risky alternatives. When markets deny riskless investment prospect his "safest" portfolio is not a mean-variance efficient portfolio and, therefore, his frustration is higher the higher is his absolute risk aversion. The good performance of mean-variance portfolio selection, regardless of the risk aversion in the riskless asset case, supports the conclusion that such obsession is responsible for the

poor performance of the mean-variance strategy in the other case. A CARA investor without such obsession may use mean-variance strategy and enjoy its simplicity and its cost saving approach. Perhaps other utility functions may exhibit different kinds of obsession for extreme values of their parameters as the CARA utility did. This fact may lead to a poor performance of mean-variance strategies even with the existence of a riskless asset. This is because other classes of utility functions may imply, for extreme values of their parameters, an obsession to other characteristics different from the safest route. As long as such an obsession leads to a tendency toward a portfolio that does not belong to the mean-variance efficient frontier, mean-variance investment strategy may have similar performance. But we may conclude, cautiously, that mean-variance investment strategy performs fairly well for investors that do not exhibit extreme tendencies for any risk property.

APPENDIX 4.A

V.I. Krylov (1962) describes a numerical method to integrate any function with respect to Gamma function,

$$\int_0^{\infty} x^{\alpha} \exp(-x) f(x) dx \quad \text{utilizing Chebyshev-Laguerra orthogonal}$$

polynomials defined as

$$L_n^{(\alpha)}(x) = (-1)^n x^{-\alpha} \exp(x) \frac{d^n}{dx^n} x^{\alpha+n} \exp(-x)$$

A quadrature formula of the highest degree of precision is given by

$$\int_0^{\infty} x^{\alpha} \exp(-x) f(x) dx = \sum_{k=1}^n A_k f(x_k) + R(f)$$

where x_k are the roots of the Laguerre polynomial

$L_n^{(\alpha)}(x)$ and

$$A_k = \frac{n! \Gamma(n+\alpha+1)}{x_k (L_n^{(\alpha)}(x_k))^2}$$

Krylov gives in Appendix C of his book the values of x_k and A_k for different levels of n , for $\alpha = 0$ and $\alpha = 1$. The integral in (4.31) can be written in the

form of

$$\int_0^{\infty} f(X) \exp(-X) dX$$

where

$$f(X) = \exp\left(-\frac{1}{2} \left(\sum_{t=1}^N (Z_t - \mu - b(X-p))^2 \right) W(Z_t - \mu - b(X-p))\right) X^{(p-1)}$$

The approximation has been carried on for $n = 10$.

APPENDIX 4.B

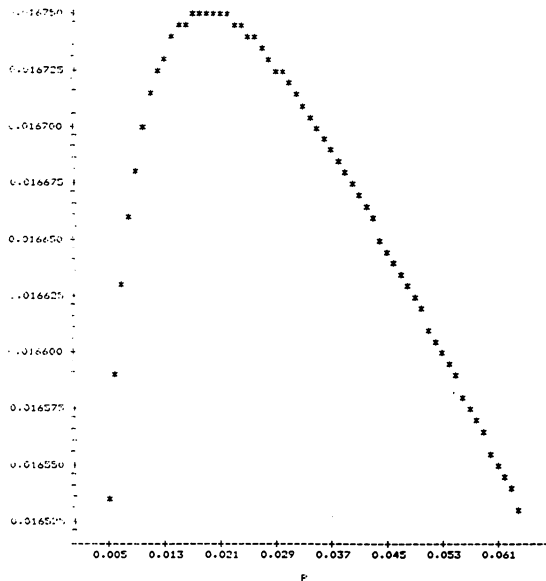
EXHIBIT 4.B.1

THE TABLE BELOW PROVIDES THE SMALLEST OF THE EIGENVALUES OF $\mathbf{M}$
AND THE POSTERIOR DISTRIBUTION OF $\mathbf{P}$

P	MINIMUM EIG- ENVALUE OF M	POSTERIOR PROB. OF P
0.005	0.00186903	0.0165366
0.006	0.00185572	0.0165907
0.007	0.00184248	0.0166300
0.008	0.00182858	0.0166596
0.009	0.00181414	0.0166823
0.010	0.00179905	0.0166999
0.011	0.00178376	0.0167134
0.012	0.00176670	0.0167241
0.013	0.00174910	0.0167327
0.014	0.00173101	0.0167385
0.015	0.00171180	0.0167432
0.016	0.00169162	0.0167465
0.017	0.00167044	0.0167487
0.018	0.00164825	0.0167500
0.019	0.00162503	0.0167504
0.020	0.00160078	0.0167502
0.021	0.00157552	0.0167493
0.022	0.00154925	0.0167479
0.023	0.00152201	0.0167460
0.024	0.00149382	0.0167437
0.025	0.00146474	0.0167410
0.026	0.00143481	0.0167380
0.027	0.00140408	0.0167347
0.028	0.00137261	0.0167311
0.029	0.00134045	0.0167273
0.030	0.00130767	0.0167232
0.031	0.00127431	0.0167190
0.032	0.00124045	0.0167146
0.033	0.00120612	0.0167100
0.034	0.00117138	0.0167053
0.035	0.00113629	0.0167005
0.036	0.00110088	0.0166955
0.037	0.00106520	0.0166904
0.038	0.00102929	0.0166852
0.039	0.00099319	0.0166800
0.040	0.00095693	0.0166746
0.041	0.00092054	0.0166692
0.042	0.00088405	0.0166636
0.043	0.00084748	0.0166581
0.044	0.00081085	0.0166524
0.045	0.00077419	0.0166467
0.046	0.00073752	0.0166410
0.047	0.00070084	0.0166352
0.048	0.00066418	0.0166293
0.049	0.00062755	0.0166233
0.050	0.00059097	0.0166175
0.051	0.00055443	0.0166116
0.052	0.00051795	0.0166056
0.053	0.00048155	0.0165996
0.054	0.00044521	0.0165936
0.055	0.00040897	0.0165875
0.056	0.00037281	0.0165814
0.057	0.00033674	0.0165753
0.058	0.00030077	0.0165692
0.059	0.00026491	0.0165631
0.060	0.000229149	0.0165570
0.061	0.000193497	0.0165508
0.062	0.000157956	0.0165446
0.063	0.000122529	0.0165385
0.064	0.000087216	0.0165323

EXHIBIT 4.R.2
THE GRAPH OF THE POSTERIOR DISTRIBUTION OF P

FIGURE
10.1



APPENDIX 4.C

THIS APPENDIX REPORTS THE PORTFOLIOS OF THE OPTIMAL INVESTMENT STRATEGY AND TWO CORRESPONDING MEAN VARIANCE EFFICIENT PORTFOLIOS FOR DIFFERENT LEVELS OF THE RISK AVERSION A

EXHIBIT 4.C.1

THE CASE WHEN ALL ASSETS ARE RISKY

A=2					
X	-0.312995	1.19884	0.142993	-0.0027074	-0.198552
	-1.7059	0.691236	0.743556	0.590534	-0.147004
XMVE	-0.322246	1.18479	0.142134	-0.0223036	-0.220274
	-1.69932	0.743376	0.728831	0.603656	-0.138639
A=4					
X	-0.143652	0.598967	0.254138	0.0219299	-0.0380621
	-0.812973	0.35086	0.414892	0.436045	-0.0821444
XMVE	-0.15536	0.582543	0.252899	-0.00234007	-0.0651751
	-0.806254	0.41587	0.397204	0.452516	-0.0719029
A=6					
X	-0.0864864	0.460179	0.291244	0.0316944	0.0171431
	-0.515935	0.233297	0.306535	0.383522	-0.0611938
XMVE	-0.0998391	0.382186	0.289749	0.0043015	-0.0135759
	-0.509144	0.306914	0.286876	0.402233	-0.0497008
A=8					
X	-0.0576177	0.30125	0.30982	0.0371941	0.0454251
	-0.367655	0.172882	0.252833	0.356852	-0.0509848
XMVE	-0.0721233	0.282168	0.308145	0.00761695	0.0121823
	-0.360828	0.252524	0.231801	0.377132	-0.0386175
A=10					
X	-0.0401607	0.242114	0.320977	0.0407871	0.0627172
	-0.278801	0.135858	0.220839	0.340656	-0.0449858
XMVE	-0.0555155	0.222235	0.319168	0.00960364	0.027617
	-0.271953	0.219932	0.198799	0.362091	-0.0319762
A=15					
X	-0.016669	0.163616	0.33587	0.0460438	0.0862861
	-0.160511	0.0852601	0.178538	0.318753	-0.0371882
XMVE	-0.0334068	0.142451	0.333841	0.0122483	0.0481641
	-0.153642	0.176545	0.154866	0.342069	-0.0231352

