ASSET PRICING MODELS

[1] CAPM

(1) Some notation:

- \( R_{it} \) = (gross) return on asset \( i \) at time \( t \).
- \( R_{mt} \) = (gross) return on the market portfolio at time \( t \).
- \( R_{ft} \) = return on risk-free asset at time \( t \).
- \( X_{it} = R_{it} - R_{ft} \) = excess return on asset \( i \).
- \( X_{mt} = R_{mt} - R_{ft} \) = excess return on the market portfolio.

\[
\begin{bmatrix}
    R_{1t} \\
    R_{2t} \\
    \vdots \\
    R_{Nt}
\end{bmatrix}_{N \times 1} \quad \begin{bmatrix}
    X_{1t} \\
    X_{2t} \\
    \vdots \\
    X_{Nt}
\end{bmatrix}_{N \times 1} = \begin{bmatrix}
    1 \\
    1 \\
    \vdots \\
    1
\end{bmatrix}_{N \times 1}
\]

- For simplicity, we assume that \((R_{t}', R_{ft}, R_{mt})\) is iid over time.

(2) Sharpe-Lintner version of CAPM

- Sharpe, 1964, Journal of Finance
1) Basic idea:

- \( \text{var}(R_{mt}) = \) risk from the market portfolio of risky asset.
- risk price = \( p \).
  - \( \rightarrow \) cost of bearing the market risk = \( p \text{var}(R_{mt}) \).
  - \( \rightarrow \) At equilibrium, cost of risk = expected gain from risk.
  - \( \rightarrow \) \( p \text{var}(R_{mt}) = E(R_{mt}) - R_{ft} \)
  - \( \rightarrow \) \( p = \frac{E(R_{mt}) - R_{ft}}{\text{var}(R_{mt})} \).

- Systematic risk of an individual asset \( i \):
  The risk of asset \( i \) due to correlation between returns on asset \( i \) and the whole risky-asset market
  - \( \rightarrow \) \( \text{cov}(R_{it}, R_{mt}) \).
- Let \( \beta_i \) be the systematic risk of an individual asset \( i \) relative to the market risk:
  \( \beta_i = \frac{\text{cov}(R_{it}, R_{mt})}{\text{var}(R_{mt})} \).
  - \( \rightarrow \) This \( \beta_i \) can be estimated by the time-series OLS on:
    \( R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it} \).
- Cost of bearing the (systematic) risk of asset \( i \):
  \( \beta_i \text{var}(R_{mt}) p = \beta_i [E(R_{mt}) - R_{ft}] \).
• Equilibrium condition:

\[ E(R_{it}) - \beta_i[E(R_{mt}) - R_{ft}] = R_{ft}, \text{ for all } i = 1, \ldots, N \]

\[ \rightarrow \quad E(R_{it}) = R_{ft} + \beta_i[E(R_{mt}) - R_{ft}] \text{ or } E(X_{it}) = \beta_i E(X_{mt}). \]

[Capital Asset Pricing Model]

2) Empirical Model

• Model:

\[ X_{1t} = \alpha_1 + \beta_1 X_{mt} + \varepsilon_{1t}; \]
\[ X_{2t} = \alpha_2 + \beta_2 X_{mt} + \varepsilon_{2t}; \]
\[ \vdots \]
\[ X_{Nt} = \alpha_N + \beta_N X_{mt} + \varepsilon_{3t}. \]

\[ \rightarrow \quad X_t = \alpha + \beta X_{mt} + \varepsilon_t, \]

where \( \alpha = (\alpha_1, \ldots, \alpha_N)' \), \( \beta = (\beta_1, \ldots, \beta_N)' \) and \( \varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})' \).

• Assume that the \( \varepsilon_t \) are iid over time with \( \text{Cov}(\varepsilon_t) = \Sigma = [\sigma_{ii}]_{N \times N} \).

• Comments.
  
  • No heteroskedasticity over time.
  
  • Reasonable if \( \varepsilon_t \) is normal.
  
  • If \( \varepsilon_t \) is t-distributed, heteroskedasticity should exist.

