7 | The Capital Asset Pricing Model

The previous handouts have been concerned with how an individual or firm could select an optimum portfolio, or a set of portfolios. If investors act as we have prescribed, then we should be able to draw on the analysis to determine how the aggregate of investors will behave and how prices and returns at which markets will clear are set. The construction of general equilibrium models will allow us to determine the relevant measure of risk for any asset and the relationship between expected return and risk for any asset when markets are in equilibrium. Furthermore, the equilibrium models are derived from models of how portfolios should be constructed, the models themselves have major implications for the characteristics of optimum portfolios.

Sharpe (1964) proposed the capital asset pricing model (CAPM) as an equilibrium theory built on the Modern Portfolio Theory. The CAPM is the simplest form of an equilibrium model. The multifactor model will be discussed in the next handout.

The validity of CAPM depends on a long list of assumptions, we list only some of them.

1. The market prices are “in equilibrium.” In particular, for each asset, supply equals demand.

2. Investors are expected to make decisions solely in terms of expected values and standard deviations of the returns on their portfolios.

3. All investors are assumed to define the relevant period in exactly the same manner and have identical expectations with respect to the necessary inputs to the portfolio decision. This is the assumption of the homogeneity of expectations.

4. Unlimited short sales are allowed. The individual investor can sell short any number of any shares.

5. Unlimited lending and borrowing at the riskfree rate are available. The investor can lend or borrow any amount of funds desired at a rate of interest equal to the rate for riskfree securities.

6. All investors choose portfolios optimally according to the principles of efficient diversification. This implies that everyone holds a tangency portfolio of risky assets as well as the risk-free asset.

7. The market rewards people for assuming unavoidable risk, but there is no reward for needless risks due to inefficient portfolio selection. Therefore, the risk premium on a single security is due to its contribution to the risk of the tangency portfolio.

These assumptions are not likely met, in fact many of the assumptions behind the CAPM are untenable. Despite the stringent assumptions and the simplicity of the CAPM model, it is of great importance and its generalized version does a good job of describing prices in the capital markets.

The Capital Market Line

If all investors have homogeneous expectations and (Assumption 3), then they will each face the same efficient frontier. The portfolio of risky assets P_i held by any investor will be identical to the portfolio of risky assets held by any other investor. If all investors hold the same risky portfolio, then, in equilibrium, it must be the market portfolio. The market portfolio is a portfolio comprising all risky assets. Let M be the Market portfolio and R an efficient portfolio. In equilibrium, all investors will end up with portfolios somewhere along the line called *the capital market line* (CML),

$$\mu_R = \mu_f + \frac{\mu_M - \mu_f}{\sigma_M}\,\sigma_R\,, \tag{7.1}$$

and all efficient portfolios would lie along this line. In (7.1) μ_f, μ_M and σ_M are constant, only σ_R and μ_R vary as we change the efficient portfolio R.

The CML shows how μ_R depends on σ_R. The slope of the CML is the Sharpe ratio of the market portfolio,

$$\frac{\mu_M - \mu_f}{\sigma_M} = \frac{\text{the risk premium of the market portfolio}}{\text{the risk of the market portfolio}} = \frac{\mu_R - \mu_f}{\sigma_R}\,.$$

The last equality is from a rewrite of (7.1). All efficient portfolios have the same Sharpe's ratio as the market portfolio. This slope can be thought of as the market price of risk for all efficient portfolios. It is the extra return that can be gained by increasing the level of risk on an efficient portfolio by one unit. The second term of (7.1) is simply the market price of risk times the amount of risk in a efficient portfolio. This term represents the part of return that is due to risk. Thus the expected return on an efficient portfolio is

Expected return = (Price of time) + (Price of risk) × (Amount of risk)

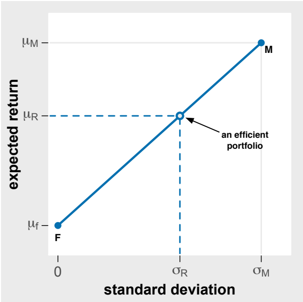
The CML is easy to derive, similar to the line (6.2). Consider an efficient portfolio that allocates a proportion w of its assets to the market portfolio and $1-w$ to the risk-free asset. Then the return

$$R = wR_M + (1-w)\mu_f = \mu_f + w(R_M - \mu_f).$$

The only restriction is that $0 \le w \le 1$. Taking expectation, we get

$$\mu_R = \mu_f + w(\mu_M - \mu_f), \qquad w = \frac{\sigma_R}{\sigma_M}. \tag{7.2}$$

The CML says that the optimal way to invest is to either (a) decide on the risk you can tolerate σ_R, where $0 \le \sigma_R \le \sigma_M$; or (b) choose a reward you desire μ_R, $\mu_f \le \mu_R \le \mu_M$. Then solve for the allocation w using the equations (7.2).



Your investment is consists of (1) w proportion in a market index fund, i.e., a fund that tracks the market as a whole; and (2) $1-w$ proportion in risk-free Treasury bills or bonds.

The CAPM Formula

Although the CML establishes the return on an efficient portfolio, it does not describe equilibrium returns on nonefficient portfolios

or on individual securities. We now turn to the development of a relationship that does so.

Suppose that there are many securities indexed by j and their returns denoted by R_j and $\mu_j = E(R_j)$. The CAPM model is

$$\mu_j = \mu_f + \beta_j(\mu_M - \mu_f), \qquad \beta_j = \frac{\sigma_{jM}}{\sigma_M^2} = \frac{\mathrm{Cov}(R_j, R_M)}{\mathrm{var}(R_M)}. \tag{7.3}$$

Thus β_j is the slope of the best linear predictor of the jth security's returns using returns of the market portfolio as the predictor variable. In fact, the best linear predictor of R_j, Based on R_M is

$$\hat{R}_j = \beta_{0,j} + \beta_j R_M, \qquad \beta_{0,j} = E(R_j) - \beta_j E(R_M).$$

Suppose that we have a bivariate time series $\{(R_{jt}, R_{Mt})^T\}$, $t = 1,\dots,n$, then, the estimated slope of the linear regression of R_{jt} on R_{Mt} is

$$\hat{\beta}_j = \frac{\sum_{t=1}^n (R_{jt} - \bar{R}_j)(R_{Mt} - \bar{R}_M)}{\sum_{t=1}^n (R_{Mt} - \bar{R}_M)^2}, \tag{7.4}$$

which is the ratio of $\hat{\sigma}_{jM}$ and $\hat{\sigma}_M^2$.

