Lecture 0: Background and Statistics

1. The Behavior of Stock Returns
   - introduction to probability distributions and basic statistics
   - definition of **returns** and compounding
2. Portfolio Mathematics
3. Return Distributions

- random variables: observations governed by a probability distribution.
  - an unknown outcome.
  - denoted with a tilde above it.

Properties/Moments/Parameters of distributions

1. The mean or expected value:
   (a) discrete random variable
   If $\tilde{X}$ takes the value X_i with probability p_i :
   $E(\tilde{X}) = \sum_{i=1}^{n} p_i X_i$

   (b) continuous random variable
   $$E(\tilde{X}) = \int_X p(X) X dx,$$
   where p(x) is the probability density function for $\tilde{X}$.
2. The standard deviation:
   (a) discrete random variable
   $\sigma^2(\tilde{X}) = E([\tilde{X} - E(\tilde{X})]^2) = \sum_{i=1}^{n} p_i (X_i - E(\tilde{X}))^2$
   $\mathrm{Std}(\tilde{X}) = \sigma(\tilde{X}) = \sqrt{\sigma^2(\tilde{X})}$
3. continuous random variable

$$\sigma^2(\tilde{X}) = E([\tilde{X} - E(\tilde{X})]^2) = \int_X p(X)(X - E(\tilde{X}))^2 dx$$
$$\text{Std}(\tilde{X}) = \sigma(\tilde{X}) = \sqrt{\sigma^2(\tilde{X})}$$

- measures of dispersion of r.v. $\tilde{X}$. Measures average variability around successive random drawings from the distribution of $\tilde{X}$ about the mean of the distribution.

4. The Normal distribution:

   Probability that any random draw from a normally distributed r.v. $\tilde{x}$ is within one standard deviation of the mean is .6826.

   $$P(E(\tilde{x}) - \sigma(\tilde{x}) <= \tilde{x} <= E(\tilde{x}) + \sigma(\tilde{x}) = 0.6826$$

   and within two standard deviations is,

   $$P(E(\tilde{x}) - 2\sigma(\tilde{x}) <= \tilde{x} <= E(\tilde{x}) + 2\sigma(\tilde{x}) = 0.9550$$

   Or, we can transform an normally distributed r.v. into units of standard deviation from its mean,

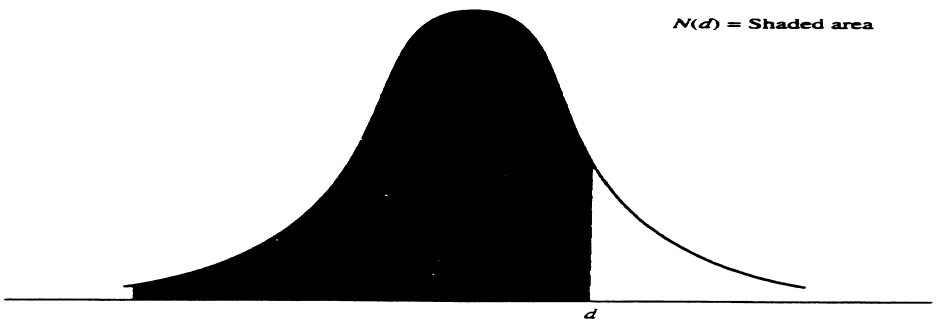
   $$\tilde{r} = \frac{\tilde{x} - E(\tilde{x})}{\sigma(\tilde{x})}$$

   This will follow the unit or "standard" normal distribution, which has mean of 0 and standard deviation of 1.

Since we know the distribution of $\tilde{r}$, then all we need to know is the mean and standard deviation of $\tilde{x}$.

$\Rightarrow$ this is why the normal distribution is completely characterized by its mean and standard deviation.

And, it is why we will use this to characterize returns on portfolios (i.e., it makes the problem simple).



5. The *Sample* Mean and Standard Deviation

- Means and standard deviations are never truly known, but must be estimated from samples of data.
- We want the population means and standard deviations, but must use sample means and standard deviations.

  ex: monthly returns on IBM (suppose are normally distributed)
- Sample mean $= \bar{x} = \sum_{t=1}^{T} x_t/T$
  where T is the sample size. This is a simple average.

- Here, each observation in the sample is weighted equally by the *frequency* of the observations. Previously, the outcomes for $\tilde{x}$ were weighted differently by their *probabilities*.
- Note: we are not assuming each time period (or month) is equally likely, but rather that sample relative frequencies *approximate* population probabilities.
- Sample variance $= s^2(x) = \sum_{t=1}^{T} \frac{(x_t - \bar{x})^2}{T-1}$
  Dividing by $T-1$ is just an adjustment (for degrees of freedom) to generate an unbiased estimate of the true population variance (discussed more later).

Since we also do not know the true distribution of returns, . . .

6. Testing for normality

- The studentized range:

$$SR = \frac{Max(x_t) - Min(x_t)}{s(x)}$$

  = range measured in units of standard deviation.
- why is this a useful statistic for the assumption of normality?
- what types of distributions would give rise to unusually high studentized ranges?

Since samples of data from a true population differ from one another, there is variation across samples that induces variation or a probability distribution across sample statistics themselves! (including the sample mean, variance, and studentize range).

Thus, these statistics are random variables themselves.

Therefore, in empirical research, inferences are made with uncertainty. We never prove a hypothesis to be true or false with certainty. We can only accept or reject a hypothesis with some degree of confidence, which we usually summarize with a probability distribution about the sample estimates.

Finance theory says:

- average returns over long periods of time are determined by risk

Two fundamental questions of finance:

- what is risk and how should we measure it?
- how much extra return do we need to be compensated for the additional risk?
  - is the 6% equity premium enough? is it too much?

In fact, economists have had trouble justifying a premium as big as 6% on the basis of risk:

- This is termed the *equity premium puzzle*.
- We will look at what determines different risks and expected returns *across* different securities, assets, regions (the "Cross-Section of Returns and Risk")
- We will look at what determines the changes in risk and expected returns over time for a given asset or security (the "Time-Series of Returns and Risk")

  Both are fundamental questions that need answers to understand why prices are what they are and why the move.

Both are fundamental in determining how to invest and applying finance theory to practice (not only investment practice).

Portfolio Mathematics

Portfolio weights:

- The portfolio weight for stock j, denoted w_j, is the fraction of a portfolio's wealth held in stock j:

$$w_j = \frac{\$ \text{ held in stock } j}{\$ \text{ value of the portfolio}}$$

- By definition, the portfolio weights must sum to 1.
- $\sum_{j=1}^{N} w_j = 1$ or in matrix notation $\mathbf{1}'\mathbf{W} = 1$, where $\mathbf{1}$ is a vector of ones and $\mathbf{W}$ is the $(N \times 1)$ vector of portfolio weights.

Portfolio returns:

- For N stocks, $\tilde{R}_p = \sum_{j=1}^{N} w_j \tilde{r}_j = W'R$.

Since the return on a portfolio is a weighted sum of the returns on the securities in the portfolio, we need to determine how the distribution of a weighted sum of r.v.'s is related to the distributions of the individual summands.

We can simplify this analysis by assuming that returns are approximately normal. This means we only have to worry about the mean and variance.

Portfolio *expected* **returns:**

- The expected return of a portfolio is the portfolio-weighted average of the expected returns of the individual stocks in the portfolio,

$$\begin{aligned} E[\tilde{R}_p] &= \sum_{j=1}^{N} E[w_j \tilde{r}_j] = E[w_1 \tilde{r}_1 + \ldots + w_N \tilde{r}_N] \\ &= w_1 E[\tilde{r}_1] + \ldots + w_N E[\tilde{r}_N] = \sum_{j=1}^{N} w_j E[\tilde{r}_j] \\ &= W' \mu \end{aligned}$$

This comes from the fact,

$$E[\alpha \tilde{x}] = \alpha E[\tilde{x}]$$

(prove yourself)

Variances and standard deviations:

Start with the following fact,

$$\begin{aligned} \sigma^2(\alpha \tilde{x}) &= \alpha^2 \sigma^2(\tilde{x}) \\ \sigma(\alpha \tilde{x}) &= |\alpha| \sigma(\tilde{x}) \end{aligned}$$

(prove yourself)

$$\sigma^2(\tilde{R}_p) = E\{[\tilde{R}_p - E(\tilde{R}_p)]^2\}$$

for n = 2,

$$= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_{12}$$

Covariances and Correlations:

- Measures the degree to which two variables (stocks) move together.

  If > 0, then the two stocks move together, if < 0, then the two stocks move in opposite directions.

- The **covariance** between two stock returns is:

  $$cov(\tilde{r}_1, \tilde{r}_2) = \sigma_{12} = E[(\tilde{r}_1 - E[\tilde{r}_1])(\tilde{r}_2 - E[\tilde{r}_2])].$$

- To compute covariances, need to know something about the *joint distribution* of returns.

- Variance is a special case of the covariance: $cov(\tilde{r}_i, \tilde{r}_i) = var(\tilde{r}_i)$.

- Covariance depends on the units of measurement, while **correlation** does not.

- **Correlation** between two returns is the covariance divided by the product of their standard deviations.

  $$corr(\tilde{r}_1, \tilde{r}_2) = \rho_{12} = \frac{cov(\tilde{r}_1, \tilde{r}_2)}{stdev_1 stdev_2} = \frac{\sigma_{12}}{\sigma_1\sigma_2}.$$

the correlation is a *scaled* covariance.

this means that all correlations are $-1 <= \rho <= 1$.

We now have everything we need to analyze portfolios and the risk-return tradeoff of investments.

Portfolio Variances and Diversification:

- Consider a two-stock portfolio:

$$\begin{aligned} var(\tilde{R}_p) &= var(w_1\tilde{r}_1 + w_2\tilde{r}_2) \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_{12} \\ &= (\text{weight}_1)^2 var_1 + (\text{weight}_2)^2 var_2 \\ &\quad +2\text{weight}_1\text{weight}_2 cov_{12} \end{aligned}$$

- We can substitute correlations for covariances to express the variance of a portfolio:

$$var(\tilde{R}_p) = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2$$

- We can use this equation to illustrate the principle of diversification:

  ex: suppose both stock variances are 0.04 and we invest equally in stocks 1 and 2. Then, the variance of this portfolio is,

$$\begin{aligned} var(\tilde{R}_p) &= (0.5)^2(0.04) + (0.5)^2(0.04) \\ &+ 2(0.5)(0.5)\rho_{12}\sqrt{0.04}\sqrt{0.04} \\ &= 0.02 + 0.02\rho_{12} \end{aligned}$$

- KEY RESULT:
  - This illustrates a key point: the variance of the resulting portfolio will always be less than the variances of each of the two individual stocks (which are 0.04 in this case), so long as the correlation, ρ_{12} between the two stocks is less than 1 (i.e., so long as they are not *perfectly positively correlated*).
  - Also, the lower the correlation, the lower the variance of the resulting portfolio.

