Chapter Summary

Chapter 3 *Theoretical Developments in Modern Finance*

# Indicates some sub-sections contain advanced material
* Indicates complete section is advanced material

Chapter 3 *Theoretical Developments in Modern Finance*

3.1 Basics of Mean-Variance Portfolio Analysis

A. The Analytical Preliminaries

Many of the essential analytical tools used in mean-variance portfolio analysis can be found in the results for linear combinations of random variables from mathematical statistics. One basic result is the following, e.g., Freund (1971, p.195):

Theorem: ***Moments of Linear Combinations of Random Variables***

If $X(1), X(2), \ldots , X(N)$ are random variables and $a_1, a_2, \ldots , a_N$ are constants and $Y = a_1 X(1) + a_2 X(2) + \ldots + a_N X(N)$ then:

$$E[Y] = \sum_{i=1}^{N} a_i \, E[X(i)]$$

$$var[Y] = \sum_{i=1}^{N} a_i^2 \, var[X(i)] + 2 \sum\sum_{i>j} a_i \, a_j \, cov[X(i), X(j)]$$

where the double sum over $i > j$ extends over all values of i and j, from 1 to N, for which $i > j$.

Derivation of $var[Y]$ requires the observation that $cov[X(i), X(j)] = cov[X(j), X(i)]$. One immediate corollary is that if $cov[X(i), X(j)] = 0$ for all i and j where $i \neq j$, i.e., the random variables are all independent, and $a_1 = a_2 = \ldots = a_N = 1/N$ then $var[Y]$ has the property:

$$\lim_{N \to \infty} var[Y] = \lim_{N \to \infty} \left\{ \sum_{i=1}^{N} a_i^2 \, var[X(i)] \right\} = 0$$

This result has applications in insurance where the random variables are policy payouts and the a_i are the fraction of the portfolio of policies attributable to policy i.

Extending these results to portfolios follows immediately from identifying the random variables as the returns on individual securities held in a given portfolio, i.e., let $X(i) = R_i$ for all i. The definition of the portfolio expected return follows:

Definition: The expected return on the portfolio $E[R_p]$ is the value weighted sum of the expected returns on the individual securities, the $E[R_i]$:

$$E[R_p] = \sum_{i=1}^{k} w_i \; E[R_i]$$

where k is the number of securities in the portfolio. To calculate the value weights, w_i:

$$w_i = \frac{\$A_i}{\sum_{i=1}^{k} A_i} \qquad \textit{where} \qquad \sum_{i=1}^{k} w_i = 1$$

with $\$A_i$ being the dollar value invested in security i and the sum over all $\$A_i$ being the total amount of money invested in the portfolio

---

As a simple example, consider having \$1 million invested in a portfolio of 2 securities, and there is \$500,000 in each security, then each $w_i = .5$. As a slightly more complicated example consider the following problem: At the beginning of the year, Joe Investor owned four securities in the following amounts: A, 100 shares; B, 400 shares; C, 200 shares; D, 200 shares. The current prices of the securities are: A = \$12.50; B = \$17.50; C = \$25; and, D = \$50. In one year's time, Joe expects the prices to be: A = \$25; B = \$20; C = \$30; and D = \$55. What is the expected return on Joe's portfolio for the year? The solution to this problem is determined by calculating the total value invested as: 100 (12.50) + 400 (\$17.50) + 200 (\$25) + 200 (\$50) = \$23,250. This permits the calculation of the value weights: w_A = 1250/23250 = .054; w_B = .301; w_C = .215; w_D = .430. The expected return on the portfolio can now be calculated as: $E[R_p]$ = .054($E[R_A]$) + .301($E[R_B]$) + .215($E[R_C]$) + .430($E[R_D]$) = .054 (1.00) + .301 (.143) + .215 (.2) + .430 (.1) = .183 (18.3%).

The other key element in the mean-variance portfolio model is the standard deviation of portfolio returns. As with calculating the risk for individual securities, calculations are done for the variance and the standard deviation is determined by taking a square root. The standard deviation, as opposed to the variance, is of the appropriate measure of risk because it is in the same units as the expected returns. However, calculations are done using the variance.

Definition: The standard deviation of portfolio returns, σ_p is the square root of the variance of portfolio returns $var[R_p] \equiv \sigma_p^2$. Various ***equivalent*** forms for the portfolio variance formula are available:

$$var[R_p] = \sigma_p^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} w_i\, w_j\, \sigma_{ij} = \sum_{i=1}^{k} w_i^2\, \sigma_i^2 + 2 \sum_{i>j}^{k} w_i\, w_j\, \sigma_{ij}$$

$$= \sum_{i=1}^{k} w_i^2\, \sigma_i^2 + 2 \sum_{j=1}^{k} \sum_{i=1,\ i>j}^{k} w_i\, w_j\, \sigma_{ij} = \sum_{i=1}^{k} w_i^2\, \sigma_i^2 + 2 \sum\sum_{i>j} w_i\, w_j\, \sigma_{ij}$$

where $cov[R_i\, , R_j] = \sigma_{ij}$. In the double sum expression, when $i=j$ the covariance is a variance. These expressions can be further manipulated by making further substitutions using the definition for ρ_{ij}, the correlation between R_i and R_j, i.e., $cov[R_i\, , R_j] = \sigma_{ij} \equiv \rho_{ij}\, \sigma_i\, \sigma_j$.

It is easiest to understand these results for the case where $k = 2$, when there are only two securities in the portfolio. In this case:

$$\sigma_p^2 = w_1^2\, \sigma_1^2 + w_2^2\, \sigma_2^2 + 2\, w_1\, w_2\, \sigma_{12}$$

$$= w_1^2\, \sigma_1^2 + (1 - w_1)^2\, \sigma_2^2 + 2\, w_1\, (1 - w_1)\, \rho_{12}\, \sigma_1\, \sigma_2$$

An end of chapter exercise is provided to illustrate the manual application of this result. Similarly for 3 assets in the portfolio:

$$\sigma_p^2 = w_1^2\, \sigma_1^2 + w_2^2\, \sigma_2^2 + w_3^2\, \sigma_3^2$$

$$+ 2\, \{w_1\, w_2\, \sigma_{12} + w_1\, w_3\, \sigma_{13} + w_2\, w_3\, \sigma_{23}\}$$

When there are k securities in the portfolio, the resulting portfolio variance will contain k variance terms and $\{k(k - 1)\}/2$ covariance terms. The substitutions using the definition for the correlation coefficient and the restriction that the sum of the value weights equals one is left as an exercise.

Having the formula for the variance of portfolio return permits the ready identification of an important special case of an optimum portfolio: the minimum variance portfolio, the portfolio that has the smallest risk in the set of all possible portfolios. This formula for this portfolio can be derived by minimizing $var[R_p]$ with respect to the choice variables, the value weights for each of the individual securities, subject to the restriction that the sum of the value weights be equal to one. In the simple case of the minimum variance portfolio for two securities, using the result that $w_1 + w_2 = 1$:

$$\sigma_p^2 = w_1^2\, \sigma_1^2 + (1 - w_1)^2\, \sigma_2^2 + 2\, w_1\, (1 - w_1)\, \sigma_{12}$$

$$\frac{d\sigma_p^2}{dw_1} = 2\, w_1\, \sigma_1^2 - 2\,(1 - w_1)\, \sigma_2^2 + 2\,(1 - 2w_1)\, \sigma_{12} = 0$$

$$w_1^* = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\, \sigma_{12}}$$

This result demonstrates that the minimum variance portfolio will be a combination of the two securities and not just be fully invested in the security with the lowest risk.

The intuition behind the portfolio diversification problem can be illustrated with the following artificial situation: assume that all the securities in the market have the same expected return of 10% and the same standard deviation of security return of 15% with the covariance between all security returns being .02. Construct an equally weighted portfolio containing N securities. While the expected return on this portfolio will be 10%, the variance of the equally weighted portfolio containing N securities will be:

$$\sigma_p^2 = \sum_{i=1}^{N} \frac{\sigma_i^2}{N^2} + 2 \sum_{i > j}^{N} \frac{\sigma_{i,j}}{N^2} < .15$$

While the expected return is a linear combination of the individual security expected returns, the result same does not apply to the variance. This property of the variance for a linear combination of random variables is an essential ingredient in the portfolio optimization models.

Recall the result stated previously where, if the random variables are uncorrelated, then the variance of an equally weighted linear combination will go to zero as N goes to infinity. What happens to the portfolio expected return and standard deviation of this portfolio as N gets large? Because it is assumed that all standard deviations are the same, there are N equal terms in the first sum and, because the covariances have been assumed to be equal, there are $N(N - 1)$ terms in the second sum and the portfolio variance reduces to:

$$\sigma_p^2 = \frac{\sigma_i^2}{N} + N(N-1)\, \frac{\sigma_{ij}}{N^2} = \frac{\sigma_i^2}{N} + \left(1 - \frac{1}{N}\right) \sigma_{ij}$$

As $N \rightarrow \infty$, the first term goes to zero and the portfolio variance is reduced to the covariance. To get this result, N must be very large. Even for portfolios containing, say, 100 securities, there is still some contribution to variance from the σ_i. As noted, when the covariance between all available securities is zero (independent returns), the standard deviation of the portfolio will go to zero as N gets large.

This simple example provides a pedagogical basis for illustrating the gains to diversification. Examining the variance of the equally weighted portfolio described above, the (σ_i^2 / N) term applies to the specific risk associated with the individual securities, where the σ_i^2 for each of individual ecurity are associated with individual firm specific risks. Because this component of portfolio variance goes to zero as the number of securities gets large, it is appropriately described as ***diversifiable risk***. The *(1 - (1/N))* σ_{ij} term is associated with the covariance between security returns. Because this source of portfolio risk does not go to zero as N goes to infinity, it is appropriately described as ***nondiversifiable risk***. It follows that the portfolio standard deviation can be decomposed into the sum of diversifiable risk and nondiversifiable risk. Hence, the risk associated with any portfolio of securities equals the sum of the diversifiable risk and nondiversifiable risk for that specific portfolio. 'Efficiently diversified' portfolios have eliminated diversifiable risk.

As securities are added to the equally weighted portfolio, the risk of the portfolio is reduced until the lower bound provided by non-diversifiable risk is reached. However, the amount of risk reduction decreases as the number of the securities in the portfolio increases to the point where there is no more firm specific risk that can be eliminated. The lower bound on portfolio risk, associated with the non-diversifiable risk, is due to the covariance between security returns. Modern portfolio theory expends considerable theoretical effort in developing the capital asset pricing model (CAPM) where individual security returns depend on a combination of the riskless interest rate and the covariance of the individual security return with the return on the market portfolio. In this model, it is the nondiversifiable risk, referred to as systematic risk, that is compensated with higher expected return. Hence, ***there is a tradeoff between systematic risk and expected return***. Because firm specific risk (unsystematic risk) can be eliminated in an efficiently diversified portfolio, the security market will not reward this source of risk with higher expected return.

B. The Optimization Model

The mean-variance portfolio optimization model is a central paradigm of modern Finance. The essence of the model is captured in the following quadratic optimization problem, e.g., Elton and Gruber (1995), Luenberger (1998):[1]

$$\min_{\{w_i\}} \quad var[R_p] = \sum_{i=1}^{k} \sum_{j=1}^{k} w_i \, w_j \, \sigma_{i,j}$$

$$subject\ to: \qquad E[R_p] = \sum_{i=1}^{k} w_i \, E[R_i] = \overline{c_n}$$

$$for \quad \overline{c_n} \in \{c_0,\, c_1,\, c_2,\, \ldots\} \quad where: \; c_0 = c_{mv} \; and: \; \sum_{i=1}^{k} w_i = 1$$

where: k is the number of risky securities or assets available for investment; $E[R_i]$ is the (conditional) expected return on security or asset i; $E[R_p]$ is the expected return on the portfolio; c_{mv} is the return on the minimum variance portfolio; and, $var[R_p]$ is the variance of portfolio return. In this model, the $\{w_i\}$ are the value weights, the fraction of the total value of the portfolio invested in each asset. Though it is conventional to develop the model under the assumptions of perfect capital markets (see

BOX), it is possible, even desirable, to impose additional restrictions on the optimization problem. One such additional restriction is $w_i \geq 0$ for all i. This restriction prevents short selling of securities. Without this restriction, all or almost all securities will be held in some form, either long or short. With the short selling restriction, the resulting optimal portfolios will have many securities that have value weights equal to zero.

INSERT Figure 3-aa What are perfect capital markets?