Betas and Security Market Line

The security market linear equation will be derived later using the CAPM (7.3). The equation is given by

$$\mu_j - \mu_f = \beta_j(\mu_M - \mu_f). \tag{7.5}$$

The line representing this equation is called *the security market line* (SML). Note that β_j in the SML is not a slope, it is a variable in the linear equation (7.5). More precisely, μ_j is a linear function of β_j with slope $\mu_M - \mu_f$. In other words,β_j is a slope in one context but is the independent variable in the different context of the SML. One can estimate β_j using (7.4) then plug this estimate into (7.5) .

The SML says that the risk premium of the jth asset is the product of its beta (β_j) and the risk premium of the market portfolio ($\mu_M - \mu_f$) . Therefore, β_j measures both the riskiness of the jth asset and the reward for assuming that riskiness. Consequently, β_j is a measure of how "aggressive" the jth asset is. By definition, the beta for the market portfolio is 1; i.e., $\beta_M = 1$. This suggest the rules,

$$\beta_j > 1 \ \Rightarrow \text{“aggressive”},$$
$$\beta_j = 1 \ \Rightarrow \text{“average risk”},$$
$$\beta_j < 1 \ \Rightarrow \text{“not aggressive”}.$$

Figure 7.1 illustrates the SML and an asset J that is not on the SML. This asset contradicts the CAPM, because according to the CAPM all assets are on the SML so no such asset exists.

If an asset like J did exist. No investor would buy it because its risk premium is too low. Consequently, the price of J would decline and its expected return would increase till it is on the SML.

Most of financial websites list betas. For example, Beta in Yahoo Finance is based on monthly returns of the recent 5 years (60 months) and the S&P 500 index as market portfolio.

The CML applies only to the return R of an efficient portfolio, while

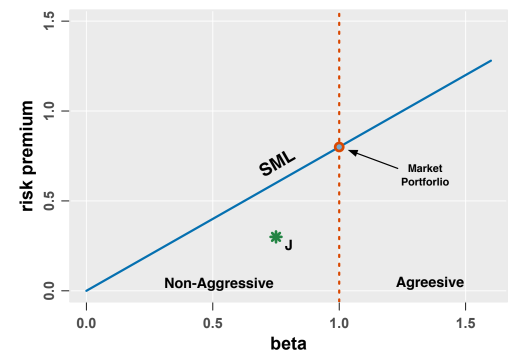


Figure 7.1: Security market line (SML) showing that the risk premium of an asset is a linear function of the asset's beta. J is a security not on the line and a contradiction to the CAPM. Theory predicts that the price of J decreases until J is on the SML. The vertical dotted line separates the nonaggressive and aggressive regions.

the SML applies applies to any asset,

$$\text{CML}: \quad \mu_R - \mu_f = \frac{\sigma_R}{\sigma_M}(\mu_M - \mu_f)$$
$$\text{SML}: \quad \mu_j - \mu_f = \beta_j(\mu_M - \mu_f)$$

If we take an efficient portfolio and consider it as an asset then both equations hold and $\sigma_R/\sigma_M = \beta_R$.

The Security Characteristic Line

The CAPM is an equilibrium relationship. The relationship does not hold over all intervals of time. High-beta stocks are expected to give a higher return than low-beta stocks because they are more risky. This does not mean that they will always give a higher return.

The security characteristic line is the model that incorporates the variation of returns over time. Let R_{jt} be the return at time t on jth asset. Similarly, let $R_{M,t}$ and $\mu_{f,t}$ be the return on the time t. The security characteristic line is a regression model,

$$R_{jt} = \mu_{f\,t} + \beta_j(R_{M\,t} - \mu_{f\,t}) + \varepsilon_{jt}\,, \tag{7.6}$$

where ε_{jt} is $N(0, \sigma_{\varepsilon_j})$. It is assumed that $\mathrm{Cov}(\varepsilon_{jt}, \varepsilon_{j't}) = 0$ for $j \neq j'$. Taking expectations in (7.6), we get

$$\mu_{jt} = \mu_{f\,t} + \beta_j(\mu_{M\,t} - \mu_{f\,t})\,,$$

which is equation (7.5) the SML at time index t. The SML gives information about returns but not about the variance of the returns. The latter is given by the security characteristic line (7.6). The characteristic line is said to be a return-generating process, it gives a probability model of the returns not just a model of their expected values.

The characteristic line implies that the following variance and covariances

$$\sigma_j^2 = \beta_j^2\sigma_M^2 + \sigma_{\varepsilon_j}^2\,, \qquad \sigma_{jj'} = \beta_j\beta_{j'}\sigma_M^2\,, \qquad \sigma_{Mj} = \beta_j\sigma_M^2\,.$$

The total risk of the jth asset is $\sigma_j = \sqrt{\beta_j^2\sigma_M^2 + \sigma_{\varepsilon_j}^2}$. The risk has two components, $\beta_j^2\sigma_M^2$ is the *market of systematic component of risk* and $\sigma_{\varepsilon_j}^2$ is the *unique, nonmarket or unsystematic component* of risk.

