Financial Econometrics

Wang Yanchu

Shanghai University of Finance and Economics

November 2024

Factor Pricing Models

- Asset pricing theory
- Capital asset pricing model (CAPM)
- Intertemporal capital asset pricing model (ICAPM)
- Arbitrage pricing theory (APT)
- Estimating and evaluating asset pricing models
- Literature: Cochrane (2005), Campbell, Lo, and Mackinlay (1997)

Asset Pricing Theory

- Basic concept in asset pricing: prices equals discounted expected payoff. Consider a simple two-period case:

$$\begin{aligned} p_t &= E(m_{t+1}x_{t+1}), \\ m_{t+1} &= f(data, parameter), \end{aligned}$$

  where p_t = asset price, x_{t+1} = asset payoff, m_{t+1} = stochastic discount factor.
- This equation should hold for all assets: stocks, bonds, options etc.

Consumption-Based Model

- Consider a simple two-period model in which a representative agent who receives exogenous wealth W_t in period t, and choose to invest τ units of asset in period t. The consumption in period t is
$$C_t = W_t - \tau P_t. \tag{1}$$
- In period $t+1$, the investor receives wealth W_{t+1}, and sells the asset and use all the payoff for consumption, thus the consumption in period $t+1$ is
$$C_{t+1} = W_{t+1} + \tau P_{t+1}. \tag{2}$$
- The agent chooses to maximize the following function
$$\max_{\tau} u(C_t) + E_t[\beta u(C_{t+1})]. \tag{3}$$
- The first order condition implies that
$$P_t u'(C_t) = E_t[\beta u'(C_{t+1}) P_{t+1}], \text{ or} \tag{4}$$
$$P_t = E_t[\beta \frac{u'(C_{t+1})}{u'(C_t)} P_{t+1}]. \tag{5}$$

Consumption-Based Model

- In eq. (4), β is the subjective discount factor - people prefer to consume right now instead of deferring the consumption in future.
- $u(\cdot)$ is the utility function. It measures utility of consumption. We make the following assumption for $u(\cdot)$:
  1. $u'(\cdot) > 0$.
  2. $u''(\cdot) < 0$.
- We often choose $u(\cdot)$ as follows:

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}, \tag{6}$$

  where γ is the risk-aversion coefficient. When $\gamma \to 1$, its limit is $u(C) = \ln C$.

Consumption-Based Model

- From eq. (4) and by introducing the stochastic discount factor $m_{t+1} = \beta \frac{u'(C_{t+1})}{u'(C_t)}$, we can write

$$p_t = E(m_{t+1} x_{t+1}), \tag{7}$$

  where $x_{t+1} = p_{t+1}$ is the payoff in period $t + 1$.
- m_{t+1} is also called marginal rate of substitution, pricing kernel, change of measure, or state-price density.
- Define R_{t+1} as the gross return

$$R_{t+1} = \frac{x_{t+1}}{p_t},$$

  we can write eq. (7) as

$$1 = E(m_{t+1} R_{t+1}). \tag{8}$$

- Now consider in eq. (8), the gross return is given by a riskfree asset, in such case, we have

$$1 = E(mR_f),$$

or

$$R_f = 1/E(m). \tag{9}$$

- If a riskfree asset is not traded, we can define $R_f = 1/E(m)$ as the "shadow" riskfree rate. It is also called "zero-beta" rate.
- Consider $u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$, take it into eq. (9), we obtain

$$R_f = \frac{1}{\beta}(\frac{C_{t+1}}{C_t})^{\gamma}.$$

Consumption-Based Model

- We know that $cov(m, x) = E(mx) - E(m)E(x)$, use this, we can write eq. (7) as
$$p = E(m)E(x) + cov(m, x),$$
or
$$p = \frac{E(x)}{R_f} + cov(m, x).$$
- The price can be decomposed into two parts:
1. The present-value discounted by a riskfree rate.
2. risk adjustment. We can show that an asset whose payoff covaries positively with consumption has a lower price. Why?

Beta-Pricing Model

- We now consider eq. (8)

$$1 = E(mR_i).$$

- Write this as

$$1 = E(m)E(R_i) + cov(m, R_i),$$

and using $R_f = 1/E(m)$, we obtain

$$\begin{aligned} E(R_i) - R_f &= -R_f\, cov(m, R_i), \text{ or} \\ E(R_i) - R_f &= -\frac{cov(m, R_i)}{E(m)}. \end{aligned} \tag{10}$$

Beta-Pricing Model

- Eq. (10) can be written as

$$
\begin{aligned}
E(R_i) - R_f &= -\frac{cov(m, R_i)}{E(m)} \\
&= \left(\frac{cov(m, R_i)}{var(m)}\right)\left(-\frac{var(m)}{E(m)}\right), \text{ or} \\
E(R_i) &= R_f + \beta_{i,m}\lambda_m, \qquad (11)
\end{aligned}
$$

  where $\beta_{i,m}$ is the regression coefficient of the return R_i on m.
- This is a beta pricing model.
- λ_m is interpreted as the price of risk and $\beta_{i,m}$ is the quantity of risk for asset i.