- Now consider a many stock portfolio:

$$\begin{aligned} Var(\tilde{R}_p) = \sigma_p^2 &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} \\ &= \underset{(1\times N)(N\times N)(N\times 1)}{W'VW} \end{aligned}$$

$\Rightarrow$ there are N variance terms and $N^2 - N$ covariance terms.

for the variance of a portfolio of 100 stocks, this means there are 100 variance terms and 9,900 covariance

terms!
(which do you think is more important for determining the variance of a portfolio? The variance or the covariance of each individual stock?)

- Computing the covariance between two portfolios (A and B), each containing N stocks:

$$\sigma_{AB} = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i^A w_j^B \sigma_{ij} = W_A' V W_B.$$

where w_i^A are the weights portfolio A assigns to the N stocks, and w_j^B are the weights portfolio B assigns to the N stocks.

*ex: Consider two portfolios of the same three stocks. Portfolio A is an equal-weighted portfolio of stocks 1, 2, and 3; Portfolio B places half its weight on stock 1 and half on stock 3. Find the covariance between the two portfolios if the Variance-Covariance matrix of the three stocks is

	1	2	3
1	0.04	0.002	0.005
2	0.002	0.05	0.003
3	0.005	0.003	0.06

$$\begin{aligned}\sigma_{AB} &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i^A w_j^B \sigma_{ij} = W_A' V W_B \\ &= 0.0192\end{aligned}$$

- We also define

$$\beta_{iP} = \frac{cov(\tilde{r}_i, \tilde{R}_p)}{\sigma^2(\tilde{R}_p)}$$

This is the risk of security i in portfolio p relative to the risk of portfolio p.

The Distribution of Stock Returns

- Important for determining underlying risks.
- Useful for evaluating and analyzing investments.
- Bachelier (1900) was first work, although Osborne (1959) got first credit.
  - began by assuming prices changes are random drawings from the same distribution.
  - this means successive price changes are independent and identically distributed (i.i.d.).
  - assume transactions are evenly spread across time (days, weeks, months) and that there are many of them.

    $\Rightarrow$ distribution of sum of iid drawings approaches a normal distribution.

We formulate things in terms of returns rather than price changes. (why?)

What units are returns in? What else do we need to know?

Whereas prices changes for a month depend on the *sum* of intermediate daily price changes, the return for a month depends on the *product* of intermediate daily returns (because of compounding).

$$1 + R_{it} = (1 + r_1)(1 + r_2)...(1 + r_T).$$

Instead of assuming successive price changes are iid, suppose successive values of $\tilde{r}_t$ are iid.

Then successive values of $ln(1 + r_t)$ are also iid.

This implies,

$$\begin{aligned} ln(1 + R_{it}) &= ln[(1 + r_1)(1 + r_2)...(1 + r_T)] \\ &= ln(1 + r_1) + ln(1 + r_2) + ... + ln(1 + r_T). \end{aligned}$$

Thus, if T is large, $ln(1 + r_t)$ will be approximately normally distributed.

$ln(1 + r_t)$ is the rate of return assuming continuous compounding for the period t.

Converts multiplicative process of compounding into an additive one. Advantage–statistical time-series properties much easier to deal with.

If returns are iid normal, then compounded returns cannot be normal, since they are multiplicative. Thus, cross-sectionally, we often assume *simple* returns are normal (since portfolios are sums of individual stocks at a point in time), but time-series wise, we assume that *continuously compounded* returns are iid normal.

This implies that simple returns are iid *lognormal*,

$$\begin{aligned} ln(1 + r_{it}) &\sim N(\mu_i, \sigma_i^2) \\ &\Rightarrow \\ E[r_{it}] &= e^{\mu_i + \sigma_i^2/2} - 1 \end{aligned}$$

$$\sigma^2(r_{it}) = e^{2\mu_i+\sigma_i^2}[e^{\sigma_i^2} - 1]$$

Added advantage: Does not violate limited liability.

Testing the Normality Assumption in Returns

1. Daily returns – Fama (1965).

- use studentized range.
- calculate likelihood under a normal distribution that a daily return would be 4 σ from its mean (about 1 in 50 years). Compare to actual data (about 4 in 5 years!).
- Skewness (scaled third moment)

$$skew = E\left[\frac{\tilde{r}_t - E[\tilde{r}_t]}{\sigma(\tilde{r}_t)}\right]^3$$

- Calculate sample skewness

$$\hat{s}kew = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\tilde{r}_t - \bar{r}}{s(\tilde{r}_t)}\right)^3$$

- Kurtosis (scaled fourth moment)

$$kurt = E\left[\frac{\tilde{r}_t - E[\tilde{r}_t]}{\sigma(\tilde{r}_t)}\right]^4$$

- Calculate sample kurtosis

$$\hat{kurt} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\tilde{r}_t - \bar{r}}{s(\tilde{r}_t)}\right)^4$$

Portfolio returns generally have negative skewness and kurtosis above 3. (i.e., negatively skewed and fatter tails than the normal distribution).

2. Monthly returns (better).

   For ease and without loss of generality (relative to the stable class of distributions), we will assume monthly returns are distributed normally.

Multivariate Normal Distribution of Returns

Let $\tilde{y}_1, ..., \tilde{y}_n$ be continuous jointly distributed r.v.'s with joint density function $f(y_1, ..., y_n)$.

If any linear combination of the y_i has a normal distribution, then the joint distribution of $\tilde{y}_1, ..., \tilde{y}_n$ is multivariate normal and vice versa.

Some properties:

1. Just like univariate normal, multivariate normal is defined completely by the mean and variance. Only here it is means of all n variables and covariance matrix of the y_n's.

2. Independence implies zero covariance. But, under multivariate normal, zero covariance also implies independence.

$$\begin{aligned}
if \sigma_{ij} &= 0 \forall i, j, \\
f(r_{1t}, ..., r_{nt}) &= f(r_{1t}) f(r_{wt}) ... f(r_{nt}) \\
\Rightarrow f(r_{it} | r_{1t}, ..., r_{i-1t}, r_{i+1t}, ..., r_{nt}) &= f(r_{it})
\end{aligned}$$

This means the conditional distribution on stock return i is the same for all combinations of the returns on other securities. That is, the conditional distribution = the marginal distribution.

3. Each r_{it} has a univariate normal distribution.

Lecture 1: Portfolio Theory and the CAPM

This lecture note covers:

- a framework for making asset allocation decisions among many assets.
- the efficient set mathematics.
- the Capital Asset Pricing Model.

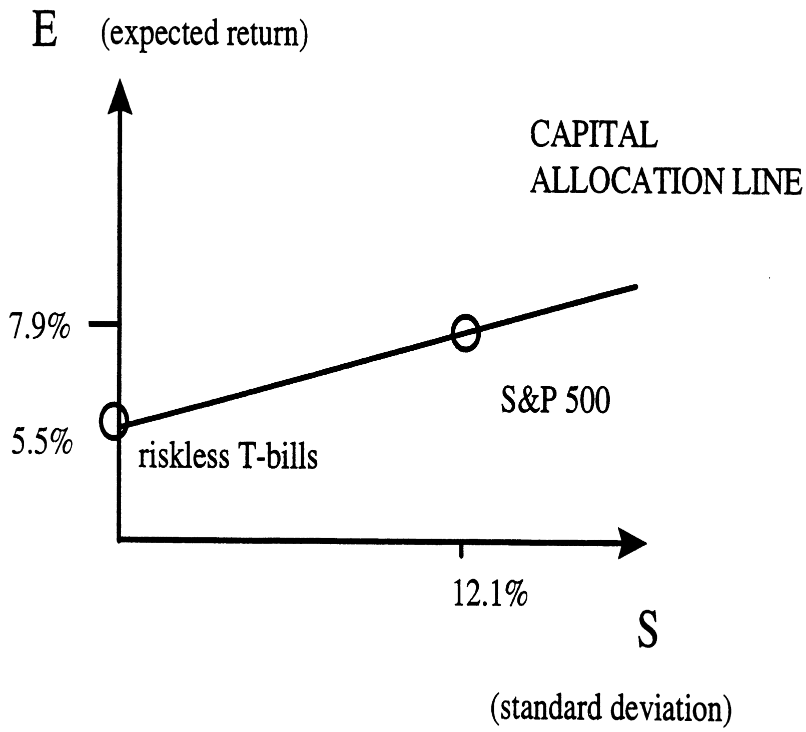
The Basic Allocation Problem

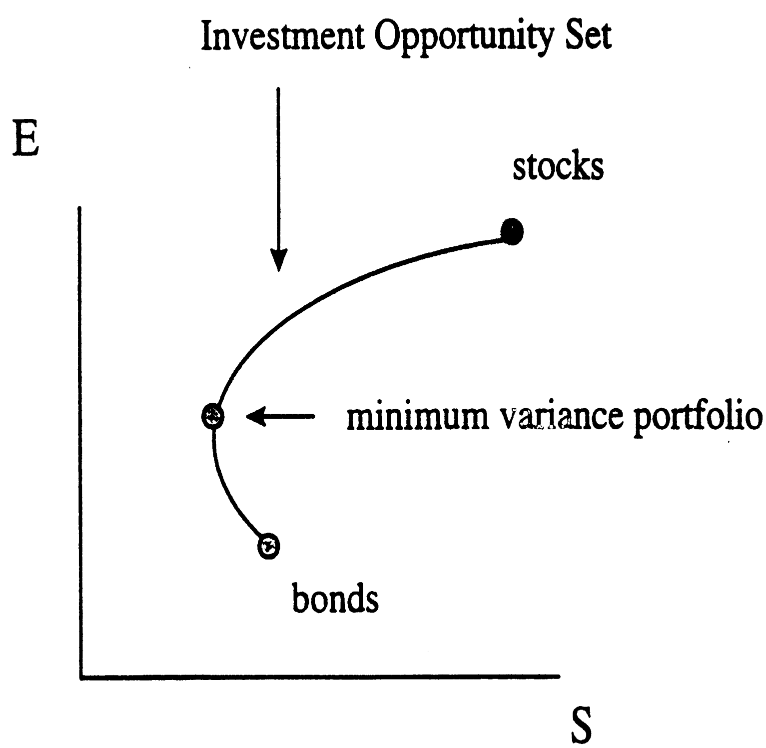
- You are wondering how to split your money among a host of available assets

