Pricing Factor Model

Dr. Jinghan Meng

FINA3351 – Spreadsheet Modelling in Finance

Roadmap for Today

1. Data Sources
2. Measuring Returns
3. Estimating CAPM
4. Estimating Multifactor Models (FF3 Factor Model)
5. Testing SML of CAPM – Fama-Macbeth regression
6. Testing SML of Multifactor Models (FF3 Factor Model)

Stock and Interest Rate Data Sources

Data Source: Stocks

- Stock and index data source: Yahoo!Finance
  - Free stock and index trading data
  - Both US market and some international trading data
  - Data frequency: daily, weekly, monthly
- Historical data is NOT allowed to download for free now.
- To (temporarily) solve the issue:
  - Excel files “GetDataFromYahooFinance_Win.xlsm” and “GetDataFromYahooFinance_Mac.xlsm” provide VBA codes that directly extract stock market data into Excel file.

Data Source: Portfolios and Pricing Factors

- Pricing factors and portfolios in multifactor pricing model: Kenneth R. French's website
  - Free to download
  - US and some international pricing factors. For example,
    - Fama-French 3 factors – market, size, book-to-market (B/M);
    - momentum, long-term reversal, E/P, CF/P, D/P, industry…
  - Data frequency: daily, weekly, monthly
  - Portfolios constructed based on these factors

Risk-free Rate

- We usually use shortest bill/note yield as the risk-free investment return (risk-free rate, or "R_f").
- Hong Kong: Hong Kong Interbank Offered Rate (HIBOR). Overnight HIBOR is preferred.

  You may find the data from government websites:

  - Hong Kong Monetary Authority: (Financial Sector → Interest Rate → Base rate)

- US and global interest rates data source: https://fred.stlouisfed.org/
  - US: 4-week Treasury-bill secondary market rate
  - Global: London Interbank Offered Rate (LIBOR)

    Secured Overnight Financing rate (SOFR) (from 2022)

How to Measure Returns

Example: 7 US Stocks

❖ Open Excel file "Lec5_FactorModel.xlsm".

❖ We enter information on sheet "Inputs" and run sub procedure `GetData` to download data into sheet "Historical_Data".

❖ **Data source**: Yahoo!Finance

❖ **Sample period**: 10-year monthly data from 2014.12 to 2024.12

❖ Run sub procedure `CopyData`. Adjusted close price of 7 stocks are pasted into Sheet "US stocks".

❖ ***Note***: dividend payments and stock splits have been adjusted.

Adjusted Close Price

- Adjusted close price is the closing price after adjustments for all applicable splits and dividend distributions.

(Yahoo!Finance help)

- **Split multipliers** are determined by the split ratio.
  - In a 2 for 1 split, the pre-split data is multiplied by 0.5.
  - In a 4 for 1 split, the pre-split data is multiplied by 0.25.
  - In a 1 for 5 reverse split, the pre-split data is multiplied by 5.

Adjusted Close Price

- **Dividend multipliers** are calculated based on dividend as a percentage of the price, primarily to avoid negative historical pricing.
  - Example:

  Feb 18 closing price is $24.95.

  A $0.08 cash dividend is distributed on Feb 19 (ex- date),

  Then all the pre-dividend data are multiplied by (1-0.08/24.95) = 0.9968.

- See a sample calculation on Yahoo!Finance help.

Continuously Compounded vs. Discrete Return

- Discrete return:

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1$$

- Continuously compounded return (log-return):

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(1 + R_t)$$

$$1 + R_t = e^{r_t}$$

  - $R_t =$ discrete return over time $t - 1$ to t
  - $r_t =$ continuously compounded return over time $t - 1$ to t
  - $S_t =$ (adjusted) price at the end of time t
- Log return is always smaller than discrete return.
- Log-return is very close to the original scale when return is small.

Log Returns

- Log return:

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(1 + R_t)$$

- How do we interpret r_t?

$\Rightarrow$ r_t is the instantaneous return.

Informal proof:

- Suppose we divide period $t-1$ to t into N time intervals. Time increment Δt is infinitely small. Then $\Delta t = \frac{1}{N}$.

- The stock return over Δt is $r_t \Delta t$. Then total return is

$$(1 + r_t \Delta t)^{N} = (1 + r_t \Delta t)^{(1/\Delta t)} \rightarrow e^{r_t}$$

as $\Delta t \rightarrow 0$.

- This is the stock return with continuously compounding.

Log Returns

- Log return:

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(1 + R_t)$$

- We will calculate the monthly returns of 7 US stocks.
- In this analysis, we assume that r_t is constant over month t.
- In the empirical analysis, log return is preferable. Why?