Mean-Variance Frontier

- We can also write eq. (11) as

$$E(R_i) = R_f - \rho_{m,i} \frac{\sigma(m)}{E(m)} \sigma(R_i),$$

  and thus

$$|E(R_i) - R_f| \leqslant \frac{\sigma(m)}{E(m)} \sigma(R_i).$$

- This implies that means and variances of any asset returns must lie in the wedge-shaped region. The boundary is called the mean-variance frontier.
- All frontier returns are perfectly correlated with the discount factor and thus we have

$$\begin{aligned} R_{mv} &= a + bm, \text{ or} \\ m &= d + eR_{mv}. \end{aligned}$$

Mean-Variance Frontier

- Given this, we can use a mean-variance efficient return to write

$$E(R_i) = R_f + \beta_{i,mv}[E(R_{mv}) - R_f].$$

- This is the classical CAPM.
- The main assumptions:
  1. MV portfolio optimization, and one-period investment horizon.
  2. Homogeneous expectations, all investors are risk-aversion.
  3. Perfect competition, and no taxes or transaction costs.
  4. No restrictions on short sales and borrowing/lending at the riskfree rate R_f.
- In this case, we obviously have $\lambda = E(R_{mv}) - R_f$.

Sharpe Ratio

- It is defined as
$$\frac{E(R_i) - R_f}{\sigma(R_i)}.$$
- If R_i is a mean-variance return, then we have
$$\left|\frac{E(R_{mv}) - R_f}{\sigma(R_{mv})}\right| = \frac{\sigma(m)}{E(m)} = \sigma(m)R_f.$$
- The Sharpe ratio of the frontier is governed by the volatility of the discount rate.
- Again use the power utility, we can show that
$$\left|\frac{E(R_{mv}) - R_f}{\sigma(R_{mv})}\right| = \frac{\sigma[(c_{t+1}/c_t)^{-\gamma}]}{E[(c_{t+1}/c_t)^{-\gamma}]} \tag{12}$$

Sharpe Ratio

- Using the lognormal assumption, i.e. assume that

$$\Delta \ln c = \ln c_{t+1} - \ln c_t \sim N(\mu, \sigma^2),$$

then we can write eq. (12) as

$$\left|\frac{E(R_{mv}) - R_f}{\sigma(R_{mv})}\right| = \sqrt{e^{\gamma^2\sigma^2(\Delta \ln c_{t+1})-1}} \approx \gamma\sigma(\Delta \ln c).$$

- The famous "equity premium" puzzle. In U.S., in the past 50 years annual stock return = 9%, volatility=16%, riskfree rate = 1%. Aggregate consumption has a mean and standard deviation about 1%. Thus we have a risk averse coefficient of 50!

Factor Pricing Models

- Factor pricing models express the stochastic discount factor as a linear function of pricing factors

$$m_{t+1} = a + b' f_{t+1},$$

  which implies

$$E(R_{t+1}) = \gamma + \beta' \lambda.$$

- All factor models are derived as specializations of the consumption-based model.

CAPM

- The CAPM, credited to Sharpe (1964) and Linter (1965), is the first, most famous, and most widely used in asset pricing.
- It ties the discount factor m to a linear one-factor function

$$m_{t+1} = a + bR^W_{t+1}.$$

- One can determine the value for a and b by the following two equations

$$\begin{aligned} 1 &= E(mR^W), \text{ and} \\ 1 &= E(m)R^f. \end{aligned}$$

- CAPM is most frequently stated in the following case

$$E(R_i) = R_f + \beta_{i,R^W}[E(R^W) - R_f].$$

ICAPM

- The ICAPM by Merton (1973) generates linear discount factor models

$$m_{t+1} = a + b' f_{t+1},$$

in which the factors are "state variables" for the investor's consumption-portfolio decisions.

- State variables are those that determine how well the investor can do in his maximization.
- The ICAPM does not tell us the identity of the state variables, and many authors use the ICAPM as an obligatory citation to theory on the way to using factors composed of ad hoc portfolios.
- Fama (1991) criticizes the abuse of ICAPM as a "fishing license".
- The ICAPM really is not quite such a expansive license.