Reducing Unique Risk The market component of risk cannot be reduced by diversification, but the unique component can be reduced by sufficient diversification. Suppose that there are N assets with returns $R_{1t}, \ldots, R_{Nt}$ for holding period t. The return of a port-

folio with weights $w_1, \ldots, w_N$ is

$$R_{Pt} = \sum_{i=1}^{N} w_i R_{it}\,.$$

Let R_{Mt} be the return on the market portfolio. According to the characteristic line model

$$R_{Pt} = \mu_{f\,t} + \beta_P(R_{M\,t} - \mu_{f\,t}) + \varepsilon_{Pt}.$$

where the portfolio beta β_P and epsilon ε_{Pt} are

$$\beta_P = \sum_{j=1}^{N} w_j\beta_j\,, \quad \text{and} \quad \varepsilon_{Pt} = \sum_{j=1}^{N} w_j\varepsilon_{jt}\,.$$

We now assume that ε_{jt}, $j = 1, \ldots, N$, are uncorrelated. The unique component of the portfolio is

$$\sigma_{\varepsilon_P}^2 = \sum_{j=1}^{N} w_j^2\sigma_{\varepsilon_j}^2\,.$$

The naive weights $w_j = 1/N$ will yield

$$\sigma_{\varepsilon P}^2 = \frac{1}{N^2}\sum_{j=1}^{N}\sigma_{\varepsilon_j}^2 = \frac{1}{N}\,\bar{\sigma}_\varepsilon^2\,, \qquad \text{where} \quad \bar{\sigma}_\varepsilon^2 = \frac{1}{n}\sum_{j=1}^{N}\sigma_{\varepsilon_j}^2\,.$$

If the variance of individual asset does not vary much the unique risk σ_{ε_P} is roughly $1/\sqrt{N}$ of a single asset.

The assumption of uncorrelation among ε_{jt} is of course not realistic. Correlation among stock can be modeled using a factor model, which will be discussed later.

More on Portfolio Theory

Suppose the Market portfolio consists of N_0 risk assets. Let w_{iM}, $i = 1, \dots, , N_0$ be the weights of these assets in the mark t portfolio. Then *the risk of Market Portfolio* is

$$\sigma_M^2 = \sum_{j=1}^{N_0} \sum_{i=1}^{N_0} w_{jM} w_{iM} \sigma_{ij} = \sum_{j=1}^{N_0} w_{jM} \left(\sum_{i=1}^{N_0} w_{iM} \sigma_{ij} \right) = \sum_{j=1}^{N_0} w_{jM} \sigma_{jM} .$$

The contribution of the jth asset to the risk of the market portfolio is $w_{jM}\sigma_{jM}$, the covariance between the return on the jth asset and the return on the market portfolio (not the asset variance σ_j^2) with the weight w_{jM} .

Derivation of the SML Consider a portfolio P that consists of asset A with weight w_A and the market portfolio with weight $(1-w_A)$. The return on this portfolio is $R_{Pt} = w_A R_{At} + (1 - w_A) R_{Mt}$. The expected return and the risk are

$$\begin{aligned} \mu_P &= w_A \mu_A + (1 - w_A)\mu_M \\ \sigma_P &= \left\{ w_A \sigma_A^2 + (1 - w_A)^2 \sigma_M^2 + 2 w_A (1 - w_A) \sigma_{AM} \right\}^{1/2} . \end{aligned}$$

The portfolio curve is given by (σ_P, μ_P) with varying w_A. The tangency portfolio is the point when $w_A = 0$, that is, the entire investment is on the tangency portfolio. Also, the derivative at the tangency portfolio is the slope of the CML, we have the following equation,

$$\frac{\mu_M - \mu_f}{\sigma_M} = \left. \frac{\partial \mu_P}{\partial \sigma_P} \right|_{w_A = 0} . \tag{7.7}$$

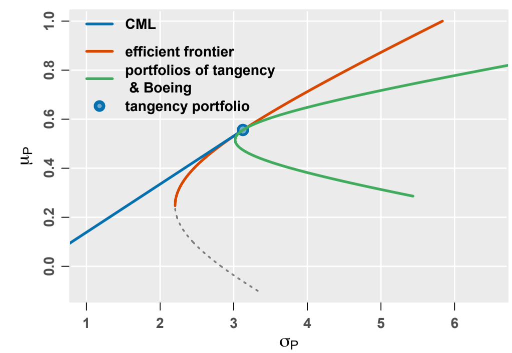


Figure 7.2: The green curve is the locus of portfolios combining Boeing stock and the tangency portfolio. The green curve intersects the efficient frontier at the tangency portfolio. Therefore, the derivative of the green curve at the tangency portfolio is equal to the slope of the CML (blue).

The derivative at RHS without evaluating at $w_A = 0$ is

$$\frac{\partial \mu_P / \partial w_A}{\partial \sigma_P / \partial w_A} = \frac{(\mu_A - \mu_M)\sigma_P}{w_A \sigma_A^2 - (1 - w_A)\sigma_M^2 + (1 - 2w_A)\sigma_{AM}} .$$

Pluggin $w_A = 0$ to the last equation, (7.7) becomes

$$\frac{\mu_M - \mu_f}{\sigma_M} = \frac{(\mu_A - \mu_M)\sigma_M}{\sigma_{AM} - \sigma_M^2},$$

which, after some algebra, gives

$$\mu_A - \mu_f = \frac{\sigma_{AM}}{\sigma_M^2}(\mu_M - \mu_f) = \beta_A(\mu_M - \mu_f) .$$

This is the SML equation of (7.5). Figure 7.3 shows examples of betas and the SML with a proxy market portfolio that consists of 20 stocks. The proxy portfolio is also used to compute the efficient frontier in Figure 7.2.

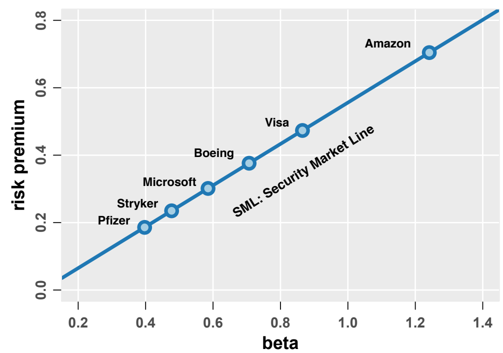


Figure 7.3: Examples of betas and the SML based on weekly returns from Jan 4, 2007 to Oct 9, 2018. The proxy market portfolio consists of 20 stocks.