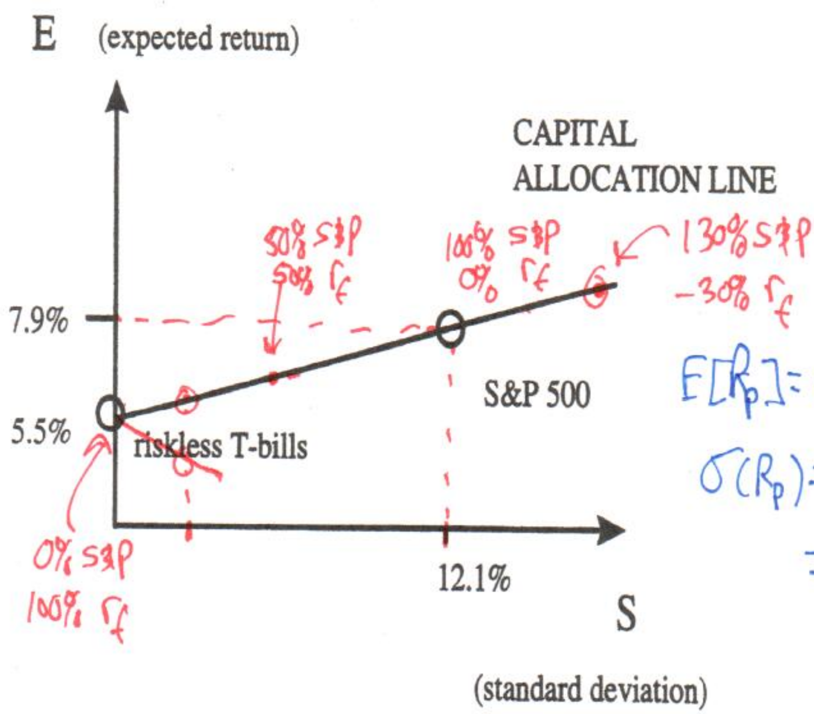
Question: How should we choose between all these investments?

- if returns are distributed under a multivariate normal, then all we care about are means and variances.
- Consider the following mean (E) and standard deviation (S) diagram of all possible assets and combinations of assets.

  1. one risky and one riskless asset.
  2. two risky assets.
  3. many risky assets.

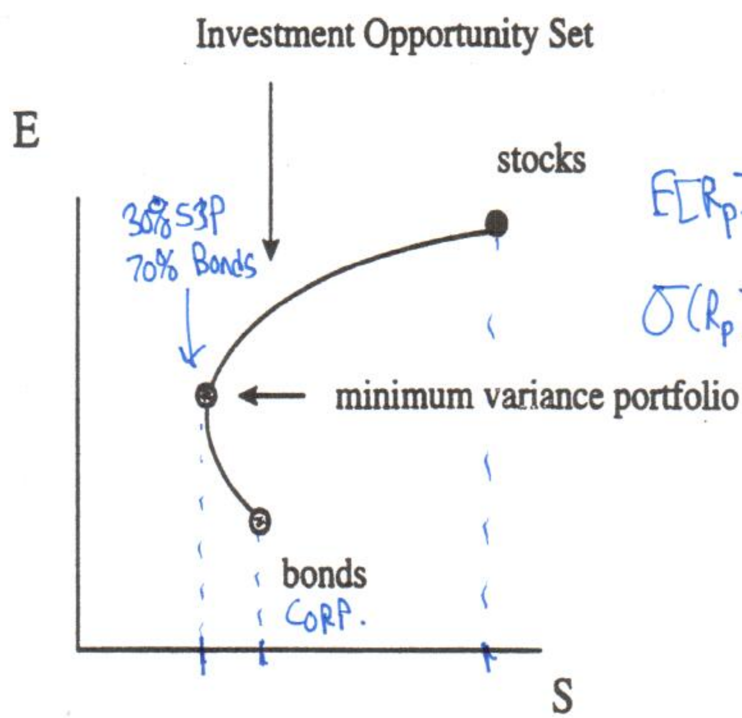

E (expected return)
CAPITAL
ALLOCATION LINE
7.9%
5.5%
S&P 500
riskless T-bills
12.1%
S
(standard deviation)


Investment Opportunity Set
E
stocks
minimum variance portfolio
bonds
S



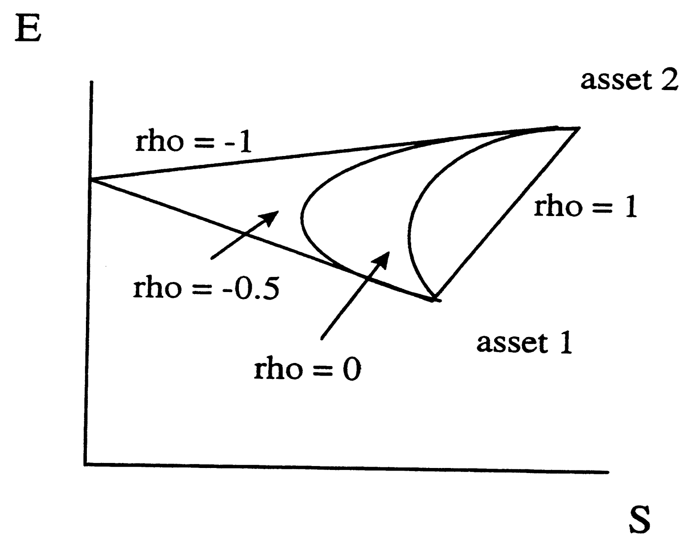
$$E[R_p] = w\,E[r_{S\&P}] + (1-w)\,r_f$$

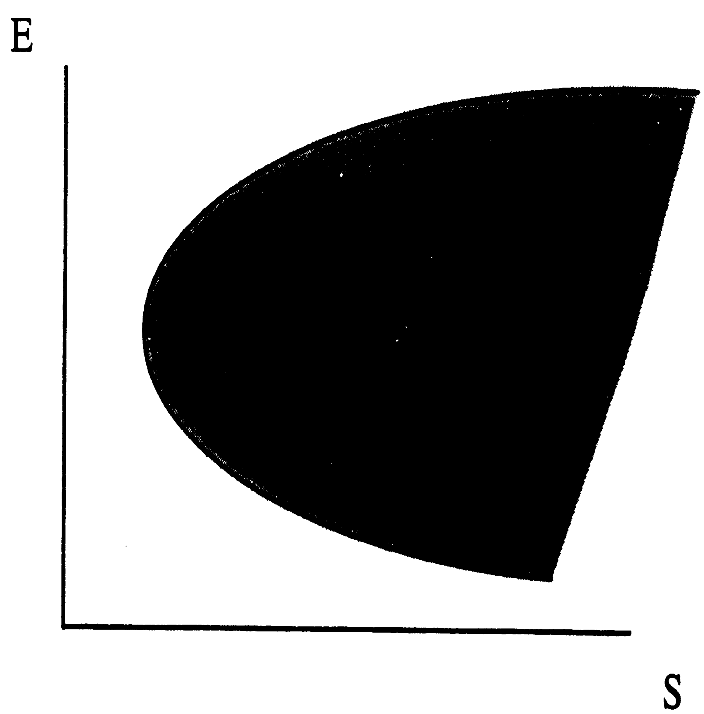
$$\sigma(R_p) = \sqrt{w^2 \sigma^2_{S\&P} + 0} = w\,\sigma_{S\&P}$$

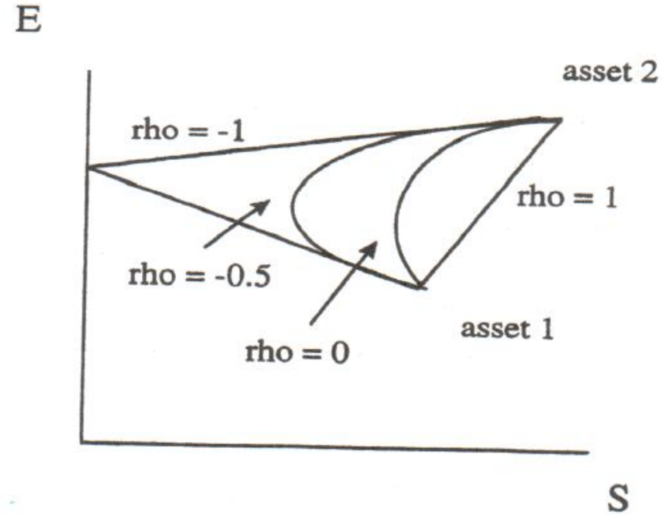


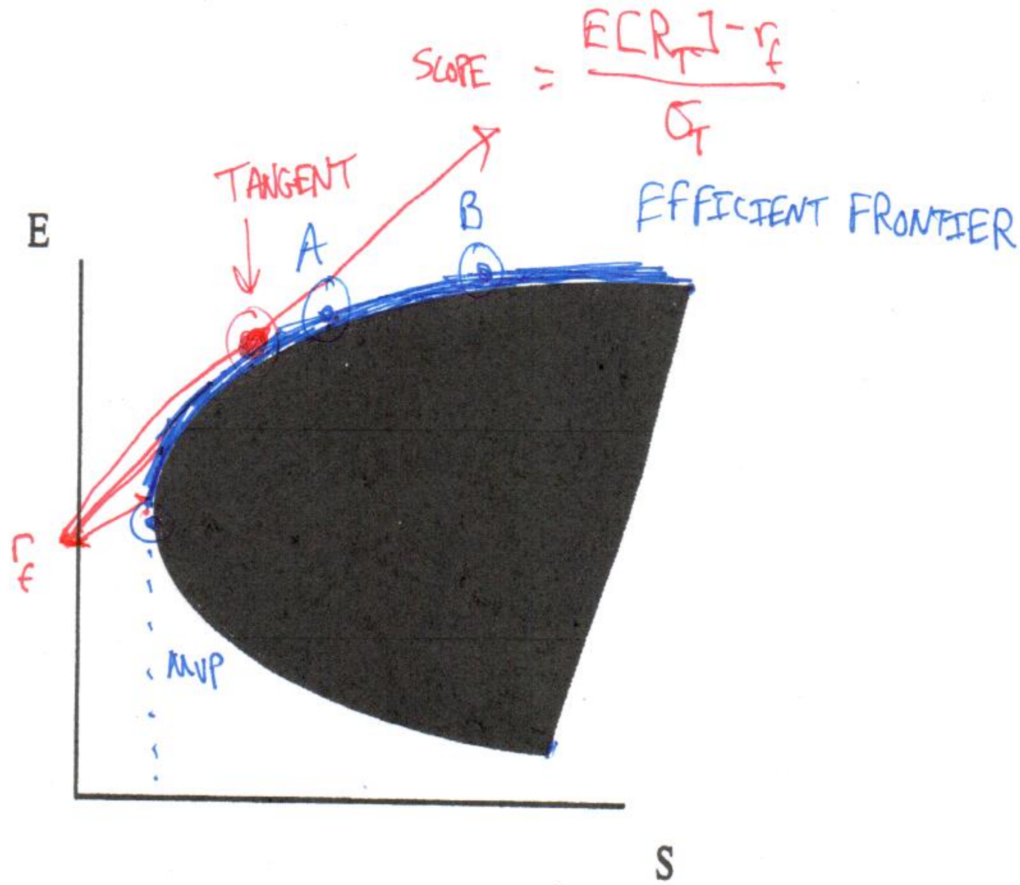
$$E[R_p] = w\,E[r_{S\&P}] + (1-w)\,E[r_B]$$

