Statistical Analysis of the CAPM

I. Sharpe–Linter CAPM

Brief Review of the Sharpe–Lintner CAPM

- The Sharpe–Lintner CAPM assumes that

(i) all investors act according to the $\mu - \sigma$ rule,

(ii) can lend and borrow any desired amount at a common risk–free rate r_f,

(iii) and exhibit perfect agreement with respect to the probability distribution of asset returns.

- Under these (key) assumptions, the market portfolio is mean–variance efficient, implying that it is characterized by weight vector

$$x_m = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f \mathbf{1}_N)}{\mathbf{1}_N' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f \mathbf{1}_N)}. \tag{1}$$

- The central equation of the Sharpe–Lintner CAPM is a direct consequence of (1) and is given by

$$\mu_i - r_f = \beta_i(\mu_m - r_f), \quad i = 1, \ldots, N, \qquad (2)$$

  where

  - r_f is the risk–free rate,
  - μ_m is the expected return of the market portfolio, and
  - $\beta_i = COV(R_i, R_m)/\sigma_m^2$, where
  - R_i is the return of asset i and R_m is the return of the market portfolio.

- Equation (2) states that there is a linear relation between the *excess* return of asset i (over the risk–free) rate and the *excess* return of the market portfolio, with zero intercept.

- Equation (2) also *implies* efficiency.

Framework for Estimation and Testing

- The CAPM relationship (2) is expressed in terms of expected values, which are not observable.

- To obtain a model with observable quantities, we describe *excess returns* using the *excess return market model*:

$$r_{it} = \alpha_i + \beta_i r_{m,t} + \epsilon_{it} \quad i = 1, \dots, N \tag{3}$$

$$E(\epsilon_{it}) = 0, \quad i = 1, \dots, N \tag{4}$$

$$E(\epsilon_{it}\epsilon_{jt'}) = \begin{cases} \sigma_{ij} & \text{if } t = t' \\ 0 & \text{if } t \neq t' \end{cases} \quad i, j = 1, \dots, N \tag{5}$$

$$E(r_{m,t}\epsilon_{i,t}) = 0, \quad i = 1, \dots, N. \tag{6}$$

- Here r_{it} is the *excess return* on asset i in period t (over risk–free rate), and $r_{m,t}$ is the *excess return* on the market portfolio in period t (over risk–free rate).

- At first glance, the market model we will be using looks similar to the Single–Index Model (SIM), but there are important differences:

  – All returns involved are *excess returns* over the risk–free rate r_f.

  – According to equation (5), the asset–specific error terms may be correlated. Thus, we allow for a non-diagonal covariance matrix, $\mathbf{\Sigma}$, of the vector $\boldsymbol{\epsilon}_t = [\epsilon_{1t}, \ldots, \epsilon_{Nt}]'$,

$$COV(\boldsymbol{\epsilon}_t) = \mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1N} & \sigma_{2N} & \cdots & \sigma_N^2 \end{bmatrix}$$

  Conditional on the excess return of the market, we then also have

$$COV(\boldsymbol{r}_t) = \mathbf{\Sigma}, \qquad (7)$$

  where $\boldsymbol{r}_t = [r_{1t}, \ldots, r_{2t}]'$.

– Note, however, that we still assume that there is no correlation over time, i.e. $E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}'_{t'}) = 0$ for $t \neq t'$, and that the covariance matrix $\boldsymbol{\Sigma}$ is constant over time.

- We will assume that the betas are constant over time.[1]

- We will also assume that the error terms follow a multivariate normal distribution, i.e.,

$$\boldsymbol{\epsilon}_t \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}). \tag{8}$$

- The Sharpe–Lintner CAPM implies that the intercept in the excess return market model is zero, i.e., $\boldsymbol{\alpha} = \mathbf{0}$.

[1]This is by no means self-evident and can, in principle, be tested using econometric techniques for detecting structural breaks. For an overview with a view towards CAPM applications, see Schmid/Trede: Finanzmarktstatistik, Springer.

- That is, a test of this model corresponds to a test of the hypothesis

$$H_0 : \alpha_i = 0, \quad i = 1, \ldots, N. \tag{9}$$

- To perform such a test, it is necessary to estimate the parameters of the model and to derive an appropriate test statistic.