[MacKinlay and Richardson (1991, JF).]
• Equilibrium condition:
\[ E(R_{it}) - \beta_i[E(R_{mt}) - R_{fi}] = R_{fi}, \text{ for all } i = 1, \ldots, N \]
\[ \rightarrow E(R_{it}) = R_{fi} + \beta_i[E(R_{mt}) - R_{fi}] \text{ or } E(X_{it}) = \beta_iE(X_{mt}). \]
[Capital Asset Pricing Model]

2) Empirical Model
• Model:
\[ X_{1t} = \alpha_1 + \beta_1 X_{mt} + \varepsilon_{1t}; \]
\[ X_{2t} = \alpha_2 + \beta_2 X_{mt} + \varepsilon_{2t}; \]
\[ \vdots \]
\[ X_{Nt} = \alpha_N + \beta_N X_{mt} + \varepsilon_{Nt}. \]
\[ \rightarrow X_t = \alpha + \beta X_{mt} + \varepsilon_t, \]
where \( \alpha = (\alpha_1, \ldots, \alpha_N)' \), \( \beta = (\beta_1, \ldots, \beta_N)' \) and \( \varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})' \).

• Assume that the \( \varepsilon_t \) are iid over time with \( \text{Cov}(\varepsilon_t) = \Sigma = [\sigma_{ii}]_{N \times N} \).
• Comments.
  • No heteroskedasticity over time.
  • Reasonable if \( \varepsilon_t \) is normal.
  • If \( \varepsilon_t \) is t-distributed, heteroskedasticity should exist.

[MacKinlay and Richardson (1991, JF).]
• Estimation:
  • SUR model with the same regressor: GLS = OLS.
    → Use MLE or OLS.
  • CAPM implies $H_0: \alpha_1 = \ldots = \alpha_N = 0$.
    → Can test these restrictions by Wald or LR.
    [See Ch. 4-5 of CLM.]
    → If $N$ is too large, the test result would be unreliable.
    [See Ahn and Gadarowski, 2004]

• Two-Pass Regression Method (Fama-MacBeth, 1973, JPE)
  • Suppose that $\beta_i$’s are known. Then, we can consider the following cross-sectional regression model for each $t$:
    \[(*) \quad X_t = e_N \gamma_1 + \beta \gamma_2 + \text{error},\]
    where $\beta = (\beta_1, \ldots, \beta_N)'$.
  • If the CAPM is correct, it should be the case that $E(\gamma_{1t}) = 0$ and $E(\gamma_{2t}) \neq 0$. 

ASSET PRICING-5
• Estimation procedure:
• STEP 1: For each i, do time-series OLS to estimate \( \beta_i (\hat{\beta}_i) \).
• STEP 2: For each t, do cross-section OLS to estimate \( \gamma_{1t} \) and \( \gamma_{2t} (\hat{\gamma}_{1t} \) and \( \hat{\gamma}_{2t} \).
• STEP 3: Compute:
\[
\hat{\gamma}_j = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{jt}; \text{var}(\hat{\gamma}_j) = \frac{1}{T(T-1)} \sum_{t=1}^{T} (\hat{\gamma}_{jt} - \hat{\gamma}_j)^2.
\]
• STEP 4: Do t-tests to check whether \( \gamma_1 = 0 \) and \( \gamma_2 \neq 0 \).

• Equivalent Procedure [Shanken, 1992, RFS]
• STEP 1: For each i, do time-series OLS to estimate \( \beta_i (\hat{\beta}_i) \).

Let \( \hat{\gamma} = (e_N, \hat{\beta}), \gamma = (\gamma_1, \gamma_2)' \) and \( \bar{X} = \frac{1}{T} \sum_{t=1}^{T} X_t \)
• STEP 2: Do OLS on \( \bar{X} = \hat{B} \gamma + \text{error} \):
\[
\hat{\gamma} = (\hat{B}'\hat{B})^{-1} \hat{B}'\bar{X};
\]
\[
\text{Cov}(\hat{\gamma}) = (\hat{B}'\hat{B})^{-1} \hat{B}' \frac{1}{T^2} \sum_{t=1}^{T} (X_t - \bar{X})(X_t - \bar{X})' \hat{B}(\hat{B}'\hat{B})^{-1}
\]
• The above covariance matrix is valid only if true \( \beta_i \)'s are used.
  → Correct form of the covariance matrix will be discussed below.
(3) Black-version of CAPM:

1) Basic Model
   • Model when there is no risk-free asset.
   • \( R_{omt} \) = return on the zero-beta portfolio associated with \( m \).
     [portfolio that has the minimum variance of all portfolios uncorrelated with \( m \)]
     \( \rightarrow \) Let \( \gamma = E(R_{omt}) \).
   • Black-Version of CAPM:
     \[ E(R_{it}) = \gamma + \beta_i[E(R_{mt})-\gamma] \rightarrow E(R_{it}) = \gamma(1-\beta_i) + \beta_i E(R_{mt}). \]