Estimating Beta and Testing the CAPM

Recall that the security characteristic line is

$$R_{jt} = \mu_{ft} + \beta_j(R_{Mt} - \mu_{ft}) + \varepsilon_{jt}\,.$$

Let $Y_{jt} = R_{jt} - \mu_{ft}$ and $Y_{Mt} = R_{Mt} - \mu_{ft}$, the excess returns on the jth security and market portfolio. Then the security characteristic line of the asset j can be written as

$$Y_{jt} = \beta_j Y_{Mt} + \varepsilon_{jt}\,. \tag{7.8}$$

This is a regression model with β_j being the slope and without an intercept because the intercept of (7.8) is 0 according to the CAPM.

Model for Estimation and Testing

A model that can be used for estimating β of an asset and testing for the CAPM model, that is, 0 intercept, is $Y_{jt} = \alpha_j + \beta_{jt} Y_{Mt} + \varepsilon_{jt}\,.$

Suppose a portfolio have N risky assets, then we can consider an excess return model of N assets, $\boldsymbol{Y}_t = (Y_{1t}, \ldots, , Y_{Nt})^T$,

$$\boldsymbol{Y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} Y_{Mt} + \boldsymbol{\varepsilon}_t\,, \tag{7.9}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)^T$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, , \varepsilon_{Nt})^T$. If $\alpha_j \neq 0$, we recognize the possibility of mispricing. When the CAPM holds, no portfolio should have a nonvanishing α so that the null hypothesis $H_0 : \boldsymbol{\alpha} = \boldsymbol{0}$. While the test is designed to validate CAPM, it can also be applied to check whether a constructed portfolio is efficient among the assets used to construct the portfolio. If it is efficient, then there should be no α for any assets that are used to construct the portfolio.

Estimation Using Regression

The OLS Estimation Equation (7.9) is N simple linear regression models in a compact form, in which all N response variables share the same regressors $(1, \boldsymbol{Y}_{Mt})$ or in data $\boldsymbol{X} = (\boldsymbol{1}, \boldsymbol{Y}_M)$, an $N \times 2$ matrix. Let $\boldsymbol{Y}_j = (Y_{j1}, \ldots, Y_{jn})^T$ be the sample of excess returns of jth asset, for the jth simple linear regression, the OLS estimator of $(\alpha_j, \beta_j)^T$ is given by $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}_j$. The distribution of $(\hat{\alpha}_j, \hat{\beta}_j)^T$ is normal with

$$E(\hat{\alpha}_j, \hat{\beta}_j)^T = (\alpha_j, \beta_j)^T \quad \text{and} \quad \text{var}(\hat{\alpha}_j, \hat{\beta}_j)^T = \boldsymbol{M}\sigma^2_{\varepsilon j} \tag{7.10}$$

where $M = (\boldsymbol{X}^T\boldsymbol{X})^{-1}$ and $\sigma^2_{\varepsilon j} = \text{var}(\varepsilon_{jt})$. In particular, $\text{var}(\hat{\alpha}_j) = m_{11}\sigma^2_{\varepsilon j}$. Thus $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \ldots, \alpha_N)^T$ a N-dimension multivariate normal as,

$$\hat{\boldsymbol{\alpha}} \sim N\big(\boldsymbol{\alpha}, m_{11}\boldsymbol{\Sigma}_\varepsilon\big)\,. \tag{7.11}$$

The OLS estimators can be written in vector-matrix form by

$$\hat{\boldsymbol{\alpha}} = \bar{\boldsymbol{Y}} - \hat{\boldsymbol{\beta}}\bar{Y}_M\,,$$
$$\hat{\boldsymbol{\beta}} = \frac{1}{\sum_{t=1}^{n}(Y_{Mt} - \bar{Y}_M)^2}\sum_{t=1}^{n}(\boldsymbol{Y}_t - \bar{\boldsymbol{Y}})(Y_{Mt} - \bar{Y}_M)\,,$$
$$\hat{\boldsymbol{\Sigma}}_\varepsilon = \frac{1}{n}\sum_{t=1}^{n}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t^T\,, \quad \hat{\boldsymbol{\varepsilon}}_t = \boldsymbol{Y}_t - \hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}Y_{Mt} \tag{7.12}$$

Connection to the MLE If we assume that $\boldsymbol{Y}_t$'s are normal, then the OLS regression estimators coincide with the MLE of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ based on the conditional distribution $\boldsymbol{Y}_t\,|\,Y_{Mt} \sim N(\boldsymbol{\alpha} + \boldsymbol{\beta}Y_{Mt}\,, \boldsymbol{\Sigma}_\varepsilon)$ with the conditional likelihood

$$\ell(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Sigma}_\varepsilon) = -\frac{n}{2}\log|\boldsymbol{\Sigma}_\varepsilon| - \frac{1}{2}\sum_{t=1}^{n}(\boldsymbol{Y}_t - \boldsymbol{\alpha} - \boldsymbol{\beta}Y_{Mt})^T\boldsymbol{\Sigma}_\varepsilon^{-1}(\boldsymbol{Y}_t - \boldsymbol{\alpha} - \boldsymbol{\beta}Y_{Mt})\,. \tag{7.13}$$

Taking partial derivatives with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ and equate them to 0, we get the set the MLE the same as those of OLS (7.12).

Interpretation of alpha. If α_j is nonzero, then the security j is mispriced, at least according to the CAPM. If $\alpha_j > 0$, then security j is underpriced; the returns are too high on average. This is an indication of the asset worth purchasing. It is essential that excess returns not returns should be used. Otherwise, the intercept of the regression equation does not estimate α, one cannot test whether α is zero by testing the intercept.