Why Do We Prefer Log Return (1)

- Log-linearization converts growth rates into additive function.
- Log-return from time 0 to T is

$$r_{0\to T} = \ln\left(\frac{S_T}{S_0}\right) = \ln\left(\frac{S_1}{S_0}\times\frac{S_2}{S_1}\times\cdots\frac{S_T}{S_{T-1}}\right) = r_1 + r_2 + \cdots + r_T$$

- Simple average of log-returns

$$\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$$

is geometric average of return:

$$e^{T\bar{r}} = e^{r_1}e^{r_2}\cdots e^{r_T} = (1+R_1)(1+R_2)\cdots(1+R_T)$$

Why Do We Prefer Log Return (2)

- If we assume that stock price S_t follows lognormal distribution, then

$$r_t = \ln(S_t) - \ln(S_{t-1})$$

follows a normal distribution.

- Here, we use the definition of lognormal distribution, and a property of normal distribution – invariant under linear transformation.

- Recall what you learned in FINA3350, assuming that stock price process is geometric Brownian motion,

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

then $S_t = S_0 e^{(\mu - 0.5\sigma^2)t + \sigma B_t}$ is lognormally distributed.

Why Do We Prefer Log Return (3)

- Data frequency transformation makes sense.
- For example, when we annualize expectation ($\bar{r}$) and standard deviation (σ) of monthly returns,

$$\text{Annual expectation} = 12\bar{r},$$

$$\text{Annual variance} = 12\sigma^2$$

- We assume that monthly returns r_t are the same over a year.
- Annualize expected return of log-return *r*:

$$E(r_{0\to 12}) = E(r_1 + r_2 + \cdots + r_{12}) = 12\bar{r}$$

- Annualize expected return of discrete return *R*:

$$\begin{aligned} E(R_{0\to 12}) &= E\left(\frac{S_1}{S_0} \times \frac{S_2}{S_1} \times \cdots \frac{S_{12}}{S_{11}} - 1\right) \\ &= E\big((1+R_1)(1+R_2)\cdots(1+R_{12}) - 1\big) \\ &\neq 12\bar{R} \end{aligned}$$

香港大學 THE UNIVERSITY OF HONG KONG

Why Do We Prefer Log Return (4)

- Annualize variance of log-return *r*:

$$\begin{aligned}\mathrm{Var}(r_{0\to 12}) &= \mathrm{Var}(r_1 + r_2 + \cdots + r_{12}) \\ &= \mathrm{Var}(r_1) + \cdots + \mathrm{Var}(r_{12}) \\ &= 12\sigma^2\end{aligned}$$

- We have assumed that market is weakly efficient and returns in each month are <u>uncorrelated</u>.

- Annualizing standard deviation of discrete return *R*:

$$\begin{aligned}\mathrm{Var}(R_{0\to 12}) &= \mathrm{Var}\big((1+R_1)(1+R_2)\cdots(1+R_{12}) - 1\big) \\ &\neq \mathrm{Var}(R_1) + \cdots + \mathrm{Var}(R_{12})\end{aligned}$$

Why Do We Prefer Log Return (5)

- Log return is used in Black-Scholes model:

$$
\begin{aligned}
c &= S\,N(d_1) - K\,e^{-rT} N(d_2) \\
p &= K\,e^{-rT}\,N(-d_2) - S\,N(-d_1) \\
\text{where}\quad d_1 &= \frac{\ln(S/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \\
d_2 &= d_1 - \sigma\sqrt{T}
\end{aligned}
$$

- *r*, *σ* are expectation and standard deviation of log returns under risk-neutral probability.

Estimating CAPM

CAPM: Overview

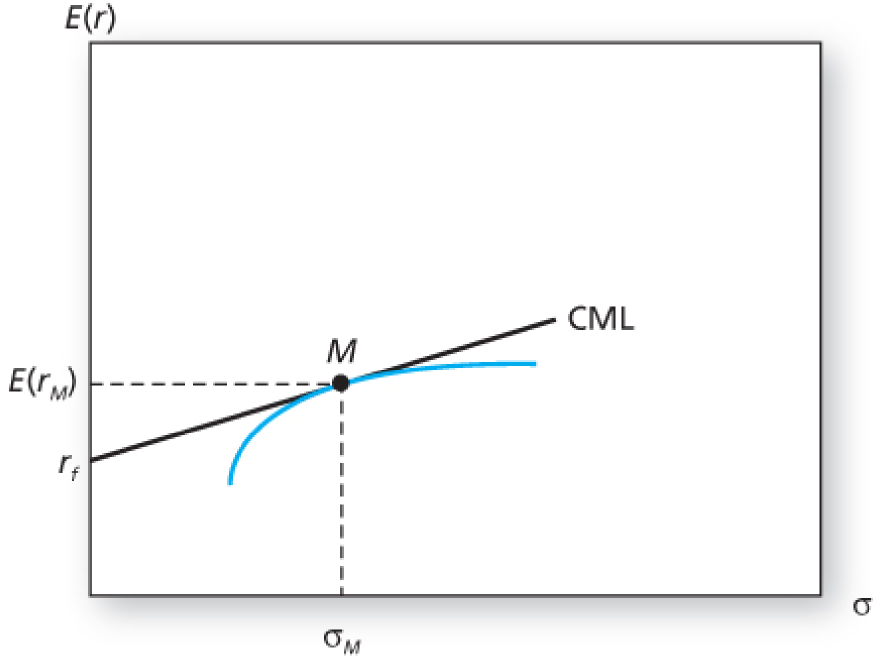
❖ The capital asset pricing model (CAPM) is one of the most influential innovations in financial theory in the latter half of the twentieth century.