- Write our *excess return market model* as

$$\begin{aligned} \boldsymbol{r}_t &= \boldsymbol{\alpha} + \boldsymbol{\beta} r_{m,t} + \boldsymbol{\epsilon}_t, \quad t = 1, \ldots, T, \\ \boldsymbol{\epsilon}_t &\overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}), \end{aligned}$$

where $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]'$, and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_N]'$.

- The density of excess returns, conditional on the market return, $r_{m,t}$, is

$$f(\boldsymbol{r}_t | r_{m,t}) = \frac{\exp\left\{-\frac{1}{2}(\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t})\right\}}{(2\pi)^{N/2} |\boldsymbol{\Sigma}|^{1/2}},$$

and the joint density is

$$f(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T | r_{m,1}, \ldots, r_{T,1}) = \prod_{t=1}^{T} f(\boldsymbol{r}_t | r_{m,t}) = \frac{\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t})\right\}}{(2\pi)^{NT/2} |\boldsymbol{\Sigma}|^{T/2}} \tag{10}$$

- To estimate the unknown parameters, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\Sigma}$, of this density, we use the method of *maximum likelihood*.

- To do so, we define the log–likelihood function, i.e., the log of the joint density viewed as a function of the unknown parameters.

- The maximum likelihood estimator is then found by maximizing this function with respect to its arguments, i.e., the unknown parameters.

- From (10), the log–likelihood function is

$$
\begin{aligned}
& \log L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) \qquad (11) \\
& = -\frac{NT}{2} \log(2\pi) - \frac{T}{2} \log|\boldsymbol{\Sigma}| \\
& -\frac{1}{2} \sum_{t=1}^{T} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t})' \boldsymbol{\Sigma}^{-1} \\
& \times (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t}),
\end{aligned}
$$

which we want to maximize with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.

- From (11), it is clear that the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are determined by minimizing

$$\begin{aligned} S &= \sum_{t=1}^{T} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t}) \\ &= \sum_{t=1}^{T} \left\{ \boldsymbol{r}_t' \boldsymbol{\Sigma}^{-1} \boldsymbol{r}_t - 2 \boldsymbol{r}_t' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\alpha} + \boldsymbol{\beta} r_{m,t}) \right. \\ &\qquad \left. + (\boldsymbol{\alpha} + \boldsymbol{\beta} r_{m,t})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\alpha} + \boldsymbol{\beta} r_{m,t}) \right\}. \end{aligned}$$

- The first order conditions are

$$\begin{aligned} \frac{\partial S}{\partial \boldsymbol{\alpha}} &= -2 \boldsymbol{\Sigma}^{-1} \sum_{t=1}^{T} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t}) = 0, \\ \frac{\partial S}{\partial \boldsymbol{\beta}} &= -2 \boldsymbol{\Sigma}^{-1} \sum_{t=1}^{T} r_{m,t} (\boldsymbol{r}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} r_{m,t}) = 0. \end{aligned}$$

- We get the standard OLS estimators.

- That is,

$$\widehat{\boldsymbol{\alpha}} = \bar{\boldsymbol{r}} - \widehat{\boldsymbol{\beta}}\bar{r}_m$$

and

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= \frac{\sum_{t=1}^{T}(\boldsymbol{r}_t - \bar{\boldsymbol{r}})(r_{m,t} - \bar{r}_m)}{\sum_{t=1}^{T}(r_{m,t} - \bar{r}_m)^2} \\
&= \frac{\sum_{t=1}^{T}(\boldsymbol{r}_t - \bar{\boldsymbol{r}})(r_{m,t} - \bar{r}_m)}{T\widehat{\sigma}_m^2} \\
&= \frac{\sum_{t=1}^{T}(r_{m,t} - \bar{r}_m)\boldsymbol{r}_t}{T\widehat{\sigma}_m^2}
\end{aligned}$$

where

$$\begin{aligned}
\bar{\boldsymbol{r}} &= \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{r}_t, \quad \bar{r}_m = \frac{1}{T}\sum_{t=1}^{T}r_{m,t}, \\
\widehat{\sigma}_m^2 &= \frac{1}{T}\sum_{t=1}^{T}(r_{m,t} - \bar{r}_m)^2.
\end{aligned}$$