The quadratic optimization problem is to determine the value weights for each security which minimize the variance of the return on the portfolio, subject to a target level of expected return. Because there is range of possible expected returns that can be chosen, ***the solution to the optimization problem will be a set of portfolios***, each with its own set of optimal weights. This set of optimal portfolios is typically referred to as the ***efficient frontier***. Other terms such as efficient set (Fama 1976), portfolio possibilities curve (Elton and Gruber 1995) and mean-variance efficient locus (Ingersoll 1987) are also used. There are a number of solution methodologies that can be used to solve quadratic optimization problems. A simple iterative method involves initially solving the minimum variance problem. The resulting optimal minimum variance weights are used to identify c_{mv}. This value is used to specify $c_1 = c_{mv} + \epsilon$, where ϵ (> 0) is specified according to the desired precision required in the solutions. Using c_1 it is now possible to solve the Lagrangian problem for the next portfolio along the frontier. This process continues for $c_2 = c_{mv} + 2\epsilon$, $c_3 = c_{mv} + 3\epsilon$ and so on until the desired efficient frontier is determined. With the short sales constraint, the maximum c_i is given by having all funds invested in the highest returning security. Without the short sales constraint, the efficient frontier can be extended indefinitely. Because the underlying problem is quadratic, there will be two 'optimal' solutions, one of which is ignored because it will have higher risk for the same expected return. Hence, the efficient frontier only contains the portfolios with high return/lowest risk.

The number of variations that have emerged from this basic model is staggering.[2] Initially, implementation of the model was impeded by the large number of parameters required to make the model operational. In addition to the k individual asset returns, $E[R_i]$, there are k variances, σ_i^2, and $\{k(k-1)\}/2$ covariances which have to be estimated from past data. Even if these parameters are available, the model is only capable of generating a set of mean-variance optimal portfolios, the efficient frontier. Additional structure is needed to select a specific portfolio from the set of optimal portfolios. Tobin (1958), Sharpe (1963, 1964) and others handled this problem by introducing a riskless asset. This permits the investor to form portfolios which combine the riskless asset with an efficient frontier portfolio. In this fashion, the investor is able to achieve the same level of expected return as that generated by an efficient frontier portfolio, again with a lower level of risk. Effectively, the addition of a riskless asset transforms the investment opportunity set from a convex function, the efficient frontier, to a set of linear functions, the capital allocation lines.

In general, where there are many possible securities available for inclusion in the portfolio, solution of the efficient set from the optimization problem is complicated. For purposes of illustration, it convenient to assume that there is only two risky securities. In this case it is possible to derive the efficient frontier directly, permitting basic concepts to be illustrated. So, assume you are considering creating a portfolio combining a stock fund composed of large stocks (S) and a bond fund (B). The statistics for these funds are: *E[R]*, large stock fund = 12%, bond fund = 5%; σ, large stock fund = 15%, bond fund = 8%. For ease of calculation, assume the correlation coefficient between the funds is zero. It is now possible to calculate $E[R_p]$ and σ_p for all possible portfolios, starting from 0% invested in the stock fund ($w_s = 0$) and going to 100% ($w_s = 1$), in increments of 20%. This produces:

w_s	$E[R_p]$	σ_p
0	5.0%	8.0%
.2	6.4%	7.07%
.4	7.8%	7.68%
.6	9.2%	9.55%
.8	10.6%	12.11%
1.0	12%	15%

Plotting these values in *{E[R],σ}* space produces the efficient frontier. From these values, it is apparent that an investment of 100% in the bond fund ($w_s = 0$) does not make sense because portfolios with values such as $w_s = .2$ and $w_s = .4$ both provide a higher portfolio expected return with a lower level of portfolio risk. Recalling that the returns were assumed to be independent, the portfolio weights, $E[R_{mv}]$ and σ_{mv} for the minimum-variance portfolio are:

$$w_s = \frac{\sigma_b^2}{\sigma_s^2 + \sigma_b^2} = \frac{.08^2}{.15^2 + .08^2} = .2215 \quad \rightarrow \quad w_b = .7785$$

It follows that $E[R_{mv}] = .0655$ and $\sigma_{mv} = .0706$.

Suppose the number of securities available for selection in the portfolio included a third fund composed of small stocks which has $E[R] = 15\%$ and $\sigma = .20$. How would you derive the efficient frontier for portfolios combining the three funds? Attempting to use the method of direct calculation that worked for the two security case is no longer possible. The two security case reduced to solving for one weight because it was possible to substitute out the other weight using the constraint that the sum of the weights equals one. Hence, there was no optimization problem to solve as there was only one effective weight. Barring trivial cases, when there are three or more securities the optimal weights have to be determined by solving the first order conditions of the optimization problem in order to identify how much of a given frontier portfolio is invested in each security. Ingersoll (1987) discusses the relevant solution procedure.

C. Capital Allocation Lines: Introducing a Riskfree Asset

The efficient frontier specifies a set of portfolios that achieve optimal combinations of the available ***risky*** assets. In order to provide a practical guide to portfolio selection, some method is required to identify a specific portfolio from the efficient frontier. Observing that the efficient frontier is a convex relationship between *E[R]* and σ for portfolios of risky assets, by introducing a ***riskfree*** asset it is possible to produce a linear set of available portfolios. In particular, when a riskfree asset is introduced a range of portfolios can be identified which are not available with risky assets alone. What is the riskfree asset? This depends on the investment horizon or the portfolio rebalancing period. The riskfree asset must be free of default risk, have no coupons to reinvest (this would create coupon reinvestment risk), have maturity equal to the investment horizon and be denominated in domestic currency. For US investors, it is typical to use the 3 month Tbill rate as the riskfree interest rate. Though it is conventional to proxy the riskfree asset with a 3 month US Treasury bill, some presentations, e.g., Damodaran (1994), argue that the US Treasury long term bond yield is appropriate. The problem of specifying the riskfree asset will be explored in more detail shortly.

INSERT FIGURE 3-a CAL with efficient frontier

The riskfree asset permits the creation of portfolios which combine the riskfree asset with a risky portfolio located on the efficient frontier (see Figure 3-a). In ($E[R]$, σ) space, the line connecting the riskfree rate with a point on the efficient frontier is referred to as a ***capital allocation line*** (CAL). (This terminology is used in Bodie, Kane and Marcus 1999). There are as many capital allocation lines as there are portfolios on the efficient frontier. Each point on a given capital allocation line defines a range of possible portfolios which combine the riskfree asset and an efficient portfolio composed of risky assets. Along a given capital allocation line connecting the riskfree rate, r, with an efficient portfolio X, the expected return ($E[R_R]$) and standard deviation (σ_R) of a portfolio combining the efficient frontier portfolio and the riskfree asset can be specified:

$$E[R_R] = w_r\, r + (1 - w_r)\, E[R_x]$$

$$\sigma_R^2 = w_r^2\, \sigma_r^2 + (1 - w_r)^2\, \sigma_x^2 + 2\, w_r\, (1 - w_r)\, \sigma_{rx}$$

$$= (1 - w_r)^2\, \sigma_x^2 \qquad \rightarrow \qquad \sigma_R = (1 - w_r)\, \sigma_x$$

The result for the variance of the portfolio follows because the riskfree asset has no risk or covariance because it is risk free.

Some care has to be taken in interpreting the weight w_r. Though the notation w is used, this weight is not associated with the w_i weights for the various efficient frontier portfolios. The w_r weight is the fraction of the portfolio in the riskfree asset and ($1 - w_r$) is the fraction held in the efficient portfolio. When $1 \geq w_r \geq 0$, this implies that the portfolio involves a positive investment in both the riskfree asset and the efficient portfolio. In Figure 3-a this condition is applicable to all the portfolios lying on the portion of the CAL between r and X. At r, $w_r = 1$ and at X on the efficient frontier $w_r = 0$. When $w_r \leq 0$, this implies that the riskfree asset is held short, i.e., the investor is borrowing at the riskless rate. This is equivalent to saying that, in addition to investment of the original capital, the investor has also borrowed money at the riskfree rate and has purchased additional units of the risky portfolio X. For example, if the investor has $1 million of original capital to invest and $w_r = -.5$ ($w_X = 1.5$), then the investor has borrowed an additional $500,000 and has used this money to purchase an additional $500,000 of X. Portfolios where $w_r \leq 0$ lie to the right of X on the CAL. The key point to recognize is that the presence of the riskfree asset permits the investor to attain portfolios which are not available using risky assets alone.

D. The Capital Market Line and Market Equilibrium

To this point, the problem of picking an individual portfolio from the set of efficient frontier portfolios has not been solved. A convex set has been replaced by a set of CAL's, each of which has a theoretically infinite number of possible portfolios. What has been demonstrated is that, with a riskfree asset, it is possible to specify ($E[R]$, σ) tradeoffs that are unattainable with risky assets alone. In order to identify the best CAL and the appropriate portfolio to select on that CAL, it is conventional to introduce the mean-varance expected utility function: $EU[R] = E[R] - b\, var[R]$. As depicted in Figure 3-b, the mean-variance EU function defines a preference ordering over the ($E[R]$, σ) space. Movements in a northwest direction indicate increasingly higher levels of expected utility. It follows that the CAL which is just tangent to efficient frontier will attain the highest level of expected utility and the tangency of that CAL with the highest EU curve will be the specific portfolio that maximizes EU. This importance of this particular CAL is recognized specifically by referring to it as the ***capital market line*** (CML).

INSERT FIGURE 3-b EU map for Mean-Variance function

As it turns out, the slope of the capital allocation line for any efficient portfolio X will be of interest in identifying the properties of the capital market line. Observing that the rise of the CAL is $E[R_x] - r$ and the run from the origin is σ_x it follows that the slope of any CAL is $(E[R_x] - r)/\sigma_x$. Using this result, the equation of the capital allocation line for X becomes:

$$E[R] = r + \frac{E[R_x] - r}{\sigma_x} \sigma_R$$

Refer to Figure 3-b describing the indifference curves in ($E[R]$, σ) space for a risk averse investor. Because utility increases as the slope of the capital allocation line increases, the rational investor will achieve the maximum level of utility by selecting the capital allocation line with maximum slope, i.e., the CAL that is just tangent to the efficient set. Under perfect market assumptions, the CAL that is just tangent to the efficient set represents the highest level of utility. This ***tangency portfolio*** is the portfolio which represents the ***market equilibrium.*** With the additional theoretical apparatus provided by the capital asset pricing model (CAPM) it can be demonstrated that the tangency portfolio associated with the capital market line is the ***market portfolio.***

To illustrate the process of identifying a specific portfolio, refer back to the stock/bond fund portfolio discussed previously. Assume that a riskfree asset is available with r = 3%. By maximizing the slope of the CAL, the weights for the tangency portfolio are given to be w_S = .561 and w_B = .439. For these weights, $E[R_M]$ = .0893 and σ_M = .0912 where M indicates the tangency portfolio. Observing that the slope of CML is (.0893 - .03)/.0912 = .65, the equation of the capital market line can be specified as: $E[R_R] = .03 + .65\ \sigma_R$. Now, suppose the investor's indifference map is given by the expected utility function: $EU[R] = E[R] - \{3.25\ var[R]\}$. Recognizing that $var[R] = ((1 - w_r)\ \sigma_M)^2$ and $E[R_R] = w_r\ r + (1 - w_r)\ E[R_x]$, the process of maximizing $EU[R]$ gives $w_r = -.096$ as the optimal holding for the riskless asset. This implies that, for the specified mean-variance expected utility function, the optimal solution involves a solution on the CML to the right of the tangency with the efficient frontier. It can be verified that alternative values of the risk aversion parameter b give: $b = 5$, $w_r = .2875$; $b = 3$, $w_r = -.01786$; and, $b = 2$, $w_r = -.78125$.

E. Criticism of Mean-Variance Portfolio Analysis

While the mean-variance portfolio model has considerable theoretical appeal, there are a number of substantive problems that arise in implementing the model. One obvious problem concerns the large number of parameters that have to be estimated. Even if this problem can be overcome, attempting to capture the gains, *out-of-sample*, has proved to be illusive, particularly when international assets are permitted to be part of the set of available securities. In practice, the use of ***ex post*** (in-sample) data to estimate the relevant ***ex ante*** (out-of-sample) parameters creates numerous problems, not the least of which is instability in both the mean and variance-covariance parameter estimates. This is especially the case where expected returns are of interest. As pointed out by Eaker, et.al. (1991): "The problem with including returns in the portfolio selection decision is that such portfolios generally perform poorly in out-of-sample tests."

The mean-variance portfolio model is a central tenet of modern Finance. Much like another central tenet, the efficient markets hypothesis, enthusiasm for the mean-variance portfolio model within Finance has evolved considerably. In recent years the model has been subjected to substantial critical scrutiny. The first wave in the assault on the mean-variance approach can be attributed to Jorion (1985, p.265), which describes the problems emphatically in the context of internationally diversified portfolios:

> Mean-variance analysis has serious shortcomings which are too often ignored ... Perhaps the most serious defect in the classical (portfolio) approach is the poor out-of-sample performance of the optimal portfolios. Performance measures always deteriorate substantially outside the sample period, and the supposedly optimal choice is sometimes dominated by a naive method....Another problem is the instability in the optimal portfolio: the proportions allocated to each asset are extremely sensitive to variations in expected returns, and adding a few observations may change the portfolio distribution completely. Also, optimal portfolios are not necessarily well diversified. Often a corner solution appears, where most of the investments are zero and large proportions are assigned to countries with relatively small capital markets and high average returns.

