Empirical Asset Pricing

Francesco Franzoni

Swiss Finance Institute - University of Lugano

Testing Asset Pricing Models: Overview

Lecture Outline

1. Testing CAPM: historical development
2. Statistical description of the asset pricing tests
3. Stochastic Discount Factor Representation and Tests
4. Conditional Asset Pricing Models
5. Statistical discipline

Relevant readings:

- Cochrane: chapters 6, 7, 8, 12, 14, 15, 16
- Huang and Litzenberger: chapter 10

1. Testing CAPM: historical development

Quick Review: The CAPM

- The Sharpe-Lintner-Mossin (SLM) version of the model assumes that the risk-free rate is available. For any asset i :

$$E(R_i) = R_f + \beta_i \left(E(R_m) - R_f \right) \tag{1}$$

where R_m is the total wealth (i.e., market) portfolio and

$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} \tag{2}$$

- The model states that the market return is mean-variance efficient

- That is: the expected return on any asset can be spanned by the expected return on the market and the risk free rate

- The Black (1972) version of the model is derived without assuming the existence of the risk-free rate:

$$E(R_i) = E(R_0) + \beta_i (E(R_m) - E(R_0)) \tag{3}$$

where R_0 is the return on a mean-variance efficient portfolio with zero covariance with R_m. It turns out that $E(R_m) - E(R_0) > 0$

Testable Predictions

- SLM version:

1. Consider the time-series regression

$$R_{it} - R_f = \alpha_i + \beta_i \left(R_{mt} - R_f \right) + \varepsilon_{it}. \tag{4}$$

Take expectations of both sides. The model predicts

(a) $\alpha_i = 0 \quad \forall i = 1 \ldots N$

2. Consider the cross-sectional regression

$$E_T\left(R_i^e\right) = a + \beta_i \lambda + d_i \gamma + u_i \qquad i = 1...N \tag{5}$$

$E_T\left(R_i^e\right)$: time-series average excess return (it is an estimate of the expected excess return)
d_i : vector of stock characteristics different from beta (e.g. size, dividend yield, volatility, etc.):

(a) $\lambda = E\left(R_m\right) - R_f > 0$

(b) $\alpha_i \equiv a + u_i = 0 \quad \forall i$ (zero pricing errors)

(c) $\gamma = 0$

Conceptual Issues in the Tests

1. The model predictions are stated in terms of *ex ante* expected returns and betas but tested on sample moments
   - The solution is to assume Rational Expectations: the realized returns are drawings from equilibrium distributions. Because of rational expectations, the *ex ante* distributions correspond to the equilibrium distributions. So, the sample moments converge to the ex-ante moments

2. CAPM holds period by period: it is a conditional model.
   How to handle non-stationarity in the moments of the *ex ante* distribution?
   - One solution is to assume that CAPM holds unconditionally too.
     Implications of this assumption:
     The conditional CAPM:

$$
\begin{aligned}
E_t\left(R^e_{i,t+1}\right) &= \beta_{it} E_t\left(R^e_{m,t+1}\right) \\
\beta_{it} &= \frac{Cov_t\left(R_{i,t+1}, R_{m,t+1}\right)}{Var_t\left(R_{m,t+1}\right)}
\end{aligned}
$$

Take unconditional expectations of both sides

$$E\left(E_t\left(R^e_{i,t+1}\right)\right) = E\left(\beta_{it}E_t\left(R^e_{m,t+1}\right)\right)$$

$$E\left(R^e_{i,t+1}\right) = Cov\left(\beta_{it}, E_t\left(R^e_{m,t+1}\right)\right) + E\left(\beta_{it}\right)E\left(R^e_{m,t+1}\right)$$

For the model to hold unconditionally, one needs to assume that

$$Cov\left(\beta_{it}, E_t\left(R^e_{m,t+1}\right)\right) = 0$$

$$E\left(\beta_{it}\right) = \beta_i = \frac{Cov\left(R_i, R_m\right)}{Var\left(R_m\right)}$$

This is not necessarily true. For example:

– Beta can change along the business cycle

– Beta can change along the firm's life-cycle

- The other solution is to model the conditional moments. We will discuss this strategy later. For now, the focus is on unconditional tests

3. *Roll's (1977) critique*: the total wealth portfolio includes non-marketable assets (e.g. human capital, private businesses, etc.). Instead, the tests typically focus on a proxy of the market that is traded (e.g. the S&P 500, the CRSP index, etc.)

   - Consequence: measurement error in the market return.
     Rejection of CAPM may depend on use of incorrect market portfolio

   - Most tests ignore the unobservability and assume proxy is mean-variance efficient

   - Also: if the true market portfolio is sufficiently correlated with the proxy (above 70%), a rejection of the proxy implies a rejection of the true portfolio (Stambaugh (1982), Kandal and Stambaugh (1987), Shanken (1987))

   - Some authors have tried to compute a broader proxy by including the return on human capital (Jagannathan and Wang (1996))

