

## Lecture 8

### CAPM

### CAPM as a Regression

- The CAPM puts structure –i.e., how investors form efficient portfolios– to Markowitz’s (1952) mean-variance optimization theory.
- The CAPM assumes only one source of systematic risk: Market Risk.
- Systematic risk:
  - (1) Cannot be diversified
  - (2) Has to be hedged
  - (3) In equilibrium it is compensated by a risk premium
- The stock market exposes investors to a certain degree to market risk.  
=> Investors will be compensated.
- The compensation will be proportional to your risk exposure.

- The data generating CAPM is:

$$R_{i,t} - r_f = \alpha_i + \beta_i (R_{m,t} - r_f) + \varepsilon_{i,t} \quad i=1,\dots,N \text{ and } t=1,\dots,T$$

$R_{i,t}$  = return on asset  $i$  at time  $t$ .

$r_f$  = return of riskless asset at time  $t$ .

$R_{m,t}$  = return on the market portfolio at time  $t$ .

$\alpha_i$  and  $\beta_i$  are the coefficients to be estimated.

$$\text{Cov}(R_{m,t}, \varepsilon_{i,t}) = 0$$

The DGP model is also called the Security Characteristic Line (SCL).

- If  $\alpha_i = 0$ , then

$$E[R_{i,t} - r_f] = \beta_i E[(R_{m,t} - r_f)] \quad (\text{This is the Sharpe-Litner CAPM.})$$

$E[(R_{m,t} - r_f)]$  is called the market risk premium: the difference between the return on the market portfolio and the return on a riskless bond.

The expected return on asset  $i$  over  $r_f$  is proportional to the market risk premium.  $\beta_i$  is the proportionality factor (sensitivity to market risk).

If  $\beta_i = 0$ , asset  $i$  is not exposed to market risk. Thus, the investor is not compensated with higher return. Zero- $\beta$  asset, market neutral.

If  $\beta_i > 0$ , asset  $i$  is exposed to market risk and  $R_{i,t} \geq r_f$ , provided that  $E[R_{m,t} - r_f] > 0$ .

Q: What is the Market Portfolio? It represents all wealth. We need to include not only all stocks, but all bonds, real estate, privately held capital, publicly held capital (roads, universities, etc.), and human capital in the world. (Easy to state, but complicated to form.)

Q: How do we calculate  $E[R_{m,t}]$  and  $r_f$ ?

- The CAPM can be represented as a relation between  $E[R]$  and  $\beta$ :

$$E[R_i] = r_f + \beta_i \lambda \quad (\text{Security Market Line=SML})$$

- The CAPM is very simple: Only one source of risk –market risk– affects expected returns.

- Advantage:

- (1) Simplicity.
- (2) It provides a good benchmark (performance evaluation, etc.)
- (3) It distinguishes between diversifiable and non-diversifiable risk.

- Disadvantages:

- (1) It is likely that other sources of risk exist.  
(Omitted variable bias in the estimates of  $\beta_i$ .)
- (2) The market portfolio is unobservable. Thus,  $R_{m,t}$  will be proxied by an observed market portfolio (say, EW-CRSP).  
(Measurement error is introduced. Roll's (1977) critique.)

## Variance Decomposition of Returns

- Using the CAPM, we can calculate the Variance of  $R_{i,t}$

$$R_{i,t} - r_f = \alpha_i + \beta_i (R_{m,t} - r_f) + \varepsilon_{i,t}$$

$$\text{Var}(R_{i,t}) = \beta_i^2 \text{Var}(R_{m,t}) + \text{Var}(\varepsilon_{i,t})$$

- We have decomposed the total risk into two components:
  - Systematic risk –related to  $R_m$
  - Unsystematic risk –related to  $\varepsilon_{i,t}$
- We can calculate the covariance between any two assets:
 
$$\text{Cov}(R_{i,t}, R_{j,t}) = \beta_i \beta_j \text{Var}(R_{m,t}) \text{ (assume } \varepsilon_{i,t} \text{ and } \varepsilon_{j,t} \text{ uncorrelated)}$$