One also must be careful. If we reject the null hypothesis $H_0 : \alpha_j \neq 0$, all we have shown is that the security j was mispriced in the past.

Testing Statistics for All α's

The null and alternate hypotheses of interest are

$$H_0 : \boldsymbol{\alpha} = \mathbf{0} \qquad \text{against} \qquad H_A : \boldsymbol{\alpha} \neq \mathbf{0}\,. \tag{7.14}$$

The tests will be based on $\hat{\boldsymbol{\alpha}}$ which has a limiting distribution (7.11) and $\hat{\boldsymbol{\Sigma}}_\varepsilon$ in (7.12) has a distribution,

$$n\,\hat{\boldsymbol{\Sigma}}_\varepsilon \sim W_N(n-2, \boldsymbol{\Sigma}_\varepsilon) \quad \text{approximately,}$$

a dimension $N \times N$ Wishart distribution with $n-2$ degrees of freedom. See Definition 7.1 in Appendix. By the projection theory, $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\alpha}}$ is independent of $\hat{\boldsymbol{\varepsilon}}_t$ and thus $\hat{\boldsymbol{\Sigma}}_\varepsilon$. Several statistics will be established based on these statistical properties of $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\Sigma}}_\varepsilon$.

Wald Statistic By definition 7.2, the Wald test statistic for testing H_0 in (7.14) based on $\hat{\boldsymbol{\alpha}}$ and its distribution (7.11) is

$$\mathscr{T}_{W,0} = \hat{\boldsymbol{\alpha}}^T\left\{\operatorname{var}(\hat{\boldsymbol{\alpha}})\right\}^{-1}\hat{\boldsymbol{\alpha}} = \frac{1}{m_{11}}\hat{\boldsymbol{\alpha}}^T\boldsymbol{\Sigma}_\varepsilon^{-1}\hat{\boldsymbol{\alpha}} \sim \chi_N^2\,.$$

This version of Wald statistic is not feasible as due to the unknown $\boldsymbol{\Sigma}_\varepsilon$. A sensible way to use the Wald test is to substitute $\boldsymbol{\Sigma}_\varepsilon$ with its estimate $\hat{\boldsymbol{\Sigma}}_\varepsilon$, however, this leads to a different distribution. That is,

$$\mathscr{T}_W = \frac{n-N-1}{nN}\,\frac{1}{m_{11}}\,\hat{\boldsymbol{\alpha}}^T\hat{\boldsymbol{\Sigma}}_\varepsilon^{-1}\hat{\boldsymbol{\alpha}} \sim F_{N,n-N-1}\,. \tag{7.15}$$

See Theorem 7.2. The F distribution is an asymptotic approximation, it is more accurate under the normal assumption of data.

Likelihood ratio statistic The likelihood ratio test statistic is defined on page 57, Handout 3. It is the discrepancy between the full model and the restricted model under H_0, namely

$$\mathrm{LRT} = 2\left\{\ell(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_\varepsilon) - \ell(\mathbf{0}, \hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\Sigma}}_{0\varepsilon})\right\} \tag{7.16}$$

where $\hat{\boldsymbol{\beta}}_0$ and $\hat{\boldsymbol{\Sigma}}_{0\varepsilon}$ are the maximizer of restricted likelihood $\ell(\mathbf{0}, \boldsymbol{\beta}, \boldsymbol{\Sigma}_\varepsilon)$, where $\ell(\cdot)$ is the likelihood defined in (7.13) if $\boldsymbol{Y}_t$ is normal.

The restricted MLE $\hat{\boldsymbol{\beta}}_0$ is simply the slope of N simple linear regressions without intercepts and $\hat{\boldsymbol{\Sigma}}_{0\varepsilon}$ is computed from the residuals of these regression models. That is,

$$\hat{\boldsymbol{\beta}}_0 = \frac{1}{\sum_{t=1}^n Y_{Mt}^2} \sum_{t=1}^n \boldsymbol{Y}_t Y_{Mt},$$

$$\hat{\boldsymbol{\Sigma}}_{0\varepsilon} = \frac{1}{n}\sum_{t=1}^n \hat{\boldsymbol{\varepsilon}}_{0t}\hat{\boldsymbol{\varepsilon}}_{0t}^T, \quad \hat{\boldsymbol{\varepsilon}}_{0t} = \boldsymbol{Y}_t - \hat{\boldsymbol{\beta}}_0 Y_{Mt}. \tag{7.17}$$

Applying (7.12), (7.13) and (7.17) into the RHS of (7.16), we get the likelihood ratio statistic

$$\mathrm{LRT} = n\left\{\log|\hat{\boldsymbol{\Sigma}}_{0\varepsilon}| - \log|\hat{\boldsymbol{\Sigma}}_\varepsilon|\right\}.$$

The LRT statistic is approximately χ^2 distributed with DF being the number of restrictions in H_0, which is N here. This LRT statistic however is based on a likelihood under the normality of $\boldsymbol{Y}_t$ which is unrealistic. A better approximation can be obtained by the following adjustment for non-normality,

$$\mathscr{T}_{\mathrm{LR}} = (n - N/2 - 2)\left\{\log|\hat{\boldsymbol{\Sigma}}_{0\varepsilon}| - \log|\hat{\boldsymbol{\Sigma}}_\varepsilon|\right\} \sim \chi_N^2. \tag{7.18}$$

The final versions of the Wald statistic and LRT are $\mathscr{T}_W$ and $\mathscr{T}_{\mathrm{LR}}$ given in (7.15) and (7.18) respectively. If $\boldsymbol{Y}_t$ are normally distributed these two statistics perform very similarly in terms of size and power. When the normality assumption fails, the LRT $\mathscr{T}_{\mathrm{LR}}$ is preferable because the χ^2 approximation holds asymptotically.