$$\sigma(R_p) = \sqrt{w^2 \sigma^2_{S\&P} + (1-w)^2 \sigma^2_B + 2w(1-w)\,\sigma_{S,B}}$$


E
asset 2
rho = -1
rho = 1
rho = -0.5
rho = 0
asset 1
S


E
S


E
asset 2
rho = -1
rho = 1
rho = -0.5
asset 1
rho = 0
S


SLOPE = $\frac{E[R_T]-r_f}{\sigma_T}$
TANGENT
B
A
EFFICIENT FRONTIER
E
r_f
MVP
S

- the investments open to us all lie on a straight line, called the CAPITAL ALLOCATION LINE
- these are all the risk-return combinations.
- to see the feasible investment set generated by combinations of a *riskless* asset and a *risky* asset, draw a straight line between them
- the slope of the CAL is the increase in expected return of the chosen portfolio per unit of additional risk (standard deviation). It is the measure of extra return per risk:

$$\text{Slope} = \frac{E[\tilde{r}_{\mathrm{P}}] - r_f}{\sigma_{\mathrm{P}}}.$$

- we can use a *utility function* to decide between them
  - let E be the expected return on an investment
  - let S be the standard deviation
- then the function $U(E, S)$ represents how good you feel about the investment (think of this as a "score of happiness")
- we will use $U(E, S) = E - \alpha S^2$, a *quadratic* utility function
  - U(E,S) is increasing in E (you like expected return)
  - U(E,S) is decreasing in S (you don't like risk)

- α is the coefficient of risk aversion: (how risk affects you)

  higher α means more risk-averse

  $\alpha = 0$ is risk-neutral

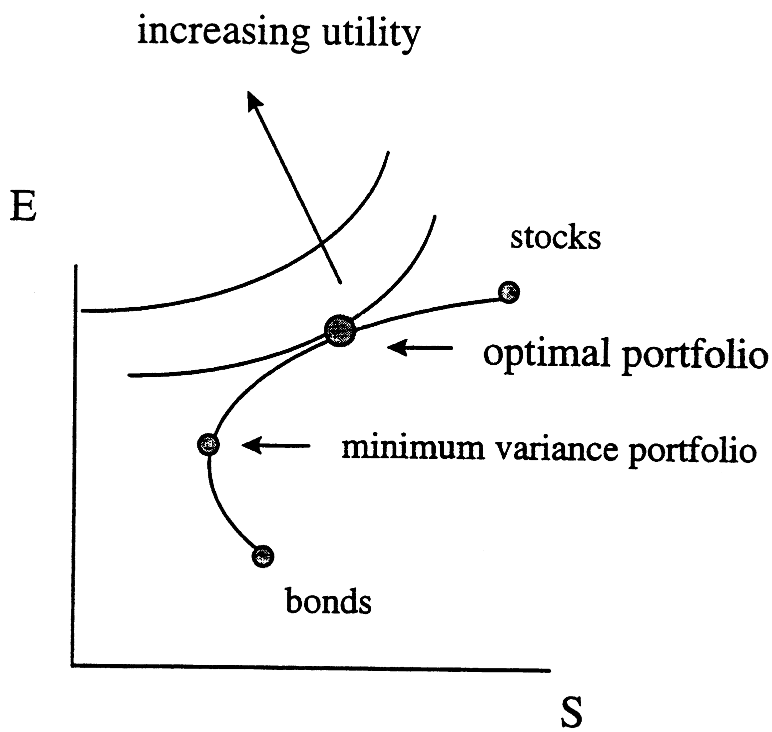
  $\alpha < 0$ is risk-loving

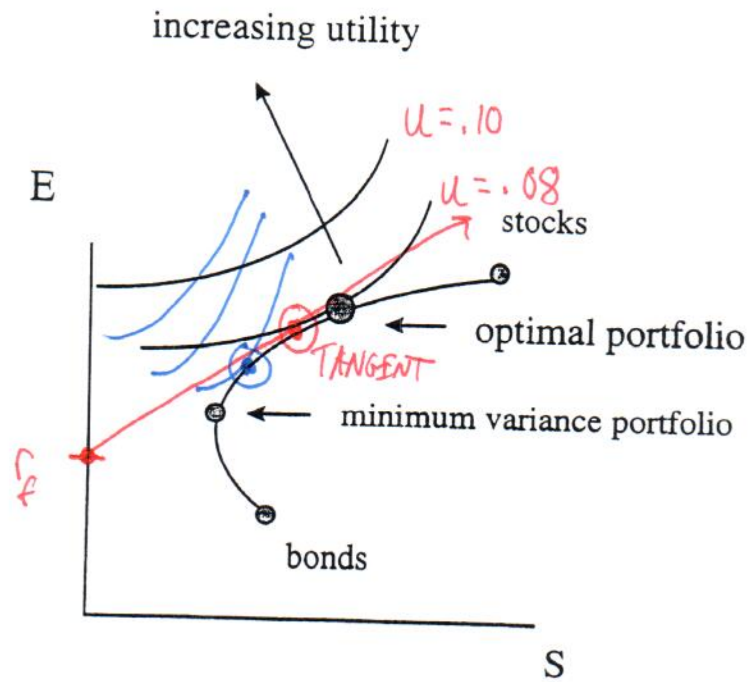
- use the utility function to make an investment decision:

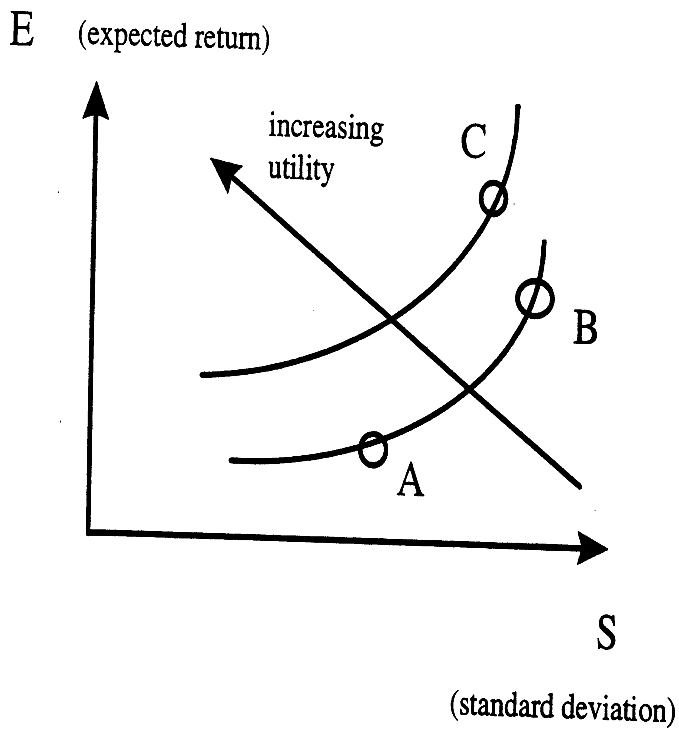
  $U(E, S) = E - \alpha S^2$

  Choose the strategy providing maximum utility

  - just drag the indifference curves down towards the investment opportunity set
  - the first point of contact is the optimal portfolio


increasing utility
E
stocks
optimal portfolio
minimum variance portfolio
bonds
S

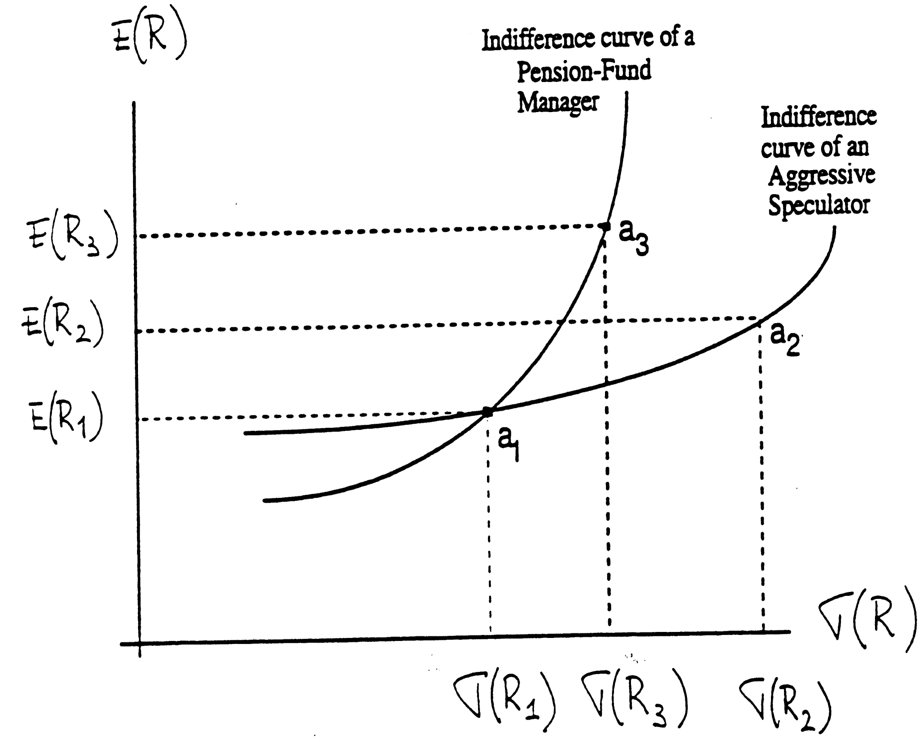

increasing utility
U=.10
U=.08
E
stocks
optimal portfolio
TANGENT
minimum variance portfolio
r_f
bonds
S



