More about the CAPM

- What is relationship between expected utility and mean-variance preferences? (D and D sect. 5.2 and app. 5.1)
- How can the CAPM be extended to situations without a risk free asset? (D and D sect. 6.3–6.5, Roll's app.)

Mean-variance versus vN-M expected utility

- In general those who maximize $E[U(\tilde{W})]$ care about the whole distribution of $\tilde{W}$.
- Will care about only mean and variance if those two characterize the whole distribution.
- Will alternatively care about only mean and variance if $U()$ is a quadratic function.

The third way to underpin mean-var assumption

- Perhaps things are so complicated that people resort to just considering mean and variance. (Whether they are vN-M people or not.)

Mean-var preferences due to distribution

- Assume that choices are always between random variables with one particular type ("class") of probability distribution.
- Could be, e.g., choice only between binomially distributed variables. (There are different binomial distributions, summarized in three parameters which uniquely define each one of them.)
- Or, e.g., only between variables with a chi-square distribution. Or variables with normal distribution. Or variables with a lognormal distribution.
- Some of these distributions, such as the normal distribution and the lognormal distribution, are characterized completely by two parameters, the mean and the variance.
- If all possible choices belong to the same class, then the choice can be made on the basis of the parameters for each of the distributions.
- Example: Would you prefer a normally distributed wealth with mean 1000 and variance 40000 or another normally distributed wealth with mean 500 and variance 10000?

- If mean and variance characterize each alternative completely, then all one cares about is mean and variance.
- Most convenient: Normal distribution, since sums (and more generally, any linear combinations) of normally distributed variables. are also normal. Most opportunity sets consist of alternative linear combinations of variables.
- Problem: Positive probability for negative outcomes. Share prices are never negative.

Mean-var preferences due to quadratic U

Assume

$$U(w) \equiv cw^2 + bw + a$$

where $a, b > 0, c < 0$ are constants. With this U function:

$$\begin{aligned} E[U(\tilde{W})] &= cE(\tilde{W}^2) + bE(\tilde{W}) + a \\ &= c\{E(\tilde{W}^2) - [E(\tilde{W})]^2\} + c[E(\tilde{W})]^2 + bE(\tilde{W}) + a \\ &= c\,\mathrm{var}(\tilde{W}) + c[E(\tilde{W})]^2 + bE(\tilde{W}) + a, \end{aligned}$$

which is a function only of mean and variance of $\tilde{W}$.

Problem: U function is decreasing for large values of W. Must choose c and b such that those large values have zero probability.

Another problem: Increasing (absolute) risk aversion.

Indifference curves in mean-stddev diagrams

- If mean and variance are sufficient to determine choices, then mean and $\sqrt{\text{variance}}$ are also sufficient.
- More practical to work with mean (μ) and standard deviation (σ) diagrams.
- Common to put standard deviation on horizontal axis.
- Will show that indifference curves are increasing and convex in (σ, μ) diagrams.
- Consider normal distribution and quadratic U separately.
- Indifference curves are contour curves of $E[U(\tilde{W})]$.
- Total differentiation:

$$0 = dE[U(\tilde{W})] = \frac{\partial E[U(\tilde{W})]}{\partial \sigma} d\sigma + \frac{\partial E[U(\tilde{W})]}{\partial \mu} d\mu.$$

Indifference curves from quadratic U

Assume $W < -b/(2c)$ with certainty in order to have $U'(W) > 0$.

$$E[U(\tilde{W})] = c\sigma^2 + c\mu^2 + b\mu + a.$$

First-order derivatives:

$$\frac{\partial E[U(\tilde{W})]}{\partial \sigma} = 2c\sigma < 0, \ \frac{\partial E[U(\tilde{W})]}{\partial \mu} = 2c\mu + b > 0,$$

Thus the slope of the indifference curves,

$$\left.\frac{d\mu}{d\sigma}\right|_{E[U(\tilde{W})] \text{ const.}} = -\frac{\frac{\partial E[U(\tilde{W})]}{\partial \sigma}}{\frac{\partial E[U(\tilde{W})]}{\partial \mu}} = \frac{-2c\sigma}{2c\mu + b},$$

is positive, and approaches 0 as $\sigma \to 0^+$.