- To find the MLE of $\boldsymbol{\Sigma}$, we make use of the following differentiation rules:

  (i) Let $\boldsymbol{X}$ and $\boldsymbol{A}$ be $n \times n$ matrices, so that

$$\text{tr}(\boldsymbol{X}\boldsymbol{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} a_{ji}. \tag{12}$$

  For *symmetric* $\boldsymbol{X}$ ($x_{ij} = x_{ji}$), we therefore have

$$\frac{\partial \text{tr}(\boldsymbol{X}\boldsymbol{A})}{\partial x_{ij}} = \begin{cases} a_{ii} & i = j \\ a_{ij} + a_{ji} & i \neq j. \end{cases} \tag{13}$$

  Hence

$$\frac{\partial \text{tr}(\boldsymbol{X}\boldsymbol{A})}{\partial \boldsymbol{X}} = \boldsymbol{A} + \boldsymbol{A}' - \text{diag}(\boldsymbol{A}),$$

  and

$$\frac{\partial \text{tr}(\boldsymbol{X}\boldsymbol{A})}{\partial \boldsymbol{X}} = 2\boldsymbol{A} - \text{diag}(\boldsymbol{A}), \tag{14}$$

  if $\boldsymbol{A}$ is also symmetric.

(ii) To find an expression for the derivative of $|\boldsymbol{X}|$, recall that

$$|\boldsymbol{X}| = x_{i1}C_{i1} + x_{i2}C_{i2} + \cdots + x_{in}C_{in}, \quad (15)$$

where C_{ij} is the cofactor of x_{ij} in $\boldsymbol{X}$, $i, j = 1, \ldots, n$.
Again for symmetric $\boldsymbol{X}$, we thus have

$$\begin{aligned} \frac{\partial \log |\boldsymbol{X}|}{\partial \boldsymbol{X}} &= \frac{1}{|\boldsymbol{X}|} \frac{\partial |\boldsymbol{X}|}{\partial \boldsymbol{X}} \\ &= 2\boldsymbol{X}^{-1} - \text{diag}(\boldsymbol{X}^{-1}). \quad (16) \end{aligned}$$

The log–likelihood function can be written as

$$\begin{aligned}\log L &= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Sigma}| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T} \text{tr}\left(\widehat{\boldsymbol{\epsilon}}'\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\epsilon}}\right) \\ &= -\frac{NT}{2}\log(2\pi) + \frac{T}{2}\log|\boldsymbol{\Sigma}^{-1}| \quad (17) \\ &\quad -\frac{1}{2}\sum_{t=1}^{T} \text{tr}\left(\boldsymbol{\Sigma}^{-1}\widehat{\boldsymbol{\epsilon}}\widehat{\boldsymbol{\epsilon}}'\right), \quad (18)\end{aligned}$$

where

– $\widehat{\boldsymbol{\epsilon}}_t = \boldsymbol{r}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}} r_{m,t}$,
– (17) uses $|A^{-1}| = |A|^{-1}$, and
– (18) uses the permutation rule $\text{tr}(ABC) = \text{tr}(BCA)$.

Thus, using (14) and (16), we require

$$\frac{\partial \log L}{\partial \boldsymbol{\Sigma}^{-1}} = \frac{T}{2}[2\boldsymbol{\Sigma} - \text{diag}(\boldsymbol{\Sigma})] - \frac{1}{2}\left[2\sum_{t=1}^{T}\widehat{\boldsymbol{\epsilon}}_t\widehat{\boldsymbol{\epsilon}}_t' - \text{diag}\left(\sum_{t=1}^{T}\widehat{\boldsymbol{\epsilon}}_t\widehat{\boldsymbol{\epsilon}}_t'\right)\right] = 0,$$

implying

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{\epsilon}}_t\widehat{\boldsymbol{\epsilon}}_t' \quad (19)$$

$$= \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{r}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}}r_{m,t})(\boldsymbol{r}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}}r_{m,t})'.$$