❖ What does the CAPM actually say?

1. The capital market line (CML) defines the individual optimal portfolios for an investor interested in the mean and variance of her optimal portfolio.
2. Given agreement between investors on the statistical properties of asset returns and on the importance of mean-variance optimization, the security market line (SML) defines the risk-return relation for each individual asset.

CAPM: CML

- Efficient frontier and capital market line (CML). M is the market portfolio.
- Market portfolio maximizes the Sharpe ratio $\frac{E(r_P)-r_f}{\sigma_P}$ for all feasible portfolios P.



CAPM: Security Market Line (SML)

- Risk premium on an individual asset i:

$$E(R_i^e) = \beta_i E(R_M^e),$$
$$\beta_i = \frac{\text{Cov}(R_i, R_M)}{\sigma_M^2}$$

where $R_i^e = R_i - r_f$ and $R_M^e = R_M - r_f$ denote excess returns.

Estimating SML – Regression

- We can use time series regressions to estimate SML:

$$R_{i,t}^{e} = \alpha_i + \beta_i R_{M,t}^{e} + e_{i,t}$$

  - $R_{i,t}^{e}$ and $R_{M,t}^{e}$ are excess returns of individual asset i and market portfolio.
  - $e_{i,t} \sim N(0, \sigma_e^2)$ is firm-specific (idiosyncratic) shock of the asset at time t, with variance σ_e^2.
- We should determine a candidate for the market portfolio M.
  - Take US market as an example.
  - One popular candidate is S&P 500 index. But it only includes 500 large stocks.
  - We choose market index used in Fama-French three factor model.

Estimating SML – Regression

- We choose market index used in Fama-French three factor model:

Rm-Rf, the excess return on the market, value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t, good shares and price data at the beginning of t, and good return data for t minus the one-month Treasury bill rate. The one-month Treasury bill rate data through May 2024 are from Ibbotson Associates. Starting from December 2024, the one-month Treasury bill rate is from ICE BofA US 1-Month Treasury Bill Index.

CRSP data source: http://www.crsp.org/products/research-products

- HKU students can register WRDS (Wharton Research Data Services) and use CRSP data for free.
- Data for risk-free rate is one-month Treasury bill rate observed **at the beginning of the month**.
- Sheet "FF3 Factor" contains Fama-French three factors and risk-free rate data downloaded from Kenneth French's data library.

Estimating SML: `LINEST` Function

- Use GE stock as an example to estimate SML.
- Sample period: 2014.12 – 2024.12
- Data frequency: monthly
- In sheet "LINEST (CAPM)", excess return is computed as

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) - R_{f,t-1}.$$

  - In empirical analysis, we should use <u>one-period lagged</u> risk-free rate.
  - We use lagged risk-free rate to proxy constant rate in CAPM.
  - Also, risk-free rate is not released as timely as stock price data.
  - Risk-free rates on Kenneth French's website are the beginning-of-month data, which are already one-month lagged.

Estimating SML: `LINEST` Function

- Regression results are generated using three ways:
  1. Data Analysis Toolpak
  2. `LINEST` function
  3. EXCEL functions (for univariate regression only):
     - `INTERCEPT` returns estimate of alpha
     - `SLOPE` returns estimate of beta
     - `RSQ` returns R-squared of the regression.
- We use two Excel functions to search for a specified observation in the data table:
  1. `INDEX` function
  2. `MATCH` function

Estimating SML: *t*-statistics

- In tab “LINEST (CAPM)”, we calculate *t*-statistics of alpha and beta.
- T-statistic for hypothesis test of alpha

$$H_0\ (\text{null}):\ \alpha = \alpha_0, H_1(\text{alternative}):\ \alpha \neq \alpha_0$$

is defined as

$$\text{t}{-}\text{stat} = \frac{\hat{\alpha} - \alpha_0}{SE(\alpha)}$$

- T-statistic for hypothesis test of beta

$$H_0\ (\text{null}):\ \beta = \beta_0, H_1(\text{alternative}):\ \beta \neq \beta_0$$

is defined as

$$\text{t}{-}\text{stat} = \frac{\hat{\beta} - \beta_0}{SE(\beta)}$$

Estimating SML: *P*-value

- We calculate p-value of α and β, based on the t-statistics for hypothesis tests.
- P-value of t-test: https://en.wikipedia.org/wiki/Student%27s_t-test
- Excel function T.DIST.2T returns the p-value of two-tailed Student's t-distribution:

$$p = \Pr(T \geq |\text{t}-\text{stat}| | H_0)$$

where T follows Student's t-distribution with degree of freedom df of the regression:

$$df = \text{No. of Observation} - \text{No. of Regressor} - 1$$

You can also find it in regression results generated by Data Analysis and LINEST function.