2) Empirical estimation and testing.
   • Empirical Model:
     \[ R_{1t} = \alpha_1 + \beta_1 R_{mt} + \varepsilon_{1t}; \]
     \[ R_{2t} = \alpha_2 + \beta_2 R_{mt} + \varepsilon_{2t}; \]
     \[ \vdots \]
     \[ R_{nt} = \alpha_n + \beta_n R_{mt} + \varepsilon_{3t}. \]
     \( \rightarrow \) \( R_t = \alpha + \beta R_{mt} + \varepsilon_t, \)
   • Black-version of CAPM implies \( H_0: (e_N-\beta)\gamma = \alpha. \)
     \( \rightarrow \) See CLM for how to test this hypothesis.
(3) When returns are heteroskedastic or autocorrelated over time.

- Estimate the parameters by GMM.
- Moment conditions:

\[
E\left(\frac{1}{X_{mt}}\right)\left(X_{it} - \alpha_i - \beta_i X_{mt}\right) = 0, \ i = 1, ..., N
\]

\[
\rightarrow E\left(\left(\frac{1}{X_{mt}}\right) \otimes \left(X_t - \alpha - \beta X_{mt}\right)\right) = 0.
\]
Multifactor Pricing Model

(1) Arbitrage Pricing Model [Ross, JET, 1976]

• Assumption 1:
  • \( R_{it} = \alpha_i + \beta_{i1}f_{1t} + \beta_{i2}f_{2t} + \ldots + \beta_{ik}f_{kt} + \epsilon_{it}, \)
    where \( f_{1t}, \ldots, f_{kt} \) are macroeconomic or portfolio factors.
  • \( R_t = \alpha + \beta_1f_{1t} + \ldots + \beta_kf_{kt} + \epsilon_t = \alpha + Bf_t + \epsilon_t, \)
    where \( B = (\beta_1, \ldots, \beta_k) \) and \( f_t = (f_{1t}, \ldots, f_{kt})'. \)
  • \( \text{Cov}(\epsilon_t) = \Sigma_{N \times N} \) (\( \epsilon_{it} \) are cross-sectionally correlated).
    → If there is no missing factor, \( \Sigma \) should be diagonal.
  • \( R_t \) and \( f_t \) are covariance-stationary and ergodic.
  • The factors in \( f_i \) are strictly exogeous:
    \( E(f_s\epsilon_{it}) = 0 \) for all \( i = 1, \ldots, N \), and all \( t \) and \( s. \)
    → \( E(f_s \otimes \epsilon_t) = 0. \)
  • The beta matrix \( B \) is of full column.
    → How could we test for \( \text{rank}(B) \)?
    → What happens if \( B \) is not full column? [See below.]

• Assumption 2 (No Autocorrelation): Assumption 1 plus
  • \( E(\epsilon_t\epsilon_s' | f_1, \ldots, f_T) = 0_{N \times N}. \)
• Assumption 3 (No Heteroskedasticity): Assumption 2 plus
  • \( \text{Cov}(\varepsilon_t \mid f_1, \ldots, f_T) = \Sigma \), for all \( t \).

• Assumption 4 (Autocorrelation in \( f_t \)):
  • Let \( \hat{\Sigma}_T \) be the Newey-West estimator of

\[
\lim_{T \to \infty} \text{Var}\left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (f_t - E(f_t))(f_t - E(f_t))' \right)
\]

• Assumption 5 (No Autocorrelation in \( f_t \)):
  • \( \hat{\Sigma}_T = \frac{1}{T} \sum_{t=1}^{T} (f_t - \bar{f})(f_t - \bar{f})' \)
    • \( \rightarrow_p \lim_{T \to \infty} \text{Var}\left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (f_t - E(f_t))(f_t - E(f_t))' \right) \)

• Ross (1976) shows that the absence of arbitrage implies:
  • \( H_o : E(R_t) = e_N \gamma_0 + B \gamma_1 = (e_N, B) \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix} \equiv B_c \gamma \). Or equivalently,
    • \( H_o^\alpha : \alpha = e_N \lambda_o + B \lambda_1 = B_c \lambda \), where \( \lambda_1 = \gamma_1 - E(f_t) \).
(2) Estimation and Testing [Ahn and Gadarowski, 2004]