- The OLS estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are unbiased and normally distributed with covariance matrices

$$
\begin{aligned}
COV(\widehat{\boldsymbol{\beta}}) &= \frac{1}{T\widehat{\sigma}_m^4} COV\left\{\sum_{t=1}^{T}(r_{m,t}-\bar{r}_m)\boldsymbol{r}_t\right\} \\
&= \frac{1}{T\widehat{\sigma}_m^4}\sum_{t=1}^{T}(r_{m,t}-\bar{r}_m)^2 COV(r_t) \\
&= \frac{1}{T\widehat{\sigma}_m^2}\boldsymbol{\Sigma}, \quad \text{and}
\end{aligned}
$$

$$
\begin{aligned}
COV(\widehat{\boldsymbol{\alpha}}) &= COV(\bar{\boldsymbol{r}} - \widehat{\boldsymbol{\beta}}\bar{r}_m) \\
&= COV\left\{\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{r}_t - \bar{r}_m\sum_{t=1}^{T}\frac{(r_{m,t}-\bar{r}_m)\boldsymbol{r}_t}{T\widehat{\sigma}_m^2}\right\} \\
&= \frac{1}{T^2\widehat{\sigma}_m^4}\sum_{t=1}^{T}[\widehat{\sigma}_m^2 - \bar{r}_m(r_{m,t}-\bar{r}_m)]^2 COV(\boldsymbol{r}_t) \\
&= \frac{1}{T^2\widehat{\sigma}_m^4}T[\widehat{\sigma}_m^4 + \bar{r}_m^2\widehat{\sigma}_m^2]\boldsymbol{\Sigma} \\
&= \frac{1}{T}\left(1+\frac{\bar{r}_m^2}{\widehat{\sigma}_m^2}\right)\boldsymbol{\Sigma}. \qquad (20)
\end{aligned}
$$

- It can be shown that $T\widehat{\boldsymbol{\Sigma}}$ has a Wishart distribution, $\mathcal{W}_N(T-2, \boldsymbol{\Sigma})$, which is a matrix generalization of the χ^2 distribution.[2]

  Moreover, $\widehat{\boldsymbol{\Sigma}}$ is independent of both $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\beta}}$.

[2] See, for example, Zellner (1971).

Testing for mean–variance efficiency ($\boldsymbol{\alpha} = \mathbf{0}$)

- We discuss two tests of the null hypothesis $\boldsymbol{\alpha} = 0$, in historical order.

- The first is a likelihood ratio (LR) test relying on asymptotic arguments,

- The second is an exact finite–sample F-test. Subsequently, the relation between the tests will be considered.

Likelihood Ratio (LR) Test

- To conduct the likelihood ratio test, we first compute the Maximum Likelihood Estimator under the null hypothesis that $\boldsymbol{\alpha} = 0$, which is a regression through the origin. Denote the corresponding estimators by $\widehat{\boldsymbol{\beta}}_0$ and $\widehat{\boldsymbol{\Sigma}}_0$. They are given by

$$\widehat{\boldsymbol{\beta}}_0 = \frac{\sum_{t=1}^{T} \boldsymbol{r}_t r_{m,t}}{\sum_{t=1}^{T} r_{m,t}^2}, \tag{21}$$

and

$$\widehat{\boldsymbol{\Sigma}}_0 = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\epsilon}}_t^0 \widehat{\boldsymbol{\epsilon}}_t^0 \tag{22}$$

$$= \frac{1}{T}\sum_{t=1}^{T} (\boldsymbol{r}_t - \widehat{\boldsymbol{\beta}}_0 r_{m,t})(\boldsymbol{r}_t - \widehat{\boldsymbol{\beta}}_0 r_{m,t})',$$

where $\widehat{\boldsymbol{\epsilon}}_t^0 = \boldsymbol{r}_t - \widehat{\boldsymbol{\beta}}_0 r_{m,t}$.

- The Likelihood Ratio Test is based on the comparison between the log–likelihood values of the unconstrained model and the constrained model.

- More precisely, the LR test statistic is given by

$$\mathcal{LR} = -2(\log L_0 - \log L_1), \tag{23}$$

where $\log L_0$ is the log–likelihood function of the constrained model, and $\log L_1$ is the log–likelihood function of the unconstrained model, each evaluated at the respective MLEs.

- The asymptotic distribution of $\mathcal{LR}$ defined in (23) is χ^2 with degrees of freedom equal to the number of parameter restrictions implied by the null hypothesis.