A=20					
X	-0.00460406 -0.101465	0.124561 0.0592805	0.343326 0.157587	0.0489294 0.307632	0.0983538 -0.0334004
XMVE	-0.0223722 -0.0945928	0.102631 0.15489	0.341165 0.132939	0.0135683 0.332075	0.0584193 -0.0187226
A=25					
X	0.0023622 -0.0660772	0.101205 0.0434226	0.347804 0.145095	0.050763 0.300891	0.105707 -0.0311718
XMVE	-0.0157595 -0.0592057	0.0787677 0.141913	0.345554 0.119798	0.0143594 0.326086	0.064565 -0.0160783
A=30					
X	0.0071621 -0.0425042	0.0856703 0.0327228	0.350791 0.136804	0.0520337 0.296365	0.110662 -0.0297069
XMVE	-0.0113548 -0.0356343	0.0628725 0.133269	0.348478 0.111045	0.0148863 0.322097	0.0686585 -0.0143169
A=40					
X	0.0131899 -0.0130613	0.0662977 0.0191887	0.354526 0.126487	0.0536824 0.290668	0.116923 -0.0279017
XMVE	-0.00585378 -0.00619691	0.0430209 0.122473	0.352129 0.100114	0.0155443 0.317115	0.073771 -0.0121171
A=50					
X	0.0168227 0.00459091	0.0547005 0.0109759	0.356769 0.120324	0.0547065 0.287227	0.120717 -0.0268337
XMVE	-0.00255597 0.0114508	0.0311201 0.116002	0.354318 0.0935608	0.0159388 0.314129	0.0768359 -0.0107983
A=60					
X	0.0192518 0.016353	0.0469807 0.00545735	0.358265 0.116227	0.0554049 0.284923	0.123264 -0.0261284
XMVE	-0.00358687 0.0232092	0.0231908 0.11169	0.355776 0.0891945	0.0162017 0.312139	0.078878 -0.00991967
A=70					
X	0.0209906 0.0247513	0.0414726 0.0014978	0.359334 0.113307	0.0559117 0.283271	0.125092 -0.0256281
XMVE	0.00121015 0.0316045	0.0175293 0.108611	0.356817 0.086077	0.0163893 0.310718	0.080336 -0.00929231
A=80					
X	0.0222968 0.0310484	0.037345 -0.00148535	0.360135 0.11112	0.0562964 0.28203	0.126468 -0.0252548
XMVE	0.0023864 0.0378991	0.0132846 0.106302	0.357598 0.0837396	0.01653 0.309653	0.0814292 -0.00882194

A=90						
X	0.023314 0.035945	0.0341367 -0.00381285	0.360759 0.109422	0.0565983 0.281062	0.127542 -0.0249657	
XHVE	0.00330104 0.0427936	0.00998393 0.104508	0.358205 0.0819221	0.0166395 0.308824	0.0822792 -0.00845618	
A=100						
X	0.0241285 0.0398616	0.0315713 -0.00567953	0.361258 0.108064	0.0568416 0.280287	0.128402 -0.0247351	
XHVE	0.00403261 0.0467085	0.00734393 0.103072	0.358691 0.0804684	0.016727 0.308162	0.0829591 -0.00816364	
A=150						
X	0.0265767 0.0516076	0.0238827 -0.0113052	0.362755 0.104	0.0575812 0.277955	0.130995 -0.0240476	
XHVE	0.00622651 0.0584488	-0.000573214 0.0987664	0.360147 0.0761088	0.0169894 0.306175	0.084998 -0.00728632	
A=200						
X	0.0278034 0.0574785	0.0200425 -0.0141328	0.363504 0.101971	0.0579566 0.276785	0.132297 -0.0237063	
XHVE	0.007323 0.0643165	-0.00453012 0.0966146	0.360874 0.0739299	0.0171206 0.305182	0.0860171 -0.00684785	
A=250						
X	0.0285403 0.0610003	0.0177398 -0.0158342	0.363954 0.100756	0.0581837 0.276083	0.13308 -0.0235023	
XHVE	0.00798074 0.0678363	-0.00690371 0.0953239	0.361311 0.0726229	0.0171993 0.304586	0.0866284 -0.00658482	
A=300						
X	0.0290319 0.0633478	0.0162053 -0.0169706	0.364253 0.0999463	0.0583359 0.275613	0.133604 -0.0233666	
XHVE	0.00841917 0.0701825	-0.00848588 0.0944635	0.361602 0.0717517	0.0172517 0.304189	0.0870358 -0.0064095	
A=400						
X	0.0296468 0.0662819	0.0142878 -0.0183934	0.364628 0.0989348	0.058527 0.275026	0.134258 -0.0231974	
XHVE	0.00896714 0.0731148	-0.0104633 0.0933881	0.361966 0.0706628	0.0173173 0.303693	0.0875451 -0.00619037	
A=500						
X	0.030016 0.0680422	0.0131377 -0.0192483	0.364853 0.0983283	0.0586421 0.274674	0.134652 -0.0230961	
XHVE	0.00929588 0.074874	-0.0116496 0.092743	0.362184 0.0700095	0.0173566 0.303395	0.0878506 -0.00605892	

EXHIBIT 4.C.2
THE RISKLESS ASSET CASE

A=2				
X	-3.21512	-0.210334	1.22857	1.31904
	0.190357	0.242778	-1.46567	0.61031
	1.05424	1.4671	-0.22127	
XMVE	-1.01745	-0.330377	1.23641	0.498964
	-0.00674814	-0.147237	-1.71726	0.872255
	0.835014	0.928276	-0.151837	
A=4				
X	-1.10756	-0.105167	0.614287	0.659521
	0.0951783	0.121389	-0.732837	0.305155
	0.52712	0.733548	-0.110635	
XMVE	-0.008726	-0.165189	0.618203	0.249482
	-0.00337407	-0.0736184	-0.858632	0.436127
	0.417507	0.464138	-0.0759186	
A=6				
X	-0.405039	-0.0701113	0.409525	0.43968
	0.0634522	0.080926	-0.488558	0.203437
	0.351413	0.489032	-0.0737567	
XMVE	0.327516	-0.110126	0.412136	0.166321
	-0.00224938	-0.0490789	-0.572421	0.290772
	0.278338	0.309425	-0.0506124	
A=8				
X	-0.0537793	-0.0525835	0.307143	0.32976
	0.0475891	0.0606945	-0.366419	0.152577
	0.26356	0.366774	-0.0553175	
XMVE	0.495637	-0.0825943	0.309102	0.124741
	-0.00168704	-0.0368092	-0.429316	0.218064
	0.208753	0.232069	-0.0379593	
A=10				
X	0.156977	-0.0420668	0.245715	0.263808
	0.0380713	0.0485556	-0.293135	0.122062
	0.210848	0.293419	-0.044254	
XMVE	0.59651	-0.0660754	0.247281	0.0997929
	-0.00134963	-0.0294474	-0.343453	0.174451
	0.167003	0.185655	-0.0303674	
A=15				
X	0.437984	-0.0280445	0.16381	0.175872
	0.0253809	0.0323704	-0.195423	0.0813746
	0.140565	0.195613	-0.0295027	
XMVE	0.731006	-0.0440503	0.146854	0.0665286
	-0.000897752	-0.0196316	-0.228969	0.116301
	0.111335	0.12377	-0.020245	