## Diversification

- The total variance of  $R_{i,t}$  is given by:  

$$\text{Var}(R_{i,t}) = \beta_i^2 \text{Var}(R_{m,t}) + \text{Var}(\varepsilon_{i,t}) \quad (\text{assume } \text{Var}(\varepsilon_{i,t}) = \sigma^2 \text{ for all } i.)$$
- Suppose we have  $N$  assets. Let's form an equally weighted portfolio. The return of this portfolio would be:

$$R_{p,t} = (1/N) R_{1,t} + (1/N) R_{2,t} + \dots + (1/N) R_{N,t} = (1/N) \sum_i R_{i,t}$$

Using the CAPM for each asset.

$$\begin{aligned} R_{p,t} &= (1/N) \sum_i (r_f + \beta_i (R_{m,t} - r_f) + \varepsilon_{i,t}) \\ &= r_f + (R_{m,t} - r_f) (1/N) \sum_i \beta_i + (1/N) \sum_i \varepsilon_{i,t} \end{aligned}$$

- Now, the variance of the portfolio is:

$$\text{Var}(R_{p,t}) = \beta^2 \text{Var}(R_{m,t} - r_f) + (1/N) \sigma^2 \quad (\beta = (1/N) \sum_i \beta_i)$$

The portfolio variance is decomposed in two parts as any other asset. The first part, due to market risk, is the same. The second part, due to unsystematic risk is smaller.

As  $N \rightarrow \infty$ , the variance due to unsystematic risk will disappear. That's diversification!

Q: How many stocks ( $N=?$ ) are needed in a portfolio to get a portfolio only exposed to market risk –i.e., to be diversified? (Campbell et al. (2000) say that  $\sigma^2$  has increased.)

## Extending CAPM

- No risk-free asset : Zero-beta CAPM. Black (1972)
- Non-traded assets – human capital: Mayers( 1972)
- Intertemporal CAPM -factors: Merton (1973)
- Dividends, taxes: Brennan (1970), Litzenberger and Ramaswamy (1979)
- Foreign exchange risk: Solnik (1974)
- Inflation: Long (1974), Friend, Landskroner and Losq (1976).
- International CAPM, PPP risk: Sercu (1980), Stulz (1981), Adler and Dumas (1983)
- Investment restrictions: Stulz (1983).

## Testing the CAPM

- In equilibrium, the CAPM predicts that all investors hold portfolios that are efficient in the expected return-standard deviation space. Therefore, the Market Portfolio is efficient. To test the CAPM, we must test the prediction that the Market Portfolio is positioned on the efficient set.
- The early tests of the CAPM did not test directly the prediction: “The Market Portfolio is efficient.” Instead, papers investigated a linear positive relationship exists between portfolio return and beta, the SML. Then, they concluded that the Market Portfolio must be efficient.

- Standard assumptions for testing CAPM:
  - Rational expectations for  $R_{i,t}$ ,  $r_f$ ,  $R_{m,t}$ ,  $Z_{i,t}$  (any other variable)  
*Ex-ante*  $\approx$  *ex-post* (-i.e., “realized” proxy for “expected”).  
 For example:  $R_{i,t} = E[R_{i,t}] + v_{i,t}$  where  $v_{i,t}$  is white noise.
  - Constant beta –at least, through estimation period.
  - Holding period is known, usually one month.
- Elton (1999) believes the R.E. standard assumption is incorrect:
  - Periods longer than 10 years where average stock market returns had  $R_m < r_f$  (1973 to 1984).
  - Periods longer than 50 years where risky long-term bonds on average underperformed  $r_f$  (1927 to 1981).

There are information events (surprises) that are highly correlated. These events are large. They have a significant long effect on the realized mean

Thus,  $R_{i,t}$  is a poor measure of  $E[R_{i,t}]$ . Misspecified model (through  $v_{i,t}$ ):  
 $v_{i,t} = I_{i,t} + \varepsilon_{i,t}$  ( $I_{i,t}$ : significant information event, a jump process?)

- Early tests focused on the cross-section of stock returns. Test SML.
    - If  $\beta_i$  is known, a cross-sectional regression with  $E[R_i]$  and  $\beta_i$  can be used to test the CAPM:  

$$E[R_i] = \gamma + \beta_i \lambda + \zeta_i \quad (\text{SML})$$
    - Test  $H_0: \lambda > 0$ . (The value of  $\gamma$  is also of importance. Why?)
- Problem: We do not know  $\beta_i$ . It has to be estimated. This will introduce measurement error: bias!
- Examples of the early tests: Black, Jensen and Scholes (1972), Fama and MacBeth (1973), Blume and Friend (1973).