As it turns out, this attack is decidedly overstated. However, the basic point remains: *ex post* estimates of expected returns, based on arithmetic or weighted average estimators, are not reliable estimates of future returns. Relative to estimates of variances and covariances, Jorion (1985), Eun and Resnick (1988) and others demonstrate that estimates of expected returns are considerably more unstable over time.

Empirically, the parameter instability problem has a number of implications. For example, *ex ante* results concerning the return on a given portfolio may vary significantly from sample to sample.

Jorion (1985) examines the out-of-sample performance of the two *ex post* optimal internationally diversified portfolios identified by Grubel (1968) and Levy and Sarnat (1974), together with two "naive" portfolios, the equally weighted and market value weighted portfolios. As measured by the Sharpe ratio, Jorion found that over the next investment horizon, the *ex ante* performance of the two mean-variance efficient portfolios was inferior to the performance of the naive equally weighted portfolio. Jorion (1985) also provides evidence that, in estimating *ex post* returns, longer sampling windows, e.g., five years for monthly data, provides superior *ex ante* forecasting when compared with shorter sampling windows, e.g., 1 year of monthly data. The difficulty with longer sampling windows is that it takes a longer time interval for the estimates to react to changing market conditions.[3]

In addition to the length of the sampling window, the type of estimator also can have a significant impact on the *ex ante* results. In applications, means and variances are estimated from the most recent data available at the time the portfolio is rebalanced. Grauer and Hakansson (1992) provide data on the behavior of the mean-variance efficient portfolio over time, using quarterly rebalancing of the portfolio, based on the most recent 40 quarters of data. The investment universe includes US equities and bonds, seven non-US equities (Canada, Japan, UK, Switzerland, Netherlands, Germany and France). The simulations try four different methods of estimating the mean and provide results over a wide range of possible investor preferences. Their results indicate the importance of mean estimation to portfolio composition. When simple historical averages are used to estimate the means, the average portfolio contains 3 risky assets or less for the "typical" range of investor preferences. When the CAPM is used, eight to the maximum possible ten risky assets is common in the average portfolio.

In addition to the estimation method used to determine the relevant parameter inputs, the presence or absence of short-selling has been found to be fundamental in assessing the performance of mean-variance efficient portfolios. Even though the early studies implicitly assumed short selling was not permitted, at least since Jorion (1985) it has been recognized that odd results can be obtained when short selling is permitted. For example, Jorion reports results for the time series properties of the optimal weight on domestic assets in the *ex post* tangency portfolio (see below). A considerable amount of short-selling is indicated at various times, as much as -2.4 times the total principal value of the portfolio at one point in 1978. For many types of investment situations, e.g., pension funds, life insurance companies, this amount of short selling would be unacceptable and unobtainable. Evidence on portfolio composition with short selling restrictions, e.g., Glen and Jorion (1993), indicates a dramatic narrowing of the number of assets held in the portfolio is likely, amplifying the concentration of a given portfolio in a small number of assets.

Somehow, proponents of the model believe that the out-of-sample prediction problems can be resolved by improving the estimation methods that are used. This still leaves the problem of identifying the appropriate portfolio from the set of mean-variance efficient portfolios. Following Sharpe and others, the efficient frontier portfolio to be selected is that portfolio associated with the capital allocation line which is just tangent to the efficient frontier, i.e., the portfolio associated with the ***capital market line***. This tangency portfolio can be determined by solving the following optimization problem, e.g., Eun and Resnick (1994):

$$\max_{\{w_i\}} \frac{E[R_p] - r}{\sigma_p} \qquad \textit{subject to}: \quad \sum_{i=1}^{k} w_i = 1$$

where r is, as before, the riskfree interest rate. On theoretical grounds, the tangency portfolio is the *ex ante* mean-variance-expected-utility optimizing risky portfolio. Even though the precise combination of riskless asset and risky tangency portfolio for any given investor requires specification of the relevant parameters for the investor's mean-variance expected utility function, the optimal risky portfolio has been determined.

Given the *ex post* estimates of the relevant means, variances and covariances, the optimality problem is solved and the resulting tangency portfolio will represent the optimal, *in-sample* portfolio. Whether this in-sample optimality translates into superior *out-of-sample* performance is an open question. The answer to this question becomes even more complex when foreign assets are admitted into the asset universe. In particular, the domestic currency return on a foreign asset depends on a combination two random variables: the return denominated in foreign currency terms; and, the change in the exchange rate. The correlation between foreign and domestic asset returns will tend to be lower than the correlations between domestic assets, making foreign assets excellent candidates for diversification. In addition, foreign assets can also provide the possibility of significantly higher returns than domestic assets. As demonstrated by Eun and Resnick (1994), the difficulties associated with estimating expected returns results in the minimum variance portfolio having generally superior *ex ante* performance compared to the tangency portfolio.

3.2 Separation, the CAPM and the Market Model

A. Two Fund Separation

The combination of mean-variance expected utility, perfect markets and the CML provides the basis for a version of the ***two fund separation property***: in market equilibrium, rational risk averse investors will hold portfolios which combine the riskfree asset with the tangency portfolio. The precise combination of the riskfree asset and the tangency portfolio will depend on the risk preferences of the individual investor. The CML result does not provide any information about how to determine the return on individual assets, or any portfolio of assets which is not efficient. The CML also does not provide specific information about the asset composition of the tangency portfolio. This information is provided by the capital asset pricing model (CAPM). If the CAPM is incorporated, then it can be shown that the tangency portfolio will be the market portfolio. In this case, the two fund separation property says that, in market equilibrium, ***rational risk averse investors will hold portfolios which combine the riskfree asset and the market portfolio***.

Two fund separation provides the theoretical basis for a persuasive and implementable investment strategy. This strategy requires a strong belief in efficient markets. If markets are efficient then the gains to individual security selection strategies, using either fundamental or technical analysis, will be illusory. The decision problem facing the rational investor is to determine what fraction of invested capital to hold in the risky market portfolio and what fraction to hold in the riskfree assets. Investors with high levels of risk tolerance will leverage up, by borrowing at the riskfree rate, and purchase more of the market portfolio. Investors with moderate to low levels of risk tolerance will have positive investment weights for both the riskfree asset and the market portfolio. Though this perfect markets result requires some adjustment to account for market imperfections, e.g., differences between lending and borrowing rates, the basic intuition survives in tact. As pointed out by Roll (1978), the main practical ambiguities lies with the specification of the riskfree asset and the market portfolio.

The CAPM provides a method for determining the expected return, $E[R_i]$, for any asset *i*, not just for portfolios on the efficient frontier. The CAPM is an *ex ante* model that can be expressed as:

$$E[R_i] = r + \{E[R_m] - r\}\, \beta_i$$

where $E[R_m]$ is the expected return on the market portfolio and β_i is a measure of the ***systematic risk*** of asset *i*. In words, the CAPM can be expressed as: the expected return on asset *i* = risk free rate + systematic risk premium for asset *i*. A key variable in the CAPM is β which is specified as:

$$\beta_i = \frac{cov[R_i, R_m]}{\sigma_m^2} \equiv \frac{\sigma_{i\,m}}{\sigma_m^2}$$

where σ_m^2 is the variance of the return on the market portfolio.

Some examples of mechanical calculations that can be done with CAPM are: assume that the rate of return on the market $E[R_m] = .15$ and $r = .05$ and $\beta_i = 1.5$ then the expected return on asset *i* is $E[R_i] = .05 + (.15 - .05)(1.5) = .20$; assume that $E[R_i] = .1$, $E[R_m] = .105$ and $\beta_i = .9$ then the riskfree rate *r* is 5.5%; assume that $E[R_i] = .2$, $E[R_m] = .15$ and $r = .10$, then $\beta_i = 2$. Beta is applicable not only for individual securities but also for portfolio of securities. The following useful result can readily be derived: the beta of a portfolio, β_p is the value weighted sum of the individual betas (the β_i's). This follows because the CAPM holds for any asset, including individual assets as well as portfolios of assets. Recognizing that the CAPM will hold for the efficient portfolios on the efficient frontier, the CAPM can be used to show that the tangency portfolio in the CML is the market portfolio.

To demonstrate this result, assume that the CAPM is true. If the tangency portfolio is the market portfolio, the CML provides the result:

$$E[R_R] = r + \frac{E[R_m] - r}{\sigma_m}\, \sigma_R$$

where *m* refers to the market portfolio. Using the result that $\sigma_R = (1 - w_f)\sigma_m$ this can be rewritten:

$$E[R_R] = r + \frac{E[R_m] - r}{\sigma_m}(1 - w_r)\, \sigma_m = r + \{E[R_m] - r\}\,(1 - w_r)$$

If the CAPM is true then it will hold for any portfolio along the CML. If the market portfolio is the tangency portfolio for the CML then $E[R] = w_r\, r + (1 - w_r)\, E[R_m]$ and evaluating β gives:

$$\beta = \frac{cov[E[R],\, E[R_m]]}{\sigma_m^{\,2}} = \frac{cov[\{w_r\, r + (1 - w_r)\, E[R_m]\},\, E[R_m]]}{\sigma_m^{\,2}}$$

$$= \frac{(1 - w_r)\, \sigma_m^{\,2}}{\sigma_m^{\,2}} = (1 - w_r)$$

Substituting this result back into the CML shows that the CAPM and CML are equivalent when the tangency portfolio is the market portfolio.

INSERT FIGURE 3-c Security Market Line

The relationship between the CAPM and the CML can be expressed in a linear form in ($E[R]$, β) space. This linear relationship is the ***security market line*** (SML). While the CAL and CML provide a linear relationship in ($E[R]$, σ) space between total risk, as measured by standard deviation, and expected return, the SML provides a linear relationship between systematic risk, as measured by β, and expected return. The equation for the SML can be derived by identifying two points on the line: the riskfree rate where $\beta_f = 0$ and $E[R_i] = r$ and the market portfolio where $\beta_m = 1$, $E[R_i] = E[R_m]$. It follows that the slope is $(E[R_m] - r)$. From this the equation for the SML can be stated: $E[R_i] = r + (E[R_m] - r)\, \beta_i$. Hence, the SML is the graphical representation of the CAPM. A useful pedagogical application of the SML is to describe whether a particular security is over or underpriced relative to its measure of systematic risk.

*B. The Capital Asset Pricing Model**

The Markowitz model of mean-variance portfolio optimization is concerned with the behavior of an individual investor selecting securities for inclusion in an optimal portfolio. Practical application of this model is complicated by the large number of parameters that have to be estimated and the associated complexity of the solutions as the number of securities is increased. The CAPM provided a theoretical mechanism for handling this problem. Using the CAPM, the problem of estimating the optimal weights for the individual assets in the tangency portfolio is replaced with a method of identifying the predetermined set of weights associated with the market portfolio. Though the notion of a 'market portfolio' is somewhat nebulous, in practice it has been interpreted to be a widely diversified value-weighted portfolio of common stocks such as the S&P 500, e.g., Damodaran (1994). This all raises the need to examine the derivation of the CAPM in more detail.

In the years since the CAPM was introduced by Sharpe (1964), Lintner (1965) and Mossin (1966), considerable effort has been given to extending and expanding the basic model (see Chapter 10). The basic CAPM, also referred to as the one factor or single index model, is derived under perfect markets assumptions. The process of extending and expanding has been largely concerned with relaxing these assumptions. Unlike the partial equilibrium approach of the mean-variance portfolio model, the CAPM is a general equilibrium model. It is this feature that permits the CAPM to go beyond the basic portfolio structure to make statements about the expected returns for individual assets. General equilibrium requires market clearing conditions for all assets to be satisfied. To accomplish this, the CAPM relies on the assumption the investors are homogeneous, possessing the same expectations about the means and variances of returns and the same investment horizon. All investors are assumed to have the same form of mean-variance expected utility function.

Much of the derivation of the CAPM follows the mean-variance optimization procedure. The homogeneity assumption is invoked after the riskfree rate is introduced and the optimality of the tangency portfolio is established. Because investors are homogeneous and market clearing is required, it must be that the tangency portfolio is the market portfolio. Fama (1976, p.274-5) describes the logical argument:

> ... a market equilibrium requires a market-clearing set of prices; a market equilibrium requires that, in aggregate, investors demand them in the proportions in which they are outstanding. Given the nature of the efficient set when there is risk-free borrowing and lending, this market-clearing condition means that a market equilibrium is not attained until the one tangency portfolio that all investors try to combine with risk-free borrowing or lending is a portfolio of all the positive variance securities in the market, where each security is weighted by the ratio of the total market value ... of all its outstanding units to the total market value of all outstanding units of securities. In short, a market equilibrium is not reached until the tangency portfolio ... is the value-weighted version of the market portfolio ... A market equilibrium – a set of security prices that clears the securities market and a value of (the risk-free rate) that clears the borrowing-lending market – requires that the tangency portfolio be the market-weighted version of the market portfolio.

This last step, the identification of the tangency portfolio with the market portfolio, follows immediately from the investor homogeneity assumption. In the derivation of the CAPM, this step is something of an afterthought.