APT

- The Arbitrage pricing theory (APT), developed by Ross (1976), starts from a statistical characterization.
- The intuition behind the APT is that the completely idiosyncratic movements in asset returns should not carry any risk prices, since investors can diversify idiosyncratic returns away by holding portfolios.
- Therefore, risk prices or expected returns on a security should be related to the security's covariance with the common components or "factors" only.
- If a set of asset returns are generated by a linear factor model

$$\begin{aligned} R_i &= a_i + \sum_{j=1}^{N} \beta_{ij} f_j + \varepsilon_i, \\ E(\varepsilon_i) &= E(\varepsilon_i f_j) = 0, \end{aligned}$$

then the discount factor can be expressed as

$$m = a + b' f.$$

APT

- The APT models the tendency of asset payoffs to move together via a statistical factor decomposition

$$R_i = a_i + \sum_{j=1}^{N} \beta_{ij} f_j + \varepsilon_i,$$

where f_j are the factors, the β_{ij} are the betas of factor loadings.

- It is usual to demean the factors, $\widetilde{f} = f - E(f)$, and thus the above equation can be written as

$$R_i = E(R_i) + \sum_{j=1}^{N} \beta_{ij} \widetilde{f}_j + \varepsilon_i.$$

- We also assume that $E(\varepsilon_i \varepsilon_j) = 0$, i.e. ε_i are uncorrelated with each other.

APT

- In that assumption, we can see that the factor structure is a restriction on the covariance matrix of returns. If only one factor,

$$\begin{aligned} cov(R_i, R_j) &= E[(\beta_i \widetilde{f} + \varepsilon_i)(\beta_j \widetilde{f} + \varepsilon_j)] \\ &= \beta_i \beta_j \sigma^2(f) + \begin{cases} \sigma^2_{\varepsilon_i}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} . \end{aligned}$$

- For multiple factors, we can describe the covariance matrix as a matrix $\beta\beta'$ plus a diagonal matrix.
- We can use this property to estimate the factor model.

APT and ICAPM

- APT and ICAPM stories are often confused.
- The biggest difference between APT and ICAPM for empirical work is in the inspiration for factors.
- The APT suggests that one start with a statistical analysis of the covariance matrix of returns and find portfolios that characterize common movement.
- The ICAPM suggests that one start by thinking about state variables that describe the conditional distribution of future asset returns.
- In practice, there is not much difference between APT and ICAPM.

Tests of CAPM

- Consider asset i, from CAPM, we know

$$E(R_i) = R_f + \beta_i [E(R_m) - R_f] \tag{13}$$

- Let the excess returns $Z_i = R_i - R_f$, and $Z_m = R_m - R_f$, the CAPM in excess return form is

$$E(Z_i) = \beta_i E(Z_m),$$

  where $\beta_i = \frac{cov(Z_i, Z_m)}{var(Z_m)}$ measures the systematic risk.

- For testing, consider the time series regression

$$Z_{it} = \alpha_i + \beta_i Z_{mt} + \varepsilon_{it}, \ \ t = 1, ..., T. \tag{14}$$

- For the error term, start with IID assumption, that is, $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}^2) = \sigma_i^2$, $cov(Z_{mt}, \varepsilon_{it}) = 0$.

Tests of CAPM

- Tests of CAPM focus on three implications:
  1. $\alpha_i = 0$.
  2. β_i completely captures the cross-sectional variation of expected excess return.
  3. The market risk premium, $E(Z_{mt})$, is positive.
- We first focus on testing the first hypothesis, later we consider the second and the third one.
- Write the time series regression (14) in vector form:

$$\underset{T\times 1}{Z_i} = \alpha_i \underset{T\times 1}{\iota} + \beta_i \underset{T\times 1}{Z_m} + \underset{T\times 1}{\varepsilon_i}. \tag{15}$$

The parameter vector $\theta = (\alpha_i, \beta_i)'$.

Tests of CAPM

- We can estimate the model via OLS regression.
- Construct matrix

$$\underset{T\times 2}{X} = (\iota \; Z_m)$$

- (15) becomes

$$\underset{T\times 1}{Z_i} = \underset{T\times 2}{X}\underset{2\times 1}{\theta} + \underset{T\times 1}{\varepsilon_i} \tag{16}$$

- Multiply both sides of (16) with X' from left

$$\underset{T\times 1}{X'Z_i} = \underset{T\times 2}{X'X}\underset{2\times 1}{\theta} + X'\underset{T\times 1}{\varepsilon_i} \tag{17}$$

- In order to obtain the estimates of θ, multiply both sides of (17) with $(X'X)^{-1}$ from left

$$\underset{T\times 1}{(X'X)^{-1}X'Z_i} = (X'X)^{-1}\underset{T\times 2}{X'X}\underset{2\times 1}{\theta} + (X'X)^{-1}X'\underset{T\times 1}{\varepsilon_i}$$

Tests of CAPM

- We get
$$\theta = (X'X)^{-1}X'Z_i,$$
since $E(X'\varepsilon_i) = 0$ by assumption.
- We assume that
$$\varepsilon_i \sim N(0, \sigma_i^2).$$
- The distribution of estimators:
$$\begin{pmatrix} \widehat{\alpha}_i \\ \widehat{\beta}_i \end{pmatrix} \sim N_2 \left(\begin{matrix} \alpha_i \\ \beta_i \end{matrix}, diag(\sigma_i^2 (X'X)^{-1}) \right).$$
- For α_i
$$\widehat{\alpha}_i \sim N(\alpha_i, \frac{1}{T}(1 + \frac{\widehat{\mu}_m^2}{\widehat{\sigma}_m^2})\sigma_i^2).$$