Findings: Positive for CAPM, though the estimated risk-free rate tended to be high.

- More modern tests focused on the time-series behavior. Test the SCL.
  - Two popular approaches:
    - Test  $H_0: \alpha_i=0$  ( $\alpha_i$  is the pricing error. Jensen's alpha.)  
(Joint tests are more efficient  $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_N=0$  (for all  $i$ ))
    - Add more explanatory variables  $\mathbf{Z}_{i,t}$  to the CAPM regression:
 
$$R_{i,t} - r_f = \alpha_i + \beta_i (R_{m,t} - r_f) + \boldsymbol{\delta} \mathbf{Z}_{i,t} + \varepsilon_{i,t}$$
 Test:  $H_0: \boldsymbol{\delta}=\mathbf{0}$ . (We are testing CAPM's specification.)

Findings: Negative for CAPM.  $\boldsymbol{\delta}$  is significant.

## Early Tests: Two pass technique

- Two pass technique:
  - First pass: time series estimation where security (or portfolio) returns were regressed against a market index,  $m$ :
 
$$R_{i,t} - r_f = \alpha_i + \beta_i (R_{m,t} - r_f) + \varepsilon_{i,t} \quad (\text{CAPM})$$
  - Second pass: cross-sectional estimation where the estimated CAPM-beta from the first pass is related to average return:
 
$$R_i = \gamma (1-\beta_i) + \lambda \beta_i + \zeta_{i,t} \quad (\text{SML for security } i)$$
 ( $\gamma$  equals  $r_f$  in the CAPM and  $E[R_{0m}]$  in the Black CAPM. While  $\lambda$  is the expected market return.
- Main problem: Measurement error in  $\beta_i$ . Solution: Measure  $\beta$ 's based on the notion that portfolio  $\beta_p$  estimates will be less affected by measurement error than individual  $\beta_i$  estimates due to aggregation.

- *Example: Black-Jensen-Scholes (1972)*

Data: 1926-1965 NYSE stocks

$R_m$  = Returns on the NYSE Index

- Start with 1926-1930 (60 months). Do pass 1 for each stock. Get  $\beta$ .
- Rank securities by  $\beta$  and form into portfolios 1-10.
- Calculate monthly returns for each of the 12 months of 1931 for the 10 portfolios.
- Recalculate betas using 1927-1931 period, and so on. (Rolling regression.)  
=> We have 12 monthly returns for 35 years = 420 monthly returns (for each portfolio).
- For the entire sample, calculate mean portfolio returns,  $\text{mean}(r_p)$ , and estimate the beta coefficient for each of the 10 portfolios. Get  $\beta_p$ .
- Do pass 2 for the portfolios (Regress  $\text{mean}(r_p)$  against  $\beta_p$ . -estimate the ex-post SML.)

Findings: Positive and significant slope for the SML. The estimated risk-free rate tended to be high.

- *Example: Fama-MacBeth (1973)*

Data: 1926-1968 NYSE stocks

$R_m$  = Returns on the NYSE Index

- Start with 1926-1929 (48 months). Do pass 1 for each stock. Get  $\beta$ .
- Rank securities by  $\beta$  and form into portfolios 1-20.
- Calculate monthly returns for each from 1930-1934 (60 months) for the 20 portfolios. Do pass 1 for portfolios. Get  $\beta_p$ .
- Do pass 2 for each month in the 1935-1938 period (48 SML).
- For each of the months in the 1935-1938 period, also estimate:  

$$R_p - r_f = a_0 + a_1 \beta_p + a_2 (\beta_p)^2 + \zeta_{i,t} \quad (\text{non-linearity?})$$