Econometric Issues in the Tests

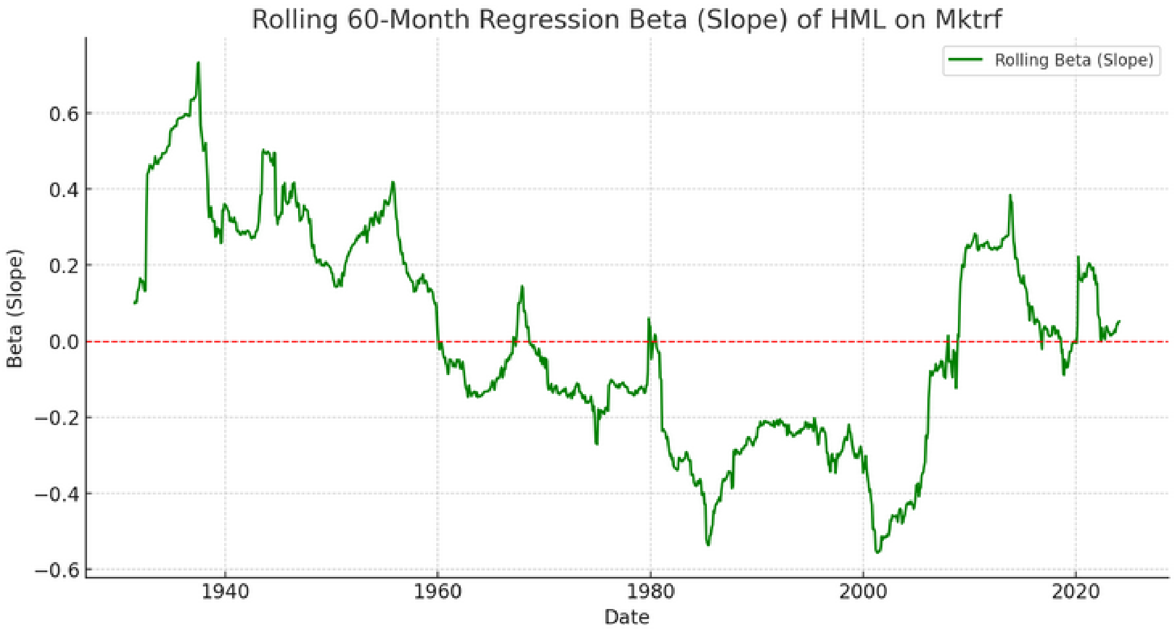
1. The errors in equation (5) are likely to be heteroskedastic and correlated across assets
   - OLS coefficients are unbiased but inefficient (OLS is not BLUE, GLS is)
   - OLS standard errors are biased
   - Possible Solutions:

   (a) Use feasible GLS...I have never seen this in practice

   (b) Use OLS to estimate the slopes and adjust the standard errors

   (c) Fama and MacBeth (1973) procedure

2. The betas in equation (5) are measured with error because they are estimates of true betas
   - Measurement error causes $\hat{\gamma}$ to be biased towards zero and CAPM rejected
   - Main Solution: Group the data into portfolios
3. The betas in equation (4) are likely to change over time
   - Betas change over the life-cycle in a non-stationary way
   - Historical solution: form portfolios according to a stationary characteristic
   - Does it work? See Franzoni (2002): Where's beta going?



4. High volatility of individual stock returns
   - For individual stocks, cannot reject hypothesis that average returns are all the same
   - St. error of mean is $\frac{\sigma_i}{T^{1/2}}$. If σ_i between 40% and 80% and $T = 60$ months, very large confidence intervals
   - Solution: form portfolios
5. Non-synchronous trading

- Small stocks react slowly to information
- Solution: Scholes and Williams (1977)

$$\hat{\beta} = \frac{\hat{\beta}^+ + \hat{\beta} + \hat{\beta}^-}{1 + 2\rho}$$

where $\hat{\beta}^+, \hat{\beta}^-$, and $\hat{\beta}$ come from the simple regression of stock returns on leads, lags, and the contemporaneous market return:

$$\begin{aligned} R^e_{it} &= \alpha^+_i + \beta^+_i R_{m,t+1} + \varepsilon^+_t \\ R^e_{it} &= \alpha_i + \beta_i R_{m,t} + \varepsilon_t \\ R^e_{it} &= \alpha^-_i + \beta^-_i R_{m,t-1} + \varepsilon^-_t \end{aligned}$$

ρ is the 1st order auto-correlation of the market return.
Note that if $\rho = 1$, it is as if you were running three times the same regression. Then, you have to divide the sum of the betas by $1 + 2\rho = 3$
If $\rho = 0$, each of the three regressions provides independent information on the reaction of stock returns to market returns. Then, the sum of the three betas does not need to be normalized.

- Another solution is due to Dimson (1979)

$$\hat{\beta} = \hat{\beta}^{+} + \hat{\beta} + \hat{\beta}^{-}$$

where the estimates come from

$$R^e_{it} = \alpha_i + \beta_i^{+} R_{m,t+1} + \beta_i R_{m,t} + \beta_i^{-} R_{m,t-1} + \varepsilon_t$$

Portfolio Grouping Approach

- Consider forming G groups of L stocks

- The measured beta for group g is

$$\hat{\beta}_g = \frac{1}{L}\sum_{j=1}^{L}\hat{\beta}_j = \frac{1}{L}\sum_{j=1}^{L}\left(\beta_j + w_j\right) = \beta_g + \frac{1}{L}\sum_{j=1}^{L} w_j$$

- So the variance of the measurement error for the portfolio beta is

$$Var\left(\frac{1}{L}\sum_{j=1}^{L} w_j\right) = \frac{1}{L}\sigma_w^2$$

  as long as the measurement errors of the individual securities are uncorrelated (classical measurement error)

- As $L \to \infty$ the measurement error goes to zero and the estimates are consistent

Grouping and Efficiency

- Grouping reduces dispersion in betas relative to ungrouped data:

$$\underbrace{Var\left(\beta_i\right)}_{Ungrouped\ Variance} = \underbrace{Var\left(E\left(\beta|G\right)\right)}_{Between-Group\ Variance} + \underbrace{E_G\left(Var\left(\beta_i|G\right)\right)}_{Average\ Within-Group\ Variance}$$

- Hence, it reduces the efficiency of the cross-sectional estimates
  - See Gagliardini, Ossola, and Scaillet (2016) for a stock-level approach to estimating risk premia that maximizes efficiency

- Need grouping procedure that maximizes across-group variation in betas and minimizes correlation with measurement error

- E.g.: group by previous estimates of betas from non-overlapping data
  - Measurement error is uncorrelated over time
  - Estimated betas are positively correlated with true betas over time

2. Statistical description of AP tests

- General focus on multi-factor asset pricing models

$$E\left(R_i^e\right) = \beta_i' \lambda \tag{6}$$

where β_i is a K-dimensional vector of multiple-regression slopes and λ is K-vector of factor risk premia

- For example: APT, ICAPM

- CAPM is the case with $K = 1$

- The methodology extends unambiguosly from the CAPM tests to tests of multifactor models

- Questions:
  - How to estimate parameters?
  - Standard errors?
  - How to test the model predictions?