Eg 7.1. In this example, the S&P 500 index is used for market portfolio and the 3-month US treasury bill for the risk-free asset. Eight stock will be analyzed, Johnson & Johnson (JNJ), Coca-Cola (KO), Mcdonald's (MCD), NVIDIA (NVDA), Oracle (ORCL), Procter & Gamble (PG), Tesla (TSLA), UnitedHealth (UNH). The data consist of weekly returns between January 1, 2011 and September 30, 2024, the sample size is 717.

```
head(Rt,2);
##                JNJ     KO     MCD    NVDA    ORCL      PG    TSLA    UNH
## 2011-01-07 -0.3508 -3.5902 -2.9544 22.7936 -1.8835 -0.4177  5.9077 3.4933
## 2011-01-14 -0.0799  0.3332 -0.4177 17.1612  0.8670  1.5843 -9.2305 5.8588
tail(Rt,2);
##                JNJ     KO    MCD    NVDA   ORCL      PG   TSLA     UNH
## 2024-09-20 -0.8250 0.3216 0.1011 -2.6373 3.6182  0.0804 3.3981 -2.9508
## 2024-09-27 -1.6956 0.2092 2.2848  4.5501 0.4395 -0.3853 8.9129  1.1843
n = dim(Rt)[1]; N = dim(Rt)[2];c(n = n, N = N)
##   n   N
## 717   8
```

The US treasury bill and bonds can be obtained using `quantmod` package's `getSymbols()` by setting the data source `src = "FRED"`. This will load data from the St. Louis Federal Reserve Bank's FRED system. Available frequencies are daily, weekly, monthly and yearly with different ticker symbols. All data are annual rates in percentage (%). We use weekly data of 3-month T-Bill as risk free asset returns in this example.

```
## 3 Month T-Bill as risk-free rates
getSymbols("WTB3MS",src = "FRED")  ## download weekly data of US 3-Mo T-Bill
Rf = as.matrix(Rf); colnames(Rf) = "Rf" ## remove 'xts' class
Rf = Rf/52;  ## convert to weekly rate
head(Rf,2);  ## Starting

##                      Rf
## 2011-01-07 5.177515e-05
## 2011-01-14 5.547337e-05

tail(Rf, 2); ## Ending

##                     Rf
## 2024-09-20 0.001723373
## 2024-09-27 0.001667899

cat("dimension:", dim(Rf))

## dimension: 717 1
```

Recall that the ticker symbol of S&P 500 index is "^GSPC", the information about this series is omitted. We will fit the CAPM to each security's excess returns, in particular the linear equation (7.8) using the the R function `lm()` (see the Regression handout).

```
Yt = apply(Rt, 2, function(u) u-Rf) ## excess  returns of each stock
YM = SnP - Rf  ## excess market return
fit = lm(Yt ~ YM);  ## Regressing Yt on YM
fit

## Call:
## lm(formula = Yt ~ YM)
##
## Coefficients:
##              JNJ     KO      MCD     NVDA    ORCL    PG      TSLA    UNH
## (Intercept)  0.0597  0.0158  0.0981  0.4827  0.0587  0.0692  0.3714  0.2239
## YM           0.5562  0.6961  0.6583  1.6324  0.9554  0.5400  1.6125  0.8862
```



```
names(fit)

##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "xlevels"       "call"          "terms"         "model"
```

Coefficients in the `fit` object are $\hat{\alpha}$'s and $\hat{\beta}$'s, the alpha and beta estimates of each security. We can also get them with the R function, `coef(fit)`.

Nvidia and Tesla have the two highest beta's implying the two most aggressive assets among the eight assets. In theory, they should have highest returns and risks. This agrees with the period of data we fit, the average and standard deviation of their excess returns are highest.

```
cat("Average excess return: ");apply(Yt,2,mean);

## Average excess return:
##    JNJ     KO    MCD   NVDA   ORCL     PG   TSLA    UNH
## 0.1634 0.1456 0.2209 0.7871 0.2369 0.1699 0.6722 0.3892

cat("Risks:");apply(Yt,2,sd)

## Risks:
##   JNJ    KO   MCD  NVDA  ORCL    PG  TSLA   UNH
## 2.216 2.498 2.465 5.950 3.431 2.276 7.574 3.491
```

Our model consists of 8 simple linear regression. Applying `summary()` to the `lm()` return object `fit` will get give us the 8 regression results such as the standard errors, tests and R-Squared. We show only one simple linear regression output, the relevant statistical results will be put together in summary tables.

```
sfit = summary(fit); sfit[[1]];  ## Get All 8 regressions; show the first
```

```
## Call:
## lm(formula = JNJ ~ YM)
##
## Residuals:
##     Min      1Q Median      3Q     Max
## -8.310 -0.996 -0.042  1.038  9.670
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0597     0.0684    0.87     0.38
## YM            0.5562     0.0302   18.42   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.83 on 715 degrees of freedom
## Multiple R-squared:  0.322,Adjusted R-squared:  0.321
## F-statistic:  339 on 1 and 715 DF,  p-value: <2e-16
```

```
names(sfit[[1]]); ## output of the first regression
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
names(sfit) = syb; ## rename the regression list
```

The return values give tests of individual α_j and β_j in `coefficients`. We can also find the matrix $\boldsymbol{M}$ of (7.10) in `cov.unscaled` and R-Squared in `$r.aquared`.