Drawing a diagram can be very helpful to illustrate utiltity curves:

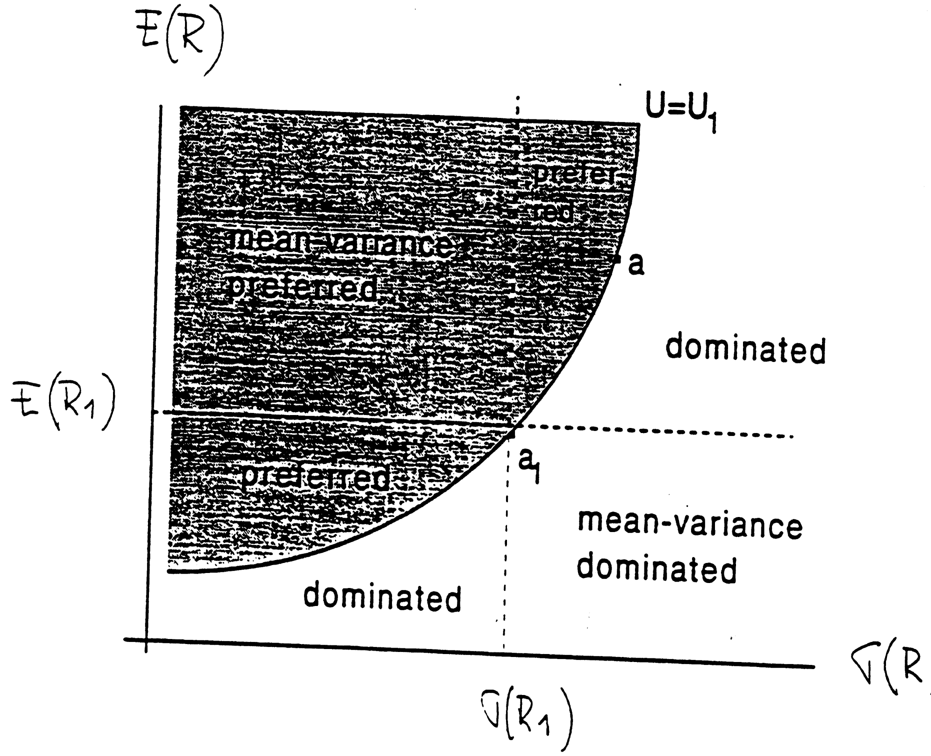
– the lines on this graph are *indifference curves:*

  * the investor is indifferent between the portfolios on any one curve
  * the investments that have different expected returns and risks, but that give the same level of utility to an investor define an indifference curve.
  * the level of risk aversion (α) of an investor determines the steepness of his/her indifference curves.
    · A pension fund manager, with high risk aver-



sion, requires a large increase in expected return (from $E(R_1)$ to $E(R_3)$) as compensation for an increase in risk (from $\sigma(R_1)$ to $\sigma(R_3)$), and be *indifferent* between the two investments 1 and 3.

- An aggressive speculator, with low risk aversion, requires a lower increase in expected return (from $E(R_1)$ to $E(R_2)$) as compensation for an increase in risk (from $\sigma(R_1)$ to $\sigma(R_2)$), and be *indifferent* between the two investments 1 and 2.
- Comparing investments (portfolios)
- If investment i has both a larger expected re-



turn and lower standard deviation than investment 1, then

· all risk averse investors will prefer investment i to investment 1, and we say that investment i is **mean variance preferred** to investment 1.

- to find the optimal portfolio, drag the indifference curve down until it just touches the CAL

- write down the maximization problem in general terms and solve it with calculus or the computer:

- We need to make an assumption about the utility function.

- As a simple example suppose we have a quadratic utility function, for one riskless and one risky asset the optimization is easy,

$$
\begin{aligned}
&\mathit{maximize} \\
&E_\omega - \alpha S_\omega^2 \;=\; (1-\omega)r_f + \omega E_P - \alpha(\omega^2 S_P^2)
\end{aligned}
$$

Take the first derivative with respect to ω, and set it equal to zero:

$$E_P - r_f - 2\alpha\omega S_P^2 = 0$$

$$\Rightarrow \boxed{\quad \omega^* \;=\; \frac{E_P - r_f}{2\alpha S_P^2} \quad}$$

For many risky assets, we need to find the equation for the feasible set. In this case, it is the equation for the efficient set of all risky investments.

This is called the efficient set mathematics.

Alternatively, instead of making a utility assumption, we could make a distributional assumption about returns. For example, if returns are mutivariate normally distributed, then only mean and variance matter $\Rightarrow$ apply mean-variance analysis.

Mean-Variance mathematics and the efficient frontier of risky assets.

(to be derived in class)

Mean-Variance mathematics and the efficient frontier of risky assets.

(to be derived in class)

$$\underset{(N\times 1)}{W_p} = \begin{bmatrix} w_{1,p} \\ w_{2,p} \\ \vdots \\ w_{N,p} \end{bmatrix} = \text{Port. weights} \qquad \underset{(N\times 1)}{R} = \begin{bmatrix} \bar{r}_1 \\ \vdots \\ \vdots \\ \bar{r}_N \end{bmatrix} = \text{mean rets.}$$

$$\underset{(N\times N)}{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \vdots & \sigma_{22} & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN} \end{bmatrix} = \text{COV. MATRIX} \qquad \underset{(N\times 1)}{\mathbb{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

PORTFOLIO

$$W'R = r_p$$
$$w'VW = \sigma_p^2$$
$$w'\mathbb{1} = 1$$

$$\min\ w'Vw$$
s.t.
$$w'R = r_p \quad (\text{e.g. } 10\%)$$
$$w'\mathbb{1} = 1$$

SOLVE:

$$\mathcal{L} = w'VW - \lambda_1(w'R - r_p) - \lambda_2(w'\mathbb{1} - 1)$$

$$\frac{\partial \mathcal{L}}{\partial w} = 2VW - \lambda_1 R - \lambda_2 \mathbb{1} = 0$$

$$\Rightarrow VW = (\lambda_1 R + \lambda_2 \mathbb{1})\tfrac{1}{2}$$

$$W^* = \tfrac{1}{2} V^{-1} (\lambda_1 R + \lambda_2 \mathbf{1})$$

$$\boxed{W^* = \tfrac{1}{2} V^{-1} (R\ \mathbf{1}) \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}}$$

$$(R\ \mathbf{1})' W = \tfrac{1}{2} \underbrace{(R\ \mathbf{1})' V^{-1} (R\ \mathbf{1})}_{A} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$A^{-1} (R\ \mathbf{1})' W = \tfrac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$A^{-1} \begin{pmatrix} R'W \\ \mathbf{1}'W \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$A^{-1} \begin{pmatrix} r_p \\ 1 \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$\boxed{W^* = V^{-1} (R\ \mathbf{1}) A^{-1} \begin{pmatrix} r_p \\ 1 \end{pmatrix}}$$

A = fundamental matrix $= (R\mathbf{1})' V^{-1} (R\ \mathbf{1})$

$$= \begin{bmatrix} R'V^{-1}R & \mathbf{1}'V^{-1}R \\ R'V^{-1}\mathbf{1} & \mathbf{1}'V^{-1}\mathbf{1} \end{bmatrix} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

$$A^{-1}A = I \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow A^{-1} = \frac{\begin{bmatrix} c & -b \\ -b & a \end{bmatrix}}{ac - b^2}$$

$$\sigma_p^2 = W'VW = \underbrace{(r_p\ 1) A^{-1} (R\mathbf{1})' \cancel{V^{-1}}}_{W'} \cancel{V} \underbrace{V^{-1} (R\ \mathbf{1}) A^{-1} \begin{pmatrix} r_p \\ 1 \end{pmatrix}}_{W}$$

$$= (r_p \; 1) A^{-1} \cancel{A} \cancel{A^{-1}} \begin{pmatrix} r_p \\ 1 \end{pmatrix}$$

$$= (r_p \; 1) A^{-1} \begin{pmatrix} r_p \\ 1 \end{pmatrix}$$

$$\boxed{\overset{*}{\sigma}_p^{2} = \frac{a - 2br_p + cr_p^2}{ac - b^2}}$$

<u>MVP</u>

$$\frac{d\sigma_p^2}{dr_p} = -2b + 2cr_p = 0 \Rightarrow r_p = b/c$$

$$r_{MVP} = b/c = \frac{1'V^{-1}R}{1'V^{-1}1} = w_{MVP}'R$$

$$= \underbrace{\left(\frac{1'V^{-1}}{1'V^{-1}1}\right)}_{= w_{MVP}} R = w_{MVP}'R$$

$$\sigma_{MVP}^2 = w_{MVP}' V w_{MVP} = \left(\frac{1'V^{-1}}{1'V^{-1}1}\right)' V \left(\frac{1'V^{-1}}{1'V^{-1}1}\right) = \frac{1}{1'V^{-1}1} = 1/c \checkmark$$

$$\sigma_{MVP}^2 = \frac{a - 2b(b/c) + c(b/c)^2}{ac - b^2} = 1/c = \frac{1}{1'V^{-1}1} \checkmark$$

$$^{*}Cov(R_p, R_{MVP}) = Cov(R_q, R_{MVP}) = Cov(R_{MVP}, R_{MVP}) = 1/c = \sigma_{MVP}^{2\,*}$$

$$\forall p, q$$

$$\sigma_{p,MVP} = w_p' V w_{MVP} = w_p' \cancel{V} \left(\frac{1'\cancel{V^{-1}}}{1'V^{-1}1}\right) = \frac{w_p'1}{1'V^{-1}1} = \frac{1}{1'V^{-1}1} \checkmark$$