Time-Series Regressions

- You can apply this approach only if factors $f_{1,...,}f_K$ are returns

- Then, the model (6) applies to factors as well:

$$E\left(f_k\right) = \lambda_k \quad k = 1...K \tag{7}$$

- So, one can re-write the model in (6) as:

$$E\left(R_i^e\right) = \beta_i' E\left(f\right) \tag{8}$$

where $E\left(f\right)$ is the vector of **expected excess returns** on the k factors

- Assume $K = 1$ for simplicity, consider the time-series regression:

$$R_{it}^e = \alpha_i + \beta_i f_t + \varepsilon_t \tag{9}$$

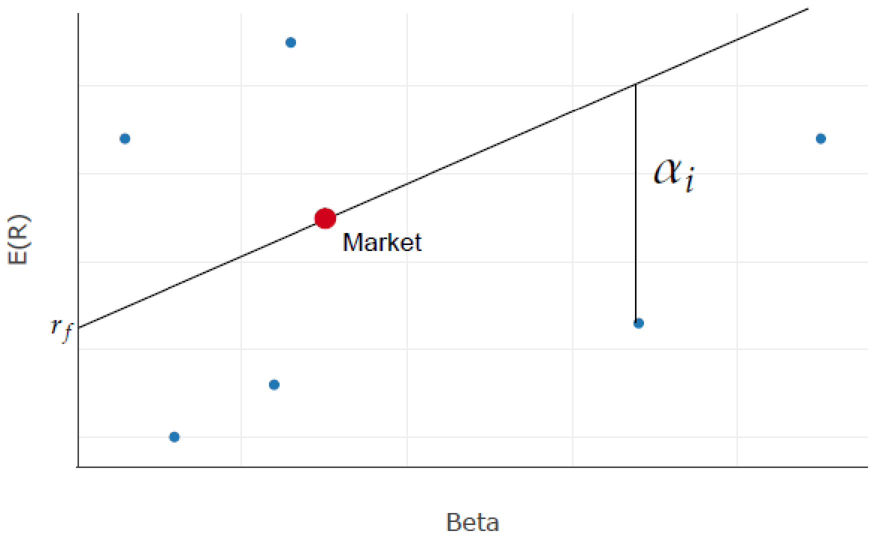
- Take expectation of each side. Then, the implication of (8) is

$$\alpha_i = 0 \quad i = 1...N$$

- You can estimate α_i by running regression (9) for each asset

- Then, use t-tests to test $\alpha_i = 0$

- In an E(R)-β space, you are pricing the factor correctly by fitting a line through the factor and allowing pricing errors on the other assets



- You want to test the hypothesis that all alphas are jointly zero

- The ε_{it} are correlated across assets with variance-covariance matrix $\mathbf{\Sigma} = E\left(\varepsilon_t \varepsilon_t'\right)$

- The asymptotically valid test for

$$H_o : \alpha = 0$$

is

$$\hat{\alpha}' \left[\widehat{Var(\hat{\alpha})} \right]^{-1} \hat{\alpha} \sim \chi^2_N$$

where α is a N-dimensional vector of pricing errors

- Intuition: reject the model if the weighted sum of the squared errors is far from zero

- In the case $K = 1$, assuming no autocorrelation and stationarity, and noting that $\hat{\alpha}_i$ contains the average ε_{it}, the test becomes

$$T \left[1 + \left(\frac{E_T(f)}{\hat{\sigma}(f)} \right)^2 \right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \sim \chi^2_N$$

where $E_T(f)$ is the sample mean of the factor (an estimate of the risk premium) and $\hat{\Sigma}$ is the sample vcov matrix of the residuals from the N regressions

Proof. (Sketch): $\widehat{Var}(\hat{\alpha})$ is $N \times N$ matrix

$$\widehat{Var}(\hat{\alpha}) = \begin{bmatrix} \hat{\sigma}^2_{\hat{\alpha}_1} & \hat{\sigma}_{\hat{\alpha}_1\hat{\alpha}_2} & \cdots & \hat{\sigma}_{\hat{\alpha}_1\hat{\alpha}_N} \\ \hat{\sigma}_{\hat{\alpha}_2\hat{\alpha}_1} & \hat{\sigma}^2_{\hat{\alpha}_2} & & \\ \vdots & & \ddots & \\ \hat{\sigma}_{\hat{\alpha}_N\hat{\alpha}_1} & & & \hat{\sigma}^2_{\hat{\alpha}_N} \end{bmatrix}$$

Focus on:

$$\begin{aligned} \hat{\alpha}_i &= \bar{R}^e_i - \hat{\beta}_i E_T(f) \\ &= \alpha_i + \beta_i E_T(f) + \bar{\varepsilon}_i - \hat{\beta}_i E_T(f) \\ &= \alpha_i + \left(\beta_i - \hat{\beta}_i\right) E_T(f) + \bar{\varepsilon}_i \end{aligned}$$

So:

$$
\begin{aligned}
Var(\hat{\alpha}_i) &= Var\left(\hat{\beta}_i\right) E_T^2(f) + Var(\bar{\varepsilon}_i) \\
&= \frac{\sigma_{\varepsilon_i}^2}{\hat{\sigma}^2(f)\, T} E_T^2(f) + \frac{1}{T}\sigma_{\varepsilon_i}^2 \\
&= \frac{1}{T}\sigma_{\varepsilon_i}^2 \left[1 + \left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2\right]
\end{aligned}
$$

Similarly:

$$
Cov\left(\hat{\alpha}_i, \hat{\alpha}_j\right) = \frac{1}{T}\sigma_{\varepsilon_i \varepsilon_j} \left[1 + \left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2\right]
$$

Hence, it follows that:

$$
Var(\hat{\alpha}) = \frac{1}{T}\left[1 + \left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2\right] \boldsymbol{\Sigma}
$$