Tests of CAPM

- μ_m is the expected market excess return

$$\widehat{\mu}_m = \frac{1}{T}\sum_{t=1}^{T} Z_{mt}.$$

- $\sigma_m^2 = var(Z_{mt})$, the market risk

$$\widehat{\sigma}_m^2 = \frac{1}{T}\sum_{t=1}^{T}(Z_{mt} - \widehat{\mu}_m)^2.$$

- $\frac{\widehat{\mu}_m}{\widehat{\sigma}_m}$ is the Sharpe ratio for market.
- σ_i^2 is the estimation error

$$\widehat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}^2.$$

Tests of CAPM

- Test of the null, $H_0 : \alpha_i = 0$

$$t - ratio = \frac{\widehat{\alpha}_i}{\sqrt{\frac{1}{T}(1 + \frac{\widehat{\mu}_m^2}{\widehat{\sigma}_m^2})\sigma_i^2}} \sim t_{T-2}.$$

The t-ratio $\rightarrow N(0,1)$ as $T \rightarrow \infty$.

- System of N assets

$$\underset{N\times 1}{\mathbf{Z}_t} = \underset{N\times 1}{\boldsymbol{\alpha}} + \underset{N\times 1}{\boldsymbol{\beta}} \underset{1\times 1}{\mathbf{Z}_{mt}} + \underset{N\times 1}{\boldsymbol{\varepsilon}_t},$$

where $E(\varepsilon_t \varepsilon_t') = \underset{N\times N}{\Sigma}$.

Tests of CAPM

- $H_0 : \boldsymbol{\alpha} = 0$.
- Run N separate OLS regressions

$$\widehat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\varepsilon}}_t \widehat{\boldsymbol{\varepsilon}}_t'.$$

$$\widehat{\boldsymbol{\varepsilon}}_t = Z_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}} Z_{mt}, \text{ and}$$

$$\widehat{\boldsymbol{\alpha}} \sim N_N(\boldsymbol{\alpha}, \frac{1}{T}(1 + \frac{\widehat{\mu}_m^2}{\widehat{\sigma}_m^2})\widehat{\Sigma}).$$

$$T\widehat{\Sigma} \sim W_N(T-2, \Sigma),$$

where $W_N(T-2, \Sigma)$ indicates that the $(N \times N)$ matrix $T\widehat{\Sigma}$ has a Wishart distribution with $(T-2)$ degrees of freedom and covariance matrix Σ. It is a multivariate generalization of the chi-square distribution.

Tests of CAPM

- $H_0 : \boldsymbol{\alpha} = 0$ against $H_A : \boldsymbol{\alpha} \neq 0$.
- The Wald test statistics

$$
\begin{aligned}
J_0 &= \widehat{\boldsymbol{\alpha}}'[var(\widehat{\boldsymbol{\alpha}})]^{-1}\widehat{\boldsymbol{\alpha}} \\
&= T[(1+\frac{\widehat{\mu}_m^2}{\widehat{\sigma}_m^2})]^{-1}\widehat{\boldsymbol{\alpha}}'\widehat{\Sigma}^{-1}\widehat{\boldsymbol{\alpha}} \sim \chi_N^2
\end{aligned}
\tag{18}
$$

- Note that J_0 depends on the asymptotic theory.

- *Theorem*
  Let the m-vector $\mathbf{x}$ be distributed $N(0, \Omega)$, let the $(m \times m)$ matrix $\mathbf{A}$ be distributed $W_m(n, \Omega)$ with $n \geqslant m$, and let $\mathbf{x}$ and $\mathbf{A}$ be independent, then

$$\frac{n-m+1}{m}\mathbf{x}'\mathbf{A}\mathbf{x} \sim F_{m,n-m+1}.$$

- Set $\mathbf{x} = \sqrt{T}[(1+\frac{\widehat{\mu}_m^2}{\widehat{\sigma}_m^2})]^{-1/2}\widehat{\boldsymbol{\alpha}}$, $\mathbf{A} = T\widehat{\Sigma}$, $m = N$, and $n = T-2$, we obtain the Exact (Finite sample) test

$$J_1 = \frac{T-N-1}{N}[(1+\frac{\widehat{\mu}_m^2}{\widehat{\sigma}_m^2})]^{-1}\widehat{\boldsymbol{\alpha}}'\widehat{\Sigma}^{-1}\widehat{\boldsymbol{\alpha}} \sim F_{N,T-N-1}. \tag{19}$$