A=20				
X	0.578488	-0.0210334	0.122857	0.131904
	0.0190357	0.0242778	-0.146567	0.061031
	0.105424	0.14671	-0.022127	
XHVE	0.798255	-0.0330377	0.123641	0.0498964
	-0.000674814	-0.0147237	-0.171726	0.0872255
	0.0835014	0.0928276	-0.0151837	
A=25				
X	0.662791	-0.0168267	0.0982859	0.105523
	0.0152285	0.0194222	-0.117254	0.0488248
	0.0843392	0.117368	-0.0177016	
XHVE	0.838604	-0.0264302	0.0989125	0.0399172
	-0.000539851	-0.0117789	-0.137381	0.0697804
	0.0668011	0.0742621	-0.012147	
A=30				
X	0.718992	-0.0140223	0.0819049	0.0879361
	0.0126904	0.0161852	-0.0977116	0.0406873
	0.0702827	0.0978064	-0.0147513	
XHVE	0.865503	-0.0220251	0.0824271	0.0332643
	-0.000449876	-0.00981579	-0.114484	0.0581503
	0.0556676	0.061885	-0.0101225	
A=40				
X	0.789244	-0.0105167	0.0614287	0.0659521
	0.00951783	0.0121389	-0.0732837	0.0305155
	0.052712	0.0733548	-0.0110635	
XHVE	0.899127	-0.0165189	0.0618203	0.0249482
	-0.000337407	-0.00736184	-0.0858632	0.0436177
	0.0417507	0.0464138	-0.00759186	
A=50				
X	0.932558	-0.00336534	0.0196572	0.0211047
	0.0030457	0.00388445	-0.0234508	0.00976496
	0.0168678	0.0234735	-0.00354032	
XHVE	0.967721	-0.00528603	0.0197825	0.00798343
	-0.00010797	-0.00235579	-0.0274762	0.0139561
	0.0133602	0.0148524	-0.0024294	
A=60				
X	0.859496	-0.00701113	0.0409525	0.043968
	0.00634522	0.0080926	-0.0488558	0.0203437
	0.0351413	0.0489032	-0.00737567	
XHVE	0.932752	-0.0110126	0.0412136	0.0166321
	-0.000224938	-0.00490789	-0.0572421	0.0290752
	0.0278338	0.0309425	-0.00506124	
A=70				
X	0.879568	-0.00600954	0.0351021	0.0376869
	0.00543876	0.00693651	-0.0418764	0.0174374
	0.0301212	0.041917	-0.006322	
XHVE	0.942359	-0.00943935	0.0353259	0.0142561
	-0.000197804	-0.00420677	-0.0490647	0.0249216
	0.0238575	0.0265222	-0.00433821	

A=80				
X	0.894622	-0.00525835	0.0307143	0.032976
	0.00475891	0.00606945	-0.0366419	0.0152577
	0.026356	0.0366774	-0.00553175	
XMVE	0.949564	-0.00825943	0.0309102	0.0124741
	-0.000168704	-0.00368092	-0.0429316	0.0218064
	0.0208753	0.0232069	-0.00379593	
A=90				
X	0.906331	-0.00467409	0.0273016	0.029312
	0.00423015	0.00539507	-0.0325705	0.0135624
	0.0234276	0.0326021	-0.00491711	
XMVE	0.955168	-0.00734171	0.0274757	0.0110881
	-0.000149959	-0.00327193	-0.0381614	0.0193834
	0.0185559	0.0206283	-0.00337416	
A=100				
X	0.915698	-0.00420668	0.0245715	0.0263808
	0.00380713	0.00485554	-0.0293135	0.0122062
	0.0210848	0.0293419	-0.0044254	
XMVE	0.959651	-0.00660754	0.0247281	0.00997929
	-0.000134963	-0.00294474	-0.0343453	0.0174451
	0.0167003	0.0185655	-0.00303674	
A=150				
X	0.943798	-0.00280445	0.016381	0.0175872
	0.00253809	0.00323704	-0.0195423	0.00813746
	0.0140565	0.0195613	-0.00295027	
XMVE	0.973101	-0.00440503	0.0164854	0.00665286
	-0.000089975	-0.00196316	-0.0228969	0.0116301
	0.0111335	0.012377	-0.0020245	
A=200				
X	0.957849	-0.00210334	0.0122857	0.0131904
	0.00190357	0.00242778	-0.0146567	0.0061031
	0.0105424	0.014671	-0.0022127	
XMVE	0.979825	-0.00330377	0.0123641	0.00498964
	-0.000067481	-0.00147237	-0.0171726	0.00872255
	0.00835014	0.00928276	-0.00151837	
A=250				
X	0.966279	-0.00168267	0.00982859	0.0105523
	0.00152285	0.00194222	-0.0117254	0.00488248
	0.00843392	0.0117368	-0.00177016	
XMVE	0.98386	-0.00264302	0.00989125	0.00399172
	-0.000053985	-0.00117789	-0.0137381	0.00697804
	0.00668011	0.00742621	-0.0012147	
A=300				
X	0.971899	-0.00140223	0.00819049	0.00879361
	0.00126904	0.00161852	-0.00977116	0.00406873
	0.00702827	0.00978064	-0.00147513	
XMVE	0.98655	-0.00220251	0.00824271	0.00332643
	-0.000044988	-0.000981579	-0.0114484	0.00581503
	0.00556676	0.0061885	-0.00101225	

A=400

X	0.978924	-0.00105167	0.00614287	0.00639521
	0.000951783	0.00121389	-0.00732837	0.00305155
	0.0052712	0.00733548	-0.00110635	

XHVE	0.989913	-0.00165189	0.00618203	0.00249482
	-0.000033741	-0.000736184	-0.00858632	0.00436127
	0.00417507	0.00464138	-0.000759186	

A=500

X	0.98314	-0.000841335	0.00491429	0.00527617
	0.000761426	0.000971112	-0.0058627	0.00244124
	0.00421696	0.00586838	-0.000885081	

XHVE	0.99193	-0.00132151	0.00494563	0.00199586
	-0.000026993	-0.000588947	-0.00686906	0.00348902
	0.00334005	0.0037131	-0.000607349	

CHAPTER V

THREE PARAMETER CAPITAL ASSET PRICING MODEL

The class of non-spherical distributions, introduced in Chapter III, allowed a derivation of a normative theory for portfolio selection as an extension of the mean-variance normative theory. Within its three-parameter framework we were able to assess the performance of mean-variance portfolio selection. As the mean-variance framework allows a derivation of asset pricing theory in a perfect competitive market, our three parameter normative theory allows a similar derivation of a capital asset pricing model (CAPM) as an extended version of the conventional Sharpe-Linter CAPM.

Two previous attempts in the literature have been made to derive a mean-variance-skewness CAPM as an extension to the Sharpe-Linter CAPM. These are of Kraus and Litzenberger (1976) and (1983). The approach in the first paper begins in postulating a mean-variance-skewness determined expected utility. Instead of providing assumptions on the return generating process and investor preference that determine the nature of the dependence of the expected utility on the moments of the portfolio return, they postulated a general form of the expected utility that depends on the first three

moments of the portfolio return distribution. Although such an approach is proper for the two-moments case it creates theoretical problems for the three-moments case.

The variance of the portfolio return is a convex function of the decision variable of the portfolio problem. Therefore, a risk aversion to variance warrants the concavity of the expected utility in the decision variables. The concavity property warrants both the existence of maxima and the sufficiency of the first order conditions for maxima. Unlike the two-moment case such a property is lost in the three-moment case. Since skewness is neither a concave or a convex function of the decision variables of the portfolio problem we cannot postulate the concavity of the expected utility as long as the skewness of the portfolio distribution is given in a general form of the skewness structure of the individual assets. Once the expected utility is not concave, the first order conditions are not sufficient for maxima. The existence of maxima is not guaranteed unless the portfolio space is closed and bounded. Kraus and Litzenberger (1976) postulated an unrestricted functional form of a mean-variance-skewness expected utility. They took the first order conditions of the portfolio problem of an individual investor to describe his equilibrium. It did not occur to the authors to comment on the sufficiency of the first order conditions. In particular, a three-moment truncated Taylor

expansion of the expected utility faces the same difficulty. Therefore, the implications of Klaus and Litzenberger (1976) approach to capital asset pricing is not correct.

Kraus and Litzenberger (1983) provided a consumption oriented three-moment CAPM in a state preference framework. They proposed the "quadratic characteristic lines" on the consumption portfolio, to be defined below, as sufficient for a three-moment CAPM. It will be proved that these conditions imply mean-variance efficiency and the Sharpe-Linter CAPM for which their three-moment CAPM is supposed to be alternative. Moreover, it will be proved that the Sharpe-Linter CAPM is robust under any functional form of dependence between security returns and the return on the consumption portfolio, provided that other conditions in the definition of the quadratic characteristic lines exist.

By restricting the return generating process to the class of non-spherical distributions, given by (3.1), we are able to derive a non-trivial three moment CAPM (not reducible to a two-moment CAPM). This is done in the second section of this chapter. According to this model, the risk premium for any given security is determined by the risk premium of the market portfolio, the risk premium of another efficient fund (not stochastically dominated fund), and by a security specific two dimensional measure of systematic risk. This measure of systematic risk is contingent upon the relation

of the second fund to the market portfolio. In particular, the second fund could be chosen as neutral to the variance of the market portfolio (zero beta portfolio). For this choice, the first component of the systematic risk is beta: the relative contribution of this security to the variance of market portfolio. The second component is the excess of the relative contribution of a given security to the skewness of the market portfolio over its relative contribution to the variance of the market portfolio standardized by the relative contribution of the zero beta portfolio to the skewness of the market portfolio.