- We usually choose a threshold of 5% for p-value. (1% and 10% are also used in practice.)
- If p-value $\leq 5\%$, we reject the null hypothesis H_0. Otherwise, we cannot reject the null hypothesis given the observations.

Estimating SML of Multifactor Models

Fama-French 3 Factor Model

- CAPM does not perform well in the empirical analysis.
- Researchers have documented the empirical evidence such as
  - Small-cap stocks have higher average returns than large-cap stocks. → Size factor
  - Value stocks (high B/M) stocks have higher average returns than growth stocks (low B/M). → B/M factor
- Based on Arbitrage Pricing Theory (APT), we may add more pricing factors into the factor model to better capture the common risk sources.
- APT specifies a pricing relationship with a number of "systematic" factors (=> **multifactor model**).

$$E\left(R_{i,t}^{e}\right) = \beta_{1,i}E\left(F_{1,t}\right) + \beta_{2,i}E\left(F_{2,t}\right) + \cdots + \beta_{k,i}E\left(F_{k,t}\right)$$

  - $F_{k,t}$ is the return at time *t* of *k*-th factor.

FF3 Factors:

- Here is the description of how the three factor returns are computed:

The Fama/French benchmark factors, Rm-Rf, SMB, and HML, are constructed from six size/book-to-market benchmark portfolios that do not include hold ranges and do not incur transaction costs.

Rm-Rf, the excess return on the market, is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate (from Ibbotson Associates).

SMB (Small Minus Big) is the average return on three small portfolios minus the average return on three big portfolios,

$$\begin{aligned} SMB = \ & 1/3\ (Small\ Value + Small\ Neutral + Small\ Growth) \\ & - 1/3\ (Big\ Value + Big\ Neutral + Big\ Growth). \end{aligned}$$

HML (High Minus Low) is the average return on two value portfolios minus the average return on two growth portfolios,

$$\begin{aligned} HML = \ & 1/2\ (Small\ Value + Big\ Value) \\ & - 1/2\ (Small\ Growth + Big\ Growth). \end{aligned}$$

FF3 Factors:

- Here is the description of how the three factor returns are computed:

The portfolios, which are constructed at the end of each June, are the intersections of 2 portfolios formed on size (market equity, ME) and 3 portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint for year t is the median NYSE market equity at the end of June of year t. BE/ME for June of year t is the book equity for the last fiscal year end in t-1 divided by ME for December of t-1. The BE/ME breakpoints are the 30th and 70th NYSE percentiles.

	Median ME	
	Small Value	Big Value
70th BE/ME percentile		
	Small Neutral	Big Neutral
30th BE/ME percentile		
	Small Growth	Big Growth

Estimating SML

- Go to tab "FF3 Reg" for the regression.
- Time series regression for stock i:

$$R_{i,t}^{e} = \alpha_i + \beta_{1,i} R_{M,t}^{e} + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + e_{i,t}$$

$e_{i,t} \sim N(0, \sigma_e^2)$ is firm-specific (idiosyncratic) shock of the asset at time t.

- From 11/2014 to 10/2024,
  - Average monthly market return is 0.97% + 0.14% = 1.11%.
  - Average size factor return is -0.013%, which is the return differential between small and large portfolios.
  - Average value factor return is -0.018%, which is the return differential between high B/M and low B/M portfolios.
- In most recent years, the average return differences in size and book-to-market portfolios have vanished. But when you examine a longer period, e.g., since 1960 till now, the return differences are significantly positive.

Testing CAPM and Multifactor Model

Testing SML: Fama-Macbeth Regression

❖ Security market line (SML) of CAPM

$$E(R_i^e) = \beta_i E(R_M^e)$$

tells us that assets with high beta risk should have high risk premium (or expected return).

❖ What to do: check if high average returns are matched by high betas.

❖ APT specifies a pricing relationship with multiple "systematic" factors (=> **multifactor model**). The SML of *K*-factor APT is

$$E(R_i^e) = \beta_{1,i}E(F_1) + \beta_{2,i}E(F_2) + \cdots + \beta_{K,i}E(F_K)$$

❖ What to do: check if all the well-diversified portfolios satisfy the SML.

❖ *Note*: APT may not hold for individual stocks.