- The likelihood ratio test. Compare the log likelihood ratio between the unconstrained and the constrained model.
- The constrained model, assume that $\alpha = 0$.
- Consider again the system of N assets

$$\underset{N\times 1}{\mathbf{Z}_t} = \underset{N\times 1}{\boldsymbol{\alpha}} + \underset{N\times 1}{\boldsymbol{\beta}} \underset{1\times 1}{Z_{mt}} + \underset{N\times 1}{\boldsymbol{\varepsilon}_t}.$$

- For the unconstrained model, we have the following estimators from the maximum likelihood method:

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}} &= \widehat{\boldsymbol{\mu}} - \widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\mu}}_m, \\
\widehat{\boldsymbol{\beta}} &= \frac{\sum_{t=1}^T (\mathbf{Z}_t - \widehat{\boldsymbol{\mu}})(Z_{mt} - \widehat{\boldsymbol{\mu}}_m)}{\sum_{t=1}^T (Z_{mt} - \widehat{\boldsymbol{\mu}}_m)^2}, \\
\widehat{\Sigma} &= \frac{1}{T}\sum_{t=1}^T (\mathbf{Z}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}} Z_{mt})(\mathbf{Z}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}} Z_{mt})',
\end{aligned}$$

where $\widehat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^T \mathbf{Z}_t$ and $\widehat{\mu}_m = \frac{1}{T}\sum_{t=1}^T Z_{mt}$.

Tests of CAPM

- The log-likelihood function for the unconstrained model is

$$\begin{aligned} L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \Sigma) &= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\Sigma| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}(\mathbf{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt})'\Sigma^{-1}(\mathbf{Z}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} Z_{mt}) \end{aligned} \tag{20}$$

- For the constrained model with $\alpha = 0$, the estimators become

$$\widehat{\boldsymbol{\beta}}^* = \frac{\sum_{t=1}^{T}\mathbf{Z}_t Z_{mt}}{\sum_{t=1}^{T} {Z_{mt}}^2}$$
$$\widehat{\Sigma}^* = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{Z}_t - \widehat{\boldsymbol{\beta}}^* Z_{mt})(\mathbf{Z}_t - \widehat{\boldsymbol{\beta}}^* Z_{mt})'.$$

- The log-likelihood function for the constrained model is

$$\begin{aligned} L^* &= -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\Sigma^*| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}(\mathbf{Z}_t - \boldsymbol{\beta}^* Z_{mt})'\Sigma^{*-1}(\mathbf{Z}_t - \boldsymbol{\beta}^* Z_{mt}). \end{aligned}$$

Tests of CAPM

- Define the log-likelihood ratio is

$$\begin{aligned} LR &= L^* - L \\ &= -\frac{T}{2}[\log|\widehat{\Sigma}^*| - \log|\widehat{\Sigma}|]. \end{aligned}$$

- Under $H_o : \alpha = 0$, we have

$$\begin{aligned} J_2 &= -2LR \\ &= T[\log|\widehat{\Sigma}^*| - \log|\widehat{\Sigma}|] \sim \chi^2_N, \end{aligned} \tag{21}$$

where χ^2_N is the chi-square distribution with N degrees of freedom.

- We have introduced three tests: J_0, J_1 and J_2. It is possible to show that J_1 is a monotonic transformation of J_2

$$J_1 = \frac{T-N-1}{N}[\exp(\frac{J_2}{T}) - 1]. \tag{22}$$

- Since the finite sample distribution for J_1 is known, we can use (22) to derive the finite sample distribution of J_2. We can make improvement for J_2 which has better finite sample properties. Define J_3 as the modified statistic, we have

$$\begin{aligned} J_3 &= \frac{T-\frac{N}{2}-2}{T} J_2 \\ &= \frac{T-\frac{N}{2}-2}{T}[\log|\widehat{\Sigma}^*| - \log|\widehat{\Sigma}|] \sim \chi^2_N. \end{aligned} \tag{23}$$

The Black Version

- In the absence of a riskfree asset, we may consider the Black version of the CAPM in eq. (**??**)

$$ER_i - ER_Z = \beta_i(ER_m - ER_Z), \tag{24}$$

where R_Z is the zero-beta portfolio return.

- For N assets, the Black version becomes

$$\underset{N\times 1}{E(\mathbf{R}_t)} = \boldsymbol{\iota}\gamma + \boldsymbol{\beta}(E(R_{mt}) - \gamma), \tag{25}$$

where γ is the zero-beta portfolio return.

- We consider the following unconstrained market model

$$\mathbf{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}R_{mt} + \boldsymbol{\varepsilon}_t. \tag{26}$$

- Again, we assume the error term follows IID.