- In our situation, this corresponds to N degrees of freedom (N is the number of assets), because the CAPM implies that $\alpha_i = 0$ for $i = 1, \ldots, N$.

- Now

$$
\begin{aligned}
\log L_1 &= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\widehat{\boldsymbol{\Sigma}}_1| \\
&\quad -\frac{1}{2}\sum_{t=1}^{T}\mathsf{tr}\left(\widehat{\boldsymbol{\Sigma}}_1^{-1}\widehat{\boldsymbol{\epsilon}\boldsymbol{\epsilon}}'\right) \\
&= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\widehat{\boldsymbol{\Sigma}}_1| \\
&\quad -\frac{1}{2}\sum_{t=1}^{T}\mathsf{tr}\left\{\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{\epsilon}}_t\widehat{\boldsymbol{\epsilon}}_t'\right)^{-1}\widehat{\boldsymbol{\epsilon}\boldsymbol{\epsilon}}'\right\} \\
&= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\widehat{\boldsymbol{\Sigma}}_1| \\
&\quad -\frac{T}{2}\mathsf{tr}\left\{\left(\sum_{t=1}^{T}\widehat{\boldsymbol{\epsilon}}_t\widehat{\boldsymbol{\epsilon}}_t'\right)^{-1}\sum_{t=1}^{T}\widehat{\boldsymbol{\epsilon}\boldsymbol{\epsilon}}'\right\} \\
&= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\widehat{\boldsymbol{\Sigma}}_1| - \frac{T}{2}\mathsf{tr}(\boldsymbol{I}_N) \\
&= -\frac{NT}{2}(\log(2\pi)+1) - \frac{T}{2}\log|\widehat{\boldsymbol{\Sigma}}_1|.
\end{aligned}
$$

- By the same line of arguments,

$$\log L_0 = -\frac{NT}{2}(\log(2\pi) + 1) - \frac{T}{2} \log |\widehat{\mathbf{\Sigma}}_0|.$$

Consequently,

$$\mathcal{LR} = T \left[\log |\widehat{\mathbf{\Sigma}}_0| - \log |\widehat{\mathbf{\Sigma}}_1| \right] \overset{asy}{\sim} \chi^2(N). \quad (24)$$

F Test

- The finite–sample F test is based on the following result:

  Result: If N–dimensional random variable X is $N(0, \Omega)$, the $N \times N$ random matrix A is Wishart(T, Ω), and X and A are independent, then

$$\frac{T - N + 1}{N} X' A^{-1} X \sim F_{N, T-N+1}, \qquad (25)$$

  i.e., the quantity $[(T - N + 1)/N] X' A^{-1} X$ has an F distribution with N degrees of freedom in the numerator and $T - N + 1$ degrees of freedom in the denominator.

- Using, in (25),

$$X = \sqrt{T}[1 + \bar{r}_m^2/\widehat{\sigma}_m^2]^{-1/2}\widehat{\alpha} \tag{26}$$

and

$$A = T\widehat{\mathbf{\Sigma}}, \tag{27}$$

and recalling the results we have for $\widehat{\boldsymbol{\alpha}}$ (in particular, normality and (20)), the statistic

$$J = \frac{T-N-1}{N}\left(1 + \frac{\bar{r}_m^2}{\widehat{\sigma}_m^2}\right)^{-1} \widehat{\boldsymbol{\alpha}}'\widehat{\mathbf{\Sigma}}^{-1}\widehat{\boldsymbol{\alpha}} \tag{28}$$

has an F distribution with N degrees of freedom in the numerator and $T-N-1$ degrees of freedom in the denominator, i.e.,[3]

$$J \sim F_{N,T-N-1}. \tag{29}$$

[3]Gibbons/Ross/Shanken (1989): A Test of the Efficiency of a Given Portfolio. *Econometrica 57*, 1121-1152.

Economic Interpretation of the CAPM F Test

- Apart from following a known finite–sample distribution, the test statistic J defined in (28) also has economic interpretation.

- Recall that the key testable implication of the CAPM is that the market portfolio is a $\mu - \sigma$ efficient portfolio.

- In the presence of a risk–free rate, this means that the market portfolio is the tangency portfolio.