The key parts of the CAPM relate to developing the relationship between the expected return on a given asset and the expected return on an efficient portfolio. This derivation requires one useful result associated with linear combinations of random variables:[4]

$$var[R_p] = \sum_{i=1}^{k}\sum_{j=1}^{k} w_i\, w_j\, \sigma_{ij} = \sum_{i=1}^{k} w_i \left(\sum_{j=1}^{k} w_j\, \sigma_{ij} \right)$$

$$= \sum_{i=1}^{k} w_i\, cov[R_i,\ R_p]$$

Given this, the derivation of the CAPM proceeds by solving the Lagrangian arising from the mean-variance portfolio model:

$$\max_{\{w_i\}}\ L = var[R_p] - 2\,\lambda_1 \left(\sum_{i=1}^{k} w_i\, E[R_i] - \bar{c}_n \right) - 2\,\lambda_2 \left(\sum_{i=1}^{k} w_i - 1 \right)$$

This optimization problem will produce k first order conditions associated with the $\{w_i\}$ together

with two additional first order conditions for the constraints to produce a system of $k+2$ equations.
For jth security, the first order condition provides:

$$\sum_{i=1}^{k} w_i \, cov[R_i, R_j] - \lambda_1 \, E[R_j] - \lambda_2 = 0$$

As the ordering of the securities is arbitrary, this result can be equated with the first order condition for the first security to obtain:

$$\sum_{i=1}^{k} w_i \, cov[R_i, R_j] - \lambda_1 \, E[R_j] = \sum_{i=1}^{k} w_i \, cov[R_i, R_1] - \lambda_1 \, E[R_1]$$

The next step involves multiplying both sides by w_j and summing over j. This affects the right and left hand sides differently. The left hand side produces:

$$\sum_{j=1}^{k} w_j \sum_{i=1}^{k} w_i \, cov[R_i, R_j] - \lambda_1 \sum_{j=1}^{k} w_j \, E[R_j] = var[R_p] - \lambda_1 \, E[R_p]$$

The right hand side produces:

$$\sum_{j=1}^{k} w_j \left(\sum_{i=1}^{k} w_i \, cov[R_i, R_1] - \lambda_1 \, E[R_1] \right) = \sum_{i=1}^{k} w_i \, cov[R_i, R_1] - \lambda_1 \, E[R_1]$$

This result follows because there is no j on the right hand side and, as a result, the sum of the weights equals one and has no impact.
Observing that the weights apply to a mean-variance efficient portfolio, i.e., R_p is on the efficient frontier, manipulating the right and left hand sides produces the result:

$$E[R_1] - E[R_p] = \frac{1}{\lambda_1} \left(\sum_{i=1}^{k} w_i \, cov[R_i, R_1] - var[R_p] \right) = \frac{1}{\lambda_1} \left(cov[R_1, R_p] - var[R_p] \right)$$

What remains is to determine λ, which is the Lagrange multiplier associated with the impact of changes in the target level of portfolio expected return on the variance of the portfolio. When there is a riskfree rate, the λ for the tangency portfolio can be determined as:

$$2 \, \lambda_1 = \frac{d \, var[R_p]}{d \, E[R_p]} = \frac{d \, var[R_p]}{d \, \sigma_p} \frac{d \, \sigma_p}{d \, E[R_p]} = 2 \, \sigma_p \left(\frac{\sigma_p}{E[R_p] - r} \right)$$

Substituting this λ result back into the prior equation and manipulating gives the CAPM.
Nothing in this derivation demonstrates that the tangency portfolio is the market portfolio. This result is obtained from the logical argument about market clearing with homogeneity of consumers. In a sense, two fund separation, where the rational investor holds combinations of the market portfolio and the riskless asset, is too strong a condition. A more appropriate result would be a partial equilibrium result where the rational investor holds combinations of a mean-variance efficient

portfolio and the riskless asset. But this would require the mean-variance efficient portfolio to be determined, falling back to the problems associated with complexity, number of parameters to estimate, estimator forecasting error and the like. That the CAPM assumptions make proponents of modern Finance queezy is apparent in the following quote from Elton and Gruber (1984, p.273):

> It is worthwhile pointing out ... that the final test of a model is not how reasonable the assumptions behind it appear but how well the model describes reality. As the reader proceeds with this chapter, he will, no doubt, find many of its assumptions objectionable. Furthermore, the final model is so simple the reader may well wonder about its validity. As we shall see, despite the stringent assumptions and the simplicity of the model, it does an amazingly good job of describing prices in capital markets.
>
> The real world is sufficiently complex that to understand it and construct models of how it works, one must assume away those complexities that, hopefully have only a small (or no) affect on its behavior. As the physicist builds models of the movement of matter in a frictionless environment, the economist builds models where there are no institutional frictions to the movement of stock prices.

These words reflect the place of the CAPM at the foundation of the positivist superstructure that is modern Finance.

*C. The Market Model**

Because the CAPM is an *ex ante* model it depends on expected returns and other unknown values, that are not directly observable or testable. The statistical representation of the CAPM which is testable is known as the ***market model***. The market model is a bivariate regression model of the form:

$$R_{i\,t} = \alpha_i + \beta_i\, R_{m\,t} + e_{i\,t}$$

where R_{it} is the observed return on asset *i* at time *t*, R_{mt} is the observed return on the market portfolio at time *t*, α_i and β_i are statistical parameters to be estimated and e_{it} is the asset specific error which is assumed to obey the statistical properties for ordinary least squares regression. For the market model the ordinary least squares (OLS) assumptions (in vector notation) are: $E[e_i] = 0$, the firm specific error has mean zero; $E[e_i\, R_m] = 0$, the firm specific risks are uncorrelated with the market return; $E[e_i\, e_j] = 0$ for *i* not equal to *j,* the firm specific risks for different securities are uncorrelated; in addition, it is assumed that the e_i are iid random variables. The additional assumption of normality of e_i facilitates hypothesis testing. With these assumptions, ordinary least squares can be used to estimate the coefficients α_i and β_i. The market model is sometimes referred to as the single index model, e.g., Elton and Gruber (1995).

Taking expected values for the market model gives for any t:

$$E[R_i] = \alpha_i + \beta_i\, E[R_m] \quad \text{because } E[e_i] = 0$$

If the CAPM is true, it follows that $\alpha_i = (1 - \beta_i)\, r$, (β_i has the same interpretation as $cov[R_i, R_m]/var[R_m]$). This interpretation for α depends on the riskless rate being a constant. While this assumption is correct in one-step-ahead decision making, it is problematic when estimating parameters in a time series. To account for changes in *r* over time, the market model is often expressed in ***risk premium form***, by subtracting the observed riskless rate, in any given period, from the observed returns:

$$R_{i\,t} - r_t = \alpha_i + \beta_i\,[R_{m\,t} - r_t] + u_{i\,t}$$

where u_i is also a firm specific risk with ordinary least squares properties. In this form, taking expectations and assuming that the CAPM holds gives $\alpha_i = 0$ and β_i with the same interpretation.

Beta is a measure of systematic or ***market risk***. It provides information on how the stock return reacted historically when the market portfolio changed. Beta estimates are reported at a number of websites, such as www.bloomberg.com. For $\beta_i > 1$ (high beta), when the return on the market portfolio changes, the return on the stock will tend to change by ***more*** than the return on the market portfolio. Stocks with higher than market betas are considered to be ***aggressive***. When the market is expected to move up, shifting into stocks with high beta is indicated. In the US market, examples of high beta stock groups occur in industries such as trucking, consumer durables, construction and air transport. For $\beta_i < 1$ (low beta) when the return on the market portfolio changes, the return on the stock will tend to change by ***less*** than the return on the market portfolio. Stocks with lower than market betas are considered to be ***defensive***. When the market is expected to move down, shifting into stocks with high beta is indicated. In the US market, examples of low beta stock groups occur in telephone stocks, utilities, breweries and food producers/distributors.

Alpha is an ***asset specific*** measure which indicates the ***excess return*** that the security earned beyond that warranted by the risk premium captured by the security' s beta. For the market model expressed in risk premium form, positive alpha indicates that the stock outperformed the market, after adjusting for systematic risk. Negative alpha indicates that the stock underperformed the market, after adjusting for systematic risk. The market model provides a useful simplification for determining the betas and alphas for a portfolio from the alphas and betas of the individual securities:

$$E[R_p] = \sum_{i=1}^{k} w_i\, E[R_i] = \sum_{i=1}^{k} w_i\, \{\alpha_i + \beta_i\, E[R_m]\}$$

$$\sum_{i=1}^{k} w_i\, \alpha_i + \sum_{i=1}^{k} \beta_i\, E[R_m] = \alpha_p + \beta_p\, E[R_m]$$

In other words, the alpha and beta for the portfolio are the value weighted sums of the individual portfolio alphas and betas.

The market model can also be used to demonstrate that portfolio diversification leads to the elimination of unsystematic or ***firm specific*** risk leaving only systematic risk as the determinant of portfolio variance. If the market model is true then the variance for individual security returns reduces to:

$$\sigma_i^2 = \beta_i^2\, \sigma_m^{\,2} + \sigma_{e_i}^{\,2}$$

Similarly, the covariance between the security returns becomes:

$$\sigma_{i,j} = \beta_i \, \beta_j \, \sigma_m^{\,2}$$

These results follow from the assumptions made in the market model. Substituting these results into the formula for the portfolio variance, σ_p^2 gives:

$$\sigma_p^{\,2} = \sum_{i=1}^{k} w_i^2 \, \beta_i^2 \, \sigma_m^{\,2} + \sum_{i=1}^{k} w_i^2 \, \sigma_{e_i}^{\,2} + \sum_{i>j} w_i \, w_j \, \beta_i \, \beta_j \, \sigma_m^{\,2}$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{k} w_i \, w_j \, \beta_i \, \beta_j \, \sigma_m^{\,2} + \sum_{i=1}^{k} w_i^2 \, \sigma_{e_i}^{\,2}$$

Observing that the variance of the market portfolio is a common term in the double sum, it is possible to do some factoring:

$$\sigma_p^{\,2} = \{\sum_{i=1}^{k} w_i \, \beta_i\} \, \{\sum_{j=1}^{k} w_i \, \beta_i\} \, \sigma_m^{\,2} + \sum_{i=1}^{k} w_i^2 \, \sigma_{e_i}^{\,2}$$

Using the result that the beta of the portfolio is the value weighted sum of the individual security betas, the following simplification is available for the portfolio variance:

$$\sigma_p^{\,2} = \beta_p^{\,2} \, \sigma_m^{\,2} + \sum_{i=1}^{k} w_i^2 \, \sigma_{e_i}^{\,2}$$

$$\sigma_p \rightarrow \beta_p \, \sigma_m \qquad as \qquad k \rightarrow \infty \qquad and \qquad w_i \Downarrow$$

The term involving β_p is the ***systematic*** or market related risk. It depends only on the composition of the portfolio and the variance of the return on the market. The second term involves only firm specific or ***unsystematic*** risks.

To show the impact of diversification on both systematic (market) risk and unsystematic (firm specific) risk, observe that the first term involving the beta of the portfolio is not much affected by increases in the number of securities in the portfolio. The second term involving the firm specific risks is directly affected by the number of securities in the portfolio. Take the case of an equally weighted portfolio where $w_i = 1/N$. In this case, as the number of the securities in the portfolio increases the firm specific risks are reduced at the rate of $(1/N^2)$, which is converging to zero quite rapidly. This follows from observing that when the firm specific risk has been eliminated then the last term on the rhs of the last equation is zero. Taking square roots and observing that the portfolio beta is the weighted sum of the individual security betas provides the required result. However, because there are a large number of securities in a portfolio, this does ***not*** mean that the firm specific risks are eliminated. Rather, the elimination of firm specific risk depends on reducing the value weights attached to each security as N increases. For example, if $w_1 = .5$ for a particular security, and this value does not change as the number of securities in the portfolio increases, then the firm specific risk associated with that security is not eliminated as the number of securities in the portfolio increases.

3.3 Decision Making Under Uncertainty

*A. Expected Utility Theory**

The study of decision making under uncertainty is a vast subject. Theoretical applications in Finance almost invariably proceed under the guise of the expected utility hypothesis: people rank random prospects according to the expected utility of those prospects. Analytically, financial decisions typically involve solving optimization problems by selecting choice variables to maximize an expected utility function subject to a budget constraint. The choice variables are typically the proportion of the initial budget to allocate to a given asset or security. In some cases, such as the basic optimal hedging problem, e.g., Poitras (2002, ch.2), the budget constraint is embedded in the argument of the utility function. In other cases, such as in optimal portfolio diversification models, the budget constraint appears as the restriction that the sum of the value weights equals one. In either event, the central concern is expected utility. As such, a key step in the optimization problem is to specify an expected utility function which captures the preference mapping of the decision maker over random outcomes.