$$= 1/c \checkmark$$

$$R_{EW} = \frac{1}{N}\sum_{i=1}^{N} r_i$$

$$R_p = R_{EW} + \delta r_i - \delta r_K$$

$$\sigma_p^2 = \sigma_{EW}^2 + \delta^2\sigma_i^2 + \delta^2\sigma_K^2 + 2\delta(1)\sigma_{EW,r_i} - 2\delta(1)\sigma_{EW,r_K} - 2\delta^2\sigma_{i,K}$$

$$\frac{\Delta\sigma_p^2}{\Delta\delta} \quad \frac{\partial\sigma_p^2}{\partial\delta} = \cancel{2\delta\sigma_i^2} + \cancel{2\delta\sigma_K^2} + 2\sigma_{EW,r_i} - 2\cancel{\delta}\sigma_{EW,r_K} - \cancel{4\delta\sigma_{iK}}$$

$$\lim_{\delta\to 0} \Rightarrow \quad 2\sigma_{EW,i} - 2\sigma_{EW,K}$$

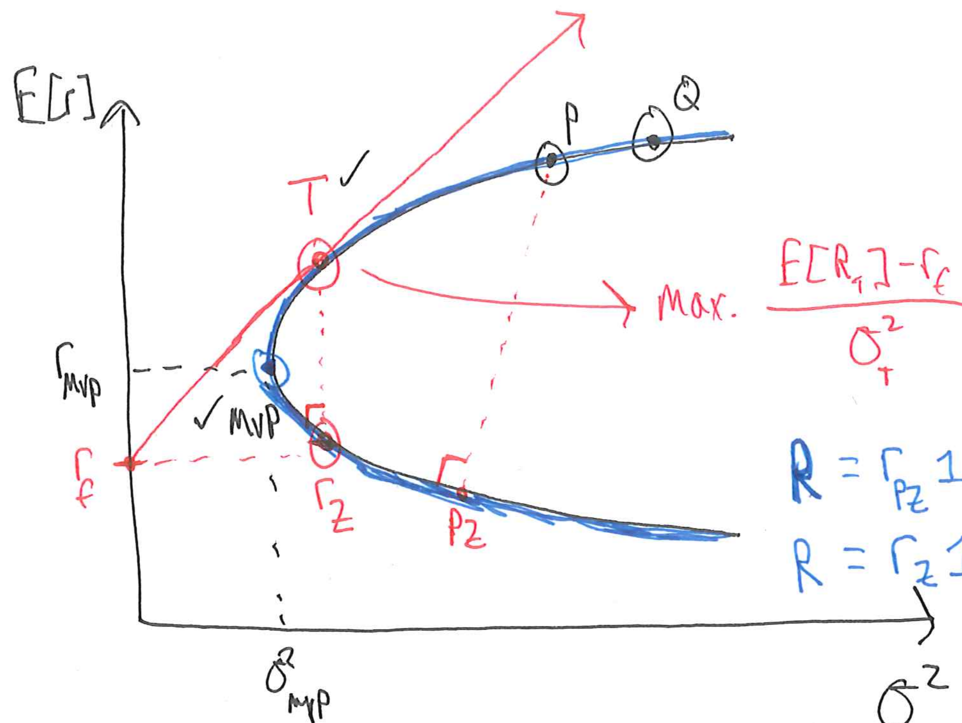
$$\Delta\sigma_p^2 \uparrow \text{ if } \quad \sigma_{EW,i} - \sigma_{EW,K} > 0$$

$$\Delta\sigma_p^2 \downarrow \text{ if } \quad \sigma_{EW,i} - \sigma_{EW,K} < 0$$

$$\sigma_{EW,i} > \sigma_{EW,K} \Rightarrow \Delta\sigma_p^2 \uparrow$$

$$\sigma_{EW,i} < \sigma_{EW,K} \Rightarrow \Delta\sigma_p^2 \downarrow$$

$$\boxed{MVP \Rightarrow \quad Cov(r_i, R_{mvp}) = Cov(r_K, R_{mvp}) \quad \forall\, i, K}$$



$$R = r_{pz}\mathbf{1} + \frac{VW}{\sigma_P^2}(r_P - r_{pz})$$

$$R = r_z\mathbf{1} + \frac{VW}{\sigma_T^2}(r_T - r_z)$$

$$R = r_f\mathbf{1} + \frac{VW}{\sigma_T^2}(r_T - r_f)$$

$$\boxed{R = r_f\mathbf{1} + \beta(r_T - r_f)}$$

(*) MVP

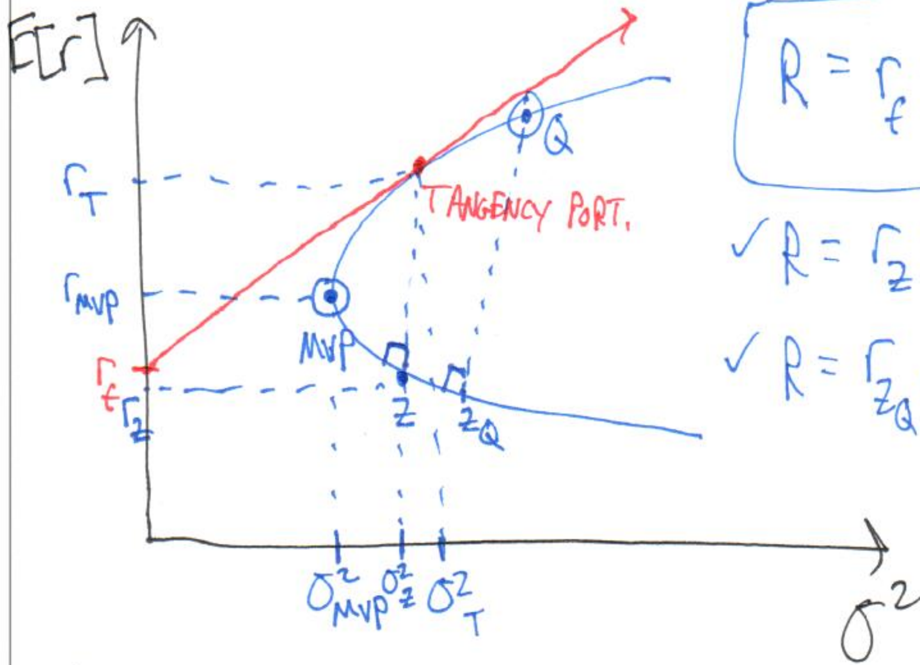
$$W'V = \mathbf{1} \Rightarrow \boxed{W = \frac{V^{-1}\mathbf{1}}{\mathbf{1}'V^{-1}\mathbf{1}} = W_{MVP}}$$

(*) TANGENCY

$$W'V = (R - r_f\mathbf{1}) \Rightarrow \boxed{W = \frac{V^{-1}(R - r_f\mathbf{1})}{\mathbf{1}'V^{-1}(R - r_f\mathbf{1})} = W_{TANG}}$$

MVP $\quad \dfrac{1}{\sigma_{i,MVP}} = \dfrac{1}{\sigma_{k,MVP}} \quad \forall i,k$ ✓

TANG $\quad \dfrac{E[r_i]-r_f}{\sigma_{i,T}} = \dfrac{E[r_k]-r_f}{\sigma_{k,T}} \quad \forall i,k$ ✓



$$\boxed{R = r_f 1 + \frac{V w_T}{\sigma_T^2}(r_T - r_f)}$$

✓ $$R = r_Z 1 + \frac{V w_T}{\sigma_T^2}(r_T - r_Z)$$

✓ $$R = r_{Z_Q} 1 + \frac{V w_Q}{\sigma_Q^2}(r_Q - r_{Z_Q}) \quad \forall Q$$

except MVP!

MVP

$$\frac{1}{Cov(r_i, R_{MVP})} = \frac{1}{Cov(r_k, R_{MVP})} \quad \forall i, k$$

$$w'V = 1 \Rightarrow \boxed{w_{MVP} = \frac{V^{-1} 1}{1' V^{-1} 1}}$$

TANGENCY

$$w'V = (R - r_f 1) \Rightarrow \boxed{w_T = \frac{V^{-1}(R - r_f 1)}{1' V^{-1}(R - r_f 1)}}$$

$$\frac{E[r_i] - r_f}{Cov(r_i, R_T)} = \frac{E[r_k] - r_f}{Cov(r_k, R_T)} \quad \forall i, k$$

Portfolio Optimization in Practice – Jorion (FAJ, 1992)

- Drawback of mean-variance analysis is that it ignore measurement error of the inputs (means, variances, covariances).
- How does this affect portfolio weights and performance?
  - use historical data to estimate inputs, which have substantial error.
  - thus, there is *estimation risk* in the analysis.
- But, MV optimization can incorporate a host of constraints (short-selling restrictions, transactions costs, liquidity concerns, turnover constraints, etc.).

How can we assess estimation risk and its impact?

- Jorion uses international bond portfolio data to illustrate benefits of diversification.
  - for a given risk level, how much can we increase expected returns?
- Table II.
- MV optimization weights heavily assets that have high average returns. What sort of problems does this in-

Table II Dollar Returns on Efficient Global Bond Portfolios With No Short Sales, 1978–1988

		Proportion Invested in						
Avg. Ret.	*Stand. Dev.*	*U.S. Doll.*	*Can. Doll.*	*Ger. Mark*	*Jap. Yen*	*Brit. Pound*	*Dutch Guilder*	*French Franc*
10.12	9.95	0.64	0.00	0.00	0.03	0.04	0.00	0.29
10.65	10.03	0.61	0.00	0.00	0.12	0.05	0.00	0.21
11.18	10.25	0.59	0.00	0.00	0.21	0.07	0.00	0.13
11.71	10.61	0.56	0.00	0.00	0.30	0.08	0.00	0.06
12.24	11.10	0.51	0.00	0.00	0.39	0.09	0.00	0.00
12.77	11.77	0.42	0.00	0.00	0.49	0.09	0.00	0.00
13.30	12.62	0.33	0.00	0.00	0.58	0.09	0.00	0.00
13.83	13.63	0.23	0.00	0.00	0.67	0.09	0.00	0.00
14.36	14.75	0.14	0.00	0.00	0.77	0.09	0.00	0.00
14.89	15.96	0.05	0.00	0.00	0.86	0.09	0.00	0.00
15.42	17.31	0.00	0.00	0.00	0.98	0.02	0.00	0.00
Max. Return/Risk								
11.96	10.82	0.55	0.00	0.00	0.34	0.09	0.00	0.02
World Index								
11.31	10.94	0.46	0.03	0.06	0.14	0.27	0.02	0.02

duce? What else does such a procedure implicitly weight on?