- It can be shown that[4]

$$J = \left(\frac{T - N - 1}{N} \right) \frac{\widehat{\theta}^{\star 2} - \widehat{\theta}_m^2}{1 + \widehat{\theta}_m^2}, \qquad (30)$$

  where $\widehat{\theta}^{\star}$ is the Sharpe ratio of the *ex post* (i.e., using the sample mean vector and the sample covariance matrix) efficient portfolio formed from the risky assets under study (including our market proxy) and $\widehat{\theta}_m$ is the Sharpe ratio of the portfolio used as a market proxy in our analysis.

- Equation (30) is particularly interesting because it uncovers what we are actually testing: We test whether our market proxy is so far away from the *ex post* efficient portfolio that we are not willing to believe that it is the *population* tangency portfolio, where the distance is measured in terms of the Sharpe Ratio.

[4] Gibbons/Ross/Shanken (1989): A Test of the Efficiency of a Given Portfolio. *Econometrica 57*, 1121-1152.

Proof of (30)

- Comparing (28) and (30), the equality between these quantities follows if we show that $\widehat{\boldsymbol{\alpha}}'\widehat{\boldsymbol{\Sigma}}^{-1}\widehat{\boldsymbol{\alpha}} = \widehat{\theta}^{\star 2} - \widehat{\theta}_m^2$.

- Let $\tilde{r} = [\bar{r}_m, \bar{r}']'$. The (sample) covariance matrix of these variables is

$$V = \begin{bmatrix} \widehat{\sigma}_m^2 & \widehat{\sigma}_m^2 \widehat{\beta}' \\ \widehat{\sigma}_m^2 \widehat{\beta} & \widehat{\Sigma} + \widehat{\sigma}_m^2 \widehat{\beta}\widehat{\beta}' \end{bmatrix}. \tag{31}$$

- We know that the efficient portfolio using the assets in $\tilde{r}$ is characterized by the weight vector

$$w = \frac{V^{-1}\tilde{r}}{1'V^{-1}\tilde{r}}, \tag{32}$$

and, thus, it has squared Sharpe ratio

$$\widehat{\theta}^{\star 2} = \frac{(w'\tilde{r})^2}{w'Vw} = \frac{(\tilde{r}'V^{-1}\tilde{r})^2}{\tilde{r}'V^{-1}\tilde{r}} = \tilde{r}'V^{-1}\tilde{r}. \tag{33}$$

- Next, it is easily checked that the inverse of (31) is

$$V^{-1} = \left[\begin{array}{cc} \widehat{\sigma}_m^{-2} + \widehat{\beta}'\widehat{\Sigma}^{-1}\widehat{\beta} & -\widehat{\beta}'\widehat{\Sigma}^{-1} \\ -\widehat{\Sigma}^{-1}\widehat{\beta} & \widehat{\Sigma}^{-1} \end{array} \right] \qquad (34)$$

- Using (34), we get by straightforward computation, and using (33),

$$\begin{aligned} \widehat{\theta}^{\star 2} &= \tilde{r}'V^{-1}\tilde{r} = [\bar{r}_m, \bar{r}']V^{-1}[\bar{r}_m, \bar{r}']' \\ &= \frac{\bar{r}_m^2}{\widehat{\sigma}_m^2} + (\bar{r} - \widehat{\beta}\bar{r}_m)'\widehat{\Sigma}^{-1}(\bar{r} - \widehat{\beta}\bar{r}_m) \\ &= \widehat{\theta}_m^2 + \widehat{\alpha}'\widehat{\Sigma}^{-1}\widehat{\alpha}, \end{aligned}$$

recalling that $\widehat{\alpha} = \bar{r} - \widehat{\beta}\bar{r}_m$.

Relation between F and LR tests

- The finite–sample F test can also be interpreted as a likelihood ratio test.