Second-order:

$$\frac{\partial^2 E[U(\tilde{W})]}{\partial \sigma^2} = 2c < 0, \ \frac{\partial^2 E[U(\tilde{W})]}{\partial \mu^2} = 2c < 0, \ \frac{\partial^2 E[U(\tilde{W})]}{\partial \mu \partial \sigma} = 0.$$

The function is concave, thus it is also quasi-concave.

Indifference curves from normally distributed $\tilde{W}$

Let $f(\varepsilon) \equiv (1/\sqrt{2\pi})e^{-\varepsilon^2/2}$, the std. normal density function. Let $W = \mu + \sigma\varepsilon$, so that $\tilde{W}$ is $N(\mu, \sigma^2)$.

Define expected utility as a function

$$E[U(\tilde{W})] = V(\mu, \sigma) = \int_{-\infty}^{\infty} U(\mu + \sigma\varepsilon) f(\varepsilon) d\varepsilon.$$

Slope of indifference curves:

$$-\frac{\frac{\partial V}{\partial \sigma}}{\frac{\partial V}{\partial \mu}} = \frac{-\int_{-\infty}^{\infty} U'(\mu + \sigma\varepsilon)\varepsilon f(\varepsilon) d\varepsilon}{\int_{-\infty}^{\infty} U'(\mu + \sigma\varepsilon) f(\varepsilon) d\varepsilon}.$$

Denominator always positive. Will show that integral in numerator is negative, so minus sign makes the whole fraction positive.

Integration by parts: Observe $f'(\varepsilon) = -\varepsilon f(\varepsilon)$. Thus:

$$\int U'(\mu + \sigma\varepsilon)\varepsilon f(\varepsilon) d\varepsilon = -U'(\mu + \sigma\varepsilon) f(\varepsilon) + \int U''(\mu + \sigma\varepsilon)\sigma f(\varepsilon) d\varepsilon.$$

First term on RHS vanishes in limit when $\varepsilon \to \pm\infty$, so that

$$\int_{-\infty}^{\infty} U'(\mu + \sigma\varepsilon)\varepsilon f(\varepsilon) d\varepsilon = \int_{-\infty}^{\infty} U''(\mu + \sigma\varepsilon)\sigma f(\varepsilon) d\varepsilon < 0.$$

Another important observation:

$$\lim_{\sigma \to 0^+} \frac{d\mu}{d\sigma} = \frac{-U'(\mu) \int_{-\infty}^{\infty} \varepsilon f(\varepsilon) d\varepsilon}{U'(\mu) \int_{-\infty}^{\infty} f(\varepsilon) d\varepsilon} = 0.$$

To show concavity of $V()$:

$$\begin{aligned} &\lambda V(\mu_1, \sigma_1) + (1-\lambda)V(\mu_2, \sigma_2) \\ = &\int_{-\infty}^{\infty} [\lambda U(\mu_1 + \sigma_1\varepsilon) + (1-\lambda)U(\mu_2 + \sigma_2\varepsilon)] f(\varepsilon) d\varepsilon \\ < &\int_{-\infty}^{\infty} U(\lambda\mu_1 + \lambda\sigma_1\varepsilon + (1-\lambda)\mu_2 + (1-\lambda)\sigma_2\varepsilon) f(\varepsilon) d\varepsilon \\ = &V(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda\sigma_1 + (1-\lambda)\sigma_2). \end{aligned}$$

The function is concave, thus it is also quasi-concave.

D&D, sect. 6.3–6.6 and Roll's appendix

Presentation here will follow Roll's appendix, pp. 158–165. Will indicate correspondence with equations in D&D.

Main points:

- Consider n risky assets, $n > 2$, no risk free asset. Then the frontier portfolio set is an hyperbola. (Mentioned without proof on p. 11 of 25 August.)
- Can derive version of CAPM without risk free asset. Important if, e.g., there is uncertain inflation.