The tests of $H_0 : \alpha_j = 0, j = 1, \ldots, N$ are summarized with the following R code.

```
## Tests for individual alpha = 0
Alpha = c()
for(i in 1:N){
    Alpha= rbind(Alpha, sfit[[i]]$coef[1,])
}
```



```
rownames(Alpha) = syb; Alpha
```

```
##      Estimate Std. Error t value Pr(>|t|)
## JNJ   0.05965    0.06843  0.8717 0.383672
## KO    0.01581    0.07276  0.2173 0.828017
## MCD   0.09812    0.07373  1.3308 0.183668
## NVDA  0.48268    0.17507  2.7570 0.005981
## ORCL  0.05873    0.09998  0.5874 0.557126
## PG    0.06920    0.07204  0.9606 0.337098
## TSLA  0.37143    0.24899  1.4918 0.136200
## UNH   0.22391    0.10722  2.0884 0.037118
```

The CAPM model is not validated for Nvidia and UnitedHealth, the tests for α_j have p-values smaller than 5%. Both α's are positive implying that the two security were underpriced.

```
## Getting R-Squared
R.Sq  = c()
for(i in 1:N){
    R.Sq = c(R.Sq, sfit[[i]]$r.sq)
}
names(R.Sq) = syb; R.Sq
```

```
##    JNJ     KO    MCD   NVDA   ORCL     PG   TSLA    UNH
## 0.3217 0.3966 0.3640 0.3843 0.3960 0.2875 0.2314 0.3291
```

The interpretation of R-Squared is the proportion of the squared risk is due to systematic risk. The R-Squared values are similar among the eight assets with Tesla being the lowest.

For β's, we will test $H_0 : \beta_j = 1$ for aggressiveness in risk.

```
## Testing H0: beta = 1
Beta = c()
for(i in 1:N){
    Beta = rbind(Beta, sfit[[i]]$coef[2,])
}
```

```
Beta[,3] = (Beta[,1]-1)/Beta[,2]; ## replace t value with tests for beta = 1
Beta[,4]= 2*(1-pnorm(abs(Beta[,3]))); ## the p-values of the test
rownames(Beta) = syb; round(Beta,4)
```

```
##      Estimate Std. Error  t value Pr(>|t|)
## JNJ    0.5562     0.0302 -14.6916   0.0000
## KO     0.6961     0.0321  -9.4649   0.0000
## MCD    0.6583     0.0325 -10.5019   0.0000
## NVDA   1.6324     0.0773   8.1847   0.0000
## ORCL   0.9554     0.0441  -1.0099   0.3126
## PG     0.5400     0.0318 -14.4656   0.0000
## TSLA   1.6125     0.1099   5.5734   0.0000
## UNH    0.8862     0.0473  -2.4053   0.0162
```

The tests for $H_0 : \beta_j = 1$ are all significant except for Oracle which will be rated "average risk" asset. NVDA and TSLA are aggressive assets and the rest of assets with significant test results are not aggressive.

The test of $H_0 : \boldsymbol{\alpha} = \mathbf{0}$, all α_i are zero, can be computed easily. We expect the test would be significant since two individual tests are significant, but one of he purposes of a test for all α_j is being able to state the overall p-value (or Type-I error).

```
alpha= Alpha[,"Estimate"]  ## alpha estimates
et = resid(fit) ## n x N matrix -- N residuals from N regressions
Sig = 1/n*t(et)%*%et ## MLE of Sigma
m11 = sfit[[1]]$cov.unscaled[1,1] ## (1,1) entry of M
wald = (n-N-1)/(n*N)*1/m11*t(alpha)%*%solve(Sig)%*%alpha  ## Wald statistic
fit0 = lm(Yt ~ YM -1) ## Restricted model, without intercept
et0 = resid(fit0)  ## residuals from fit0
Sig0 = 1/n*t(et0)%*%et0  # Sigma0_hat
lr = (n-N/2-2)*(log(det(Sig0))-log(det(Sig))) # LRT
cat("Wald test:"); c(statistic = wald, p.value = 1-pf(wald, N, n-N-1))
```

```
## Wald test:
## statistic   p.value
## 2.5933995 0.0084897
```



```
cat("Likelihood ratio test:"); c(statistic = lr, p.value = 1-pchisq(lr, N))
```

```
## Likelihood ratio test:
##  statistic    p.value
## 20.5356682  0.0084887
```

Wald's test and the likelihood ratio test have the same p-values. As we mentioned earlier, an equilibrium model depends only on one factor, the market portfolio, is not realistic. We will see equilibrium models with more factors in the next handout.

One last note, any or all of the quantities β, σ_ε^2, σ_M^2, μ_M and μ_f could depend on time t. However, it is generally assumed that the betas and σ_ε^2 of the assets as well as σ_M^2 and μ_M of the market are independent of t so that these parameters can be estimated assuming stationarity of the time series of returns.

Appendix

Definition 7.1. Wishart distribution. Suppose $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ be $i.i.d.$ $k \times 1$ random vectors and $\boldsymbol{x}_i \sim N_k(\mathbf{0}, \boldsymbol{\Omega})$. The $k \times k$ random matrix $\mathbf{S} = \sum_{i=1}^m \boldsymbol{x}_i \boldsymbol{x}_i^T$ is said to have the Wishart distribution with m degrees of freedom, denoted by $W_k(m, \boldsymbol{\Omega})$.

If $\mathbf{S} \sim W_k(m, \boldsymbol{\Omega})$, then its (i,i)th diagonal entry normalized by the (i,i)th diagonal entry of ω_{ii}, $s_{ii}/\omega_{ii} \sim \chi^2_m$, the chi-square distribution with m degrees of freedom.

Theorem 7.1. *If a $k \times k$ matrix $\mathbf{S} \sim W_k(m, \boldsymbol{\Omega})$ and $\boldsymbol{C}$ is a $k \times j$ matrix, then $\boldsymbol{C}^T \boldsymbol{S} \boldsymbol{C} \sim W_j(m, C^T \boldsymbol{\Omega} \boldsymbol{C})$.*

Definition 7.2. Wald tests. Let $\boldsymbol{\theta}$ be a $k \times 1$ parameter and its estimator $\hat{\boldsymbol{\theta}}$ is such that $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \to_D N_k(\mathbf{0}, \boldsymbol{V})$, where $\mathbf{V} = n\,\mathrm{var}(\hat{\boldsymbol{\theta}})$. Let $\mathbf{C}$ be a $k \times j$ constant matrix, $j \leq k$. In testing the hypotheses

$$H_0 : \mathbf{C}^T \boldsymbol{\theta} = \boldsymbol{c} \qquad \text{against} \qquad H_A : \mathbf{C}^T \boldsymbol{\theta} \neq \boldsymbol{c},$$

the Wald test statistic is

$$\text{Wald.Statistic} = (\mathbf{C}^T \hat{\boldsymbol{\theta}} - \boldsymbol{c})^T \{\mathbf{C}^T (\mathrm{V}/n) \mathbf{C}\}^{-1} (\mathbf{C}^T \hat{\boldsymbol{\theta}} - \boldsymbol{c}) \sim \chi^2_j \,.$$