- Let \( Z_t = (1, f_i')' \).
- Let \( \Xi = \lim_{T \to \infty} Var\left( \frac{1}{\sqrt{T}} \Sigma_{t=1}^T Z_t \otimes \epsilon_t \right) \).
- Let \( \hat{\Xi}_1 \) be the Newey-West estimator of \( \Xi \) using OLS residuals instead of \( \epsilon_t \). Then, it is consistent under Assumption 1.
- Under Assumption 2,
  \[
  \hat{\Xi}_2 \equiv \frac{1}{T} \Sigma_{t=1}^T \left( Z_t Z_t' \otimes \epsilon_t \epsilon_t' \right) \to_p \Xi.
  \]
- Under Assumption 3,
  \[
  \hat{\Xi}_3 \equiv \left( \frac{1}{T} \Sigma_{t=1}^T Z_t Z_t' \right) \otimes \left( \frac{1}{T} \Sigma_{t=1}^T \epsilon_t \epsilon_t' \right) \equiv \hat{\Delta}_{ZZ} \otimes \hat{\Sigma} \to_p \Xi.
  \]
Two-Pass Estimation of lambdas:

- Let $\hat{\Lambda} = (\hat{\alpha}, \hat{B})$ be the OLS estimator of $\alpha$ and $B$.
- Let $A$ be any $N \times N$ positive definite matrix.
- Then, a two-pass estimator of $\gamma$ is given:

$$\hat{\gamma}_{TP} = \left( \hat{B}_c' A \hat{B}_c \right)^{-1} \hat{B}_c' A \hat{\alpha}.$$ 

If $A = I_N$, the TP estimator is the Fama-MacBeth estimator.

$$\text{Cov}(\hat{\gamma}_{TP}) = \left( \hat{B}_c' A \hat{B}_c \right)^{-1} \hat{B}_c' A \hat{\Omega} A \hat{B}_c \left( \hat{B}_c' A \hat{B}_c \right)^{-1},$$

where $\hat{\Omega} = (\hat{\lambda}_1 \Delta_{ZZ}^{-1} \otimes I_N) \hat{\Xi} (\Delta_{ZZ}^{-1} \hat{\lambda}_* \otimes I_N)$, $\hat{\lambda}_* = (1, -\hat{\lambda}_1)'$, and $\hat{\lambda}_1$ is any consistent estimator of $\lambda_1$.

- Asymptotically optimal choice of $A = \left( \hat{\Omega} \right)^{-1}$.

Let $\hat{\lambda}_{OMD}$ be the optimal TP estimator using $\left( \hat{\Omega} \right)^{-1}$.

Estimation of gammas:

- $\hat{\gamma}_{TP} = \left( \hat{B}_c' AB_c \right)^{-1} \hat{B}_c' A \hat{R} = \hat{\lambda}_{TP} + J \bar{f} = \begin{pmatrix} \hat{\lambda}_{0,TP} \\ \hat{\lambda}_{1,TP} + \bar{f} \end{pmatrix},$

where $J = \begin{pmatrix} 0_{1 \times k} \\ I_k \end{pmatrix}$. 

ASSET PRICING-12
• Under Assumption 4,
\[
\text{Cov}(\hat{\gamma}_{TP}) = \text{Cov}(\hat{\lambda}_{TP}) + \frac{1}{T} J \hat{\Sigma}_F J'.
\]

• Under Assumption 5,
\[
\text{Cov}(\hat{\gamma}_{TP}) = \text{Cov}(\hat{\lambda}_{TP}) + \frac{1}{T} J \hat{\Sigma}_F J'.
\]

• Model Specification test:
  \[
  Q_{OMD} \equiv \frac{T - N + 1}{N - 1 - k} \left( \hat{\alpha} - \hat{B}_c \hat{\lambda}_{OMD} \right)' \hat{\Omega}^{-1} \left( \hat{\alpha} - \hat{B}_c \hat{\lambda}_{OMD} \right) \rightarrow \chi^2(N - 1 - k) \frac{N - 1 - k}{N - 1 - k}.
  \]

• Alternative test:
  • Assume:
  \[
  \alpha = e_N \lambda_0 + B \lambda_1 + S \lambda_2,
  \]
  where S contains firm specific variables such as firm sizes or book values.
  • \( H_o \) and \( H_o^\alpha \) imply \( \lambda_2 = 0 \).
  • For detailed test procedures, see Ahn and Gadarowski.