The version mentioned in the second point is important to understand Roll's main text, and much of the CAPM literature.

Market portfolio plays important role also in that version of the model, even though it is *not* equal to the risky part of everyone's portfolio.

Differentiation of vectors and matrices

(See chapter 23 of the math manual by Sydsæter, Strøm and Berck.)

- Derivative of a (scalar) function with respect to an $n \times 1$ vector is the $n \times 1$ vector of derivatives with resp. to each element.
- Derivative of a (scalar) function with respect to an $1 \times n$ vector is the $1 \times n$ vector of derivatives with resp. to each element.
$$x = (x_1, \ldots, x_n) \Longrightarrow \frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right).$$
- Derivative of scalar product: a and x both are $n \times 1$:
$$\frac{\partial}{\partial x}(a' \cdot x) = a'. \quad (' \text{ denotes transpose.})$$
(Generalizes scalar $\partial(b \cdot y)/\partial y = b$.)
- Derivative of $m \times 1$ vector w.r.t. $n \times 1$ vector is $m \times n$ matrix of derivatives.
- Derivative of product of matrix and vector: A is $m \times n$, x is $n \times 1$:
$$\frac{\partial}{\partial x}(Ax) = A.$$
- Derivative of quadratic form: A is $n \times n$, x is $n \times 1$:
$$\frac{\partial}{\partial x}(x'Ax) = x'(A + A').$$
(Generalizes scalar $\partial(b \cdot y^2)/\partial y = 2b \cdot y$.)
- Symmetric version of the same: $A = A'$ is $n \times n$ symmetric:
$$\frac{\partial}{\partial x}(x'Ax) = 2x'A.$$

Derivation of frontier portfolio set

(Roll uses "the efficient portfolio frontier" or just the "efficient set.")

$$\text{Defined by } \min_X \sigma_p^2 \text{ for any expected } r_p.$$

- Exist n risky assets (securities).
- $X_p = \|x_{ip}\|$, the $n \times 1$ vector of portfolio weights.
- Must have $X_p'\iota \equiv \Sigma_{i=1}^n x_{ip} = 1$ (where $\iota \equiv (1, 1, \ldots, 1)'$).
- Fundamental data, exogenous in minimization problem, are
  - $R = \|r_i\|$, $\quad n \times 1$ vector of mean rates of return
  - $V = \|\sigma_{ij}\|$, $\quad n \times n$ cov. matrix of rates of return

  These can be either population values (from a probability distribution) *or they can be sample product moments* (from a frequency distribution) (Roll, p. 159). Important for argument in main text of Roll's article.
- Mean of (r.o.r. of) portfolio p is $r_p = X'R = \Sigma_{i=1}^n x_i r_i$.
- Variance of (r.o.r. of) portfolio p is $\sigma_p^2 = X'VX = \Sigma_{i=1}^n \Sigma_{j=1}^n x_i x_j \sigma_{ij}$.
- Covar. of (r.o.r. of) 2 pf.s is $\sigma_{p_1 p_2} = X_{p_1}' V X_{p_2} = \Sigma_{i=1}^n \Sigma_{j=1}^n x_{ip_1} x_{jp_2} \sigma_{ij}$.
- Use matrix notation in solution for frontier portfolio set:
- For any value of r_p: Choose X to obtain minimum σ_p^2.
- Lagrangian $L = X'VX - \lambda_1(X'R - r_p) - \lambda_2(X'\iota - 1)$.
- F.o.c.: $\partial L/\partial X = 2VX - \lambda_1 R - \lambda_2 \iota = 0$. (DD6.8)

- F.o.c. imply $VX = \frac{1}{2}(\lambda_1 R + \lambda_2 \iota)$, which implies (A.6):

$$X = \frac{1}{2} V^{-1} (R \;\; \iota) \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}. \quad \text{(DD6.11a)}$$