In this framework, the Sharpe-Linter CAPM provides a correct pricing for a given security only under two conditions: (i) When the zero beta portfolio provides neither a premium or a discount over the riskfree rate. Recall that the zero beta portfolio contributes only to the skewness of the market portfolio but not to its variance. If such a portfolio neither offers a premium or a discount over the riskfree rate, then the market investor (the investor that holds the market portfolio) is neutral toward skewness. In this case, the zero beta portfolio and the riskfree assets are priced identically. (ii) When the relative contribution of a given security to the variance of the market portfolio is identical to its contribution to the skewness of the market portfolio. While the first condition implies the consistency

of all pricing all assets with the Sharpe-Linter CAPM, the second condition may hold only for a subset of securities. The direction of the divergence of our model from the Sharpe-Linter CAPM depends on both whether the market investor exhibits preference or aversion toward skewness and on the relative magnitudes of the contribution of a given security to the variance and skewness of the market portfolio. When the riskless asset does not exist, it is proved that the capital market allows the creation of a third fund neutral to both the variance and the skewness of the market portfolio. This fund plays the role of the riskfree asset in the previous case. It is the analog of the zero-beta portfolio in Black (1973) version of CAPM.

5.1 Quadratic characteristic lines and the CAPM

Kraus and Litzenberger (1983) provided conditions on the probability distribution of assets returns sufficient for a two-moment and a three-moment consumption-oriented CAPM. For a two-moment CAPM, their sufficient conditions require a capital market that achieves an ex-ante Pareto optimality and security returns that follows linear characteristic lines on the consumption portfolio. Their sufficient conditions for a three-moment CAPM require the security returns to follow quadratic characteristic lines and a capital market that achieves an ex-ante Pareto optimality. It will be proved that their sufficient conditions for a two-moment CAPM are a special case of the necessary and sufficient conditions, provided by Ross (1978), for a set of assets returns to exhibit two funds separation. As long as the variance of the security returns exists, the class of two funds separating distributions admits mean-variance portfolio selection. It will be proved also that the quadratic characteristic lines, proposed by Kraus and Litzenberger (1983) to be sufficient for a three-moment CAPM, together with a capital market equilibrium condition imply mean-variance efficiency and the Sharpe-Linter CAPM.

Therefore, such conditions are trivial for a three-moment CAPM.

It will be proved also that the Sharpe-Linter CAPM is robust under weaker conditions than those that define the quadratic characteristic lines on the consumption portfolio. First, it will be proved that the dependence between the individual security return and the return on the consumption portfolio may take any functional form. The fact that the consumption portfolio is the only risk factor common across securities, the restrictions on the security specific residuals in the definition of the linear characteristic lines (or the quadratic characteristic lines) given below, and the capital market equilibrium imply the Sharpe-Linter CAPM. Second, any market efficient portfolio may replace the consumption portfolio without influencing the above results. In order for a non-trivial three-moment CAPM (not reduceable to a two-moment CAPM) to take place, the security returns must exhibit a three funds separation. Conditions that characterize such distributions are given in Ross (1978).

A: Linear Characteristic Lines

According to Kraus and Litzenberger (1983), a set of security returns,

$$\{Z_i\}_{i=1}^n,$$

follow linear characteristic lines on the consumption portfolio if:

$$Z_i = \mu_i + \beta_i [R_C - E(R_i)] + \epsilon_i \quad (1)$$

where

$$\mu_i = E(Z_i)$$

and

$$\beta_i = \frac{\text{COV}(Z_i, R_C)}{\text{Var}(R_C)}$$

and

$$E[\epsilon_i | R_C] = 0$$

In Appendix 5.A it is proved that the capital market equilibrium, the linear characteristic lines, and the existence of the risk free asset imply:²

$$\beta_i = \frac{\mu_i - R_F}{E(R_C) - R_F} \quad (2)$$

On the other hand, a vector of assets returns, $Z = (Z_1, \dots, Z_n)$, exhibits two funds separation if there exist two mutual funds, X_1, X_2 , such that for any portfolio $\hat{X}$ there exist λ such that:

$$X = \lambda X_1 + (1-\lambda)X_2$$

with

$$EU(X'Z) \geq EU(\hat{X}'Z)$$

where U is any monotone non-decreasing utility function. Concave separation refers to the case when U is concave. Ross (1978) provided conditions that characterize all distributions that imply two funds separation. The following theorem provides such conditions for concave separation.

Theorem 1: A vector of security returns, Z , exhibits a concave two funds separation if, and only if, there exist two random variables y and v , a vector b and two portfolios X_1, X_2 such that:³

$$(i) \quad Z_i = y + b_i v + \epsilon_i \quad i = 1, 2, \dots, n$$

$$E(\epsilon_i) = 0$$

$$(ii) \quad E[\epsilon_i | y + tv] = 0 \quad \text{for any } t \text{ and } i = 1, 2, \dots, n$$

$$(iii) \quad \epsilon'X_1 = \epsilon'X_2 \equiv 0 \quad \text{almost everywhere (a.e.)}$$

Ross also proved that any efficient portfolio for a two funds separating distribution is mean-variance efficient, as long as its variance exists.⁴ Given the above theorem, the definition of linear characteristic lines, and the implication of the capital market equilibrium given by equation (2), we can prove the following theorem.

Theorem 2: A capital market equilibrium with security returns that follow linear characteristic lines relative to any market efficient portfolio implies two funds separation.

Proof: Equation (1) may be written as:

$$Z_i = \mu_i - \beta_i [E(R_C) - R_F] + \beta_i (R_C - R_F) + \epsilon_i$$

utilizing equation (2) we have:

$$Z_i = \mu_i - \frac{\mu_i - R_F}{E R_C - R_F} [E(R_C) - R_F] + \beta_i (R_C - R_F) + \epsilon_i$$

or:

$$Z_i = R_F + \beta_i (R_C - R_F) + \epsilon_i \quad (3)$$

Equation (3) is a special case of the form of condition (i) in theorem 1. If we let $y \equiv R_F$ a.e., $v = R_C - R_F$ and $b_i = \beta_i$, then the two forms are identical. Since $y \equiv R_F$ a.e., condition (ii) is identical to the restriction $E[\epsilon_i | R_C] = 0$. We need to show that there exist two funds that satisfy (iii) of theorem 1. Denote by index 0 the riskfree asset and recall that such an asset satisfies equation (3) with $b_0 = 0$ and $\epsilon_0 \equiv 0$ a.e. If we let X_C be the consumption portfolio and $X_F = (1, 0, \dots, 0)'$, then we can prove that these portfolios satisfy (iii) of theorem 1. By the construction of X_F we have $\epsilon' X_F \equiv 0$. To prove that $\epsilon' X_C \equiv 0$, multiply equation (1) by X_{Ci} and add over i :

$$R_C = E(R_C) + \left(\sum_{i=1}^n \beta_i X_{Ci} \right) [R_C - E(R_C)] + \epsilon' X_C \quad (4)$$

But

$$\sum_{i=1}^n X_{ci} \beta_i = \sum_{i=1}^n \frac{X_{ci} \text{COV}(Z_i, R_c)}{\text{Var}(R_c)} =$$

$$\sum_{i=1}^n \sum_{j=1}^n \frac{X_{ci} X_{cj} \text{COV}(Z_i, Z_j)}{\text{Var}(R_c)} = 1$$

This fact, together with equation (4), completes the proof.

Q.E.D

Since two funds separation is achieved if, and only if, assets returns are restricted to (i), (ii), and (iii) of theorem 1, a two-moment CAPM is achieved if, and only if, security returns are restricted to such conditions and the second moments of security returns exist. It will be proved that even quadratic characteristic lines imply two-funds separation.

B: Quadratic Characteristic Lines

According to Kraus and Litzenberger (1983), a set of security returns follow quadratic characteristic lines relative to the return on the consumption portfolio if;

$$Z_i = \mu_i + b_{1i} [R_c - E(R_c)] + b_{2i} \{[R_c - E(R_c)]^2 - \sigma_c^2\} + \epsilon_i \quad (5)$$

where

$$E(\varepsilon_i) = E[\varepsilon_i | R_C] = 0 \quad i = 1, 2, \dots, n$$

and

$$u_i = E(Z_i)$$

and

$$\sigma_C^2 = E[R_C - E(R_C)]^2$$

Krause and Litzenberger (1983) proved that a market equilibrium condition and quadratic characteristic lines imply a three-moment CAPM. It will be shown that such conditions are trivially sufficient for a three-moment CAPM since they imply a two-moment CAPM.