The Black Version

- Compare the constrained model (25) and the unconstrained model (26), we have

$$\boldsymbol{\alpha} = (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma. \tag{27}$$

- The test is now more complicated since the parameters $\boldsymbol{\beta}$ and γ enter in a nonlinear fashion.
- We may test the model using ML method. For the unconstrained model, we have the following estimators

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}} &= \widehat{\boldsymbol{\mu}} - \widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\mu}}_m, \\
\widehat{\boldsymbol{\beta}} &= \frac{\sum_{t=1}^T (\mathbf{Z}_t - \widehat{\boldsymbol{\mu}})(Z_{mt} - \widehat{\boldsymbol{\mu}}_m)}{\sum_{t=1}^T (Z_{mt} - \widehat{\boldsymbol{\mu}}_m)^2}, \\
\widehat{\Sigma} &= \frac{1}{T}\sum_{t=1}^T (\mathbf{Z}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}} Z_{mt})(\mathbf{Z}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\boldsymbol{\beta}} Z_{mt})'.
\end{aligned}$$

Note that now $\boldsymbol{\mu}$ is the mean of real returns instead of excess return, $\boldsymbol{\mu} = \frac{1}{T}\sum_{t=1}^T \mathbf{R}_t$, $\mu_m = \frac{1}{T}\sum_{t=1}^T R_{mt}$.

The Black Version

- For the constrained model, the maximum likelihood estimators are

$$\widehat{\gamma^*} = \frac{(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*)'\widehat{\Sigma}^{*-1}(\widehat{\boldsymbol{\mu}}-\widehat{\boldsymbol{\beta}}^*\mu_m)}{(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*)'\widehat{\Sigma}^{*-1}(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*)}$$

$$\widehat{\boldsymbol{\beta}}^* = \frac{\sum_{t=1}^T(\mathbf{R}_t-\widehat{\gamma}^*\boldsymbol{\iota})(R_{mt}-\widehat{\gamma}^*)}{\sum_{t=1}^T(R_{mt}-\widehat{\gamma}^*)^2},$$

$$\widehat{\Sigma}^* = \frac{1}{T}\sum_{t=1}^T[\mathbf{R}_t - \widehat{\gamma}^*(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*) - \widehat{\boldsymbol{\beta}}^* R_{mt}][\mathbf{R}_t - \widehat{\gamma}^*(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*) - \widehat{\boldsymbol{\beta}}^* R_{mt}]'.$$

- Note the above three equations should be solved simultaneously.
- We may obtain the estimator by iterating over these three equation until convergence.
- The estimators provided by the unconstrained model can be used as initial inputs.

The Black Version

- The null $H_0 : \boldsymbol{\alpha} = (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma$, against $H_A : \boldsymbol{\alpha} \neq (\boldsymbol{\iota} - \boldsymbol{\beta})\gamma$.
- We may construct the likelihood ratio test as follows

$$J_4 = T[\log|\widehat{\Sigma}^*| - \log|\widehat{\Sigma}|] \sim \chi^2_{N-1}. \tag{28}$$

- Similarly, we can also improve the finite sample performance of J_4 by making the following adjustment:

$$J_5 = (T - \frac{N}{2} - 2)[\log|\widehat{\Sigma}^*| - \log|\widehat{\Sigma}|] \sim \chi^2_{N-1} \tag{29}$$

- Another test, which assumes that γ is known, the exact maximum likelihood estimators suggested by Kandel (1984) and Shanken (1986).

- Suppose that γ is known, for the unconstrained model, we thus have

$$\mathbf{R}_t - \gamma\boldsymbol{\iota} = \boldsymbol{\alpha} + \boldsymbol{\beta}(R_{mt} - \gamma) + \boldsymbol{\varepsilon}_t, \tag{30}$$

- The maximum likelihood estimators for the unconstrained model are

$$\begin{aligned}\widehat{\boldsymbol{\alpha}} &= \widehat{\boldsymbol{\mu}} - \gamma\boldsymbol{\iota} - \widehat{\boldsymbol{\beta}}(\mu_m - \gamma), \\ \widehat{\boldsymbol{\beta}} &= \frac{\sum_{t=1}^{T}(\mathbf{R}_t - \widehat{\boldsymbol{\mu}})(R_{mt} - \widehat{\mu}_m)}{\sum_{t=1}^{T}(R_{mt} - \widehat{\mu}_m)^2},\end{aligned}$$

and

$$\widehat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}[\mathbf{R}_t - \widehat{\boldsymbol{\mu}} - \widehat{\boldsymbol{\beta}}(R_{mt} - \widehat{\mu}_m)][\mathbf{R}_t - \widehat{\boldsymbol{\mu}} - \widehat{\boldsymbol{\beta}}(R_{mt} - \widehat{\mu}_m)]'.$$

The Black Version

- The estimators of $\widehat{\boldsymbol{\beta}}$ and $\widehat{\Sigma}$ do not depend on γ, but $\widehat{\boldsymbol{\alpha}}$ does.
- The log-likelihood function value is

$$L = -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\widehat{\Sigma}| - \frac{NT}{2}, \tag{31}$$

which does not depend on γ too.
- Constraining $\boldsymbol{\alpha}$ to be zero, the constrained estimators are