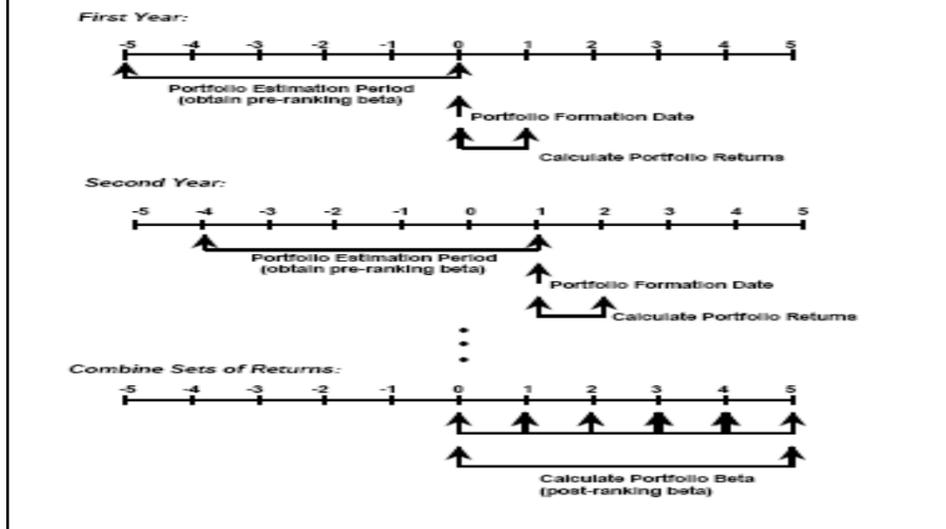
$$R_p - r_f = a_0 + a_1 \beta_p + a_2 (\beta_p)^2 + a_3 RV_p + \zeta_{i,t} \quad (\text{idiosyncratic risk?})$$
 (where  $RV_p$  = average residual variance of the portfolio.)

- Repeat above steps many times  
=> We get 390 estimates of above parameters ( $a_0, a_1, a_2, a_3$ ).
- Do t-tests for ( $a_0, a_1, a_2, a_3$ ).

Findings: Positive for CAPM. But,  $a_0$  was significantly greater than the mean of the risk-free interest rate. No evidence for misspecification.

- Difference between BJS and FM:

- In BJS,  $\beta$  and average returns were computed in the same  $t$  period.
- In FM,  $\beta$  in  $t$  were used to predict returns in  $t+1$ .



- FM has become a staple in applied finance.

- Very simple. No need to estimate SE in pass 2.
- It can be easily adapted to introduce additional risk measures – P/E, Size, B/M, Leverage, etc.
- If coefficients are constant over time, it is equivalent to a FE panel estimation.

- General Issues:

(1) Portfolios: each beta is estimated with error. If the estimation errors are uncorrelated across stocks, a portfolio reduces estimation error and improves second pass regression. The estimators are biased, but consistent.

(2) Sorting by Beta: Random portfolios have a beta close to 1. The sorting preserves some cross-sectional variation for the second pass.

(3) Rolling Regression: To reduce the bias in estimation error, estimate a lot of betas!

- The two pass procedure has some fixable problems:
    - Standard errors can be adjusted (Shanken (1992)).
    - Simultaneous estimation (pass 1 and 2) (Cochrane (2001)).
    - Serial correlation in residuals: Use Newey-West SE.
  
  - Roll (1977): Only testable CAPM implication: MVE of  $R_m$ .
    - Any *ex-post* MVE portfolio used as the index will exactly satisfy the SML by the sample  $\beta$  time series mean returns. (Just math.)
    - Unobservable Market Portfolio. Proxy problem.
  
  - Counter points to Roll (1977)
    - Stambaugh (1982): Added bonds and real estate to stock index.
    - Shanken (1987): If correlations between observable stock index and true global index exceeds .7, CAPM is rejected.
- Roll and Ross (1994): Even when proxy is not far from frontier, CAPM can be rejected.

## CAPM “Anomalies”

- Monday dummy (-): French (1980)
- January dummy (+): Roll (1983), Reinganum (1983)
- E/P (+): Basu (1977), Ball (1978), Jaffe, Keim and Westerfield (1989)
- Firm Size (-): Banz (1981), Basu (1983)
- Long-Term Reversals (-): DeBondt and Thaler (1985)
- B/M (+): Stattman (1980), Rosenberg, Reid and Lanstein (1985)
- D/E (Leverage) (+): Bhandari (1988)
- Momentum (+): Jegadeesh (1990), Jegadeesh and Titman (1993)
  
- *Beta is Dead?*: Fama and French (1992). When the negative correlation between firm size and beta is accounted for, the relation between beta and returns disappears.