Theorem 3: A capital market that achieves equilibrium with security returns that follow quadratic characteristic lines implies a two-moment CAPM.

Proof: It is sufficient to show that the assumptions of this theorem imply that security returns satisfy (i), (ii), and (iii) of theorem 1. Let X_C , as before, denote the consumption portfolio and multiply equation (5) by X_{Ci} and add over i :

$$R_C = E(R_C) + b_{1C} [R_C - E(R_C)] + b_{2C} \{ [R_C - E(R_C)]^2 - \sigma_C^2 \} + \varepsilon' X_C \quad (6)$$

where

$$b_{jc} = \sum_{i=1}^n b_{ji} X_{ci} \quad \text{and} \quad j = 1, 2$$

Solving equation (6) for $\{[R_c - E(R_c)]^2 - \sigma_c^2\}$ we have:

$$\{[R_c - E(R_c)] - \sigma_c^2\} = \frac{(1 - b_{1c})}{b_{2c}} [R_c - E(R_c)] - \frac{1}{b_{2c}} \epsilon' X_c$$

Substituting this result back in equation (5) implies:

$$Z_i = \mu_i + (b_{1i} + \frac{1 - b_{1c}}{b_{2c}} b_{2i}) [R_c - E(R_c)] + (\epsilon_i - \frac{1}{b_{2i}} \epsilon' X_c) \quad (7)$$

or

$$Z_i = \mu_i + \gamma_i [R_c - E(R_c)] + \delta_i \quad (8)$$

Moreover, we can prove that: (i) $\delta' X_c = 0$ and

(ii) $E[\delta_i | R_c] = 0$. Multiply equation (8) by X_{ci} and add over i we have:

$$R_c = E(R_c) + \gamma_c [R_c - E(R_c)] + \delta' X_c \quad (9)$$

But:

$$\gamma_c = \sum_{i=1}^n X_{ci} \gamma_i = \sum_{i=1}^n X_{ci} (b_{1i} + \frac{1 - b_{1c}}{b_{2c}} b_{2i}) = 1$$

This with equation (9) imply $\delta'X_C = 0$. Also we have:

$$E[\delta_i | R_C] = E[\epsilon_i | R_C] - \frac{1}{b_{2C}} \sum_{i=1}^n X_{Ci} E[\epsilon_i | R_C] = 0 \quad (10)$$

Equation (8) can be written as:

$$Z_i = \mu_i - \gamma_i [E(R_C) - R_F] + \gamma_i (R_C - R_F) + \delta_i \quad (11)$$

The equilibrium condition (iii) of footnote 2, equation (8), and equation (10) imply that:

$$\gamma_i = \frac{\mu_i - R_F}{E(R_C) - R_F}$$

Substituting this in equation (11) we have:

$$Z_i = R_F + \gamma_i [R_C - E(R_C)] + \delta_i \quad (12)$$

According to theorem (2), this implies the conditions (i), (ii) and (iii) of theorem 1. Q.E.D

Again, any market efficient portfolio can replace the consumption portfolio without affecting the results. Observe further that any return generating process of the form:

$$Z_i = \mu_i + b_{1i} [R_C - E(R_C)] + b_{2i} [h(R_C) - E(h(R_C))] + \epsilon_i \quad (13)$$

where:

$$E(\epsilon_i) = E[\epsilon_i | R_C] = 0$$

leads to a two-moment CAPM. This is because if we multiply equation (13) by X_{ci} , add over i , and solve for $[h(R_C) - E(h(R_C))]$ we have

$$[h(R_C) - E(h(R_C))]\epsilon'X_C = \frac{1 - b_{1c}}{b_{2c}} [R_C - E(R_C)] - \frac{1}{b_{2c}} \epsilon'X_C$$

Substituting this result back in (13) leads to a form equivalent to equation (12) with the same properties. The function $h(R_C)$, in particular, could be a polynomial with any rank.

Unless the security returns exhibit a non-trivial three funds separation, a non-trivial three-moment CAPM will not take place. Conditions necessary and sufficient for a three funds separation were provided by Ross (1978). In our setting, besides the return on the consumption portfolio (or the market portfolio), an additional risk factor common across securities must be introduced (as an extension of (i) of theorem 1). Such a risk factor must not be determined by the return on the consumption portfolio. Furthermore, the security specific residuals are required to satisfy an extended version of conditions (ii) and (iii) of theorem 1. Namely, the residuals must exhibit a fair game property to any linear combination of the two common risk factors. There must also exist an

additional fund, linearly independent of the consumption portfolio, that completely diversifies the security specific residuals. As long as security returns are generated by one common risk factor with the above restrictions on the residuals and a riskfree asset exists, a capital market equilibrium implies the Sharpe-Linter CAPM.

5.2 Mean-variance-skewness CAPM

The joint distribution of assets returns (3.1) gives the return on any individual security as a linear combination of a spherical random variable and a non-spherical random variable. Moreover, its structure is closed under the linear combination of individual securities. Namely, the portfolio return itself is a linear combination of a spherical noise and a non-spherical random variable. This is because the return on any portfolio X is given as:

$$X'Z = X'\mu + X'b \cdot y + X'\epsilon$$

Where $X'\mu + X'\epsilon$ is the spherical component of the portfolio distribution whereas $(X'b)y$ is the non-spherical part. The non-spherical nature of each security i is completely determined by the parameter b_i . For a given spherical noise of security i , the larger b_i is, the larger is the skewness of this security. It is also true that the larger $X'b$ is, for a given mean and variance, the larger is the skewness of the portfolio distribution. Moreover, there is a one to one relationship between $X'b$ and the skewness of the portfolio distribution. With distribution (3.1) an investor who decides on the mean, the variance, and the skewness of the portfolio distribution, simultaneously decides on the value of the non-sphericity parameter $X'b$, the mean and the variance of the portfolio distribution.

However, any one of these decisions determines completely the portfolio distribution. The fact that the portfolio choice for an individual investor, in distribution (3.1), collapses to a choice of the first three moments of the portfolio distribution, leads to a mean-variance-skewness capital assets pricing model. In such a model, the risk premiums of any three efficient funds together with a two dimensional measure of systematic risk for an individual security, determine its risk premium. Moreover, the measure of the systematic risk of an individual security reflects its contribution to the variance and the skewness of the market portfolio.

Chapter III describes the solution of the portfolio problem for an individual investor, regardless of his preference, for two cases: First when markets offer a riskless asset and second when markets deny a riskless security. The two cases will be treated separately in this section also. The standard assumption of a perfect competitive market is postulated. In particular, this assumption implies homogenous expectations and no restrictions on short sales.

5.3.1 The case with riskless assets:

Equation (3.21), in chapter III, describes the equilibrium decision of an individual investor facing a

riskless asset and n risky assets that follow the joint distribution (3.1). This condition may be written as:

$$VX - \lambda_1(u - r_f e) - (\lambda_3 + \sigma_y^2 b^1 X) b = 0 \quad (5.15)$$

or:

$$\sum_{j=1}^n v_{ij} X_j - \lambda_1(u_i - r_f) - (\lambda_3 + \sigma_y^2 b^1 X) b_i = 0$$

If we let $\gamma_3 = [(\lambda_3 + \sigma_y^2 b^1 X)/E^{1/3}(y^3)]$, $s_i = E^{1/3}(z_i - u_i)^2$,

$r_m = \sum_{j=1}^n z_j X_j$, and $\sigma_{im} = \text{cov}(z_i, r_m) = \sum_{j=1}^n V_{ij} X_j$, then the

equilibrium of an individual investor is given by:

$$\lambda_1(Ez_i - r_f) + \gamma_3 s_i - \sigma_{im} = 0 \quad \forall i=1,2,\dots,n \quad (5.16)$$

Any efficient portfolio of risky securities must satisfy (5.16). To demonstrate this, consider an arbitrary portfolio $\hat{X}$ that satisfies the equilibrium condition of an individual investor with $r_p = \hat{X}'Z$ and $\hat{X}'e = 1$. If we multiply (5.16) by $\hat{x}_i$ and add over i we have:

$$\lambda_1(E(r_p) - r_f) + \gamma_3 s_p - \sigma_{pm} = 0$$

This fact implies that the expected return on any security is determined by the relationship of an individual security and any two efficient funds. Once we have any two efficient funds, we can solve for λ_1 , and γ_3 .