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}^* &= \frac{\sum_{t=1}^{T}(\mathbf{R}_t - \gamma\boldsymbol{\iota})(R_{mt}-\gamma)}{\sum_{t=1}^{T}(R_{mt}-\gamma)^2}, \\
\widehat{\Sigma}^* &= \frac{1}{T}\sum_{t=1}^{T}[\mathbf{R}_t - \gamma(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*) - \widehat{\boldsymbol{\beta}}^* R_{mt}][\mathbf{R}_t - \gamma(\boldsymbol{\iota}-\widehat{\boldsymbol{\beta}}^*) - \widehat{\boldsymbol{\beta}}^* R_{mt}]',
\end{aligned}$$

and the value of log-likelihood function is

$$L^*(\gamma) = -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log|\widehat{\Sigma}^*(\gamma)| - \frac{NT}{2}, \tag{32}$$

which now depends on γ.

The Black Version

- The logarithm of the likelihood ratio is as follows:

$$\begin{aligned} LR(\gamma) &= L^*(\gamma) - L \\ &= -\frac{T}{2}[\log|\widehat{\Sigma}^*(\gamma)| - \log|\widehat{\Sigma}|]. \end{aligned} \tag{33}$$

- The value of γ that minimizes $LR(\gamma)$ will be the value which maximizes $L^*(\gamma)$ and thus be the maximum likelihood estimator of γ.
- Furthermore, we can calculate $LR(\gamma)$ in an expression as

$$\begin{aligned} LR(\gamma) &= -\frac{T}{2}\log[(\frac{\widehat{\sigma}_m^2}{(\widehat{\mu}_m - \gamma)^2 + \widehat{\sigma}_m^2})\widehat{\boldsymbol{\alpha}}(\gamma)'\Sigma^{-1}\widehat{\boldsymbol{\alpha}}(\gamma) + 1] \\ &= -\frac{T}{2}\log\{(\frac{\widehat{\sigma}_m^2}{(\widehat{\mu}_m - \gamma)^2 + \widehat{\sigma}_m^2})[\widehat{\boldsymbol{\mu}} - \gamma\boldsymbol{\iota} - \widehat{\boldsymbol{\beta}}(\mu_m - \gamma)]'\widehat{\Sigma}^{-1} \\ &\quad \times[\widehat{\boldsymbol{\mu}} - \gamma\boldsymbol{\iota} - \widehat{\boldsymbol{\beta}}(\mu_m - \gamma)]\}. \end{aligned} \tag{34}$$

The Black Version

- Maximizing equation (34) is equivalent to maximize G where

$$G = \left(\frac{\widehat{\sigma}_m^2}{(\widehat{\mu}_m - \gamma)^2 + \widehat{\sigma}_m^2}\right)[\widehat{\boldsymbol{\mu}} - \gamma\boldsymbol{\iota} - \widehat{\boldsymbol{\beta}}(\mu_m - \gamma)]'\widehat{\Sigma}^{-1}[\widehat{\boldsymbol{\mu}} - \gamma\boldsymbol{\iota} - \widehat{\boldsymbol{\beta}}(\mu_m - \gamma)]. \tag{35}$$

- We can show that there are two solutions to $\partial G/\partial\gamma = 0$, and these are the real roots of the quadratic equation

$$H(\gamma) = A\gamma^2 + B\gamma + C. \tag{36}$$

- The expression for A, B, and C are omitted here. It is given on page 202 in CLM.
- We may calculate the variance of γ and thus inferences concerning the value of γ are possible.

Size of Tests

- Example

Table: Finite-sample size of tests of the CAPM using large sample test statistics

N	T	J_0	J_2	J_3
10	60	0.170	0.096	0.051
	120	0.099	0.070	0.050
	180	0.080	0.062	0.050
	240	0.072	0.059	0.050
	360	0.064	0.056	0.050
20	60	0.462	0.211	0.057
	120	0.200	0.105	0.051
	180	0.136	0.082	0.051
	240	0.109	0.073	0.050
	360	0.086	0.064	0.050

Taken from CLM Chapter 5, Table 5.1. The exact finite-sample size is presented for tests with a size of 5% asymptotically. N is the number of assets and T is the number of monthly time-sries observations.

- CAPM also implies that a linear relation between expected returns and market risk premium: betas completely explains the cross section of expected returns.

$$\begin{aligned} E(Z_i) &= \beta_i E(Z_m) \\ &= \beta_i \lambda, \end{aligned} \tag{37}$$

where $\lambda = E(Z_m)$, the market risk premium.