- Schwert (2002): Anomalies get weaker after publication of paper.
- Explanations for the anomalies:
  - *Technical explanations*: Roll's critique, data snooping, out-of-sample performance (Schwert (2002)), measurement error (anomalous variables correlated with  $\beta$ 's), survivor bias, sensitivity to data frequency (no anomalies for annual data), Berk's (1995) critique.
    - => There are no real anomalies.
  - *Multiple risk factors*: The CAPM suffers from omitted variables. For example, small firms bear a higher risk of distress; they are less likely to survive bad economic conditions (Chan and Chen (1991).)
    - => Anomalous variables proxy additional risk factors.
  - *Irrational investor behavior*: Investors overreact to news, etc.
    - => Behavioral finance.

## Time Series Tests

- The CAPM null hypothesis is  $H_0: \alpha=0$ .
- We have two types of tests:
  - Assuming a distribution: MLE
    - MLE advantage: consistent, BAN estimators.
    - In small samples, estimates can be very bad and inefficient.
  - No distribution assumption: GMM
    - GMM: consistent, robust to heteroscedasticity and distributional assumptions
    - It relies on asymptotic inferences. In small samples, the behavior of the estimators is a question mark

• **MLE (QMLE) Approach** (Multivariate Normality)

- $\alpha_i$  and  $\beta_i$  are firm specific parameters. Joint estimation provides no benefits.
- SUR estimation to adjust for cross-correlation. Efficiency gain only.
- We need to specify  $\Sigma$  -or to estimate it with White (1980) or Newey-West (1987).
- The distribution of  $\alpha$  and  $\beta$  is given by ( $Z_m$ : market index.)

$$\hat{\alpha} \sim N\left(\alpha, \frac{1}{T} \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right] \Sigma\right)$$

$$\hat{\beta} \sim N\left(\beta, \frac{1}{T} \left[1 + \frac{1}{\hat{\sigma}_m^2}\right] \Sigma\right)$$

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

where

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

• We can use a *Wald Test* (using the unconstrained MLE estimation), a *LM Test* (using the constrained estimation), and a *LR Test* (using both estimations) to test  $H_0: \alpha=0$ .

• Gibbons, Ross, Shanken (1989): Use a *Wald Test*,  $J_1$ . Since they assume normality, they have a small sample distribution:  $J_1 \sim F_{N, T-N-1}$

$$J_1 = \left(\frac{T-N-1}{N}\right) \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}.$$

It turns out that  $\hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} = (\mu_q / \sigma_q)^2 - (\mu_p / \sigma_p)^2$  –squared SR difference between the tangency portfolio and the market portfolio.

• Jobson and Korkie (1982): Use a *LR Test*,  $J_3$ .

$$J_3 = [(T-N/2-2)/N] [\ln|\hat{\Sigma}^*| - \ln|\hat{\Sigma}|] \sim \chi^2_N$$

$[(T-N/2-2)/N]$ : a small sample correction.

Note: It turns out,  $J_1$  is a monotonic transformation of  $J_3$ .  $J_1$  is an LR test!

Findings: Overall, negative for CAPM. CLM's Table 5.3 reports a  $J_1$  with a .020 p-value for the whole 1965-1994 sample. Stronger rejections for subsamples. That is, pricing errors,  $\alpha_i$ , are jointly different from zero.

Tests have size and power problems.

#### • **Size and Power of Multivariate ML Tests**

##### • **Size** (Type I error)

What happens when we use large-sample theory when sample size is not large.

CLM report size problems:

- Problem is bad for small samples, especially for the Wald test.
- Fixing T, as N grows large, the problem becomes worse.
- The Jobson and Korkie (1982) adjustment works well.

Note: Finite sample adjustments can be very important.

##### • **Power** (The probability of rejecting $H_0$ when $H_1$ is true.)

To study power issues, we need an alternative DGP (CLM assume four different Sharpe ratios  $q$  (tangent portfolio) and the size of the test (5%).