- Here: X expressed as function of $R, V, \lambda_1, \lambda_2$.
- Want instead: X expressed as function of R, V, r_p.
- Roll shows (p. 160) this can be done, provided that not all assets have identical expected returns, and that a risk free portfolio cannot be constructed from the set of n risky assets.
  - Define $A \equiv (R \;\; \iota)' V^{-1} (R \;\; \iota)$, a 2×2 matrix.
  - Then we shall prove that

$$X = V^{-1}(R \;\; \iota) A^{-1} \begin{pmatrix} r_p \\ 1 \end{pmatrix}.$$

- Proof: Premultiply equation (A.6) by $(R \;\; \iota)'$; use the definition of A:

$$(R \;\; \iota)' X = \frac{1}{2} (R \;\; \iota)' V^{-1} (R \;\; \iota) \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \frac{1}{2} A \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}.$$

  Premultiply by A^{-1}:

$$A^{-1} (R \;\; \iota)' X = \frac{1}{2} A^{-1} A \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}.$$

  Plug this into equation on top of page; use $(R \;\; \iota)' X = (r_p \;\; 1)'$ (which follows from definition of r_p and X):

$$X = V^{-1} (R \;\; \iota) A^{-1} (R \;\; \iota)' X = V^{-1} (R \;\; \iota) A^{-1} \begin{pmatrix} r_p \\ 1 \end{pmatrix}.$$

Constants which characterize the frontier portfolio set

- Matrix A summarizes useful information about R and V:

$$A = \begin{pmatrix} r_1 & \cdots & r_n \\ 1 & \cdots & 1 \end{pmatrix} V^{-1} \begin{pmatrix} r_1 & 1 \\ \vdots & \vdots \\ r_n & 1 \end{pmatrix} \equiv \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

where $a = R'V^{-1}R, \;\; b = R'V^{-1}\iota, \;\; c = \iota' V^{-1}\iota.$

(The scalar constants a, b, c are called B, A, and C, respectively, in D&D, p. 108.)

- Will also use

$$A^{-1} = \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}.$$

- Proof of this is straightforward:

$$A \cdot A^{-1} = \frac{1}{ac - b^2} \begin{pmatrix} ca - b^2 & 0 \\ 0 & -b^2 + ac \end{pmatrix} = I_2.$$

Roll's Corollaries 1 and 2 (partly)

- Can now show that frontier portfolio set is parabola in (σ^2, r) diagram, thus hyperbola in (σ, r) diagram.
- Plug formula for frontier portfolio vector (as function of r_p, cf. middle of previous page) into general formula for portfolio variance (Roll's equation (A.11)):

$$\sigma_p^2 = X'VX = (r_p \;\; 1)A^{-1}(R \;\; \iota)'V^{-1}VV^{-1}(R \;\; \iota)A^{-1}(r_p \;\; 1)'$$

$$= (r_p \;\; 1)A^{-1}AA^{-1}(r_p \;\; 1)' = (r_p \;\; 1)A^{-1}(r_p \;\; 1)'$$

$$= (r_p \;\; 1)\frac{1}{ac - b^2}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}(r_p \;\; 1)' = \frac{a - 2br_p + cr_p^2}{ac - b^2}. \quad \text{(DD6.17)}$$

- Minimization w.r.t. r_p gives the minimum variance portfolio:

$$-b + cr_p = 0 \Longrightarrow r_p = \frac{b}{c}.$$

Roll's Corollaries 3 and 3A

- For every frontier portfolio, except the min-variance portfolio, exists another frontier portfolio with an uncorrelated rate of return.
- (The importance will become clear later.)
- Proof: Use general covariance formula, set equal to zero, solve.
- Call the uncorrelated portfolio X_z, z for "zero-beta."
- Covariance is (use similar transformations as for variance above):

$$X_z' V X_p = (r_z \;\; 1)\frac{1}{ac-b^2}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}\begin{pmatrix} r_p \\ 1 \end{pmatrix} \quad \text{(DD6.18)}$$

  which is equal to zero if and only if $r_z = (a - br_p)/(b - cr_p)$.