- Eq. (37) answers the question why average returns vary across assets: expected returns of an asset should be high if that asset has high beta.
- This gives the idea to run a cross-sectional regression:

1. First find estimates of the betas from time-series regressions.

$$Z_{it} = a_i + \beta_i Z_{mt} + \varepsilon_{it}. \tag{38}$$

2. Second estimate the market risk premium from a regression across assets of average returns:

$$E_T(Z_i) = \beta_i \lambda + \alpha_i. \tag{39}$$

Fama-MacBeth Procedure

- Fama-MacBeth (1973) suggest an alternative procedure for running cross-sectional regressions.
  1. Estimate β_i for each asset i from the time series regression.
  2. For each time point $t = 1, 2, ..., T$, estimate (39), giving the T estimates of α_{it} and λ_t.
  3. Define $\alpha_i = E[\alpha_{it}]$ and $\lambda = E[\lambda_t]$. Test $H_0 : \alpha_i = 0$ and $\lambda > 0$.

Fama-MacBeth Procedure

- Define

$$\widehat{\alpha}_i = \frac{1}{T}\Sigma_{t=1}^{T}\widehat{\alpha}_{it}$$
$$\widehat{\lambda} = \frac{1}{T}\Sigma_{t=1}^{T}\widehat{\lambda}_t.$$

- The sampling error for these estimates are

$$\sigma^2(\widehat{\alpha}_i) = \frac{1}{T^2}\Sigma_{t=1}^{T}(\widehat{\alpha}_{it} - \widehat{\alpha}_i)$$
$$\sigma^2(\widehat{\lambda}) = \frac{1}{T^2}\Sigma_{t=1}^{T}(\widehat{\lambda}_t - \widehat{\lambda}).$$

- Under null, we have

$$\frac{\widehat{\alpha}_i}{\sigma(\widehat{\alpha}_i)} \sim N(0,1)$$
$$\frac{\widehat{\lambda}}{\sigma(\widehat{\lambda})} \sim N(0,1).$$

Fama-MacBeth Procedure

- We can include additional variables in (39) and test their impact.

$$Z_t = \alpha_i + \beta_i \lambda + \underset{1\times 1}{\gamma_{2t}} \underset{N\times 1}{S_t} + \eta_t,$$

If CAPM is correct, only β_i matters for pricing. $H_0 : \gamma_2 = E[\gamma_{2t}] = 0$.

- S_t may include other stock characteristics: size, dividend yield, Book to market ratio, etc.

Fama-MacBeth Procedure

- The previous formula for calculating the standard error of $\widehat{\lambda}$ assumes that λ_t is not autocorrelated.
- We can extend the idea to estimate $\sigma^2(\widehat{\lambda})$ by allowing the autocorrelation:

$$\sigma^2(\widehat{\lambda}) = \frac{1}{T} \sum_{j=-\infty}^{+\infty} cov(\widehat{\lambda}_t, \widehat{\lambda}_{t-j}).$$

Fama-MacBeth Procedure

- It is natural to test whether all the pricing errors are zero.
- Denote by $\boldsymbol{\alpha}$ the vector of pricing errors across assets. We could estimate the covariance matrix of the sample pricing errors by

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}} &= \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{\alpha}}_t, \\
cov(\widehat{\boldsymbol{\alpha}}) &= \frac{1}{T^2}\sum_{t=1}^{T}(\widehat{\boldsymbol{\alpha}}_t-\widehat{\boldsymbol{\alpha}})(\widehat{\boldsymbol{\alpha}}_t-\widehat{\boldsymbol{\alpha}})'.
\end{aligned}$$

- Then use the test

$$\widehat{\boldsymbol{\alpha}}'cov(\widehat{\boldsymbol{\alpha}})^{-1}\widehat{\boldsymbol{\alpha}} \sim \chi^2_{N-1}.$$

Problems with CAPM

- CAPM holds period by period. When we estimate beta by time series regression, we implicitly assume that beta is constant. This is not necessarily true:
  1. Beta can change along the business cycle.
  2. Beta can change along the firm's life cycle.
- One solution is to test the model conditionally:

$$Z_{i,t} = \alpha_{i,t} + \beta_{i,t} Z_{m,t} + \varepsilon_{it} \tag{40}$$

Problems with CAPM

- Roll's (1977) critique: The total wealth portfolio includes non-marketable assets (human resource, goodwill etc.). Instead, the tests typically rely on a proxy of tradable assets (SP 500 index etc.).
- Consequence:
  1. Measurement error in the market return. Rejection of CAPM may be due to the incorrect measure of market portfolio.
  2. Most tests ignore the unobservability and assume that the proxy is mean-variance efficient.
- Solution: Some authors try to compute a broad proxy by including the return on human capital (Jagannathan & Wang, 1996).

Conclusions

- We discuss several tests for classical CAPM in a one-period, unconditional framework.
- The method can be extended to test conditional versions of CAPM, i.e. Chapter 12.
- The empirical results show evidence against CAPM.
- How to interpret the unfavorable evidence remains a question.
- All in all, CAPM is a good start point in learning asset pricing models.