So:

$$\widehat{Var}\left(\hat{\alpha}\right) = \frac{1}{T}\left[1 + \left(\frac{E_T\left(f\right)}{\hat{\sigma}\left(f\right)}\right)^2\right]\hat{\Sigma}$$

Then:

$$\hat{\alpha}'\left[\widehat{Var}\left(\hat{\alpha}\right)\right]^{-1}\hat{\alpha} = T\left[1 + \left(\frac{E_T\left(f\right)}{\hat{\sigma}\left(f\right)}\right)^2\right]^{-1}\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}$$

You only need to prove that the distribution is a chi-squared. For that you can use standard theorems on the limit distribution of squared residuals. ∎

- Gibbons, Ross, and Shanken (GRS, 1989) provide a small sample test assuming joint normality of the ε_{it}

$$\frac{T-N-1}{N}\left[1+\left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2\right]^{-1}\hat{\alpha}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\alpha}\sim F_{N,\,T-N-1}$$

- Using efficient set algebra, one can prove that

$$\left(\frac{\mu_q}{\sigma_q}\right)^2=\left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2+\hat{\alpha}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\alpha} \tag{10}$$

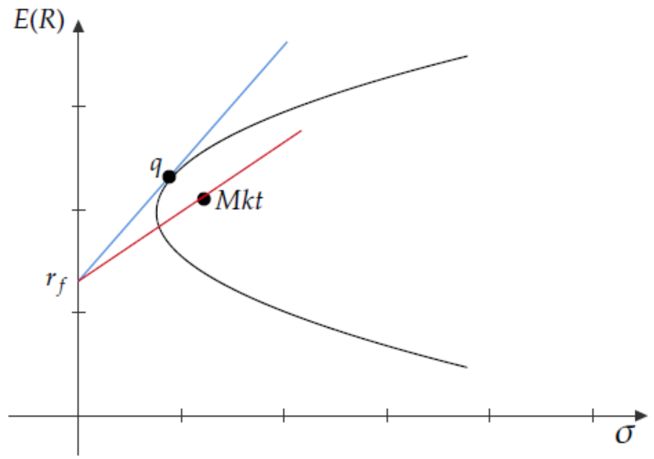
where $\frac{\mu_q}{\sigma_q}$ is the Sharpe ratio of the tangency portfolio and $\frac{E_T(f)}{\hat{\sigma}(f)}$ is the Sharpe ratio of the factor

- Then, the GRS statistic can be rewritten as

$$\frac{T-N-1}{N}\frac{\left(\mu_q/\sigma_q\right)^2-\left(E_T(f)/\hat{\sigma}(f)\right)^2}{1+\left(\frac{E_T(f)}{\hat{\sigma}(f)}\right)^2}\sim F_{N,\,T-N-1}$$

- In this form, the interpretation is: Reject the null if the factor is far from the tangency portfolio on the ex-post efficient frontier

- In other words: reject CAPM if the market is far from the ex-post efficient frontier



- In case of $K > 1$ factors, the GRS statistic is

$$\frac{T-N-K}{N}\frac{\left(\frac{\mu_q}{\sigma_q}\right)^2 - E_T(f)'\,\hat{\Omega}^{-1}E_T(f)}{1+E_T(f)'\,\hat{\Omega}^{-1}E_T(f)} \sim F_{N,T-N-K}$$

where

$$\hat{\Omega} = \frac{1}{T}\sum_{i=1}^{T}\left[f_t - E_T(f)\right]\left[f_t - E_T(f)\right]'$$

is the sample vcov matrix of the factors and

$$E_T(f)' \hat{\Omega}^{-1} E_T(f)$$

is the multifactor equivalent of the squared Sharpe ratio of the factor

Proof of Equation (10)

- Let q be a portfolio on the mean variance efficient frontier. Let μ be the vector of expected excess returns on the $N+1$ assets in the market (i.e. N test assets and 1 factor). Hence, q has minimum variance for given expected return e

$$\begin{aligned} Min \quad & q'Vq \\ s.t.\ q'\mu &= e \end{aligned}$$

- The Lagrangean for this problem is

$$L = q'Vq + 2\lambda\left(e - q'\mu\right)$$

which gives first order conditions with respect to q and λ (the Lagrange multiplier)

$$\begin{aligned} 2Vq - 2\lambda\mu &= 0 \\ e - q'\mu &= 0 \end{aligned}$$

- Hence, the frontier portfolio q has to satisfy

$$q = \lambda V^{-1}\mu \tag{11}$$

- Using Equation (11), the variance of q is therefore

$$\sigma_q^2 = q'Vq = \lambda^2 \mu' V^{-1} \mu$$

- The squared expected excess return of q is

$$\mu_q^2 = \left(q'\mu\right)^2 = \lambda^2 \left(\mu' V^{-1} \mu\right)^2$$

- We can then compute the squared Sharpe ratio of q as

$$\frac{\mu_q^2}{\sigma_q^2} = \mu' V^{-1} \mu \tag{12}$$

- Now, let us remember that this market is composed of $N+1$ assets, where the first asset is the factor f

- The vector of returns of the N assets can be written as

$$R_t = \alpha + \beta f_t + \varepsilon_t,$$

where α and β are N-dimensional vectors capturing the alphas and betas of the N assets with respect to the factor f. Also, ε is the idiosyncratic component of returns, which has variance-covariance matrix equal to $\mathbf{\Sigma}$

- Hence, the vector μ can be written as

$$\mu = \begin{bmatrix} \mu_f \\ \alpha + \beta \mu_f \end{bmatrix} \tag{13}$$

- For the same reason, the variance-covariance matrix V can be written as

$$V = \begin{bmatrix} \sigma_f^2 & \beta' \sigma_f^2 \\ \beta \sigma_f^2 & \beta \beta' \sigma_f^2 + \mathbf{\Sigma} \end{bmatrix} \tag{14}$$