CLM find substantial in power:

- For a fix N, power improves with T
- For a fix T, power decreases with N.
- Advice: Keep N small, around 10.

### • GMM Approach

Excess return:  $(R_{i,t} - r_f) - \alpha_i + \beta_i (R_{m,t} - r_f) = \varepsilon_{i,t}$

Instrument:  $[1 (R_{m,t} - r_f)] = \mathbf{X}_t'$

Moment condition:  $E[\mathbf{X}_t \otimes \varepsilon_t] = 0$ . ( $\otimes$  = Kroenecker product)

Sample moment:  $\mathbf{g}_T = (1/T) \sum_t (\mathbf{X}_t \otimes \varepsilon_t)$

From this simple GMM setup, GMM is equivalent to MLE (OLS under normality. Advantage: a robust covariance matrix can be estimated.

$$\mathbf{V} = \mathbf{D}_0' \mathbf{S}_0^{-1} \mathbf{D}_0$$

where  $\mathbf{D}_0 = E[\delta \mathbf{g}_T / \delta \boldsymbol{\theta}]$  ( $\boldsymbol{\theta} = (\alpha, \beta)$ )

$$\mathbf{S}_0 = \sum_t E[(\mathbf{X}_t \otimes \varepsilon_t) (\mathbf{X}_t \otimes \varepsilon_t)']$$

Standard J test:

$$J_T = T \mathbf{a}_{\text{GMM}}' [\mathbf{R}(\mathbf{D}_T' \mathbf{S}_T^{-1} \mathbf{D}_T)^{-1} \mathbf{R}]^{-1} \mathbf{a}_{\text{GMM}} \sim \chi^2_N$$

where  $\mathbf{R} = (1 \ 0) \otimes \mathbf{I}_N$  and  $\mathbf{a}_{\text{GMM}}$  is the GMM estimator of  $\alpha$ .

MacKinley and Richardson (1991) illustrate the bias in standard CAPM tests when non-normality (Student-t) is present.

## Multifactor Models (APT)

- Let's generalize the one-factor model to a k-factor model. Assume:

$$R_{i,t} - r_f = \alpha_i + \beta_{1,i} f_{1,t} + \beta_{2,i} f_{2,t} + \dots + \beta_{k,i} f_{k,t} + \varepsilon_{i,t} \quad (\text{APT-1})$$

-  $f_{i,t}$  is the underlying zero-mean risk factor  $i$ .

- Slopes ( $\beta_k$ 's): "factor loadings" or "sensitivity to risk factors."

- All factors are uncorrelated with  $\varepsilon_{i,t}$ .

- In an ideal (APT) world:  $\text{Cov}(f_{1,t}, f_{2,t}) = 0$ .

- When no arbitrage opportunities exist:

$$E[R_i - r_f] = \sum_k \beta_{i,k} \lambda_k + \zeta_i \quad i=1, \dots, N \quad (\text{APT-2})$$

where  $\lambda_k$  is the risk premium associated with the factor  $k$  and  $\zeta_i$  satisfies:

$$\lim_{N \rightarrow \infty} \sum_i \zeta_i^2 < \infty \quad (\text{Arbitrage pricing condition.}) (\text{APT-3})$$

- APT-1 to APT-3 describe the Arbitrage Pricing Theory (APT).

APT-1 represents the DGP for returns.

APT-2 is a cross-sectional relation for expected returns and risk premia.

- If the factors are not necessarily mutually uncorrelated and  $f_{1,t} = R_{m,t} - r_f$ , the model is called ICAPM (Merton's (1973) Intertemporal CAPM).

- We will make not distinction between ICAPM and APT. We will also assume exact factor pricing. That is:

$$E[R_i - r_f] = \sum_k \beta_{i,k} \lambda_k \quad i=1, \dots, N$$

**Classic References:** Ross (1986), Chamberlain and Rothschild (1983), Ingersoll (1984), Khan and Sun (2001).

Q: What are these other factors? APT provides no guidance regarding the factors. There are a few ways of identifying the factors:

– *Economic arguments* (observable factors)

- Chen, Roll and Ross (1986), Elton, Gruber and Mei (1994) use macro economic variables
- Fama and French (1993) use firm characteristics + market index

– *Statistical arguments* (unobservable factors)

- Lehmann and Modest (1988), Elton, Gruber, and Michaely (1990) use factor analysis for stocks and bond, respectively.
- Connor and Korjczyk (1988), Jones (2001) use Principal Components.