- Also weights heavily on assets with attractive (low) covariance properties. What sort of problems does this induce?
- These studies are typically done *in sample*. What sort of concerns might you have, and what can be done to alleviate these?
- How can we gauge these errors and what can be done?
- The problem is, we are dealing with random *samples* of data.
- One way to address this is to run **simulations**.

Simulations:

1. assume a distribution for the returns on assets (e.g., multivariate normal).
2. compute the parameters of this distribution using historical data (means, variances, covariances).
3. draw a random sample of N (number of assets) returns from this distrubtion T (number of time periods) times.

ACTUAL DATA

In-Sample $\Rightarrow \underset{(N\times 1)}{R}, \underset{(N\times N)}{V}$

$$\Rightarrow W_T^* = \frac{V^{-1}R}{1'V^{-1}R} \checkmark$$

① SIMULATED DATA

$\Rightarrow$ Compute $\underset{(N\times 1)}{R^S}, \underset{(N\times N)}{V^S}$

$$\Rightarrow W_T^S = \frac{V^{S-1}R^S}{1'V^{S-1}R^S}$$

* COMPARE

$W_T^{*\prime} R$ vs. $W_T^{S\prime} R$ why not R^S?

$W_T^{*\prime} V W_T^*$ vs. $W_T^{S\prime} V W_T^S$ not $\cancel{V^S}$

4. estimate from these simulated returns a new vector of sample average returns and new sample covariance matrix.

5. form optimal portfolios from this random sample.

6. repeat many times (1,000 - 10,000) so that distribution of optimal portfolio or efficient set is approximated with enough precision.

7. examine how distant simulated portfolios are from estimated portfolios (use graph or some “distance” metric). Of course, estimated will always outperform simulations in sample.

8. often use the Gibbons, Ross, and Shanken (1989) F-test statistic:

$$
\begin{aligned}
F &= \frac{T(T-N-1)}{N(T-2)} \frac{\theta_*^2 - \theta_p^2}{1+\theta_p^2} \\
\theta_i &= \frac{E[\tilde{r}_i] - r_f}{\sigma_i}
\end{aligned}
$$

   where portfolio * is the optimal portfolio and portfolio p is the benchmark. High value of F means the benchmark is not likely to be efficient. Follows and F-distribution.

9. can also conduct out-of-sample analysis.

10. which is easier, MVP or tangency portfolio estimation and why? (hint: think of sampling standard error

4. estimate from these simulated returns a new vector of sample average returns and new sample covariance matrix.

5. form optimal portfolios from this random sample.

6. repeat many times (1,000 - 10,000) so that distribution of optimal portfolio or efficient set is approximated with enough precision.

7. examine how distant simulated portfolios are from estimated portfolios (use graph or some "distance" metric). Of course, estimated will always outperform simulations in sample.

* 8. often use the Gibbons, Ross, and Shanken (1989) F-test statistic:

SHARPE RATIO^2 OF TANGENCY PORT. IN THE DATA

$$F = \frac{T(T-N-1)}{N(T-2)} \frac{\theta_*^2 - \theta_p^2}{1+\theta_p^2}$$

SHARPE RATIO^2 OF SIMULATED WEIGHTS OF TANG. PORT APPLIED TO ACTUAL (DATA) R & V.

$$\theta_i = \frac{E[\tilde{r}_i] - r_f}{\sigma_i}$$

where portfolio * is the optimal portfolio and portfolio p is the benchmark. High value of F means the benchmark is not likely to be efficient. Follows and F-distribution.

9. can also conduct out-of-sample analysis.

10. which is easier, MVP or tangency portfolio estimation and why? (hint: think of sampling standard error

measures on first vs. second moments (i.e., on means vs. variances and covariances)).

11. Can relax constraints on the optimization to see how this affects estimation error. What happens when relax short-selling contraint?

12. What are the pros and cons of relaxing constraints?

Figure A Statistically Equivalent Global Bond Portfolios with No Short Sales, 1978 – 1988

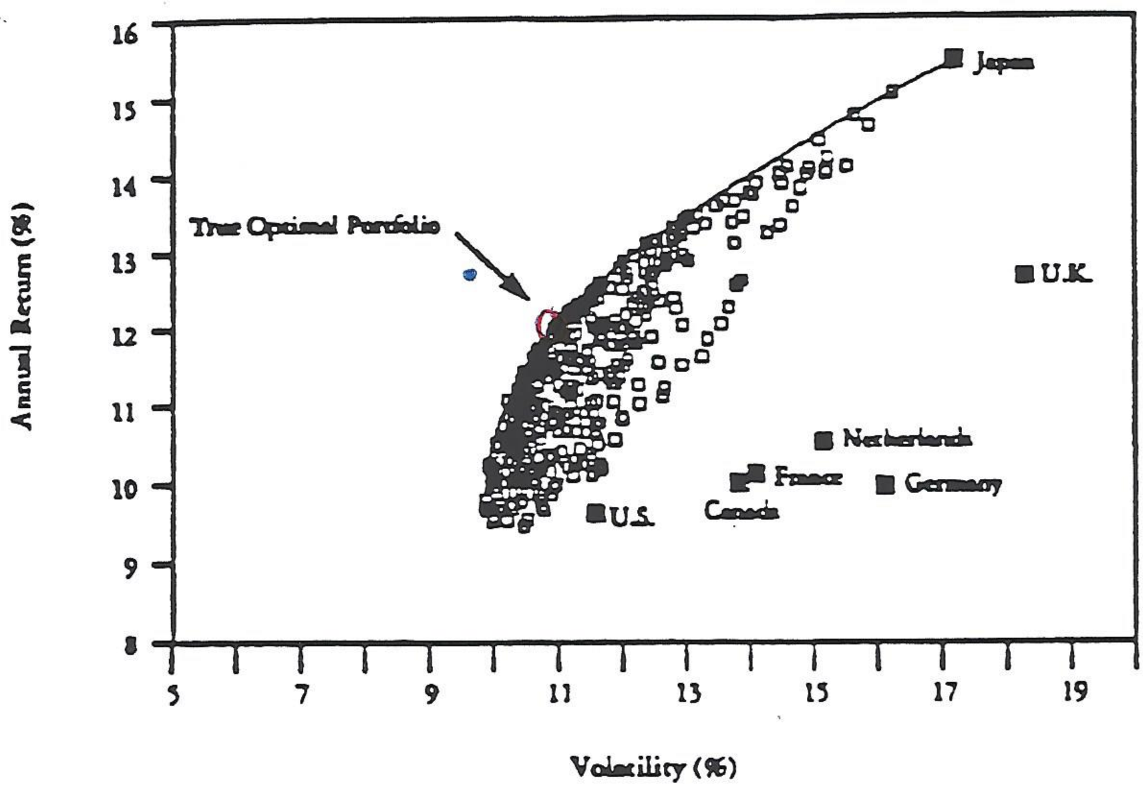
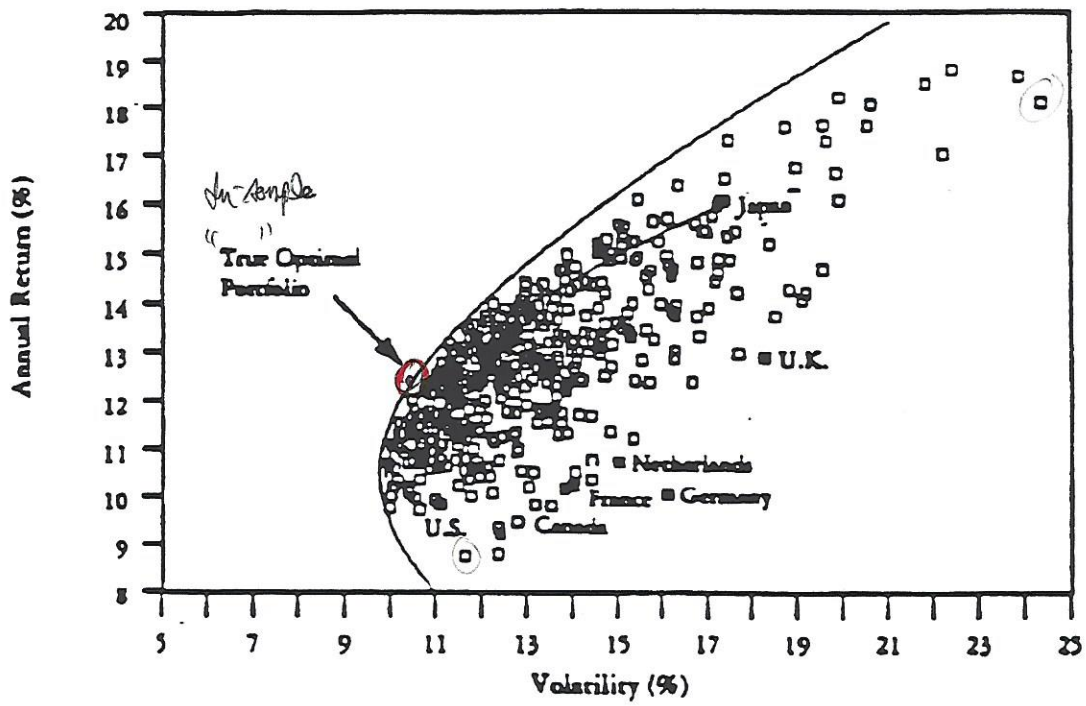


Figure C Statistically Equivalent Global Bond Portfolios with Short Sales, 1978 – 1988



The Fundamental Question

Q: Why do some securities earn higher returns on average than others?

A: Finance theory says, because of risk, but ...