Substituting the solution for λ_1 and γ_3 in (5.16) provides a relationship between any individual security expected return and these funds. Recall that $r_m = \Sigma x_i z_i$ and $r_p = \Sigma \hat{x}_i z_i$ and multiply (5.16) first by x_i and add over i and then by $\hat{x}_i$ and add over i . This leads to the following system:

$$\lambda_1 E(r_m - r_f) - \gamma_3 S_m - \sigma_m^2 = 0$$

$$\lambda_1 (E r_p - r_f) - \gamma_3 S_p - \sigma_{pm} = 0$$

or equivalently:

$$E(z_i - r_f) = k_1 S_i + k_2 \sigma_{im} \quad (5.17)$$

and

$$\begin{aligned} E(r_m - r_f) &= k_1 S_m + k_2 \sigma_m^2 \\ E(r_p - r_f) &= k_1 S_p + k_2 \sigma_{pm} \end{aligned} \quad (5.18)$$

Solving (5.18) for k_1 and k_2 and substituting in (5.17) leads to:

$$\begin{aligned} E(z_i - r_f) &= \left(\frac{S_p \sigma_{im} - S_i \sigma_{pm}}{S_p \sigma_m^2 - \sigma_{pm}^2} \right) E(r_m - r_f) + \left(\frac{\sigma_m^2 S_i - \sigma_{im} S_m}{S_p \sigma_m^2 - \sigma_{pm}^2} \right) \\ &\quad E(r_p - r_f) \end{aligned}$$

If we let $\beta_i = \sigma_{im}/\sigma_m^2$ and $\beta_p = (\sigma_{pm}/\sigma_m^2)$, then we have:

$$E(z_i - r_f) = \left(\frac{S p \beta_i - S_i \beta_p}{S_p - \beta_p S_m} \right) E(r_m - r_f) + \left(\frac{S_i - \beta_i S_m}{S_p - \beta_p S_m} \right) E(r_p - r_f) \quad (5.19)$$

Equation (5.19) describes the equilibrium condition of an individual investor. It relates the expected return on any individual security to any arbitrary efficient funds. For a general equilibrium in a capital market to take place, at least two funds must satisfy (5.19) for all investors. The capital market will be in equilibrium only if the market portfolio is an efficient portfolio for all investors. Let r_m in (5.19) denote, from now on, the return on the market portfolio. For (5.19) to describe a general equilibrium relationship, portfolio p must be efficient for all investors. We may choose portfolio p with any value of β_p and in particular we may choose such a portfolio with $\beta_p = 0$.⁹ If we let $\gamma_i = S_i/S_m$ and $\gamma_p = S_p/S_m$ then we have:

$$E(z_i - r_f) = \beta_i E(r_m - r_f) + (\gamma_i - \beta_i) \frac{E(r_p - r_f)}{\gamma_p} \quad (5.20)$$

⁹There always exists an efficient portfolio with zero correlation with the market portfolio. Let X_p be an efficient portfolio. Then X_p must be given as:

$$X_p = \lambda_1 V^{-1}(\mu - r_f e) + \lambda_2 \bar{V}^{-1}b \quad \text{and} \quad e'X_p = 0$$

X_p will be a zero beta portfolio if, and only if:

$$\begin{aligned} \sigma_{pm} &= X_p' V X_m = \lambda_1 E(r_m - r_f) + \lambda_2 S_m |E|^{1/3} (\gamma^3) = 0 \\ e'X_p &= \lambda_1 e' \bar{V}^{-1}(\mu - r_f e) + \lambda_2 e' \bar{V}^{-1}b = 1 \end{aligned}$$

The above system has a unique solution for λ_1 and λ_2 . Therefore the zero beta portfolio, relative to the market portfolio, is unique. Observe that any efficient portfolio has a counterpart portfolio with a return uncorrelated to its return

Observe that β_i is the contribution of the individual security to the variance of the market portfolio standardized by the variance of the market portfolio. It is the well known measure of the systematic risk in the conventional Sharpe-Linter CAPM. γ_i , on the other hand, is the contribution of the individual security to the market portfolio skewness standardized by the market's skewness. To see the implication of this model and its difference from the conventional CAPM we have to discriminate between two cases:

(a) The risk premium on the zero beta portfolio, $E(r_p - r_f)$, is positive: This implies that the capital market considers skewness as risk and hence bad. Therefore, a premium is charged on skewed securities. This is because the variance of portfolio p is completely diversifiable by its definition. This portfolio contributes only to the skewness of the market portfolio. A positive risk premium on such a portfolio implies that the average market investor dislikes skewness. For this case, if $\gamma_i - \beta_i > 0$, the conventional CAPM underestimates the risk premium of security i . The fact that $(\gamma_i - \beta_i) > 0$ means that the variance of the market portfolio is less sensitive to security i than is the skewness of the market portfolio. The divergence from the Sharpe-Linter CAPM depends on the difference $(\gamma_i - \beta_i)$, as well as, the price of the skewness risk of the zero beta portfolio, namely

$(E r_p - r_f)/\gamma_p$. On the other hand, when the skewness of the market portfolio is less sensitive to security i than the variance of the market portfolio ($(\gamma_i - \beta_i) < 0$), the Sharpe-Linter CAPM overestimates the risk premium of such a security.

(b) The risk premium of the zero beta portfolio is negative. This implies that the average market investor exhibits preference for skewness and therefore discounts the risk premium on a security that exhibits positive contribution to the skewness of the market portfolio on one hand, while its contribution to its variance is zero on the other. Here, the results of the previous case work in opposite directions. Namely, the larger is the sensitivity of the market portfolio skewness to a given security, relative to the sensitivity of the variance of the market portfolio to the same given security, the lower its risk premium. In this case the Sharpe-Linter CAPM overestimates the risk premium of such a security.

Observe that the Sharpe-Linter CAPM is a good approximation either when $E(r_p) - r_f$ is close to zero or when $\gamma_i - \beta_i$ is close to zero. The first implies that the market investor is neutral toward skewness. He neither charges risk premium for skewness nor does he offer a discount for such a property. The latter case implies

that both the variance of his portfolio and its skewness have the same sensitivity for security i . Since his risk premium (or discount) is proportional to this sensitivity, the conventional Sharpe-Linter systematic risk measure is good enough to determine the risk premium on this security.

The results of this section are in contrast to those of Kraus and Litzenberger (1976) and (1983). They argued, based on a truncated Taylor expansion of the utility function, that investors exhibit preference for skewness and thus they concluded that the conventional Sharpe-Linter CAPM overestimates the risk premium of securities. In a previous section it is shown that their conclusion regarding the skewness preference is not correct. The contribution of skewness to the expected utility is ambiguous and depends on both the utility function and the joint distribution of assets returns. According to our analysis, the market price of skewness has an ambiguous sign. It is an empirical question depending on whether the zero beta portfolio provides a positive or negative risk premium. Moreover, the direction and the magnitude of the divergence of the conventional Sharpe-Linter CAPM, for an individual security, depends on the sensitivity of the variance and the skewness of the market portfolio to that individual security.

5.3.2 The case with no riskless asset:

In this case the equilibrium of an individual investor is given by equation (3.11.b)

$$VX - \lambda_1 \mu - \lambda_2 e - (\lambda_3 + \sigma_Y^2 b' X) b = 0 \quad (3.11.b)$$

Following the notation of the previous section, let

$r_m = X'Z$ be the return on any efficient portfolio, γ_3 , S_i , σ_{im} are the same as they were defined in the previous subsection. The equilibrium condition (3.12) can be written as:

$$\sigma_{im} - \lambda_1 E(z_i) - \lambda_2 - \gamma_3 S_i = 0 \quad (5.20)$$

Equation (5.20), as in the previous subsection, holds for individual securities as well as portfolios. Denote by n , and p any two efficient portfolios with returns r_n , and r_p . Equation (5.20) must hold true for portfolios m , n , and p . This leads to the following system of equations:

$$\begin{aligned} \sigma_m^2 - \lambda_1 E(r_m) - \lambda_2 - \gamma_3 S_m &= 0 \\ \sigma_{nm}^2 - \lambda_1 E(r_n) - \lambda_2 - \gamma_3 S_n &= 0 \\ \sigma_{pm}^2 - \lambda_1 E(r_p) - \lambda_2 - \gamma_3 S_p &= 0 \end{aligned} \quad (5.21)$$