(Difficult economic interpretation of factors.)

Q: Which approach is better: FA or PC? Open question.

– *Consulting firms* (BARRA, DFA, etc.).

### Data Requirements for CAPM vs APT Regressions

- CAPM regression:  $R_i$ ,  $R_m$ , and  $r_f$
- APT usual regression:  $R_i$ ,  $r_f$ ,  $f_1$  (usually,  $R_m$ ), ..., and  $f_K$

- Q: How many factors should we include ( $K=?$ )?

Again, APT provides no guidance regarding  $K$ .

Note: The APT's generality can be seen as a weakness.

## Testing APT

- Tests of APT: Based on the idea that the pricing error represents diversifiable risk and, in the absence of arbitrage, the pricing error is strongly bounded as the number of assets increases.
- An Ideal test for APT:  
A test of APT should follow a three-stage process:
  - (i) The returns equations are estimated  $\Rightarrow$  get factor loadings  $\beta_{k,i}$ 's.
  - (ii) Conditional on the  $\beta_{k,i}$ 's,  $E[R_i - r_f]$  are estimated to get  $\zeta_i$ 's.
  - (iii) Conditional on the  $\zeta_i$ 's, the pricing condition is tested.  
 $\Rightarrow$  A test of arbitrage pricing is a test of the behavior of the pricing errors as the number of assets increases without bound.  
(Complicated thing to do.)

- Usually, the tests of APT (ICAPM) are based on assuming exact factor pricing, a k-factor model and testing if the alpha is zero:  

$$H_0: \alpha_i = 0$$
 (pricing errors are zero -actually, if LLN applies).
  - We fit –usually with the two pass procedure- the APT model:  

$$R_{i,t} - r_f = \alpha_i + \beta_{1,i} f_{1,t} + \beta_{2,i} f_{2,t} + \dots + \beta_{k,i} f_{k,t} + \varepsilon_{i,t}$$
 (Applied Rule: Always run OLS with intercept).
  - The  $\beta_{k,i}$ 's are also of interest (what factors are priced)
- Q: Can we price fixed income securities, real estate, or derivatives with the CAPM/APT approach?

Findings: Negative for exact factor pricing. Pricing errors,  $\alpha_i$ , are different from zero.  $\beta$  is significant (factors are priced!). K no larger than 5.

- *Example: Chen, Roll and Ross' (1986)*

Factors:

- growth in industrial production (MP).
- change in expected inflation (DEI).
- unexpected inflation (UI).
- unexpected changes in risk premiums (UPR).
- unexpected changes in term structure slope (UTS).
- Also include EV and VW market indexes in the regressions.

Methodology: Fama-MacBeth's two pass estimation of the slopes.

Findings: Statistical significant slopes for MP, UI, and UPR (this factors are priced!). Market indexes not significant.

MP	DEI	UI	UPR	UTS	EWI	VWI	Const
13.59	-0.13	-0.63	7.21	-5.21			4.12
(3.56)	(-1.64)	(-1.98)	(2.59)	(-1.69)			(1.36)
14.01	-0.13	-0.85	8.13	-5.02	5.02		6.41
(3.77)	(-1.67)	(-2.54)	(2.86)	(-1.58)	(1.22)		(1.85)
11.76	-0.12	-0.80	8.27	-5.91		-2.40	10.71
(3.05)	(-1.60)	(-2.38)	(2.97)	(-1.88)		(-0.63)	(2.76)

• *Example: Fama and French (1993)*

(1) Factors: The independent variables

Mimicking ‘portfolios’ for B/M and Size + Market Index

- *B/M (HML) factor* => risk factor related to “distressed stocks”
  - defined as return on High B/M stocks minus Low B/M (i.e., a portfolio long high B/M stocks and short low B/M stocks.)
- *Size (SMB) factor* => risk factor related to “size”
  - defined as return on High Size stocks minus Low Size (i.e., a portfolio long big firm stocks and short small firm stocks.)