- Roll also proves (p. 163) that one of $((\sigma_p, r_p), (\sigma_z, r_z))$ is on upper half of hyperbola, the other is on the lower half, i.e.: $r_z < (b/c) < r_p$ or $r_p < (b/c) < r_z$.
- Then he shows (second part of Corollary 3A): In (σ, r) diagram, a tangent to the frontier portfolio set at (σ_p, r_p) intersects the r axis at r_z.

Roll's proof of Corollary 3.A

Proof of second part of Corollary 3.A (pp. 163–164) is misleading.

The slope of a tangent to the efficient set in the (σ, r)-diagram is the derivative of the *square root* of equation (A.11), i.e., of σ_p, not of σ_p^2.

We need to take the *inverse* of this derivative, since the derivative itself would be the slope to the tangent when r_p is on the horizontal axis, while we want σ_p on the horizontal axis.

We find

$$\frac{d\sigma_p}{dr_p} = \frac{1}{\sqrt{ac-b^2}} \cdot \frac{-2b+2cr_p}{2\sqrt{a-br_p+cr_p^2}}.$$

The second part of Corollary 3.A states that a tangent (with a slope equal to the inverse of this expression) through the point (σ_p, r_p), which is the location of some (any) efficient portfolio, should intersect the vertical axis at r_z, which is known from (A.15),

$$r_z = (a - br_p)/(b - cr_p).$$

Before we know the value of r at the intersection, let us just call it r_i, which must satisfy $(r_p - r_i)/\sigma_p = (d\sigma_p/dr_p)^{-1}$.

We can solve this equation for r_i and plug in the square root of (A.11) for σ_p and the expression we just found for the derivative.

Some terms cancel out and we find that $r_i = (a - br_p)/(b - cr_p)$, which is the expression for r_z from (A.15).

Roll's Corollary 5

- All frontier portfolios can be written as linear combination of any pair of (different) frontier portfolios.
- Follows from the linearity of the frontier portfolio equation

$$X = V^{-1}(R \;\; \iota)A^{-1}\begin{pmatrix} r_p \\ 1 \end{pmatrix},$$

  where all terms on the RHS are constants, except the r_p.

- If X_{p1} and X_{p2} are frontier portfolios, and we want the frontier portfolio with expected rate of return equal to $r_3 = \alpha r_1 + (1-\alpha)r_2$, then we compose the portfolio

$$X_{p3} = V^{-1}(R \;\; \iota)A^{-1}\begin{pmatrix} \alpha r_1 + (1-\alpha)r_2 \\ 1 \end{pmatrix} =$$

$$\alpha V^{-1}(R \;\; \iota)A^{-1}\begin{pmatrix} r_1 \\ 1 \end{pmatrix} + (1-\alpha)V^{-1}(R \;\; \iota)A^{-1}\begin{pmatrix} r_2 \\ 1 \end{pmatrix} = \alpha X_{p1} + (1-\alpha)X_{p2}.$$

- Also for linear combination of many frontier portfolios.
- If all weights are between zero and one, then weighted sum will have expected rate of return between the extremes of which the sum consists.

Derivation of the zero-beta CAPM

A version of the CAPM in which there is no riskless asset? From Roll's appendix the derivation of the main results is quite straight forward:

- No riskless asset $\Rightarrow$ everyone holds portfolio on upper half of the hyperbola known as the *frontier portfolio set.*
- The market portfolio is a convex combination of the portfolios of all agents.
- Corollary 5: Any convex combination of portfolios on the upper half of the hyperbola, is itself efficient, i.e., located on the upper half of the hyperbola. $\Rightarrow$ The market portfolio is efficient.
- Can then do exactly the same reasoning as in D&D (pp. 104–105) with the little and the big hyperbola, this time with the slope $(r_m - r_z)/\sigma_m$ instead of $(r_m - r_f)/\sigma_m$ (which D&D call $(\bar{r}_M - r_f)/\sigma_M$ on p. 104).
- Conclusion: See p. 115–116 of D&D, in particular equation (6.29),

$$E(\tilde{r}_j) = E(\tilde{r}_{ZC(M)}) + \beta_{Mj}[E(\tilde{r}_M) - E(\tilde{r}_{ZC(M)})].$$