- To see this, first note that for the unconstrained MLE of $\boldsymbol{\beta}$, denoted by $\widehat{\boldsymbol{\beta}}_1$,

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_1 &= \frac{\sum_t (r_{m,t} - \bar{r}_m)\boldsymbol{r}_t}{T\widehat{\sigma}_m^2} = \frac{\sum_t r_{m,t}\boldsymbol{r}_t - \bar{r}_m \sum_t \boldsymbol{r}_t}{T\widehat{\sigma}_m^2} \\
&= \frac{\overline{r_m^2}}{\widehat{\sigma}_m^2}\widehat{\boldsymbol{\beta}}_0 - \frac{\bar{r}_m \bar{\boldsymbol{r}}}{\widehat{\sigma}_m^2} \\
&= \frac{\overline{r_m^2}}{\widehat{\sigma}_m^2}\widehat{\boldsymbol{\beta}}_0 - \frac{\bar{r}_m}{\widehat{\sigma}_m^2}(\bar{\boldsymbol{r}} - \widehat{\boldsymbol{\beta}}_1 \bar{r}_m) - \frac{\bar{r}_m^2}{\widehat{\sigma}_m^2}\widehat{\boldsymbol{\beta}}_1 \\
&= \frac{\overline{r_m^2}}{\widehat{\sigma}_m^2}\widehat{\boldsymbol{\beta}}_0 - \frac{\bar{r}_m}{\widehat{\sigma}_m^2}\widehat{\boldsymbol{\alpha}} - \frac{\bar{r}_m^2}{\widehat{\sigma}_m^2}\widehat{\boldsymbol{\beta}}_1.
\end{aligned}
$$

Rearranging and using the basic identity

$$
\widehat{\sigma}_m^2 = \overline{r_m^2} - \bar{r}_m^2 = \frac{1}{T}\sum_{t=1}^{T} r_{m,t}^2 - \bar{r}_m^2,
$$

shows that

$$\widehat{\boldsymbol{\beta}}_0 = \widehat{\boldsymbol{\beta}}_1 + \frac{\bar{r}_m}{\widehat{\sigma}_m^2 + \bar{r}_m^2}\widehat{\boldsymbol{\alpha}}. \tag{35}$$

Inserting (35) into $\widehat{\boldsymbol{\Sigma}}_0$ (see equation (22)) and noting that the normal equations (12) and (12) imply

$$\sum_{t=1}^{T}(\boldsymbol{r}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}}_1 r_{m,t})' \left(1 - \frac{\bar{r}_m r_{m,t}}{\bar{r}_m^2 + \widehat{\sigma}_m^2}\right) \widehat{\boldsymbol{\alpha}} = 0,$$

we arrive at

$$\widehat{\boldsymbol{\Sigma}}_0 = \widehat{\boldsymbol{\Sigma}}_1 + \left(\frac{\widehat{\sigma}_m^2}{\bar{r}_m^2 + \widehat{\sigma}_m^2}\right) \widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\alpha}}'. \tag{36}$$

- The *Sherman–Morrison formula for the determinant* is as follows: For nonsingular $\boldsymbol{A}$ and conformable vectors $\boldsymbol{u}$ and $\boldsymbol{v}$

$$|\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}'| = |\boldsymbol{A}|(1 + \boldsymbol{v}'\boldsymbol{A}^{-1}\boldsymbol{u}). \tag{37}$$

- Formula (37) can be shown as follows.

- Consider first the case $\boldsymbol{A} = \boldsymbol{I}$.

- Then, since

$$\begin{pmatrix} \boldsymbol{I} & \boldsymbol{u} \\ \boldsymbol{0} & 1 + \boldsymbol{v}'\boldsymbol{u} \end{pmatrix}$$
$$= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{v}' & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}' & \boldsymbol{u} \\ \boldsymbol{0} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{v}' & 1 \end{pmatrix},$$

we have

$$\det(\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}') = 1 + \boldsymbol{v}'\boldsymbol{u}. \qquad (38)$$

- Next, recalling that $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$,

$$\begin{aligned} \det(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}') &= \det(\boldsymbol{A})\left(\boldsymbol{I} + \boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}'\right) \\ &= \det(\boldsymbol{A})\left(\boldsymbol{I} + (\boldsymbol{A}^{-1}\boldsymbol{u})\boldsymbol{v}'\right) \\ &= \det(\boldsymbol{A})\left(1 + \boldsymbol{v}'\boldsymbol{A}^{-1}\boldsymbol{u}\right). \end{aligned}$$

When this formula is applied to (36), we obtain

$$|\widehat{\mathbf{\Sigma}}_0| = |\widehat{\mathbf{\Sigma}}_1| \left[1 + \frac{\widehat{\sigma}_m^2}{\bar{r}_m^2 + \widehat{\sigma}_m^2} \widehat{\boldsymbol{\alpha}}' \widehat{\mathbf{\Sigma}}_1^{-1} \widehat{\boldsymbol{\alpha}}\right]$$