❖ Fama-Macbeth regression was introduced in 1973 and is still widely used in testing pricing factor models.

Testing SML: Fama-Macbeth Regression

- Steps to test SML of CAPM:
  1. Determine a candidate for the market portfolio *M*.
  2. Determine individual assets in the regressions.
  3. First-pass regression: For each of the assets in question, determine the asset beta, $\beta_i = \frac{\mathrm{Cov}(R_i, R_M)}{\sigma_M^2}$.
  4. Second-pass regression: Regress the mean excess returns of the assets on their respective betas

$$\overline{R_i^e} = \gamma_0 + \gamma_1 \beta_i$$

If the CAPM in its descriptive format holds, the second-pass regression should be the SML, that is $\gamma_0 = 0$, and $\gamma_1 = E(R_M^e)$.

Testing SML

- Use US market as an example.
- Excess return of the market portfolio is proxied by MKT-RF is the market excess return from Kenneth French's website.
- Arbitrage pricing theory (APT) should hold for all well-diversified portfolios.
- We use 5 × 5 portfolios formed on size (market value) and book-to-market as an example to test FF3 model.
- Tab "CAPM FF25_Portfolios_5x5" contains monthly return data of 25 portfolios.

http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/25_Portfolios_5x5_CSV.zip

Testing SML – 25 Size and B/M Portfolios

- Here is the construction of portfolios and how the portfolio returns are calculated, given on French's website:

Construction:	The portfolios, which are constructed at the end of each June, are the intersections of 5 portfolios formed on size (market equity, ME) and 5 portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoints for year t are the NYSE market equity quintiles at the end of June of t. BE/ME for June of year t is the book equity for the last fiscal year end in t-1 divided by ME for December of t-1. The BE/ME breakpoints are NYSE quintiles.
Stocks:	The portfolios for July of year t to June of t+1 include all NYSE, AMEX, and NASDAQ stocks for which we have market equity data for December of t-1 and June of t, and (positive) book equity data for t-1.

Testing SML – First-Pass Regression

- Go to tab “CAPM First-pass reg” for first-pass regression in three sample periods:
  - Full sample (1926.7 – 2024.12)
  - Before 1964 (1926.7-1963.12)
  - Since 1964 (1964.1-2024.12)
- First-pass regression: For each of the assets in question, determine the asset beta, $\beta_i = \frac{\text{Cov}(R_i, R_M)}{\sigma_M^2}$.
- We compute the monthly average return, alphas, betas, and R-squared of 25 portfolios.

Testing SML – Second-Pass Regression

- Go to tab "CAPM Second-pass reg" for second-pass regression results.
- Second-pass regression: Regress the mean excess returns of the assets on their respective betas

$$\overline{R_i^e} = \gamma_0 + \gamma_1 \beta_i$$

If the CAPM in its descriptive format holds, the second-pass regression should be the SML, that is $\gamma_0 = 0$, and $\gamma_1 = E(R_M^e)$.

- We regress average excess returns of 25 portfolios on their betas.

Testing SML – Second-Pass Regression

❖ For the full sample period, CAPM is not bad:

Average market excess return	0.0069
Intercept (γ0)	0.0041
Slope (MKT-RF), γ(M)	0.0040
R-squared	0.0974
t-stat (H0: γ0=0)	1.3216
t-stat (H0: γ(M)=market average excess return)	-1.1587
df (regression)	23
P-value, intercept (H0: γ0=0)	0.1993
P-value, slope MKT-RF (H0: γ(M)=market average excess return)	0.2585

❖ Intercept (γ_0) has mean 0.41%, with p-value 0.199. Alpha is not significantly different from 0.

❖ Slope (γ_1) has mean 0.40%, with p-value 0.2585 for hypothesis test of whether slope is equal to average market excess return (with monthly mean 0.69%).

Slope is not significantly different from market risk premium. This is desirable result!

❖ R-squared is 9.74%. Not bad.

Testing SML – Second-Pass Regression

- Before 1964 (1926.7-1963.12), CAPM performs quite well:

Average market excess return	0.0086
Intercept (γ0)	0.0026
Slope (MKT-RF), γ(M)	0.0065
R-squared	0.3010
t-stat (H0: γ0=0)	0.9388
t-stat (H0: γ(M)=market average excess return)	-1.0063
df (regression)	23
P-value, intercept (H0: γ0=0)	0.3576
P-value, slope MKT-RF (H0: γ(M)=market average excess return)	0.3248

- Intercept (γ_0) has mean 0.26%, with p-value 0.358. Alpha is not significantly different from 0.

- Slope (γ_1) has mean 0.65%, with p-value 0.325 for hypothesis test of whether slope is equal to average market excess return (with monthly mean 0.86%).

Slope is not significantly different from market risk premium. This is desirable result!