If we solve (5.21) for λ_1 , λ_2 , and γ_3 and substitute their values in (5.20) we have the following:

$$\begin{aligned}
E(z_i - r_n) &= \frac{(S_i - S_n)(\beta_p - \beta_n) - (\beta_i - \beta_n)(S_p - S_n)}{(S_n - S_m)(\beta_p - \beta_n) - (S_p - S_n)(1 - \beta_n)} \\
&\quad E(r_m - r_n) \\
&+ \frac{(S_m - S_n)(\beta_i - \beta_n) - (\beta_i - \beta_n)(1 - \beta_n)}{(S_n - S_m)(\beta_p - \beta_n) - (S_p - S_n)(1 - \beta_n)} \\
&\quad E(r_p - r_n) \tag{5.22}
\end{aligned}$$

For (5.22) to describe a general equilibrium in the capital market, the portfolio m , n , and p must be efficient for all investors. Since the market portfolio must be efficient, it can be chosen as one of these funds. Let r_m denote the return on this portfolio. The other two portfolios can be chosen, relative to this portfolio, in an arbitrary way. Since the riskfree asset, if it exists, is neutral to the risk of the market portfolio, we may choose portfolio n with such a property. Namely, we may choose portfolio n such that $\beta_n = S_n = 0$. As is the case of a riskfree asset, this portfolio has no contribution either to the variance or the skewness of the market portfolio. If the risk free asset and portfolio n exist in the capital market, both have to earn the same rate: the riskfree rate. In this case there is one degree of freedom to choose a zero beta portfolio. We may choose a portfolio p as another zero beta portfolio with non-zero skewness.¹⁰ With this choice we

¹⁰To see how such a choice is feasible, let X_p , X_n be any two arbitrary portfolios. Any efficient portfolio

have the following capital assets pricing model:

$$E(z_i - r_n) = s_i E(r_m - r_n) + \gamma_i - s_i E\left(\frac{r_p - r_n}{\gamma_p}\right) \quad (5.23)$$

Equation (5.23) is similar to (5.20). The only difference is in the fact that r_n replaces the riskfree rate. When markets deny a riskfree security, they allow the construction of a security with zero relative risk, relative to the market portfolio. Such a security plays the role of the riskfree asset in pricing other securities in the capital market. In fact, portfolio r_n resembles the zero beta portfolio of the extended version of the conventional Sharpe-Linter CAPM provided by Black (1973).

must satisfy $X = \lambda_1 V^{-1} \mu + \lambda_2 V^{-1} e + \lambda_3 V^{-1} b$ with $X'e = 1$. Our choice is equivalent to the following:

$$\sigma_{nm} = \sigma_{pm} = X_n' V X_m = X_p' U X_n = \lambda_1 E(r_m) + \lambda_2 + (\lambda_3 / E^{1/3}(Y^3)) S_m = 0 \quad (i)$$

$$S_n = (X_n' b / E^{1/3}(Y^3)) = \lambda_1 b' V^{-1} \mu + \lambda_2 b' V^{-1} e + \lambda_3 b' V^{-1} b = 0 \quad (ii)$$

$$e' X_n = e' X_p = \lambda_1 e' V^{-1} \mu + \lambda_2 e' V^{-1} e + \lambda_3 e' V^{-1} b = 1 \quad (iii)$$

$\lambda_1, \lambda_2, \lambda_3$ that solves the system of three equations determines portfolio n . On the one hand, any portfolio p must satisfy only equations (i) and (iii) but not equation (ii). Therefore, either one of the i can be chosen arbitrarily. This shows the possibility of finding two efficient portfolios corresponding to any efficient portfolio. One is neutral to its total risk and the other is neutral to its variance risk.

APPENDIX 5.A

² The first order condition for an expected utility is given by:

$$EU'(X'Z)[z_i - R_F] = 0 \quad \text{for } i = 1, 2, \dots, n \quad (i)$$

Let $R_e = Z'X_e$ be the return on any market efficient portfolio (portfolio that satisfies the first order condition for all investors). Equation (i) can be written as:

$$EU'(R_e)[Z_i - \mu_1 + \mu_1 - R_F] = \text{COV}(U'(R_e), Z_i) + (\mu_1 - R_F)EU'(R_e) = 0$$

or

$$\mu_i - R_F = - \frac{\text{COV}(U'(R_e), Z_i)}{EU'(R_e)} \quad (ii)$$

But (i) holds for individual securities and for portfolios. Therefore, equation (ii) and its analog for the market efficient portfolio imply:

$$\frac{\mu_i - R_F}{E(R_e) - R_F} = \frac{\text{COV}(U'(R_e), Z_i)}{\text{COV}(U'(R_e), R_e)} = \Delta_j \quad (iii)$$

If we let $R_e = R_C$ and utilize equation (1):

Equation (5.3) is an analog form of equation (15) and (16) of Klaus and Litzenberger (1983) with only one difference. The difference is that R_e is replaced by R_c , the return on the consumption portfolio whose returns spans the aggregate consumption. The Pareto Optimality condition in Kraus and Litzenberger's framework implies that only dated claims on the aggregate consumption are required to determine the efficient allocation of consumption across individuals. For a given aggregate consumption, C , Pareto optimal allocation is achieved by maximizing the expected utility of an individual given the expected utilities of others. It is the same as $\max_K \sum_K A_K E U_K$ subject to the constraint that the sum of the individual consumptions at time t is equal to the aggregate consumption at that time, $\sum_K C_{Kt} = C_t$. This implies that for a given aggregate consumption:

$$A_K u'_K(R_K) = \lambda(C) \quad (1v)$$

Substituting this in (i) implies that $\lambda(C)$ satisfies (11i) and therefore $\lambda(R_C)$ also satisfies (11i). That is:

$$\frac{u_1 - R_F}{E(R_C) - R_F} = \frac{\text{COV}(\lambda(R_C), Z_1)}{\text{COV}(\lambda(R_C), R_C)} = A_J$$

Utilizing equation (5.1) and the fair game property:

$$E[\varepsilon_1/R_C] = 0 \text{ implies } \text{COV}(\varepsilon_1, \lambda(R_C)) = 0,$$

we have:

$$\Delta_1 = \beta_1 + \frac{\text{COV}(\epsilon_1, \lambda(R_C))}{\text{COV}(R_C, \lambda(R_C))} = \beta_1$$

Observe that this result is true regardless of the specific definition of β_1 in the definition of the linear characteristic lines, as $\text{COV}(R_C, Z_1)/\text{Var}(R_C)$.

CONCLUSIONS

This dissertation provides a three parameter framework for portfolio analysis which serves two main objectives. First, it provides an analytical setting to challenge the mean-variance portfolio selection and evaluate its practical optimality. For the first time, mean-variance portfolio selection is challenged and evaluated in a framework that implies its theoretical inconsistency with expected utility maximization. Second, for contingencies under which mean-variance framework is not valid (or practically valid), this framework provides an alternative normative theory for portfolio selection that preserves a great deal of the mathematical tractability of mean-variance portfolio selection. It also provides a three moment capital asset pricing model alternative to the mean-variance Sharpe-Linter capital asset pricing model for the same contingencies. While there has been no attempt to propose an extension for the normative theory of portfolio selection, Kraus and Litzenberger (1983) provided sufficient conditions on security returns for a three moment capital asset pricing model. It is proved that these conditions, in fact, imply the mean-variance Sharpe-Linter capital asset pricing model which they intended to replace.

Mean-variance investment strategies are evaluated for constant absolute risk averse investors. An optimization premium, which reflects the maximum fee that an investor is willing to pay in order to replace his mean-variance portfolio with the optimal one, is derived as a function of the measure of absolute risk aversion and the assets distribution parameters. Utilizing monthly returns on ten securities this optimization premium is also operationalized. When a riskless asset exists, such an optimization premium does not exceed 0.05% (0.05 penny on an invested dollar) regardless of the degree of risk aversion. But when a riskless asset is not available, mean-variance portfolio selection exhibits a fair performance only up to a certain level of absolute risk aversion. When the degree of absolute risk aversion is very high the opportunity cost of mean-variance strategies escalates in a linear fashion with the absolute risk aversion. This is because a constant absolute risk averse investor tends to choose what he perceives the safest strategy, regardless of the fortune of its alternatives, as his risk aversion increases. While the safest strategy--the riskless asset--is mean-variance efficient in the first case, it is not mean-variance efficient in the latter. This, in fact, explains the poor performance of mean-variance strategies for investors with "radical" risk aversion in the latter case. For such radical investors, no investment

prospect can compete with the safest strategy. No matter what fortune a competing strategy may strike, it will be avoided by such an investor even when its additional risk is very small. Such a behavior is uncommon, indeed, among investors.