- Using the formula for the inverse of a partitioned matrix (see, e.g., Greene, Econometric Analysis), we can obtain the inverse of V as

$$V^{-1} = \begin{bmatrix} \frac{1}{\sigma_f^2} + \beta' \boldsymbol{\Sigma}^{-1} \beta & -\beta' \boldsymbol{\Sigma}^{-1} \\ -\boldsymbol{\Sigma}^{-1} \beta & \boldsymbol{\Sigma}^{-1} \end{bmatrix} \tag{15}$$

- Using Equations (13) and (15), after some tedious algebra, we can re-write Equation (12) as

$$\frac{\mu_q^2}{\sigma_q^2} = \frac{\mu_f^2}{\sigma_f^2} + \alpha' \boldsymbol{\Sigma}^{-1} \alpha,$$

which completes the proof

Cross-Sectional Regressions

- This is the only approach available when the factors are not returns
- Two-pass methodology:
  1. Estimate β_i from time-series regressions with multiple factors if $K > 1$

     $$R_{it}^e = a_i + \beta_i' f_t + \varepsilon_{it} \quad t = 1 \ldots T \tag{16}$$

     Note that if the AP model in (6) is correct

     $$a_i = \beta_i' \left(\lambda - E\left(f\right)\right) \tag{17}$$

     which is not equal to zero in general
  2. Regress average returns on estimated β_i

     $$E_T\left(R_{it}^e\right) = \beta_i' \lambda + \alpha_i \quad i = 1 \ldots N \tag{18}$$

     where E_T denotes a time-series average over the T observations in the sample.
     Based on equations (16) and (17), under the null, the pricing errors α_i's probability limit is

(assume that $\hat{\beta} \to \beta$ and $\hat{\lambda} \to \lambda$):

$$\begin{aligned} \text{Plim } \alpha_i &= \text{Plim } (E_T(R^e_{it}) - \beta_i'\lambda) \\ &= \text{Plim } (a_i + \beta_i' E_T(f_t) + E_T(\varepsilon_{it}) - \beta_i'\lambda) \\ &\overset{H_0}{=} \text{Plim } (\beta_i'(\lambda - E(f)) + \beta_i' E_T(f_t) + E_T(\varepsilon_{it}) - \beta_i'\lambda) \\ &= \text{Plim } (E_T(\varepsilon_{it}) + \beta_i'(E_T(f_t) - E(f_t))) = 0 \end{aligned} \tag{19}$$

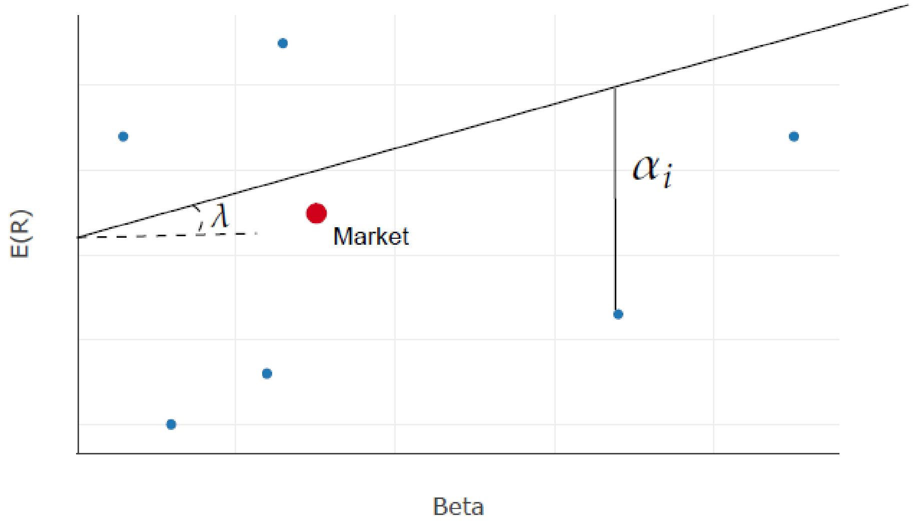
So, based on the expression in the last step of Equation (19), the vcov matrix of α, which is the $N \times 1$ vector of residuals from regression (18), is

$$E\left(\alpha\alpha'\right) = \frac{1}{T}\left(\Sigma + \beta\Omega\beta'\right)$$

where β is a $N \times K$ matrix of factor loadings for the N assets on the K factors, and Ω is the vcov of the K factors

- $\hat{\alpha}$ is the vector of fitted residuals, which under the null, have zero expectation

- In this case, you are fitting a line through the expected returns, trying to minimize pricing errors, estimating a risk premium for the factor, and allowing a non-zero pricing error for the factor



- One can test the AP model with

$$\hat{\alpha}' \left[Var\left(\hat{\alpha}\right)\right]^{-1} \hat{\alpha} \sim \chi^2_{N-K}$$

- Using standard OLS results ($Var\left(\hat{u}\right) = M Var\left(u\right) M$, where $\hat{u}$ are the fitted residuals and $M = I - X(X'X)^{-1}X'$ is the residual making matrix. In our case $u = \alpha$ and $X = \beta$), we can compute the vcov of $\hat{\alpha}$ as

$$Var\left(\hat{\alpha}\right) = \frac{1}{T}\left(I - \beta\left(\beta'\beta\right)^{-1}\beta'\right)\Sigma\left(I - \beta\left(\beta'\beta\right)^{-1}\beta'\right)$$

where $\beta\Omega\beta'$ cancels out from the vcov matrix

- Notice that these formulas assume that β are known. See next for corrections that take into account that β is estimated

Shanken Correction

- In fact, β_i are estimates, not the true parameters
- Hence, need to account for sampling error in β_i, when computing standard errors
- Shanken's (1992) correction :

$$\sigma^2\left(\hat{\lambda}_{ols}\right) = \frac{1}{T}\left[\left(\beta'\beta\right)^{-1}\beta'\Sigma\beta\left(\beta'\beta\right)^{-1}\left(1+\lambda'\Omega^{-1}\lambda\right)\right] + \frac{1}{T}\Omega \tag{20}$$

where Ω is the vcov of the factors

- You have a multiplicative $\left(1+\lambda'\Omega^{-1}\lambda\right)$ correction and an additive correction $\frac{1}{T}\Omega$
- Is the multiplicative correction important?
- Consider CAPM. In annual data: $(\lambda_m/\sigma_m)^2 = (0.08/0.16)^2 = 0.25$
  The correction is important.
  In monthly data $(\lambda_m/\sigma_m)^2 = 0.25/12 \approx 0.02$.
  The multiplicative correction is not important

- The additive correction is likely to be important

- $\frac{1}{T}\Omega$ is the variance of the mean of the factor, which is non-negligible in case f is R_m

Time-Series vs. Cross-Section

- How are two approaches different?