Expected utility calculations involve taking expectations, which are conventionally modeled using statistical properties of random variables. This may involve the explicit introduction of probability densities. There is a profound connection between the choice of a specific probability distribution and the risk aversion properties required of the expected utility function, e.g., Heaney and Poitras (1994). As for the utility component of the expected utility function, even before von Neumann and Morgenstern (1947), it had been recognized that choosing over risky prospects is decidedly different than the textbook model of consumer choice under certainty. As is well known, von Neumann and Morgenstern made a seminal contribution by proposing a set of axioms governing choice under uncertainty. Accepting that the axioms are difficult to reject lends strong support to the von Neumann and Morgenstern approach.[5]

The key construct of the axiomatic approach to decision making under uncertainty is the 'linear choice function over risky prospects', better known as the ***expected utility function***:

$$EU[x] = \sum_{j=1}^{S} \theta_j \; U_j[x]$$

where: $EU[x]$ is the expected utility of x; S is the number of possible futures states of the world; θ_j is the probability that state j will occur, where the sum of the θ's across all $j=1$ to S states is equal to one; and, $U[x_j]$ is the utility associated with the amount of x received in state j. While there are a number of possible selections for x, in what follows variables such as terminal wealth, portfolio return or terminal profit will typically be used. (The implications of this selection will be examined shortly.) The EU function ranks risky prospects with an ordering that is unique up to a positive linear transformation. This is different than the case where just $U[x_j]$ is being maximized, as in the theory of consumer choice in microeconomics where the amounts are know with certainty. In this case, the U ranking is unique up to a positive monotonic transformation.

While the ***difference between a monotonic and linear transformation*** may not seem too significant, demonstrating the difference does provide some insight into the mechanics of the expected utility calculation. Consider the case where there is a single good, x. Two prospects A and

B have to be ranked with $EU[x]$, prospect A is a lottery with two possible outcomes of either 0 units of x or 16 units of x, with each event equally likely, $\theta_1 = \theta_2 = .5$. This lottery is being compared with prospect B which is a certain outcome of 7 units of x. Let the utility function be $U[x] = x$, the expected utility is then $EU[A] = .5(16) + .5(0) = 8 > 7 = EU[B]$. Now consider what happens when a square root transformation is applied to the x: $EU[A] = .5(4) + .5(0) = 2 < \surd 7 = EU[B]$. Though the utility ranking is unchanged under the monotonic transformation of taking a square root, $U[16;4] > U[7;\surd 7] > U[0]$, the EU ranking did change. However, when $U[x] = a + bx$, then the expected utility calculation gives: $EU[A] = .5(a + b(16)) + .5(a + b(0)) = a + .5\, b(16) > a + b(7) = EU[B]$ for all positive linear transformations, $b > 0$.

Given the basic specification of the EU function, developing the empirical implications of models based on EU can be tricky because EU is a construct that cannot be directly observed. In addition to the difficulties of specifying $U[x_j]$ encountered in microeconomics (e.g., satiation points and externalities) it is also not apparent how to determine the θ_j. The axiomatic foundation cannot say much more than that the probabilities are available. Davidson (1991) and others emphasize the importance of the approach taken to identifying the probabilities in the expected utility calculation. In modern Finance, it is conventional to proceed, either explicitly or implicitly, along the lines of Ingersoll (1987, p,30): "We assume throughout the discussion that the economic agents making the decisions know the true objective probabilities of the relevant events. This is not in the tradition of using subjective probabilities but will suffice for our purposes." In assuming that the probabilities are objective, Ingersoll follows the approach used in rational expectations modeling in economics.

Davidson (1991, p.132) develops the philosophical ***implications of assuming that the probabilities are objective***:

> The objective probability environment associated with the rational expectations hypothesis presumes not only that probability distributions regarding historical phenomena have existed, but also that the same probabilities which determined past outcomes will continue to govern future events. In the context of forming expectations which do not exhibit persistent errors, it holds that time averages calculated from past data will converge with the statistical averages computed from any future time series. Knowledge of the future merely involves projecting averages based on past or current realizations to forthcoming events. The future is merely the statistical reflection of the past and economic actions are in some sense timeless. There can be no ignorance of upcoming events for those who believe the past provides reliable, unbiased, statistical information (price signals) regarding the future, and this knowledge can be obtained if only one is willing to spend the resources to examine the past.

Ergodicity provides the basis for empirically extracting objective probabilities from past data (see sec. 1.3). As such: "In the ergodic circumstances of objective probability distributions, probability is knowledge, not uncertainty". This is strong stuff. If correct, it means that the core theories of modern Finance deal with the uncertainty of future outcomes in a fashion that could be called into question. As such, predictions derived from the core models require, at the least, careful interpretation.

Taking probabilities as objectively determined proceeds beyond the intentions of the developers of the axiomatic approach to decision making under uncertainty where probabilities are subjectively determined. In the case of von Neumann-Morgenstern (1947), the subjective probabilities are given by assumption while in Savage (1954) the subjective probability assessments can be deduced. Yet, objective probabilities are implied by the general equilibrium formulation needed to derive models of security prices or returns, such as the CAPM. General equilibrium models often proceed by assuming that expectations are homogeneous or that individual agents are homogeneous, both of which are facilitated by objective probabilities. If probabilities are subjective then it is necessary to assume investor heterogeneity. However, if heterogeneity is the case, then key results, such as those concerning the properties of the market portfolio, are undermined. As a consequence, essential prescriptions, such as two fund separation, require reinterpretation.

In the absence of objectively known probabilities or, less stringently, probabilities estimated from past data under the assumption of ergodicity, it is still possible to use the techniques of decision making under uncertainty to develop useful prescriptions and predictions. In the world of Old Finance, for example, the theoretical requirements needed to satisfy general equilibrium concerns were of little use. The decision problems typically encountered were partial equilibrium. The ad hoc theoretical results applied to investors and traders confronted with a parametric world of atomistic competition where their activities did not impact prices. In this process, the expected utility function can still be an invaluable analytical tool. A number of useful results can be readily derived by applying an essential tool from functional analysis: the Taylor series expansion (see sec. 5.2). In addition, much of the intuition of the mean-variance portfolio optimization model can still be exploited, albeit in the framework of subjective estimates of risk and expected return.

Given the probabilities, analytical solutions derived using *EU[x]* will depend of the selection of the argument x. In the ***theory of consumer choice*** under certainty covered in basic microeconomics, e.g., Henderson and Quandt (1980), x refers to two or more goods, e.g., $U[q_1, q_2]$. Analysis of the consumer choice problem for a two good world proceeds by maximizing $U[q_1, q_2]$ subject to a budget constraint, $y = p_1 q_1 + p_2 q_2$, where y is the initial endowment level and p and q refer to the price and quantity of the commodity. Solving the first order conditions for ***the Lagrangian problem*** (*L)* gives the results:

$$L = U[q_1,q_2] + \lambda(y - p_1 q_1 - p_2 q_2)$$

$$\textit{Solving the foc} \;\rightarrow\; \frac{\dfrac{\partial U[q_1,q_2]}{\partial q_1}}{\dfrac{\partial U[q_1,q_2]}{\partial q_2}} \equiv \frac{U_1'}{U_2'} = \frac{p_1}{p_2} \qquad\qquad \frac{\partial L[q_1,q_2;y]}{\partial y} = \lambda$$

By specifying a functional form for U together with specific prices and an endowment level, it is possible to solve the optimal quantities of q_1 and q_2. Varying prices will permit demand functions to be determined and so on. Using appropriate mathematical complications, the results generalize in a natural way to n commodities. Useful results are obtained when q_1 and q_2 refer to income and leisure or consumption in two different time periods.

Despite the use of U there is a significant change in orientation when attention shifts from the certainty model of consumer choice to decision problems involving *EU[x]*. Instead of being concerned with allocating a fixed income endowment among a number of goods, concern shifts to a one good (plus numeraire) world of allocating the endowment between current consumption $C(0)$ and the purchase of capital assets that generate the endowment in the next period. As discussed in Chapter 10, there is an assumed separation between the generation of future income from the asset investment decision and from future labor income, with the implications of the latter usually being suppressed. The initial endowment, $W(0)$, is conventionally referred to as initial wealth. For the simple two period case where there is only one capital asset that has a random return, then terminal wealth, $W(1)$, is specified:

$$W(1) = \{W(0) - C(0) - A(0)\}(1 + r) + A(0)(1 + R(1))$$
$$= \{W(0) - C(0) - A(0)\}(1 + r) + \{A(0) + \pi(1)\}$$

where $A(0) = P(0)Q(0)$ is the value of the capital asset purchased at $t=0$, r is the riskless return on holding the numeraire (cash), $C(0)$ is the consumption at $t=0$. Unless the consumption decision of immediate interest, it is conventional to omit consideration of $C(0)$.

Given this, because the choice variable is $A(0)$, the amount of the capital asset to purchase at $t=0$, it follows that the decision problem can be specified with terminal wealth $EU[W(1)]$, asset returns, $EU[R(1)]$, or profit, $EU[\pi(1)]$. Though there is no substantive differences between the selection of x in the simple case, the level of $W(0)$ can be important in some cases making the selection of $EU[W(1)]$ appropriate in those cases. In other cases, such as the important mean-variance optimization model, attention focuses on the portfolio returns and how to allocate the initial wealth among k assets. This leads to the use of $EU[R(1)]$, where $R(1)$ refers to the return on the asset portfolio which is, in turn, dependent on the individual asset returns. The use of $EU[\pi(1)]$ is common in problems involving derivative securities, such as the selection of the optimal amount of a forward contract to establish in order to hedge a spot position. In this case the profit on the forward contract is of interest.

Many of the basic properties of EU in the portfolio selection problem can be illustrated by ***a simple two state example*** (Mossin 1973, p.16-9). Setting $r=0$ and $C(0) = 0$ for convenience, assume that the value of the asset in state 1 is zero with probability θ_1 and in state 2 is $2A(0)$ with probability $\theta_2 = (1 - \theta_1)$. Taking the initial endowment to be $W(0) = 100$, these payoffs define a range of outcomes. For example: if $A(0) = 0$, then $W(1) = 100$; if $A(0) = 20$, then $W(1) = \theta_1 (80) + (1 - \theta_1)(140)$; and if $A(0) = 60$, then $W(1) = \theta_1 (40) + (1 - \theta_1)(220)$. In general:

$$EU[W(1)] = \theta_1 U[W(0) - A(0)] + (1 - \theta_1) U[W(0) + 2A(0)].$$

The optimization problem is to choose the amount of the risky asset to purchase in order to maximize the expected value of terminal wealth. Assuming that the initial price, $P(0)$, is not affected by the quantity of the asset purchased, $A(0)$, it is sufficient to solve this problem using the first order condition (foc) (dEU/dA) and evaluating the second derivative to ensure the solution is a maximum.

Evaluating the foc for the simple problem gives:

$$\frac{dEU}{dA} = \theta_1 \left\{ \frac{dU[W(0)-A]}{dW} \frac{dW}{dA} \right\} + (1 - \theta_1) \left\{ \frac{dU[W(0)+2A]}{dW} \frac{dW}{dA} \right\}$$

$$= -\theta_1 \; U'[W(0)-A] \; + \; 2 \; (1 - \theta_1) \; U'[W(0)+2A]$$

Solving for an optimal value of A requires a specific functional form for U to be assumed. For example, let $U[W(1)] = \ln[W(1)]$. It follows that:

$$\frac{dEU}{dA} = \frac{-\theta_1}{W(0)-A} + \frac{2\;(1-\theta_1)}{W(0)+2A} = 0$$

$$\rightarrow \quad A^* = \frac{2 - 3\;\theta_1}{2} W(0) = \frac{E[R_A]}{2} W(0)$$

where $E[R_A] = \{E[W(1)] - W(0)\} / A(0)$. This illustrates an important property of the log utility function: that the proportion of initial wealth invested in the risky asset, ($A(0)/W(0)$), is constant and independent of the amount of initial wealth.

B. Cost of Risk and Risk Aversion Properties *

The expected utility function is a useful tool for analyzing the problem of determining the cost of risk. The solution to this problem would be useful in assessing whether to buy insurance or to invest in a risky capital project. While there are a number of possible methods to extract the cost of risk, consider the following solution. Let the expected value of terminal wealth be: $E[W(t+1)] = \Omega$. Observe that Ω is a parameter which permits the ***certainty equivalent*** income of a risky prospect to be defined as $\Omega - C$, where C is the cost of risk. It follows from the expected utility axioms that the cost of risk, C, can be calculated as the difference between the expected value of the risky prospect and the associated certainty equivalent income:

$$U[\Omega - C] = \sum_{i=1}^{S} \theta_i \; U[W_i] = EU[W(t+1)]$$

It is now possible to expand $U[\Omega - C]$ in a Taylor series and estimate the cost of risk by manipulating the first and second order approximations.