- R-squared is 30%. CAMP performs better in the early period.

Testing SML – Second-Pass Regression

- Since 1964 (1964.1-present), CAPM performs disappointing:

Average market excess return	0.0058
Intercept (γ0)	0.0108
Slope (MKT-RF), γ(M)	-0.0031
R-squared	0.0526
t-stat (H0: γ0=0)	3.6348
t-stat (H0: γ(M)=market average excess return)	-3.2732
df (regression)	23
P-value, intercept (H0: γ0=0)	0.0014
P-value, slope MKT-RF (H0: γ(M)=market average excess return)	0.0033

- Intercept (γ_0) has mean 1.08% with p-value of 0.0014, significantly positive.

- Slope (γ_1) has mean -0.31% with p-value of 0.0033 for hypothesis test of whether slope is equal to average market excess return (with mean 0.58%).

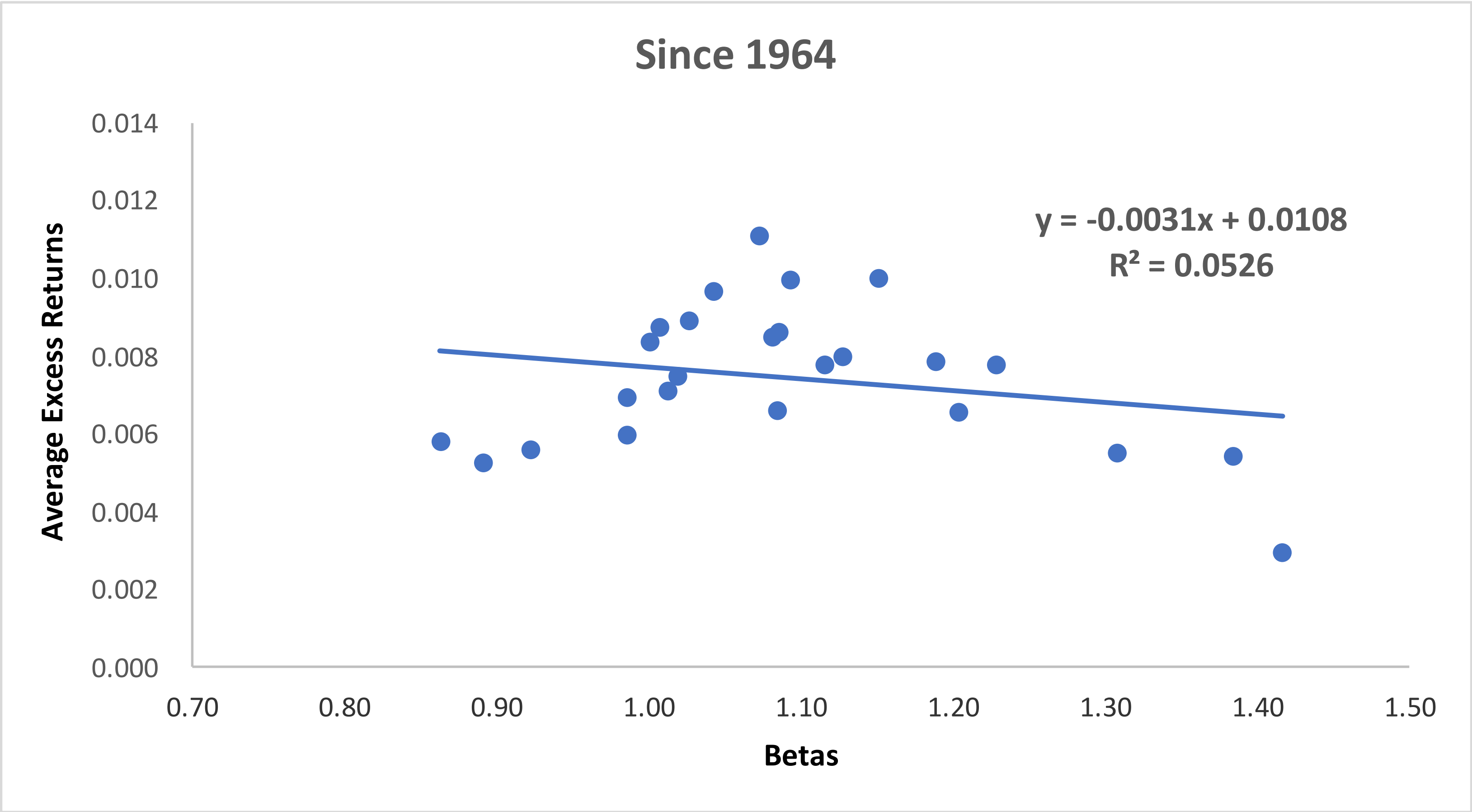
Slope is negative! This violates the CAPM.

- R-squared is only 5.26%.

- Let's take a closer look at size and B/M portfolios.