ASSET PRICING-13
• Empirical Suggestions from Ahn and Gadarowski
  • Need to check persistency of factors.
    → When factors follow unit root or near-unit-root processes, the TP estimators are unreliable.
  • Do not use too many assets (N). 25 or fewer would be appropriate.
  • Testing autocorrelation in time-series OLS residuals and factors are important. Heteroskedasticity-robust Q tests are more reliable than autocorrelation-robust Q tests.
  • The Q tests generally have low power.
  • The t-tests based on nonoptimal TP estimators are more reliable than those based on the optimal TP estimators.

(3) What happens if the beta matrix $B$ is not of full column?

• When?
  • Some factors are in fact not the determinants of returns.
    → Kan and Zhang (1999, JF) call such factors “useless factors”.
    → The columns of $B$ corresponding to useless factors are zero vectors.
    → rank($B$) < $k$. 
• Consequences?
  • The TP estimator is severely biased.
  • The price of a useless factor would appear to be significant.

• How to find out correct factors?
  • Conner and Korajczyk (1993, JF)

• What happens if some unimportant factors are used?
  → The TP estimators are severely biased.
  
  [Kan and Zhang, 1999, JF]
  → Important to test how many factors to be used.

• How many factors to be used?
  • When candidate factors are all observed:
    • Connor and Korajczyk (1993, JF).
Question:
- Suppose we do not observe factors. Wish to estimate factors.
- How many factors?

Methods:
- MLE assuming factors are iid standard normal over time:
  Large T and Small N. [See Campbell, Ch. 6.4].
- Jones (2001, JFE): Large N and small T.
- Bai (2003, ECON), and Bai and Ng: Both N and T are large.
Stochastic Discount Factor Model

- There are many asset pricing models (CAPM, APT, consumption-based CAPM, intertemporal equilibrium model):
  → See Ch. 8 of CLM.
- A particular asset pricing model typically implies:
  \[ E[R_t m_t(f_t|\delta)] = e_N, \]
  where \( m(f_t|\delta) \) is a scalar function of factors and a parameter vector \( \delta \),
  \( e_N \) is the normalized price vector, and \( m(f_t|\delta) \) is called “stochastic discount factor” (SDF).
  → For linear factor models, \( m(f_t|\delta) = \delta_1 + f_t'\delta_2 \), where \( \delta = (\delta_1,\delta_2')' \).
- The parameter vector \( \delta \) and model specification can be tested by GMM.

- Let \( w_t(\delta) = R_t m_t(\delta) - e_N \); then, \( E[w_t(\delta)] \) is the pricing error.
- If no pricing error, then, \( E[w_t(\delta)] = 0_{N \times 1} \).

- HJ-distance (Hansen and Jagannathan, 1997):
  - \( HJ(\delta) = \sqrt{E[w_t(\delta)]'G^{-1}E(w_t(\delta))} \), where \( G = E(R_t R_t') \).
  - Correct model has \( HJ(\delta) = 0 \).
  - For a given \( \delta \), this measure equals the maximum pricing error generated by a given asset pricing model.
• Standard GMM test for no pricing error:
  
  * Let $S_T$ be a consistent estimator of $\lim_{T \to \infty} \text{Cov}\left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t(\delta) \right)$.

  [by White (1980) or Newey and West (1987).]

  * Let $\hat{\delta}_{GMM}$ be the optimal GMM estimator.

  * Then, under the hypothesis of no pricing error,
    
    $J_T = T w_T(\hat{\delta}_{GMM})' S_T^{-1} w_T(\hat{\delta}_{GMM}) \to_d \chi^2(N - p),$

    where $w_T(\delta) = \frac{1}{T} \sum_{t=1}^{T} w_t(\delta)$, and $p$ is the # of parameters in $\delta$.

• Jagannathan-Wang test:
  
  * $HJ_T(\delta) = \sqrt{w_T(\delta)' G_T^{-1} w_T(\delta)}$.

  * Let $\hat{\delta}_{HJ}$ is the minimizer of $HJ_T(\delta)$.

  * Under the hypothesis of no pricing error,
    
    $SHJ_T = T \times [HJ_T(\hat{\delta}_{HJ})]^2 \to \text{weighted } \chi^2.$

  * JW (1996) provides a simulation method to compute p-value for this statistic.

  * JW conjecture that this HJ test would have better finite sample properties, because non-optimal GMM often has better finite sample properties than optimal GMM.
  • For linear factor models, the JW method has extremely poor finite sample properties, poorer than optimal GMM, especially when N is large.
  • Better to use optimal GMM!