More precisely, expanding the function $U\Omega - C]$ around Ω the first order approximation is:

$$U[\Omega - C] = U[\Omega] + U'[\Omega] \; (\Omega - C - \Omega) = U[\Omega] - U'[\Omega] \; C$$

Similarly, a second order approximation for the function $U[W(t+1)]$ can provide:

$$U[W(t+1)] = U[\Omega] - U'[\Omega] \; (W(t+1) - \Omega) + \frac{1}{2} U''[\Omega] \; (W(t+1) - \Omega)^2$$

$$\rightarrow \quad EU[W(t+1)] = U[\Omega] + \frac{1}{2} \; U''[\Omega] \; var[W(t+1)]$$

Using $U[\Omega\text{-}C]=EU[W(t+1)]$ and manipulating gives:

$$C = -\frac{U''[\Omega]}{2\ U'[\Omega]}\ var[W(t+1)] \quad \rightarrow \quad \frac{C}{W(t)} = -\frac{U''[\Omega]\ W(t)}{2\ U'[\Omega]}\ var[1 + R]$$

where $W(t+1) = W(t)\ (1 + R)$. This demonstrates theoretically that the cost of risk will vary across utility functions. This result also provides theoretical measures of the cost of risk. The measures of absolute risk aversion, $A[W] = -\{U''/U'\}$, and relative risk aversion, $R[W] = -\{U''\ W(t)\}/U' = W(t)\ A[W]$ are now textbook concepts, e.g., Elton and Gruber (1995). The risk tolerance function is defined to be $T[W] = (1 / A[W])$.

The notions of ***relative and absolute risk aversion***, $R[W]$ and $A[W]$, provide tools for evaluating the properties of various possible functional forms that could be used to model the $EU[W]$ function. Following Ingersoll (1987, p.39-40), the most widely used functions belong to the hyperbolic absolute risk aversion (HARA) or linear risk tolerance class. These functions can be represented, in general form, as ***the HARA utility function***:

$$U[W] = \frac{1 - \gamma}{\gamma}\left(\frac{\alpha\ W}{1 - \gamma} + b\right)^{\gamma} \qquad b > 0$$

where α, γ and b are parameters and the function is defined for $(b + \{\alpha W/(1 - \gamma)\}) > 0$. It can be verified by evaluating the first and second derivatives of the HARA function and solving for the risk tolerance function that $T[W] = (W/(1 - \gamma)) + (b/\alpha)$. It follows that the level of risk tolerance is linear with $T[W]$ increasing for $\gamma > 1$ and decreasing for $\gamma < 1$. Recalling that $T[W] = 1/A[W]$, the measure of absolute risk aversion will be increasing for $\gamma < 1$ and decreasing for $\gamma > 1$.

Specific functional forms for the HARA class can be derived by assuming values for the parameters of the HARA utility function. For example, the theoretically important quadratic utility function is specified using $\gamma = 2$. Setting $b=1/\alpha$ for convenience, evaluating the HARA function for this case gives:

$$U[W] = W - \frac{\alpha^2}{2}\ W^2 - \frac{1}{\alpha^2} \qquad \rightarrow \qquad U[W] = W - c\ W^2$$

It follows that the quadratic utility function possesses increasing absolute risk aversion ($A[W] = \{2c / (1 - 2cW)\} \rightarrow A'[W] = \{4c^2 / (1 - 2cW)^2\} > 0$)and increasing relative risk aversion ($R[W] = \{2cW / (1 - 2cW)\} \rightarrow R'[W] = \{2c / (1 - 2cW)^2\} > 0$). These properties of quadratic utility imply that investors will reduce both the dollar amount and the percentage amount invested in risky assets as wealth increases. This result is, seemingly, contrary to generally observed behavior: don't the richest people hold the bulk of risky assets?

The properties of the quadratic utility function create a potential quandary for modern portfolio theory. This is because quadratic utility functions provide a possible and expedient theoretical rationale for the use of ***mean-variance expected utility*** functions, a fundamental building block used to develop the Markowitz mean-variance portfolio optimization model and the capital asset pricing model. The other possible and expedient rationale is to assume that the relevant random variables, the security returns, are normally distributed. To see how quadratic utility leads to mean-variance expected utility, take expectations of the quadratic $U[W]$ and manipulate to get $EU[W] = E[W] + c\{var[W] + E[W]^2\}$. It follows that assuming quadratic utility leads to mean-variance expected utility. Given the theoretical importance of mean-variance models, at least since Mossin (1973) there have been efforts to rescue quadratic utility. For example, by allowing the parameter c to vary as wealth changes it is possible to specify a quadratic utility function that has, say constant absolute or relative risk aversion.

Though quadratic utility has a direct correspondence to mean-variance expected utility functions, it is a natural progression to consider the implications for asset pricing and portfolio selection models of assuming other possible specifications of the utility function. Three popular models are the log utility function, $U[W] = ln[W]$, the power utility or isoelastic utility function, $U[W] = \{W^\gamma / \gamma\}$ and the negative exponential utility function, $U[W] = -exp[-\alpha W]$. In the general HARA function: negative exponential utility is specified with $\gamma = -\infty$ and $b = 1$; power utility is specified with $\gamma < 1$ and $b = 0$; and log utility is specified with $\gamma = b = 0$ (with L'Hospital's rule being needed to make the requisite derivations in this case). It follows that the negative exponential utility function possesses constant absolute risk aversion and decreasing relative risk aversion; while both the power utility and log utility functions have decreasing absolute risk aversion and constant relative risk aversion. The presence of different utility functions with different risk aversion properties begs the question: which theoretical specification is most descriptive of investor behavior?

It is difficult to obtain empirical evidence that directly addresses the absolute and relative risk aversion properties of a 'representative investor'. The risk aversion measures are applicable to the behavior of a given investor at different wealth levels. Evidence on the holdings of risky assets for different investors at different wealth levels are only suggestive. For example, Blume and Friend (1975) examined Federal Reserve data on the financial holdings of consumers and found evidence in favor of constant relative risk aversion (decreasing absolute risk aversion). Recent research has found that risk aversion varies across individuals due to a range of factors such as age, income, even gender, e.g., Donkers et al (2001), Jianakoplos and Bernasek (1998). Estimates of aggregate risk aversion also vary over time, with the state of the market and where the investor is situated in their life cycle. Many studies employ complicated joint hypotheses to arrive at specific estimates of an aggregate risk aversion coefficient. For example, Hahm (1998) estimates the coefficient of relative risk aversion to be 1.25.

*C. Expected Utility and Moment Preference**

Given the theoretical importance of the mean-variance expected utility function, the relationship between moment preference and expected utility has received considerable academic attention. Important topics have included: the conditions under which mean-variance analysis is consistent with maximizing expected utility, e.g., Kroll, Levy and Markowitz (1984), Bell (1995); and, the implications of introducing skewness preference into the mean-variance framework, e.g., Kraus and Litzenberger (1976), Poitras and Heaney (1999). Brockett and Kahane (1992) among others, have shown that there is not a direct correspondence between the derivatives of the expected utility function and moments of the return distribution. The implication is that maximization of a function defined over moments, such as mean-variance or mean-variance-skewness, may not give the same solution as directly maximizing expected utility. Yet, Meyer (1987), Ormiston and Quiggin (1994) and others demonstrate that the conditions on the random variables sufficient for mean-variance rankings to provide solutions consistent with expected utility rankings are relatively weak.

As discussed in numerous sources, e.g., Loistl (1976), Levy and Markowitz (1979), Poitras and Heaney (1999), the relationship between expected utility and moment preference objective functions can be motivated using a Taylor series expansion (see sec. 5.2) of $U[W]$, the decision maker's utility function for wealth $U[W]$, evaluated at the expected value for terminal wealth $[\Omega]$ $(E[W(t+1)] = \Omega)$:

$$U[W(t+1)] = U[\Omega] + U'[\Omega](W(t+1) - \Omega) + \frac{U''[\Omega]}{2!}(W(t+1) - \Omega)^2$$

$$+ \frac{U'''[\Omega]}{3!}(W(t+1) - \Omega)^3 + \ldots..$$

Exploiting this type of expansion requires certain technical conditions be satisfied. For example, convergence of the series within the interval of interest is needed.[6] In addition, desirable properties for utility functions require: $U'[W] > 0$, non-satiation; $U''[W] < 0$, risk aversion; and, $U'''[W] > 0$, preference for positive skewness.

With relatively weak distributional restrictions, e.g., Hassett et al. (1985), the Taylor series representation of $U[W]$ can be transformed into an approximation for a general expected utility function based on the moments of the conditional distribution for $W(t+1)$. The relevant approximation is derived by taking conditional expectations at time t and ignoring terms associated with moments higher than the second, for a mean-variance approximation, and moments higher than the third, for a mean-variance-skewness approximation. Taking expectations for the mean-variance and mean-variance-skewness case gives:

$$EU_{MV}[W(t+1)] \equiv EU_{MV} = U[\Omega] + 0 + \frac{U''[\Omega]}{2!} var[W(t+1)]$$

$$= U[\Omega] - b\ var[W(t+1)]$$

$$EU_{MVS}[W(t+1)] \equiv EU_{MVS} = U[\Omega] + 0 + \frac{U''[\Omega]}{2!} var[W(t+1)] + \frac{U'''[\Omega]}{3!} skew[W(t+1)]$$

$$= U[\Omega] - b\ var[W(t+1)] + c\ skew[W(t+1)]$$

where $var[W(t+1)]$ is the variance of terminal wealth, $skew[W(t+1)]$ is the skewness or centralized third moment for terminal wealth. Restrictions imposed by assuming risk aversion and positive skewness preference permit the coefficients in EU_{MVS} to be immediately signed as $b, c > 0$. Further restrictions on b and c, as well as the admissible range of W, can be derived by taking further derivatives of the Taylor series expansion and invoking Jensen's inequality. Setting $c=0$ permits the mean-variance-skewness moment preference function to be reduced to the mean-variance function, EU_{MV}.

What are the implications of introducing this additional skewness term into the moment preference objective function? Currently, limited information is available comparing solutions from mean-variance and mean-variance-skewness approximations. Information about such comparisons would be relevant for a range of decision making situations, especially those involving skewness altering securities such as options and insurance. The few studies that do compare the mean-variance and mean-variance-skewness objective functions illustrate some confusion as to the implications of introducing skewness, e.g., Prakash et al. (1996), Horowitz (1998), Poitras and Heaney (1999). Studies examining the impact of skewness on the asset pricing and portfolio theory problems include Kraus and Litzenberger (1976), Sears and Trennepohl (1983), Lim (1989), Simaan (1993). Combining of securities into portfolios almost certainly reduces the skewness of the portfolio relative to the value weighted sum of the individual asset skewness values.

3.4 Financial Engineering with Derivative Securities

A. The Derivative Security Renaissance

The Securities Act (1933, as amended 2000) makes specific reference to both options and futures contracts written on securities. Extending the scope of the Securities Act to include derivative securities captured the growth and importance of these contracts. The substantive increase in the practical importance of derivative securities has been paralleled by a corresponding increase in the theoretical and empirical analysis of these contracts within modern Finance. To this point, modern Finance has been associated primarily with the elements of modern portfolio theory, i.e., the EMH, the CAPM, the Markowitz portfolio optimization model and related concepts. Yet, in addition to these developments, modern Finance has also been intensely concerned with the development of financial engineering. This aspect of modern Finance is sufficiently distinct from the security pricing and portfolio optimization aspect that there has been little loss in content from treating it separately for pedagogical purposes. However, the practical implications of derivative securities for security analysis and portfolio strategy cannot be overlooked and the subject has to be addressed at some point.

Derivative securities are contingent claim contracts that have payoffs that depend on the future value of some commodity, where the ‘commodity’ is broadly interpreted to include securities. Conventionally, the set of derivative securities is defined to be options, futures, forward and swap contracts, though there are also numerous hybrids that embed a contingent claim, e.g., convertible preferred stock, callable bonds. When added to a portfolio of securities, derivative securities can permit the investor to significantly alter the distribution of terminal wealth. For example, forming a portfolio that contains a common stock together with a put option on the stock produces a truncated payoff distribution for the portfolio that is bounded below. The payoff on this portfolio replicates the payoff of another portfolio that combines a call option on that stock with an investment in bonds. In other words, the availability of derivative securities opens up the practical possibility of strategy replication providing the basis for a range of new securities products that can be financially engineered. The practical implementation of these notions has created the hugely successful financial engineering industry.

The practical developments of the financial engineering industry have been accompanied by a corresponding revolution in modern Finance. This revolution has resulted in the development of a corpus of theory and empirical evidence that lies outside the stream of the developments associated with modern portfolio theory. This is possible because modern portfolio theory is concerned with the valuation of the underlying commodity while derivative security analysis is concerned with the valuation and application of contracts written on the commodity. Attempts to integrate the two streams are difficult because the notions of derivative security analysis are derived from a distinct approach to valuation. In particular, derivative securities are largely priced using absence-of-arbitrage, an approach that is concerned with the relationship between current prices. Modern portfolio theory, on the other hand, is concerned with the expected values and variances of prices that are not currently known.

To observe that derivative securities have revolutionized modern Finance is, at the same time, both an understatement and a misleading statement. The statement is misleading because revolutions in academics usually involve the removal of one body of theory and its replacement with another. In this case, modern Finance has been expanded to encompass a new body of theories and empirical results that are treated separately from the other body of theory and empirical results. In addition, in terms of intellectual history there has been a rough correspondence in the development of modern portfolio theory and derivative security pricing theory. Important contributors in the derivative pricing stream, such as Robert Merton and Fischer Black, were also important contributors to the modern portfolio stream. The practical successes that were facilitated by the derivative pricing theories of academics implicitly gave credence to the security pricing and portfolio management theories that were being advanced by the same group of academics.