The three parameter framework of portfolio selection, proposed in this dissertation, is consistent with a two dimensional measure for risk. In our framework, the risk measure in the mean-variance framework--the variance of the portfolio return--is broken down into two components: a spherical variance and a non-spherical variance. Only the spherical variance is disliked by all risk averse investors. In such a framework, the portfolio problem translates to a quadratic programming problem. The investor decides on his two dimensional measure for risk, and maximizes his expected utility. In such a framework, it is feasible to derive the efficient set of portfolios for all risk averse investors as a function of the joint distribution parameters. Moreover, the efficient set is estimatable from observable assets returns. While the mean-variance efficient set provides an efficient portfolio for any given mean of portfolio return, our efficient set provides an efficient portfolio for any given pair of mean and skewness parameter. The mean-variance efficient set is a subset of the derived three dimensional efficient set.

The three parameter portfolio selection provides a derivation of a capital asset pricing model in a perfect market. The riskfree asset, the market portfolio, and a portfolio neutral to its variance determine the set of all market efficient portfolios. A two dimensional measure of systematic risk, the risk premium on the market portfolio and the risk premium (or discount) on the third portfolio determine the risk return trade-off in the capital market. Since the risk of the market portfolio is completely determined by its variance and skewness pair, and since the third portfolio contributes nothing to the variance of the market portfolio, the third portfolio offers risk premium when skewness is disliked by the market investor, but, on the other hand, it must offer a discount if the market investor exhibits preference for skewness. While many authors argued that skewness is preferred by any risk averse investor, it is shown, by a counter example, that the attitude of risk averse investors toward skewness is ambiguous. The two dimensional measure of the systematic risk, for a given security, is its contribution for the variance and its contribution for the skewness of the market portfolio. Whether the Sharpe-Linter CAPM underestimates or overestimates the expected return of an individual security depends on two factors: (i) The size of its contribution to the variance of the market portfolio relative to its

contribution to the skewness of the same portfolio; (ii) the attitude of the market investor toward skewness. The Sharpe-Linter CAPM prices, a given security, correctly only under two conditions: (i) When the relative contribution of such a security to variance and the skewness of the market portfolio is identical, or (ii) when the market investor neither places premium or discount on skewness. It is also proved that when the riskless asset does not exist, capital markets allow the construction of an asset neutral to both the variance and the skewness of the market portfolio. Such a portfolio plays the role of the riskfree asset in pricing other securities.

REFERENCES

1. Aivazian, V.A., Callen, J.L., Krinsky, I., & Kwan, C.Y. (1983) "Mean Variance Utility Functions and the Demand for Risky Assets: An Empirical Analysis Using Flexible Functional Forms." Journal of Financial and Quantitative Analysis, December.
2. Arrow, K. (1965) Aspects of the Theory of Risk Bearing. Helsinki.
3. Borch, K. (1969) "A Note on Uncertainty and Indifference Curves." Review of Economic Studies, January.
4. _____. (1974) "The Rationale of the Mean-Standard Deviation Analysis: Comment." American Economic Review, January.
5. Chamberlain, G. (1983) "A Characterization of the Distributions that Imply Mean Variance Utility Functions." Journal of Economic Theory, June.
6. Dempster, A.D. (1969) Elements of Continuous Multivariate Analysis. Addison-Wesley.
7. Elderton, W.P. & Johnson, N.L. (1969) Systems of Frequency Curves. Griffin.
8. Feldstein, M. (1969) "Liquidity Preference and Portfolio Selection." Review of Economic Studies, January.
9. Fogiel, M. (1983) Handbook of Mathematical Formulas. Tables. Functions. Graphs. Transformations. Research and Development Association.
10. Franklin, P. (1944) Methods of Advanced Calculus. McGraw-Hill Company.
11. Genedenko, B.V. (1962) The Theory of Probability. Chelsea Publishing Company.
12. Johnson, N.L. (1969) Distributions in Statistics, Vols. 1,2,3. Houghton Mifflin Company.

13. Kelker, D. (1970) "Distribution Theory of Spherical Distributions and a Location Scale Parameter Generalization." Sankhya Ser. A32.
14. Kendall, M. & Stuart, A. (1977) The Advanced Theory of Statistics, Vol. 1. Macmillan Publishing Company.
15. Kalberg, J. & Zemba, W. (1979) "On the Robustness of the Arrow-Pratt Risk Measure." Economic Letters, Vol. 2, pp. 21-26.
16. Kraus, A. & Litzenberger, R. (1976) "Skewness Preference and the Valuation of Risky Assets." The Journal of Finance, September.
17. _____. (1983) "On the Distributional Conditions for a Consumption-Oriented Three Moments CAPM." The Journal of Finance, December.
18. Kroll, V., Levy, H. & Markowitz, H. (1984) "Mean-Variance Versus Direct Utility Maximization." Journal of Finance, March.
19. _____. (1984) "Clarification of the Second Criterion in Mean-Variance Versus Direct Utility Maximization." Journal of Finance, September.
20. Krylov, V. (1962) Approximate Calculation of Integrals. Macmillan Publishing Company.
21. Levy, H. & Hanoch, G. (1969) "Efficiency Analysis of Choices Involving Risk." The Review of Economic Studies 36, pp. 335-346.
22. Levy, H. (1974) "The Rationale of the Mean Variance Standard Deviation: Comment." American Economic Review, June.
23. Levy, H. & Markowitz, H. (1979) "Approximating Expected Utility by a Function of the Mean and the Variance." American Economic Review, June.
24. Markowitz, H. (1959) Portfolio Selections: Efficient Diversification of Investments. New York, New Haven.
25. _____. (1952) "Portfolio Selection." Journal of Finance, March.

26. Ohlson, J. (1975) "The Asymptotic Validity of Quadratic Utility as the Trading Interval Approaches Zero." In Stochastic Optimization Models in Finance, Zemba, W.T. and Vickson, R. (Eds.). Academic Press.
27. Owen, J. & Rabinowitch, R. (1983) "On the Class of Elliptical Distributions and their Applications to the Theory of Portfolio Selection." Journal of Finance, June.
28. Ord, J.K. (1972) Families of Frequency Curves. Griffin.
29. Pratt, W. (1964) "Risk Aversion in the Small and in the Large." Econometrica, January.
30. Pully, L. (1981) "A General Mean Variance Approximation to Expected Utility for Short Holding Periods." Journal of Financial and Quantitative Analysis, Vol. 16.
31. _____. (1984) "Comment on: Mean Variance Versus Direct Utility Maximization." Journal of Finance, September.
32. Pyle, D. & Turnovsky (1970) "Safety First and Expected Utility Maximization in Mean-Standard Deviation Portfolio Analysis." The Review of Economics and Statistics, 52, pp. 75-81.
33. Ross, S. (1976) "The Arbitrage Theory of Capital Asset Pricing." Journal of Economic Theory, Vol. 13, pp. 341-360.
34. _____. (1978) "Mutual Fund Separation in Financial Theory--The Separating Distributions." Journal of Economic Theory, Vol. 17, 254-286.
35. Roy, A. (1952) "Safety First and the Holding of Assets." Econometrica, July.
36. Samuelson, P. (1970) "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances, and Higher Moments." Review of Economic Studies, 37.
37. Scott, R. & Horvath, P. (1980) "On the Direction of Preference for Moments of Higher Order than the Variance." Journal of Finance, September.

38. Stiglitz, E. & Cass, D. (1970) "The Structure of Investor Preference and Asset Returns and Separability in Portfolio Allocations: A Contribution to the Pure Theory of Mutual Funds." Journal of Economic Theory 2, pp. 122-160.
39. Straussen, V. "The Existence of Probability Measures with Given Marginals." Annals of Mathematical Statistics, 36, pp. 423-439.
40. Tobin, J. (1958) "Liquidity Preference as Behavior Toward Risk." Review of Economic Studies, 25.
41. Tsiang, C. (1972) "The Rationale of the Mean Variance Analysis, Skewness Preference, and the Demand for Money." American Economic Review, June.
42. . (1974) "The Rationale of the Mean Variance Standard Deviation Analysis: Reply and Errata for Original Article." American Economic Review, June.
43. Vickson, R. (1975) "Separation in Portfolio Theory." In Stochastic Optimization Models in Finance, Zemba, W. & Vickson, R. (Eds.). Academic Press.
44. Zemba, W.T. & Vickson, R.G. (Eds.). (1975) Stochastic Optimization Models in Finance. Academic Press.