1. TS can be applied only if factors are returns
   - In CAPM, it fits a line through the pricing errors by forcing a zero pricing error on R_m
   - Test: $\alpha_i = 0$ for all i
2. CS only alternative when factors are not returns
   - In CAPM, it minimizes all pricing errors, by allowing some error on R_m
   - Historically CS has been used to test for characteristics in the cross-section of returns (especially using Fama-MacBeth approach)

Fama-MacBeth (1973) Methodology

- It is a three-pass procedure:

1. Obtain β_{it} from time-series regressions, using data up to $t-1$
2. At each date t, run a cross sectional regression

$$R^e_{it} = \beta'_{it}\lambda_t + \alpha_{it} \qquad i = 1...N$$

   and obtain a time series of $\hat{\lambda}_t$ and $\hat{\alpha}_{it}$ $\quad t = 1...T$

3. Finally, obtain full-sample estimates as time series means

$$\hat{\lambda} = \frac{1}{T}\sum_{\iota=1}^{T}\hat{\lambda}_t$$
$$\hat{\alpha}_i = \frac{1}{T}\sum_{\iota=1}^{T}\hat{\alpha}_{it}$$

and use the standard error of the mean

$$\sigma\left(\hat{\lambda}\right) = \frac{1}{T^{1/2}}\left(\underbrace{\frac{\sum_{t=1}^{T}\left(\hat{\lambda}_t-\hat{\lambda}\right)^2}{T-1}}_{\widehat{Var(\hat{\lambda}_t)}}\right)^{1/2}$$

$$\sigma\left(\hat{\alpha}_i\right) = \frac{1}{T^{1/2}}\left(\underbrace{\frac{\sum_{t=1}^{T}\left(\hat{\alpha}_{it}-\hat{\alpha}_i\right)^2}{T-1}}_{\widehat{Var(\hat{\alpha}_{it})}}\right)^{1/2}$$

- One can test for zero pricing errors using the quadratic form

$$\hat{\alpha}' cov\left(\hat{\alpha}\right)\hat{\alpha} \sim \chi^2_{N-1}$$

where $\hat{\alpha}$ is the vector of stacked $\hat{\alpha}_i$ and

$$cov\left(\hat{\alpha}\right) = \frac{1}{T^2}\sum_{t=1}^{T}\left(\hat{\alpha}_t-\hat{\alpha}\right)\left(\hat{\alpha}_t-\hat{\alpha}\right)'$$

- There are two main advantages to this approach:

  1. It allows for time-varying betas
  2. It computes standard errors by getting around the problem of heteroskedastic and correlated errors: it exploits sample variation in $\hat{\lambda}_t$ and $\hat{\alpha}_{it}$

3. SDF Representation and Tests

Equivalence of Representations

- In general one can always go from a Stochastic Discount Factor (SDF) representation

$$\begin{aligned} E\left(R_{it} m_t\right) &= 1 \\ m_t &= a - b f_t \end{aligned} \tag{21}$$

  where m_t is the SDF and f_t is a K-vector, to an expected return-beta representation

$$\begin{aligned} E\left(R_{it}\right) &= \lambda_0 + \lambda' \beta \\ \beta &= Cov\left(R_{it}, f_t\right) E\left(f_t f_t\right)^{-1} \end{aligned}$$

  and vice versa

- Here, we show the direction SDF$\longrightarrow$E(R)-β, see Cochrane ch. 6 for the proof of the other direction

- Start from Equation (21) and replace m_t into the first equation

$$\begin{aligned} E\left(R_{it}\left(a-bf_t\right)\right) &= 1 \\ aE\left(R_{it}\right)-bE\left(R_{it}f_f\right) &= 1 \\ aE\left(R_{it}\right) &= 1+bE\left(R_{it}f_t\right) \\ aE\left(R_{it}\right) &= 1+bE\left(R_{it}\right)E\left(f_t\right)+bCov\left(R_{it},f_t\right) \\ E\left(R_{it}\right)\left(a-bE\left(f\right)\right) &= 1+bCov\left(R_{it},f_t\right) \qquad (22) \end{aligned}$$

Note that by pricing the risk-free rate you get

$$\begin{aligned} R_f E\left(m\right) &= 1 \\ R_f &= \frac{1}{a-bE\left(f_t\right)} \end{aligned}$$

So, by replacing R_f into equation (22), you get

$$E\left(R_{it}\right)=R_f+bR_fCov\left(R_{it},f_t\right)$$

which can be expressed in E(R)-β form

$$\begin{aligned} E\left(R_{it}\right) &= \lambda_0 + \beta_{if}\lambda_1 \\ \beta_{if} &= \frac{Cov\left(R_{it,}f_t\right)}{Var\left(f_t\right)} \\ \lambda_1 &= bR_f Var\left(f_t\right) \\ \lambda_0 &= R_f \end{aligned} \tag{23}$$