Poitras (2002) refers to the Renaissance in derivative securities that can, roughly, be dated from the creation of the Chicago Board Options Exchange in 1973 and the introduction of a range of exchange and OTC traded derivative securities over the following decade. Historically, derivative securities trading, especially options trading, has been the subject of considerable criticism and legislative sanction due to the potential for speculative abuses. The last quarter of the 20th century is remarkable in the breadth and depth of derivative security trading. Securities markets both in the US and globally have embraced these new products. The financial engineering industry has become an important profit center for many of the largest firms in the securities industry. Within academic Finance, the virtues of derivative securities are expounded in introductory investment texts and advanced courses. The importance of financial engineering has permitted a proliferation of advanced graduate programs with titles such as Masters in Financial Engineering.

Yet, the Renaissance in derivative securities has had its blemishes. Due to a significant number of high profile and expensive losses, trading of derivative securities attracted considerable attention during the 1990s (e.g., Poitras 2002, ch.1). The list of companies involved is striking, as is the size of the losses. From Barings Bank to Gibson's Greetings to Sumitomo Corporation, from Long Term Capital Management to Proctor and Gamble to Orange County, individual firm losses ranging from hundreds of millions to billions of dollars have been reported. Such events induce a state of uneasiness among policy makers, corporate managers, investment professionals, even academics. While it is tempting to draw glib generalizations about the apparent misunderstanding of risk management practices, closer inspection reveals a decidedly more complicated situation. In some cases, the relevant lessons that could be learned cannot be convincingly determined, due to the veil of corporate secrecy surrounding specific events. In cases where the activities and motivations of the participants can be precisely determined, it seems that different debacles raise different types of quandaries. Upon closer inspection, it seems that some so-called debacles were not debacles at all.

Large losses associated with derivative security trading are not unique to the 1990s. Even though the largest losses in absolute terms have happened more recently, this is consistent with the increasing use, availability and complexity of derivative products. This has produced an evolution in the types of problems which are arising. Since the early 1970s, there has been a progressive relaxation in the US of restrictions on derivative security trading, many of which had originated in the anti-speculation atmosphere of the post-Depression era. In conjunction with this relaxation, there has been an almost bewildering expansion in the variety of derivative securities being traded, both on the OTC markets and on the futures and options exchanges. From financial commodities to energy to equities to currencies, it is difficult to keep track of the rapid progress which has been and is being made in the development and application of derivative securities.

B. The History of Portfolio Insurance

Despite the role the derivative securities have played in the evolution of modern Finance, this book is not about the pricing of derivative securities. This subject is covered to great depth in other sources, e.g., Poitras (2002). Rather practical applications of derivative securities to security analysis and investment strategy are of interest. In particular, derivative securities facilitate the implementation of risk management strategies that fall under the general title of portfolio insurance. The different forms of portfolio insurance illustrate the replication properties of derivative securities and how concepts from financial engineering can be used to guide portfolio allocation decisions. In other words, it is the 'investment strategy' implications of derivative securities that will be examined. The 'security analysis' of derivative securities would require discussion of pricing with arbitrage methods, a subject that would take the discussion too far afield.

Though portfolio insurance techniques were popularized during the 1980's, heuristic forms of portfolio insurance have been used for decades. For example, a form of portfolio insurance can be achieved with the systematic use of order placement strategies, such as stop-loss and limit orders which have been acceptable market practice at least since the 19th C. These types of trading dependent strategies suffer from the defect of being "path dependent", an undesirable property of insurance schemes. In addition to trading related techniques, option replication strategies using stock/bond combinations were also likely in use, though in the realm of proprietary management practices. These techniques also suffer from the defect of path dependence and, in the absence of 'Greek' information, would probably have been imprecise (Poitras 2002, ch.9). The application of option replication to specifying dynamically traded stock/bond portfolios was not of academic interest until much later, after the development of the Black-Scholes formula.

As for the history of insurance related financial products, some of the insurance schemes of the late 17th and 18th century did offer payouts based on specific outcomes associated with joint stock performance. Being introduced prior to the development of actuarial science, these insurance schemes were more like gambling than insurance. In more recent history, Benninga and Blume (1985) report the selling of insurance against investment losses in the UK as early as 1956. In the US, Gatto et al. (1980) report on portfolio insurance plans offered to individuals by both the

Harleysville Mutual Insurance Company and Prudential Insurance Company of America. Brennan and Schwartz (1987) observe that the Harleysville plan was the first without any element of mortality insurance. Academically, Brennan and Schwartz (1976) were the first to make the connection between the potential for integrating insurance and equity returns. Leland, O'Brien, Rubinstein and Associates were important proponents in the marketing of dynamically traded option replication strategies to institutional clients.

The explosion in the use of the various types of portfolio insurance techniques can be traced to the introduction of exchange trading in options. Liquid options markets made possible the implementation of numerous portfolio insurance strategies. Even more strategies were permitted with the development of futures and options markets for stock indices. Analytical contributions based on Black-Scholes resulted in further portfolio insurance strategies being introduced. Many "alternative paths to portfolio insurance" (Rubinstein 1985) were proposed and implemented. The widespread use of dynamically traded portfolio insurance techniques has been identified as an important contributing factor in the Oct. 1987 stock market "crash", e.g., Tosini (1988). Academic understanding of notions associated with portfolio insurance have expanded considerably since the early work by Leland (1980) and Rubinstein and Leland (1981). The 1987 "crash" provided a textbook illustration of the inadequacies of the academically inspired option replication strategies; sizable unexpected losses were experienced by investors holding what were expected to be "insured" portfolios.

One of the fundamentals driving institutions to use dynamic trading strategies was the absence of risk management products with maturities and other characteristics that captured the time profile of their particular risk exposures. Since the crash, an array of OTC and exchange traded risk management products have been introduced which greatly enhance the ability to implement path independent strategies. Included in the list of such new products would be: long dated exchange traded option products, such as LEAPS for individual stocks and longer dated index options and equity swaps. Despite these improvements, the bulk of contract liquidity on both the exchanges and OTC is still concentrated in short dated contracts (see Table 3-z). The relative absence of strict mark-to-market rules in OTC contracts provides a strong incentive to use short dated contracts.

An important element in the modern Renaissance in derivative securities was the emergence of trading in stock index futures (see Table 3-z). Sufficient liquidity in these futures contracts has facilitated the trading of futures options on these indexes. The first stock index futures contract, based on the Value Line Index, was introduced in Feb. 1982 on the KCBT. The most important stock index futures contract, the S&P 500 traded on the CME/IMM, was introduced shortly thereafter in April 1982. A raft of stock index futures contracts has appeared since that time, starting with the introduction of the NYSE Composite on the NYFE in May 1982 and the Major Market Index on the CBT in 1984. More recently, there has been the introduction of foreign indexes traded on US exchanges, such as the Nikkei 225 on the CME. This has been accompanied by the trading of domestic equity indexes on futures markets around the world, including markets in Japan, Hong Kong, Holland, Australia, England, France, Germany, Switzerland, and Canada. Another recent development has been the start of trading in the DJIA index futures in October 1997. The slow pace associated with the introduction of the DJIA was not due to a lack of interest in such a contract. On the contrary, perceiving considerable demand, the CBT had attempted to introduce a DJIA contract as early as July 1984. However, these plans were thwarted by Dow Jones and Company which initiated legal action to prevent trading of the contract. What ensued was a process lasting over a dozen years, ending with the CBT eventually introducing DJIA futures and options contracts.

INSERT TABLE 3-z Stock Index Futures and Option Prices
INSERT TABLE 3-zz NYSE Program Trading Report 11/14,2003

C. Dynamic Portfolio Insurance

The basic mechanics of "path independent" portfolio insurance can be isolated from the put-call parity arbitrage condition for a non-dividend paying stock: $S + P = C + X e^{-rt^*}$. Because the concern is portfolio insurance, S refers to the price of a portfolio of stocks (instead of an individual stock), X is the exercise price (strike price), t^* is the time to expiration measured as the fraction of a year remaining to expiration, P is the price of a put written on the portfolio with exercise price X and time to expiration t^*, C is the call price written on the portfolio with the same X and t^* as the put, and r is the riskless interest rate. Dividends have been ignored for simplicity of exposition. As stated, put-call parity provides two path independent insurance strategies. One strategy is $S + P$, buy puts against the portfolio. If S is an index portfolio, relevant exchange traded puts may be available. Another strategy is $C + X e^{-rt^*}$, buy calls and invest the remainder in appropriately dated bonds. Again, if the portfolio is an index portfolio, exchange traded calls may be available. One important advantage of this strategy is that transactions costs in bond markets are typically lower than transactions costs for stocks and the bond portfolio can be actively managed, e.g., by riding the yield curve, to earn potentially higher returns than the $S + P$ approach.

While the path independent strategies have some desirable features, there are some drawbacks. One disadvantage is the inability to accurately replicate insurance for portfolios that do not track an index for which there are traded options, i.e., the relevant portfolio options are not available. Constructing a portfolio of options using options on the individual stocks will be more expensive and there is the possibility that not all stocks will have traded options. Using index options as surrogates for the portfolio options eliminates the potential for gains from individual security selection. Combining index options with options on individual stocks raises the problem of finding the appropriate combination of these options to replicate the payout on the desired portfolio. Another disadvantage is that the maturity dates for options may not be long enough to match the portfolio' s investment horizon, i.e., there is insufficient "time invariance". This requires options positions to be rolled forward which is more expensive and has pricing risk.

To handle these types of problems, dynamic trading strategies have been developed that involve actively trading portfolios composed of stocks and bonds in order to replicate the payoff on an insured stock portfolio. Such strategies are intuitively appealing to large institutional investors such as pension funds and insurance companies that already hold stock/bond portfolios that are actively managed. These strategies can be illustrated by substituting the Black-Scholes formula into the put-call parity condition:

$$S + P = S\,N[d_1] - X e^{-rt^*} N[d_2] + X e^{-rt^*} = S\,N[d_1] + X e^{-rt^*} (1 - N[d_2]) = w_1 S + w_2 X e^{-rt^*}.$$

where *N[d]* is the cumulative normal distribution evaluated at d with d_1 and d_2 as specified in the Black-Scholes formula, e.g., Poitras (2002, p.441). The weights w_1 and w_2 indicate the proportions of the portfolio held in stock and bonds in order to achieve insurance with an exercise price of X and time to maturity of t^*. Unlike the portfolio optimization models, the weights here will not sum to one, as the relationship is derived to equate values on the rhs and lhs. The sum of the weights will be close to one but not equal to one unless the put value is zero.

From a practical perspective, it is important for the potential portfolio insurer to identify why dynamic replication strategies, i.e., strategies dynamically replicating a call option payoff using stock/bond positions, should be used. Related to this are subsidiary issues concerning how to replicate and when to replicate. In this vein, large fund managers would consider the liquidity needed to establish large enough positions using derivatives and whether there are suitable X and expiration dates available. For example, while a well-diversified fund (e.g., an index fund) could make use of options or futures written on the appropriate index, funds targeted at non-systematic risk are more likely to be obligated to use dynamic replication strategies. However, even a well-diversified fund may find that available expiration dates on traded derivatives are not long enough, i.e., sufficient "time convexity" cannot be achieved. Because the dynamic replication strategies can be designed to theoretically achieve almost any desired expiration date and exercise price, this provides another reason for the use of these strategies.

To illustrate the use of dynamic replication where dividends are paid on the portfolio, consider the creation of a synthetic put option for an index portfolio. Given that the dividend yield on the index is q, Hull (1987, p. 204) shows that the delta of a European put on the index is:

$$\Delta_P = \exp\{-qt^*\}\ [N[d_1] - 1] = \exp\{-qt^*\}\ (\Delta_c - 1)$$

where:

$$d_1 = \frac{\ln\{\frac{S}{X}\} + (r - q + \frac{\sigma^2}{2})t^*}{\sigma\sqrt{t^*}}$$

Assuming that $S = 300, X = 290, r = .09, q = .03, \sigma = 0.25$ and $t^* = .5$, evaluation of the delta of the put gives $\Delta = -0.322$. It follows that if dynamic replication of a put is being used that 32.2% of the index fund should be sold and invested in (riskfree) fixed income securities. From the properties of the put delta (Poitras 2002, p.486), as the value of the index fund drops, the delta of the put will become more negative, indicating that a larger proportion of the index fund has to be sold, i.e., a large fraction of the portfolio will be invested in fixed income securities. A similar result would hold where the value of the index was increasing. In this case the delta of the put would be less negative, indicating that fixed income securities should be sold to purchase more units of the index fund. In this case, the proportion of the portfolio invested in the index fund would increase.