Theorem 7.2. *Suppose that the k random vector $\boldsymbol{y} \sim N_k(0, \boldsymbol{\Omega})$ and the $k \times k$ random matrix $\mathbf{S} \sim W_k(m, \boldsymbol{\Omega})$ with $m > k$. If $\boldsymbol{y}$ and $\mathbf{S}$ are independent then*

$$\frac{m-k+1}{k} \boldsymbol{y}^T \mathbf{S}^{-1} \boldsymbol{y} \sim F_{k, m-k+1} \,.$$

Eg 7.2. Validating CAPM. Based on the model $\boldsymbol{Y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} Y_{Mt} + \varepsilon_t$, where $\boldsymbol{Y}_t$ is $N \times 1$ vector. The null hypothesis of interest

$$H_0 : \boldsymbol{\alpha} = 0 \,.$$

The OLS estimator $\hat{\boldsymbol{\alpha}}$ has a mean $\boldsymbol{\alpha}$ and variance $\mathrm{var}(\hat{\boldsymbol{\alpha}}) = m_{11} \boldsymbol{\Sigma}_\varepsilon$. Applying Definition 7.2 with $\boldsymbol{\theta} = \boldsymbol{\alpha}$, $j = k = N$, $\boldsymbol{C} = \mathbf{I}_k$, $\boldsymbol{c} = 0$ and $\boldsymbol{V}/n = \mathrm{var}(\hat{\boldsymbol{\alpha}})$, the infeasible Wald test statistic is

$$\mathscr{T}_{W,0} = \hat{\boldsymbol{\alpha}}^T (m_{11} \boldsymbol{\Sigma}_\varepsilon)^{-1} \hat{\boldsymbol{\alpha}} = \frac{1}{m_{11}} \hat{\boldsymbol{\alpha}}^T \boldsymbol{\Sigma}_\varepsilon^{-1} \hat{\boldsymbol{\alpha}} \sim \chi^2_N \,,$$

with $m = n-2$, the feasible Wald test statistic is

$$\mathscr{T}_W = \frac{n-N-1}{N} \hat{\boldsymbol{\alpha}}^T (n m_{11} \hat{\boldsymbol{\Sigma}}_\varepsilon)^{-1} \hat{\boldsymbol{\alpha}} = \frac{n-N-1}{nN} \frac{1}{m_{11}} \hat{\boldsymbol{\alpha}}^T \hat{\boldsymbol{\Sigma}}_\varepsilon^{-1} \hat{\boldsymbol{\alpha}} \sim F_{N, n-N-1} \,.$$

Eg 7.3. Continue Eg.7.2 continued. Suppose $\boldsymbol{W}$ is an $N \times k$ portfolio weight matrix. To test the k portfolios of assets with excess returns $\boldsymbol{Y}_t$ for the CAPM is testing the hypotheses

$$H_0 : \boldsymbol{W}^T \boldsymbol{\alpha} = \mathbf{0} \qquad \text{against} \qquad H_A : \boldsymbol{W}^T \boldsymbol{\alpha} \neq \mathbf{0} \,.$$

The LSE $\boldsymbol{W}^T \hat{\boldsymbol{\alpha}}$ has $\mathrm{var}(\boldsymbol{W}^T \hat{\boldsymbol{\alpha}}) = m_{11} \boldsymbol{W}^T \boldsymbol{\Sigma}_\varepsilon \boldsymbol{W}$ estimated by $\widehat{\mathrm{var}}(\boldsymbol{W}^T \hat{\boldsymbol{\alpha}}) = m_{11} \boldsymbol{W}^T \hat{\boldsymbol{\Sigma}}_\varepsilon \boldsymbol{W}$. Since $n \hat{\boldsymbol{\Sigma}}_\varepsilon \sim W_N(n-2, \boldsymbol{\Sigma}_\varepsilon)$, by Theorem 7.1,

$$n \widehat{\mathrm{var}}(\boldsymbol{W}^T \hat{\boldsymbol{\alpha}}) = n m_{11} \boldsymbol{W}^T \hat{\boldsymbol{\Sigma}}_\varepsilon \boldsymbol{W} \sim W_k(n-2, m_{11} \boldsymbol{W}^T \boldsymbol{\Sigma}_\varepsilon \boldsymbol{W}) \,.$$

By Theorem 7.2 with $\boldsymbol{y} = \boldsymbol{W}^T \hat{\boldsymbol{\alpha}}$, $\boldsymbol{S} = n \widehat{\mathrm{var}}(\boldsymbol{W}^T \hat{\boldsymbol{\alpha}})$ and $m = n-2$,

$$\frac{n-k-1}{k} (\boldsymbol{W}^T \hat{\boldsymbol{\alpha}})^T \left\{ n m_{11} \boldsymbol{W}^T \hat{\boldsymbol{\Sigma}}_\varepsilon \boldsymbol{W} \right\}^{-1} (\boldsymbol{W}^T \hat{\boldsymbol{\alpha}}) \sim F_{k, n-k-1}$$

$$\Longrightarrow \frac{n-k-1}{nk} \frac{1}{m_{11}} (\boldsymbol{W}^T \hat{\boldsymbol{\alpha}})^T \left\{ \boldsymbol{W}^T \hat{\boldsymbol{\Sigma}}_\varepsilon \boldsymbol{W} \right\}^{-1} (\boldsymbol{W}^T \hat{\boldsymbol{\alpha}}) \sim F_{k, n-k-1} \,.$$