Testing the SDF Representation

- The SDF representation can be tested using GMM on the moment conditions given by the pricing errors

$$E\left[R_{it} m_t(\delta)\right] = 1$$

where

$$m_t(\delta) = a - b f_t$$

- Given that we have N assets, define the pricing errors

$$w_t(\delta) = R_t m_t(\delta) - 1_N$$

where R_t is an N-vector of asset returns and 1_N is an N-vector of ones

- The N moment conditions for the vector of pricing errors $w_t(\delta)$ are

$$E\left[w_t(\delta)\right] = 0$$

- The GMM estimates the parameters in δ so as to minimize a quadratic form in the N moment conditions

$$E\left[w_t(\delta)\right]' A E\left[w_t(\delta)\right] \tag{24}$$

where A is a weighting matrix

- Assume that $w_t(\delta)$ is i.i.d. over time

- In this case, the optimal weighting matrix by Hansen and Singleton (1982) reduces to

$$A = \left[Var\left(w_t(\delta)\right)\right]^{-1}$$

which is also called **second-stage** weighting matrix, because you estimate it by computing pricing errors from a first stage in which the weighting matrix is typically the identity

- However, the optimal quadratic form in (24) cannot be used as a metric to make comparisons across different asset pricing models
  - If a model contains more noise, $Var\left(w_t(\delta)\right)$ is larger

– Then the quadratic form is smaller just because of the noise, and not because of smaller pricing errors (This argument echos Fama and French's (1993) defense against the rejection of their model by the GRS test)

- The solution proposed by Hansen and Jagannathan (1994) is to use

$$A = \left[E\left(R_t R_t\right)\right]^{-1}$$

which is the matrix of second moments of returns

- The advantage is that this matrix remains the same across different specifications of the AP model and allows comparisons of different models

- Also, the authors show that the square root of the resulting quadratic form, called the *Hansen-Jagannathan (H-J) distance*, is the pricing error of the most mispriced portfolio among the N assets by a given AP model

- They derive the distribution of this statistic, which is in general non-standard

- This approach is used in the Conditional CAPM application of Jagannathan and Wang (1996)

4. Conditional Asset Pricing

Conditional Asset Pricing (Cochrane ch. 8)

- In the case of the expected return-beta representation, we have seen above that a conditional model does not imply an unconditional model with the same factor in general
- The same conclusion holds for the SDF representation
- The coefficients in the SDF are time-varying
- Consider, for example, the SDF for CAPM

$$m_{t+1} = a_t + b_t R^m_{t+1}$$

- The pricing statement is also conditional on time t information

$$1 = E_t \left[\left(\underbrace{a_t + b_t R^m_{t+1}}_{m_{t+1}} \right) R_{it+1} \right]$$

- In general, this statement does not imply an unconditional pricing statement. Let's try. Take unconditional expectations of each side:

$$
\begin{aligned}
1 &= E\left[a_t R_{it+1} + b_t R^m_{t+1} R_{it+1}\right] \\
&= E\left[a_t\right] E\left[R_{it+1}\right] + Cov\left(a_t, R_{it+1}\right) \\
&\quad + E\left[b_t\right] E\left[R^m_{t+1} R_{it+1}\right] + Cov\left(b_t, R^m_{t+1} R_{it+1}\right)
\end{aligned}
$$

- Thus, you can get an unconditional model only if the covariance terms are zero:

$$
1 = E\left[\left(\underbrace{E\left[a_t\right]}_{a} + \underbrace{E\left[b_t\right]}_{b} R^m_{t+1}\right) R_{it+1}\right]
$$

which is **not** in general the case

Hansen and Richard Critique

- Testing conditonal models may not be enough to do things correctly
- To be sure to test the correct model, one needs to account for all the relevant conditioning information that investors use
- One needs to know investors' information set
- This is arguably impossible
- Hence, according to Hansen and Richard conditional factor models are not testable!
- It resonates with the Roll Critique

A partial solution: scaled factors

- Because investors' information changes over time, the parameters a_t and b_t in the SDF change
- One can try to model this time-variation by using a set of L instruments z_t, which are variables that plausibly enter investors' information set

$$\begin{aligned} a_t &= a' z_t \\ b_t &= b' z_t \end{aligned}$$

- Suppose, for example, 1 factor and 1 instrument

$$\begin{aligned} m_{t+1} &= a_t + b_t f_{t+1} \\ &= a_0 + a_1 z_t + (b_0 + b_1 z_t) f_{t+1} \\ &= a_0 + a_1 z_t + b_0 f_{t+1} + b_1 z_t f_{t+1} \end{aligned}$$

- In place of a conditional 1-factor model, one obtains an unconditional 3-factor model, with factors $(z_t, f_{t+1}, z_t f_{t+1})$ and fixed coefficients

- $z_t f_{t+1}$ are scaled factors

- In general, from K factors and L instruments one gets $(K + 1) \times (L + 1)$ unconditional factors (including a constant)

- This is a partial solution, because one could be omitting some relevant variable from z_t that is instead in investors' information set (Hansen and Richard Critique)

A brief history of Conditional Asset Pricing

- For a period of time, conditional factor models seemed to provide a solution to explain the failure of CAPM
  - Jagannathan and Wang (1996, JF) used a conditional CAPM with human capital to explain the size anomaly
  - Lettau and Ludvigson (2001, JF) used a conditional CAPM to explain the book-to-market anomaly (see later for the anomalies)
- However, Lewellen and Nagel (2006) put an end to this literature by showing that the empirical success of conditional asset pricing models was based on false premises

Lewellen and Nagel, 2006

- In E(R)-β representation, the conditional CAPM says

$$
\begin{aligned}
E_t\left(R^e_{it+1}\right) &= \beta_{it}\gamma_t \\
\beta_{it} &= \frac{Cov_t\left(R_{it+1}, R^m_{t+1}\right)}{Var_t\left(R^m_{t+1}\right)} \\
\gamma_t &= E_t\left(R^m_{t+1}\right) - R_{f,t}
\end{aligned}
$$

- Take unconditional expectations

$$
\begin{aligned}
E\left(R^e_{it+1}\right) &= E\left(\beta_{it}\gamma_t\right) \\
&= E\left(\beta_{it}\right)E\left(\gamma_t\right) + Cov\left(\beta_{it}, \gamma_t\right) \\
&= \bar{\beta}_i\gamma + Cov\left(\beta_{it}, \gamma_t\right) \qquad (25)
\end{aligned}
$$