INSERT Tables 3-x and 3-y Examples of Portfolio Insurance

There are a number of different methods of approaching the portfolio insurance problem. In addition to strategies already examined such as combining stocks with purchased puts or dynamically replicating such positions, the remaining strategies attempt to achieve the same objectives by using derivative positions as surrogates for stock ownership, on the presumption that there are execution advantages (e.g., greater liquidity and lower transactions costs) to using derivatives. For example, instead of owning stock, it is possible to form portfolios composed of purchased call options and fixed income securities. Another strategy would alter the dynamic replication strategies by substituting short futures positions for the stock sales required when the value of the stock position declines. However, if this approach is used, for example, to insure an index fund, it is important to recognize that the number of stock index futures to be shorted for a given long index fund will be different than in the case where the stock position is being sold directly (and invested in bonds). In addition, there is the mechanical problem of calculating the dollar equivalency hedge ratio for the futures and cash positions.

Table 3-x has a helpful tabular presentation of how various methods of portfolio insurance would perform to insurance a stock index across a range of index levels. The first table gives payoff for the path independent strategy, $S + P$. Comparing the distribution of S with $S + P$ reveals, in a simple tabluar form, what Bookstaber and Clark (1983) have examined in much more general detail using distributional plots. The addition of a put transforms the symmetric S distribution to a positively skewed $S + P$ distribution. In terms of the return distribution, the additional cost of the put will result is a lower mean value for the $S + P$ distribution. The second table gives the results from applying the dynamic replication portfolio insurance strategy derived from $S + P = S\,N[d_1] + X\,e^{-rt^*}(1 - N[d_2])$.

In practice, dynamic replication faces substantive implementation issues. Trading cannot be conducted under the perfect markets, continuous trading assumptions required for the Black-Scholes formula to capture the price of the option. Nevertheless, assuming that the Black-Scholes assumptions apply, permits the decomposition of $S + P$ into the exact holdings of stocks and bonds to hold in order to precisely replicate the $S + P$ payoff. As the stock index level falls away from $S = 100$, the stock index position will be continuously reduced to the point where the stock position is nearly zero at $S = 59.87$. Similarly, as the stock index rises, the bond is sold to the point where at $S = 162.89$, there are no funds left in bonds. From this, it is apparent how dynamic portfolio insurance strategies, if applied by a large enough fraction of market traders, would amplify market movements.

In practice, dynamic trading strategies have to deal with the realities of discrete trading. Rules have to be determined about how large a movement in S is required before the rebalancing decision is executed. There are a number of possible methods of specifying a rebalancing trigger value. The example in Table 3-y assumes that the trigger value is 5%. From this point, the tabular presentation method can only provide an accurate picture of the distribution of weights, and the associated impact on portfolio value. For example, upside movements of S will produce increasing weights for S which lag the continuously rebalanced weights, resulting in a slight reduction in portfolio value. A similar result happens for downside movements of S where the reduction in S weights lacks the continuously rebalanced weights, again resulting in a slight reduction in portfolio value. Hence, the simple introduction of discrete rebalancing results in a deterioration of the performance of the dynamic replication strategy.

As it turns out, the discretely rebalanced case has considerably more complications than can be captured in one table. Being ***path dependent***, the terminal portfolio value can take a range of values, depending on the particular time path realized by S. For the ***path independent*** cases, $S + P$ and continuous rebalancing, the distribution of portfolio value can be determined precisely because the terminal portfolio value does not depend on the particular time path realized by S. This does not happen with discrete rebalancing. For example, a price path which starts at 100 and goes to 95 generates a rebalancing involving a sale of S to produce a weight change of .638 to .495. If the next step is back to 100, the rebalancing involves the weight returning from .495 to .638. The resulting portfolio value will now be less than a portfolio value along a price path where S was unchanged and no rebalancing happened.

Table 3-y also provides results for the case where a stock/bond portfolio is created using the dynamic portfolio weights but no rebalancing is done along the time path. This is another type of path independent strategy. Though not immediately apparent from the tabular presentation, the distribution of the portfolio value for the no rebalancing case is not unlike the S distribution. Unlike the dynamically traded portfolio, the no rebalancing distribution retains the symmetric shape of the S distribution, though there is less dispersion due to the presence of a long investment in the riskless asset.[7] In all of this, the no rebalancing and static portfolio insurance cases are being unfavorly compared with the discrete rebalancing case because there are no transactions costs factored into the various calculations. In the limit, i.e., continuous rebalancing with transactions costs, the dynamic strategies may produce infinite losses. In practice, there will be a tradeoff between rebalancing frequency, the various transactions and execution changes and the terminal value of the insured portfolio. Wider rebalancing frequencies will permit greater deviation from the path independent static portfolio insurance case, but this loss of precision will be balanced out with a savings in transactions costs due to reduced trading frequency.

As discussed in Poitras (2002, p.523-4), there is nothing unique about a portfolio of domestic stocks. The notions of portfolio insurance can be applied to any commodity. One useful extension involves insuring the domestic currency value of a foreign bond position. Much as with dynamic portfolio insurance for stocks, dynamic portfolio insurance for foreign bonds can be derived using put-call parity for currency options. The objective is to dynamically trade a portfolio composed of domestic bonds and foreign bonds in order to achieve the same payout as a path independent portfolio composed of a foreign bond plus a currency put option. If the exchange rate increases, the value of the domestic currency rises relative to foreign currency, then the dynamic strategy involves selling foreign bonds and buying domestic bonds. If the exchange rate deteriorates, the domestic bond is sold in favour of buying the foreign bond. As before, the Black-Scholes formula for a call can be substituted into the put-call parity condition to derive the appropriate portfolio weights.

To see this consider the path independent value of a portfolio which contains a foreign currency bond which has the domestic currency value protected with a currency put option. The associated dynamic replication portfolio can new be derived:

$$V = S \exp\{-r_f t^*\} + P = C + X \exp\{-r t^*\}$$

$$= S \exp\{-r_f t^*\} N[d_1] - X \exp\{-r t^*\} N[d_2] + X \exp\{-r t^*\}$$

$$= S \exp\{-r_f t^*\} N[d_1] + X \exp\{-r t^*\} (1 - N[d_2])$$

In this formulation, $S\ exp\{-r_f t^*\}$ is the domestic currency value of the foreign bond position and $X\ exp\{-rt^*\}$ is the domestic currency value of the domestic bond. More precisely, $exp\{-r_f t^*\}$ is the foreign currency value of a continuously compounded zero coupon bond which matures to one unit of domestic currency; multiplying this foreign bond price by S, the time t spot exchange rate expressed in domestic direct terms, converts the foreign bond price to units of domestic currency. Similarly, $X\ exp\{-rt^*\}$ is the domestic currency value of a continuously compounded zero coupon bond which matures to the number of units of domestic currency reflected in the exercise price for the currency option.

As with portfolio insurance for stocks, portfolio insurance for foreign bonds involves dynamic trading of the position. When the value of domestic currency rises relative to foreign currency, S will fall and the dynamic strategy requires selling a portion of the foreign bond and using the funds to purchase domestic bonds. The dynamic replication formulation identifies the precise amount of foreign bonds which have to be sold in order to maintain the same payout as the path independent portfolio $[S\ exp\{-r_f t^*\} + P]$. While much the same theoretically, dynamic replication for foreign bonds can differ in practice. Unlike the dynamic strategies for stock portfolios, which can suffer from inaccurate replication due to illiquidity in the underlying stocks, the cash markets involved in the dynamic insurance for foreign bonds are typically liquidity. The foreign exchange market, as well as the domestic and foreign bond markets, are unlikely to be subject to the types of pricing discontinuities which precipitated the October 1987 market break.

QUESTIONS

1.* a) For the simple expected utility example of Sec. 3.3.A, solve for the optimal investment in the risky asset ($A(0)^*$) when the utility function has the quadratic form: $U[W(1)] = W(1) - b\ W(1)^2$. Demonstrate that $A(0)$ decreases with the amount of initial wealth $W(0)$, i.e., the greater the initial wealth level, the smaller is the investment in the risky asset.
b) Evaluate the second derivative of the general *EU* function and derive the conditions required for solution to be a maximum.

2.* By evaluating the first derivatives for $A[W]$ and $R[W]$, solve for the relative and absolute risk aversion properties of the log utility, negative exponential and power utility functions.

3. To see how to use this formula to evaluate the variance of portfolio return, consider the following information given about three securities, A, B and C:

Stock	Expected Return	Standard Dev.	Correlation Coefficients		
			A	B	C
A	14%	12%	1.00	0	.20
B	16%	15%	0	1.00	.60
C	12%	10%	.2	.60	1.00

(a) If the market value of a portfolio is compsed of 30% of A and 70% of C, what is the expected return and standard deviation of the portfolio?

(b) If the portfolio is 21% in A, 30% in B and 49% in C what is the expected return and standard deviation of the portfolio?

(c) If a portfolio contains only A and C, what combination of A and C has the lowest possible standard deviation? What is the expected return and standard deviation on this portfolio?

4. Create a numerical example applicable to the *trente demoiselles* to illustrate how the basic notions of portfolio diversification apply to this case. To construct this example define the risk associated with one individual and carry this through to the case of a portfolio of life annuities written on the lives of thirty individuals. This will involve decomposing the basic risk into parts and assessing the impact of increasing the number of lives affecting the portfolio. (Hint: see the discussion of diversification in sec. 3.1).

5. Given the following information about four risky securities:

$$\sigma_1 = .2 \qquad \sigma_2 = .2 \qquad \sigma_{1,2} = 0 \qquad E[R_1] = .05 \qquad E[R_2] = .15$$

$$\sigma_3^2 = .04 = \sigma_4^2 \qquad \sigma_{1,3} = \sigma_{1,4} = \sigma_{2,3} = \sigma_{2,4} = \sigma_{3,4} = 0 \qquad E[R_3] = .20 = E[R_4]$$

a) Derive the value weights (w_1, w_2, w_3, w_4), mean and variance of the minimum variance portfolio which combines these four securities.

b) How do the value weights (w_1, w_2, w_3, w_4), mean and variance of the minimum variance portfolio change if $\rho_{1,2} = \rho_{1,3} = \rho_{2,3} = ... = \rho_{3,4} = .5$?

6. Assume the standard Capital Asset Pricing Model is true, e.g., there is unrestricted lending or borrowing at the same riskfree rate of interest. From your stock broker your are able to obtain the following information about two stocks:

	Expected Return	Standard Deviation	Correlation with the Market Portfolio
Stock (A)	.18	.30	.4
Stock (B)	.30	.60	.75

The **variance** of the return on the market is .16

a) Compute the β's for each security? (5 points)
b) What is the expected return on the market portfolio? (5 points)
c) What are the equations for the Capital Market Line and the Security Market Line? (10 points)

7. Assuming the market model holds, explain why unsystematic risk is diversifiable while systematic

risk is not diversifiable. Be sure to identify the relevant assumptions which are being made about the error terms and where these assumptions are used in your derivation. What is the standard deviation of a fully diversified portfolio?

8. a) A long stock position can be "protected" by buying a put. How can the payoff on this portfolio of a stock and option be replicated using "dynamic hedging" strategies involving portfolios which combine only stock and bond positions?

b) Describe the various forms of portfolio insurance. How would these various forms of portfolio insurance perform in the face of discontinuous movements in equity prices such as the July 2002 market break?

NOTES

1 The σ_{ij} term is interpreted as being a covariance when $i \neq j$ and as a variance when $i = j$. Because the basic optimization problem is quadratic, it follows that the optimal solutions will take the form of an ellipse or a parabola. Consider the case where the $\{w_i\}$ are restricted to be non-negative, then the solution will be an ellipse. At any given target level of expected return, there will be two values of σ which solve the optimization problem. In evaluating the solutions, it is conventional to ignore the optimal solution which has the higher level of σ and consider only the portfolios which have the lowest σ.

2. Markowitz (2000) reviews the historical development of the model.

3. Such is the reason for using moving sampling windows instead of using all the data available. For example, 100 years of monthly data produces estimates of the arithmetic average which would not be affected by an additional observation. Hence, the optimal weights would not change over time.

4. The general approach in the following discussion is adapted from Fama (1976).

5. The axiomatic approach to choice under uncertainty has produced a considerable number of studies. Accessible and brief overviews are available in various sources, e.g., Henderson and Quandt (1980, Sec. 3.8), Layard and Walters (1979, ch.13), Ingersoll (1987, ch.1) Duffie (1988, ch.1, sec. 5). A more advanced and complete treatments is available in Fishburn (1982). The following discussion presumes that the student has already had an introductory exposure to expected utility theory.

6 Further discussion of issues related to the general properties of a Taylor series expansion for approximating a general expected utility function can be found in Loistl (1976). Hassett et al. (1985) examine specific types of problems with the Taylor series which arise where skewness is involved. Brockett and Kahane (1992) discuss the connection between preference for moments and expected utility rankings of risky prospects, arguing that $U'' < 0$ and $U''' > 0$ are not related to variance

avoidance or skewness preference. Poitras and Heaney (1999) illustrate a theoretical difficulty that can arise when optimizing a mean-variance-skewness objective function.

7. This observation provides a window into the various complications that non-linear payoffs, such as options, can have for mean-variance optimization analysis.