Thus, (24) may be written as

$$\begin{aligned} \mathcal{LR} &= T \log \frac{|\widehat{\mathbf{\Sigma}}_0|}{|\widehat{\mathbf{\Sigma}}_1|} \\ &= T \log \left[1 + \frac{\widehat{\sigma}_m^2}{\bar{r}_m^2 + \widehat{\sigma}_m^2} \widehat{\boldsymbol{\alpha}}' \widehat{\mathbf{\Sigma}}_1^{-1} \widehat{\boldsymbol{\alpha}}\right] \\ &= T \log \left[\frac{N}{T - N - 1} J + 1\right], \end{aligned}$$

where J is the F–statistic given by (28), or, equivalently,

$$J = \frac{T - N - 1}{N} \left[\exp\left\{\frac{\mathcal{LR}}{T}\right\} - 1\right], \qquad (39)$$

which, as (39) is a monotonic transformation of $\mathcal{LR}$, shows that J may also be interpreted as a likelihood ratio test.

- As the F–test based on (28) is exact, it is, for realistic sample sizes, clearly preferable compared to the likelihood ratio test relying on asymptotic arguments.

- However, for the zero–beta version, an exact test is much more difficult to obtain, and it may be useful to consider what is lost when relying on asymptotic arguments.

Finite–sample size of likelihood ratio test for nominal size 5%

$N =$ number of assets,
$T =$ sample size

For example, for $N = 10$, and $T = 60$, the critical value for a LRT with asymptotic size 5% is 18.307, which corresponds to a critical value of the exact F–test of

$$\begin{aligned} c_F &= \frac{T-N-1}{N}\left(\exp\left\{\frac{LRT}{T}\right\}-1\right) \\ &= \frac{49}{10}\left(\exp\left\{\frac{18.307}{60}\right\}-1\right) = 1.748. \end{aligned}$$

The actual size of the LRT in this situation is therefore

$$1 - F^{cdf}(1.748; 49, 10) = 0.096.$$

T	$N = 10$	$N = 20$	$N = 40$
60	0.096	0.211	**0.805**
120	0.070	0.105	0.275
180	0.062	0.082	0.164
240	0.059	0.073	0.124
360	0.056	0.064	0.092

- The table shows that the finite–sample size of the tests is larger than the asymptotic size of 5%.

- As a consequence, the large–sample tests will reject the null hypothesis too often.

Roll's Critique

- Roll (1977)[5] emphasizes that tests of the CAPM really only reject the mean–variance efficiency of the market proxy we use in the test (recall equation (30)).

- This implies that the CAPM is essentially untestable, because "*the theory is not testable unless the exact composition of the true market portfolio is known and used in the tests. This implies that the theory is not testable unless all individual assets are included in the sample*".

[5]R. Roll (1977). A Critique of the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics 4*, 129-176.

- Roll argues that using a proxy for the market portfolio is subject to two difficulties: "*First, the proxy itself might be mean–variance efficient even when the true market portfolio is not. This is a real danger since every sample will display efficient portfolios that satisfy perfectly all of the theory's implications. (...) On the other hand, the chosen proxy may turn out to be inefficient; but obviously, this alone implies nothing about the true market portfolio's efficiency*".

- Thus, what we essentially test is a joint hypothesis: The CAPM *and* the hypothsis that the portfolio used in the tests as the market proxy is the true market portfolio.

- Clearly, it is extremely difficult to measure the "market portfolio", because this entity can, in principle, include not just traded financial assets, but also consumer durables, real estate, and human capital.

- On the other hand, often our interest is not to test the CAPM (i.e., efficiency of the market portfolio) but simply whether a specific portfolio is mean–variance efficient within a given universe of assets.

References

- Campbell/Lo/MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press: Princeton.
- E. F. Fama and K. R. French (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, 18, 25–46.
- F. Schmid and M. Trede (2006(?)). Finanzmarktstatistik, Springer, Kapitel 7.
- Zellner, A. (1971). *Introduction to Bayesain Inference in Econometrics.* New York: John Wiley & Sons.