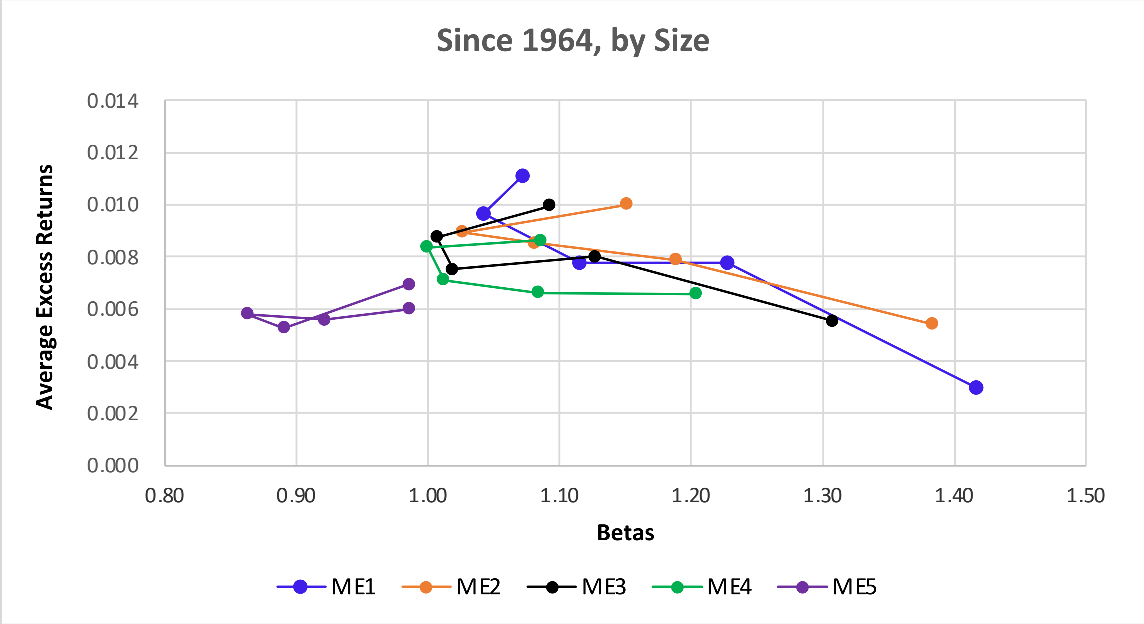
Testing SML – Second-Pass Regression

- Since 1964 (1964.1-present), CAPM performs disappointing:



Testing SML – Second-Pass Regression

- Since 1964 (1964.1-present), CAPM performs disappointing:

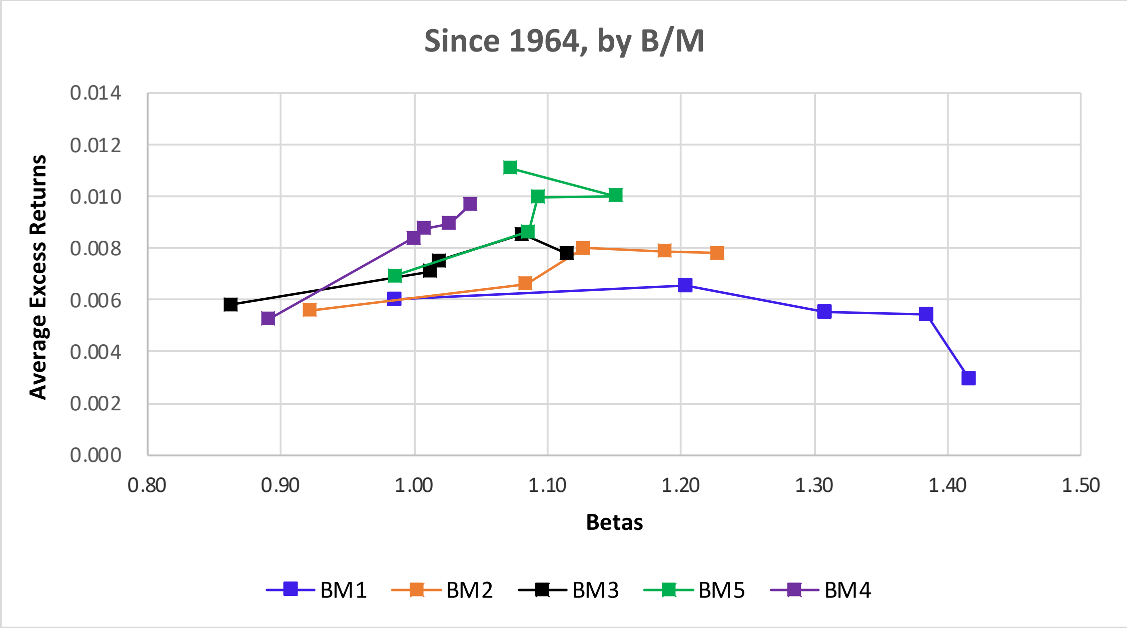


- When we check size effect, ME5 (largest) portfolio has lowest beta and lowest average return. ME1 (smallest) portfolio tends to have highest beta and highest average return.