- By definition, the unconditional alpha in a time-series regression has the following Plim

$$
\text{Plim } \alpha^u_i = E\left(R^e_{it+1}\right) - \beta^u\gamma \qquad (26)
$$

where $\beta^u = \frac{Cov(R_{it+1}, R^m_{t+1})}{Var\left(R^m_{t+1}\right)}$ is the unconditional beta and $\gamma = E\left(\gamma_t\right)$

- Replacing (25) into (26) gets a clear expression for the unconditional alpha

$$\alpha^u = \gamma(\bar{\beta} - \beta^u) + Cov(\beta_t, \gamma_t)$$

- Under mild assumptions, you have that $\bar{\beta} = \beta^u$. So, that

$$\alpha^u = Cov(\beta_t, \gamma_t)$$

- Lettau and Ludvigson's (2001) point is that β_t varies along the business cycle with the equity premium γ_t and this variation is enough to explain unconditional alphas
  - So, firms that have higher betas during recessions (when the equity premium is high)–i.e., value stocks–have higher unconditional alphas: the so-called value premium

- Lewellen and Nagel (2006) argue that this covariation is just not large enough

- They compute conditional alphas directly using higher frequency data and show that they are not zero

What were previous papers missing?

- L&N argue that these papers only test the *qualitative* implications of the conditional models. They do not test the constraints imposed by the theory

- In other words, they treat some coefficients as free parameters, which allows them to get more explanatory power in the cross-sectional regressions

- Concretely, take Equation (25)

$$E\left(R^e_{it+1}\right) = \bar{\beta}_i \gamma + Cov\left(\beta_{it}, \gamma_t\right) \tag{27}$$

- The scaled-factor approach in the E(R)-β representation boils down to expressing the time-varying beta as a linear function of some conditioning variable

$$\beta_{it} = \beta_i + \delta_i z_t \tag{28}$$

where δ_i and β_i are estimated in time-series regressions of $R_{i,t+1}$ on $R_{m,t+1}$ and $z_t \times R_{m,t+1}$

- Replace (28) into (27)

$$E\left(R^e_{it+1}\right) = \bar{\beta}_i \gamma + \delta_i Cov\left(z_t, \gamma_t\right) \tag{29}$$

- Lettau and Ludvigson test the asset pricing model in (29) by a cross-sectional regression of returns on $\bar{\beta}_i$ and δ_i

$$R^e_{it+1} = c_0 \bar{\beta}_i + c_1 \delta_i$$

  treating c_1 as a free parameter, whereas the theoretical restriction is $c_1 = Cov\left(z_t, \gamma_t\right)$

- They achieve a high R-squared and small pricing errors because they use an additional degree of freedom that they should not have used

- Lewellen and Nagel show that, when it is imposed, the restriction is actually rejected in the data and the R-squared drops substantially

5. Statistical Discipline

- From efficient set mathematics: Mean variance efficiency of a factor implies a $E(R) = \beta' E(R_q)$ representation, where β are betas on the factor and $E(R_q)$ is the mean return on the factor q, where q is a portfolio on the mean-variance frontier constructed from the (sample or population) moments of the N assets

Proof. Take a portfolio x of the N assets. You have that

$$\beta_{x,q} = \frac{Cov(R_x, R_q)}{Var(R_q)} = \frac{Cov(x'R, q'R)}{Var(q'R)} = \frac{x'Var(R)q}{q'Var(R)q} = \frac{x'Vq}{q'Vq} \tag{30}$$

where x is the vector of weights for the portfolio, q is the vector of weight for a mean-variance efficient portfolio, and V is the variance-covariance matrix of the N assets

To obtain a mean-variance efficient portfolio, it has to be the case that q solves the problem

$$\begin{aligned} & Min\ q'Vq \\ s.t.\ :\ & q'\mu = e \end{aligned}$$

where e is a fixed level of expected excess return and μ is the vector of expected excess returns for the N assets. The Lagrangean for the problem is

$$\mathcal{L} = q'Vq + 2\lambda\left(e - q'\mu\right)$$

and the first order conditions are

$$\begin{aligned} Vq - \lambda\mu &= 0 \\ q'\mu &= e \end{aligned}$$

from the first order conditions we obtain

$$Vq = \lambda\mu$$

which we can use in equation (30):

$$\begin{aligned} \beta_{x,q} &= \frac{x'Vq}{q'Vq} \\ &= \frac{\lambda x'\mu}{\lambda q'\mu} \end{aligned}$$

Rearranging and simplifying λ:

$$x'\mu = \beta_{x,q} q'\mu$$

or:

$$E\left(R_x\right) = \beta_{x,q} E\left(R_q\right)$$

■

- This theorem holds whatever distribution one uses to compute moments: objective, subjective, or ex-post (sample) distribution
- Hence: in the sample, there exists an $E(R) = \beta' E\left(R_q\right)$ representation based on sample moments
- That is: there is an ex-post mean-variance efficient portfolio that prices all assets
- Danger of ”*Fishing for Factors*”! Finding the portfolio that in-sample works
- Limit to possibility of out-of-sample tests: international data dirty; need to wait for 30 years
- Discipline: a factor needs to be economically motivated
- Cochrane's view of AP

- The Stochastic Discount Factor derives from the first order conditions for intertemporal consumption

$$E_t \left(m_{t+1} R^e_{i,t+1} \right) = 0$$

where

$$m_{t+1} = \frac{u'\left(c_{t+1}\right)}{u'\left(c_t\right)} = 1 - b' f_{t+1}$$

- Different AP models can be derived from different specifications of SDF

- In any case, sensible factors must be somehow related to consumption growth. E.g.: return on wealth portfolio (CAPM), state variables of hedging concern (ICAPM)

- It does not mean, however, that covariance of factors with consumption growth implies fully rational investors (see Kozak, Nagel, and Santosh, 2018)

Data Snooping

- Related problem
- If you search the data for significant relationships over and over again, you will find something (with 5% chance)
- In principle, you should test model on different data set
- Objectively difficult
- Need to go to the data with economically founded models