In June of each year  $t$ , break stocks into:

Two size groups: Small / Big (below/above median)

Three B/M groups: Low (bottom 30%) / Medium / High (top 30%)

Compute monthly VW returns of 6 size-B/M portfolios for the next 12 months

Factor-mimicking portfolios: zero-investment

Size: **SMB** =  $1/3(SL+SM+SH) - 1/3(BL+BM+BH)$

B/M: **HML** =  $1/2(BH+SH) - 1/2(BL+SL)$

(2) Portfolio construction (the dependent variable).

- Construct 25 stock portfolios
  - In June of each year  $t$ , stocks are sorted by size (current Market Equity) and (independently) by B/M (as of December of  $t-1$ )
  - Using NYSE quintile breakpoints, all stocks are allocated to one of 5 size portfolios and one of 5 B/M portfolios
  - From July of  $t$  to June of  $t+1$ , monthly VW returns of 25 size-B/M portfolios are computed
- Construct 7 bond portfolios
  - 2 Government portfolios: 1-5 years, 6-10 years maturity
  - 5 corporate bond portfolios: Aaa, Aa, A, Baa, below Baa

Estimation and testing methodology:

(3) Time series Regression

$$R_{i,t} = \alpha_i + \beta_{1i}R_{m,t} + \beta_{2,i}SMB_t + \beta_{3i}HML_t + \varepsilon_{i,t}$$

=> get  $b_{k,i}$  ( $= \beta_{ki}$  estimate),  $SE(b_{k,i})$ , and  $R^2$ . Also, intercept.

(3') Fama-MacBeth methodology.

Pass 1: Time series

$$R_{i,t} = \alpha_i + \beta_{1i}R_{m,t} + \beta_{2,i}SMB_t + \beta_{3i}HML_t + \varepsilon_{i,t} \Rightarrow \text{get } b_{k,i}$$

Pass 2: Cross sectional

$$\text{Mean}(R_i) = b_{0i} + \gamma_m b_{1i} + \gamma_{SMB} b_{2i} + \gamma_{HML} b_{3i} + \zeta_{i,t}$$

Findings: The three-factor regressions produces significant coefficients on all three factors: positive coefficients for Market, SMB and HMB.

That is, firms in distress (high B/M) have had high average returns. The  $R^2$  values are close to 1 for most portfolios. Overall rejection of CAPM.

Reject  $\alpha_i=0$ .

- More factors could be added: Carhart (1997) adds a momentum factor.
- Two additional popular factors: credit spreads, interest rates.

• Q: What is wrong with the FF approach?

- *Roll's critique is still valid*: The three factors are proxies. Moreover, what factors are SMB and HML proxies for?

- *No theory*: Loosely based on APT, but APT provides no specific factors.

- *Data mining bias*: Big problem. Size and B/M have been around for a while (remember the CAPM anomalies): FF factors are based on a lot of previous research. (Type I error could be big.)

- *Measurement error (Size IV?)*: Size and beta are highly correlated. Since size is measured precisely, and beta is estimated with large measurement error, size may well be an IV for beta.

## Things to come: The equity premium

- The equity premium,  $E[R_{m,t} - r_f]$ , is defined as the return on a broad stock index over a the return on a safe bond market investment.
- To estimate the equity premium we need a long historical sample because stock returns are noisy.
  - With 100 years of data and 15% standard deviation of returns per year, the standard error of the estimate is 1.5%
- But over a long period, it is plausible that the equity premium changes. This makes averages not representative.
- U.S. Estimates:
  - 1802 - 1998 (Siegel): 4.10%
  - 1871 - 1999 (Shiller): 5.75%
  - 1889 - 2000 (Mehra-Prescott): 6.92%

- These estimate of the U.S. equity premium,  $E[R_{m,t} - r_f]$ , are very high.
- “Very high”: No reasonable economic model can justify such a big premium for holding stocks versus a risk-free asset.
- This difference, called the “Equity Premium Puzzle,” has generated thousands of papers during the past 25 years. (Mehra and Prescott (1985) started this literature. There is even a Handbook of the Risk Premium (2007), edited by Mehra!)
- There are two possible (but not mutually exclusive) ways of rationalizing this puzzle:
  - Economic models are not realistic.
  - The equity premium is a statistical fluke.