- If you run Fama-Macbeth regression on size portfolios only, CAPM performs well. (Result is not included.)

Testing SML – Second-Pass Regression

❖ Since 1964 (1964.1-present), CAPM performs disappointing:



❖ When we check value effect, five portfolios' betas and average returns are not sorted as we expect. Higher beta doesn't require higher returns.

❖ If you run Fama-Macbeth regression on B/M portfolios only, CAPM performs very disappointing. (Result is not included.)

Testing SML of Multifactor Models

Testing SML: Fama-Macbeth Regression

- Steps to test SML of CAPM:
  1. Determine a candidate for the market portfolio *M*.
  2. Determine individual assets in the regressions.
  3. First-pass regression: For each of the assets in question, determine the beta coefficients of all the factors,
  $$R_{i,t}^e = \alpha_i + \beta_{M,i} R_{M,t}^e + \beta_{SMB,i} SMB_t + \beta_{HML,i} HML_t + e_{i,t}$$
  4. Second-pass regression: Regress the mean excess returns of the assets on three betas
  $$\overline{R_i^e} = \gamma_0 + \gamma_M \beta_{M,i} + \gamma_{SMB} \beta_{SMB,i} + \gamma_{HML} \beta_{HML,i} + \epsilon_i$$

If the FF3 in its descriptive format holds, the second-pass regression should be the SML, that is $\gamma_0 = 0$, and $\gamma_M, \gamma_{SMB}, \gamma_{HML}$ are significantly nonzero.

Otherwise, if the coefficient is zero, the factor cannot help explain the asset returns and should not be included in the factor model.

Testing SML – First-Pass Regression

- In the test of CAPM, we find size and value effects that cannot be explained by the market factor.
- Go to tab "FF3 First-pass reg" for first-pass regression in three sample periods:
- Time series regression:

$$R_{i,t}^{e} = \alpha_i + \beta_{M,i} R_{M,t}^{e} + \beta_{SMB,i} SMB_t + \beta_{HML,i} HML_t + e_{i,t}$$

- Take sample 1964.1-present as an example.
  - Average monthly market return is 0.58% + 0.36% = 0.94%.
  - Average size factor return is 0.17%, which is the return differential between small and large portfolios.
  - Average value factor return is 0.28%, which is the return differential between high B/M and low B/M portfolios.

Testing SML – Second-Pass Regression

- Go to tab "FF3 Second-pass reg" for second-pass regression results:

  $\overline{R_i^e} = \gamma_0 + \gamma_M \beta_{M,i} + \gamma_{SMB} \beta_{SMB,i} + \gamma_{HML} \beta_{HML,i} + \epsilon_i$, for portfolio i

- For the full sample period,
  - Intercept (γ_0) has mean 1.7% with very small p-value , significant.
  - Slopes (γ_{M}, γ_{SMB}, γ_{HML}) are all significant. All three pricing factors require significant risk premium, therefore, important in pricing assets.
  - R-squared increases from 9.74% (CAPM) to 67.6% (FF3).

Testing SML – Second-Pass Regression

- Go to tab "FF3 Second-pass reg" for second-pass regression results.
- Before 1964 (1926.7-1963.12),
  - Intercept (γ_0) has mean 2.4% with p-value 0.0007, still significant.
  - Market risk premium, and value premium are significant, but size premium is not.
  - R-squared increases from 30.1% (CAPM) to 61.1% (FF3).

Testing SML – Second-Pass Regression

- Go to tab "FF3 Second-pass reg" for second-pass regression results.
- Since 1964,
  - Intercept (γ_0) has mean 1.2% with p-value 0.0047, still significant.
  - Market risk premium is not significant, while size premium, and value premium are both significant.
  - R-squared increases from 5.26% (CAPM) to 62.4% (FF3).
  - The improvement of FF3 model is especially significant after 1964.

Extension

- On French's website, you can find portfolios formed based on other variables, such as operating profitability (OP), investment (Inv), E/P, CF/P, D/P, momentum, reversal, accruals, market beta, net share issues, daily variance, and daily residual variance, industry category.
- More than 400 pricing factors have been found after APT model was introduced, especially after Fama and French introduced three factor model.
- Fama and French's new paper suggests to add OP and Inv as well (FF5).
- There are still many discussions about how many and which pricing factors should be added to the model.

References

- Fama, E., MacBeth, J., 1973. Risk, return and equilibrium: Empirical tests. Journal of Political Economy 81, 607-636.
- Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.
- Fama, E., French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1-22.