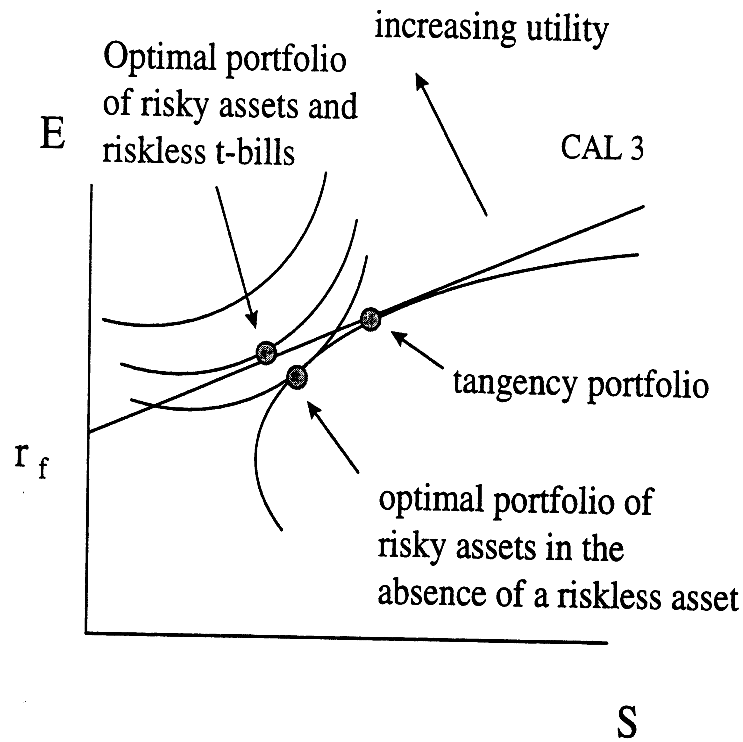
- what is risk?
- how much extra return do we require to bear the risk?

The CAPM answers these questions, and has many important applications:

- valuation
  - how much is a company worth?
  - is this stock undervalued?
- mutual fund performance evaluation
- event studies
- mispricing and market efficiency debate

The CAPM

Since every investor would hold some combination of the risk-free asset and the tangency portfolio (separation principle).



– In order to identify the tangency portfolio, we showed that it had to have the following property:

$$\frac{E[\tilde{r}_i] - r_f}{cov(\tilde{r}_i, \tilde{R}_T)} = \frac{E[\tilde{r}_j] - r_f}{cov(\tilde{r}_j, \tilde{R}_T)} \quad \forall i, j$$

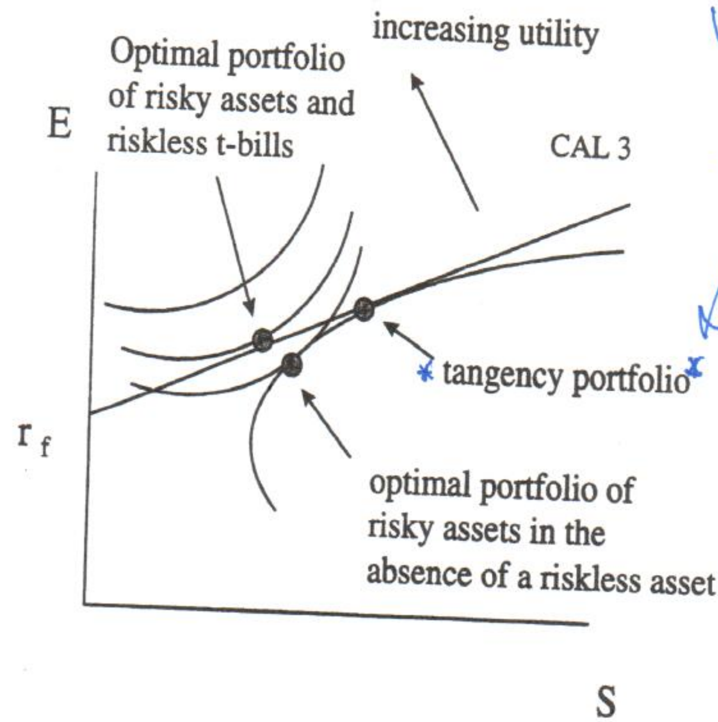
– Simply put, all assets must have an equal risk premium-to-covariance ratio or "reward-to-risk" ratio in the tangency portfolio.

The CAPM

Since every investor would hold some combination of the risk-free asset and the tangency portfolio (separation principle).

$w_T = \frac{V^{-1}(R - r_f 1)}{1' V^{-1}(R - r_f 1)}$

$\Rightarrow \frac{E[r_i]-r_f}{cov(r_i, R_T)} = \frac{E[r_k]-r_f}{cov(r_k, R_T)}$



– In order to identify the tangency portfolio, we showed that it had to have the following property:

$$\frac{E[\tilde{r}_i] - r_f}{cov(\tilde{r}_i, \tilde{R}_T)} = \frac{E[\tilde{r}_j] - r_f}{cov(\tilde{r}_j, \tilde{R}_T)} \quad \forall i, j$$

– Simply put, all assets must have an equal risk premium-to-covariance ratio or "reward-to-risk" ratio in the tangency portfolio.

– Since this holds for all assets, it must also hold for the tangency portfolio itself,

$$\Rightarrow \frac{E[\tilde{r}_i] - r_f}{cov(\tilde{r}_i, \tilde{R}_T)} = \frac{E[\tilde{R}_T] - r_f}{cov(\tilde{R}_T, \tilde{R}_T)} = \frac{E[\tilde{R}_T] - r_f}{var(\tilde{R}_T)}$$

This then can be rearranged to derive a relation between an asset's expected return and risk,

$$E[\tilde{r}_i] = r_f + \frac{cov(\tilde{r}_i, \tilde{R}_T)}{var(\tilde{R}_T)}(E[\tilde{R}_T] - r_f)$$

– The relevant measure of risk, here, is the covariance between the asset's returns and the returns on the tangency portfolio.
  * Intuition: In economics, it is the *marginal cost* of goods that determines their prices, not their total or average cost.
  * Likewise, the *marginal variance* or covariance determines the additional risk of an investment, and therefore its price (here, expressed as returns not dollars).

– Or, $\frac{cov(\tilde{r}, \tilde{R}_T)}{var(\tilde{R}_T)} = \beta$ (it is the slope from a regression

– Since this holds for all assets, it must also hold for the tangency portfolio itself,

$$\Rightarrow \frac{E[\tilde{r}_i] - r_f}{cov(\tilde{r}_i, \tilde{R}_T)} = \frac{E[\tilde{R}_T] - r_f}{cov(\tilde{R}_T, \tilde{R}_T)} = \frac{E[\tilde{R}_T] - r_f}{var(\tilde{R}_T)}$$

This then can be rearranged to derive a relation between an asset's expected return and risk,

$$R = r_f 1 + \frac{Vw}{\sigma_T^2}(r_T - r_f)$$

(N×1) (N×1) (N×1)

$$E[\tilde{r}_i] = r_f + \frac{cov(\tilde{r}_i, \tilde{R}_T)}{var(\tilde{R}_T)}(E[\tilde{R}_T] - r_f)$$

$= \beta_{i,T}$

– The relevant measure of risk, here, is the covariance between the asset's returns and the returns on the tangency portfolio.

* <u>Intuition:</u> In economics, it is the *marginal cost* of goods that determines their prices, not their total or average cost.
* Likewise, the *marginal variance* or covariance determines the additional risk of an investment, and therefore its price (here, expressed as returns not dollars).

– Or, $\frac{cov(\tilde{r}, \tilde{R}_T)}{var(\tilde{R}_T)} = \beta$ (it is the slope from a regression of $\tilde{r}$ on $\tilde{R}_T$),

$$\Rightarrow E[\tilde{r}_i] = r_f + \beta_i(E[\tilde{R}_T] - r_f)$$

– Since everyone holds the same *risky* portfolio (the tangency portfolio) and since the supply of risky assets must equal the demand, this implies . . .

– **Punchline of the CAPM:** THE TANGENCY PORTFOLIO IS THE MARKET PORTFOLIO.

$\Rightarrow$ CAPM equation: $E[\tilde{r}_i] = r_f + \beta_i(E[\tilde{R}_M] - r_f)$.

Why do we hold the market portfolio?

– follows directly from the separation principle

Assume:

– returns are distributed multivariate normal

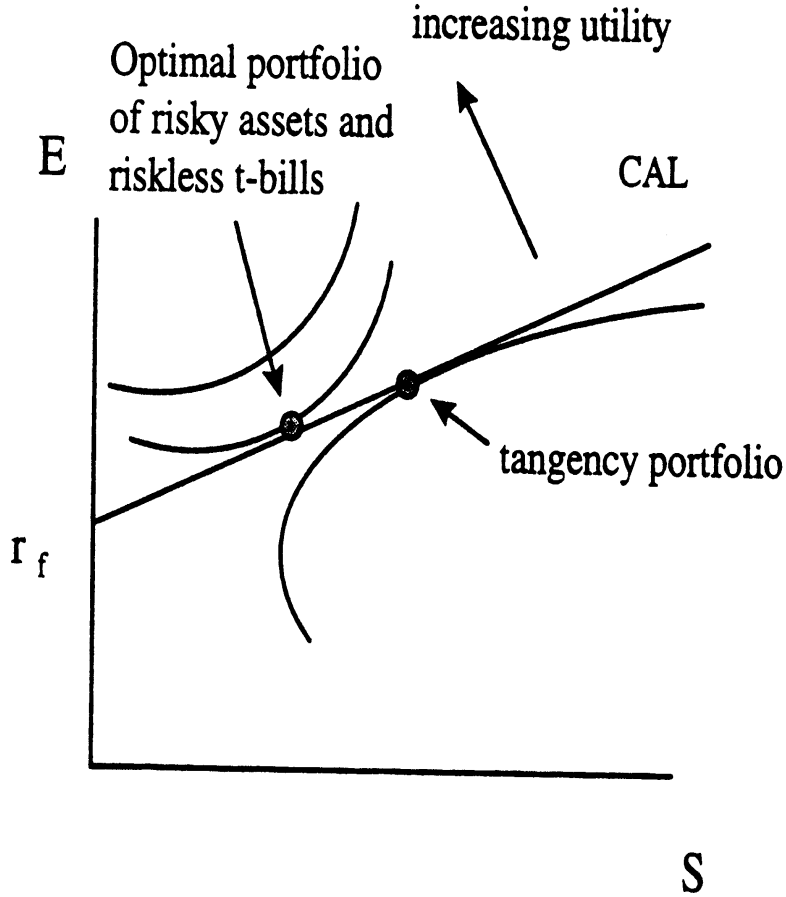
– everyone uses mean-variance analysis

– we all have the same estimates of expected returns, variances etc...(homogeneous beliefs)

Then:

– the proportion of one risky asset to another is the *same* in everybody's portfolio

– supply = demand $\Rightarrow$ those proportions must be the proportions of the total market made up by each asset



– the tangency portfolio *is* the market portfolio

– thus, we refer to the CAL as the **Capital Market Line** (CML).

$\Rightarrow$ since risk is covariance with what we hold, it must be *covariance with the market portfolio.*