Introduction to Finance for Data Scientists

Session 3: Market Efficiency

Jean-Edouard Colliard and Johan Hombert

HEC Paris, 2022

Road map

The CAPM

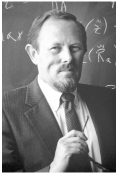
Market efficiency

Are markets efficient?

Issues for data scientists

The CAPM

another Nobel winning idea



William Sharpe (1990)

The Capital Asset Pricing Model (CAPM)

- **CAPM:** Under some assumptions, asset i's expected return is[1]

$$\boxed{E(r_i) = r_f + \beta_i \times (E(r_m) - r_f)}$$

  - r_f: risk-free rate
  - r_m: return on market portfolio (portfolio of all risky assets traded in capital markets: stocks, bonds, etc.)
  - $\beta_i = Cov(r_i, r_m)/V(r_m)$: asset i's **market beta** is asset i's systematic risk measured by its exposure to market risk, i.e., its comovement with the market portfolio
  - $E(r_m) - r_f$: market risk premium
  - $\beta_i \times (E(r_m) - r_f)$: asset i's **risk premium**, depends on asset i's systematic risk not idiosyncratic risk

[1] See the "bonus slides" for assumptions and heuristic proof of CAPM

The CAPM

- Intuition of the CAPM

  Risk-averse investors dislike assets with high market beta because these assets add to the variance of the market portfolio

  $\Rightarrow$ High β_i $\Rightarrow$ High systematic risk $\Rightarrow$ Investors require high risk premium (i.e., high expected return) on these assets

- The CAPM applies to individual assets and to portfolios
  - Beta of a portfolio = weighted average beta of assets in the portfolio
  - Note that the CAPM also applies to the market portfolio, which has a beta equal to one

The CAPM

- Historical returns of main asset classes, 1960–2017

	Moments of return distribution (in % per year)		
	S.D.	Mean	Beta
Stocks	17.3	11.3	1.5
Real estate	19.3	12.3	1.3
Corporate bonds	8.4	7.8	0.5
Commodities	24.9	8.7	0.1

E.g., commodities have low mean return despite high S.D. because they have low beta

The CAPM in data

- Consider time-series data, $t = 1, \ldots, T$

  r_{it}: return on some asset (or portfolio) i in period t (e.g., month)

  r_{mt}: market return in period t

  r_{ft}: risk-free return in period t (known at $t-1$)

- Consider the following time-series linear regression, called the "market model" for asset i's return

$$r_{it} - r_{ft} = \alpha_i + \beta_i \left(r_{mt} - r_{ft} \right) + \epsilon_{it}$$

- Estimate α_i and β_i using OLS

  $\beta_i = \dfrac{Cov(r_{it} - r_{ft}, r_{mt} - r_{ft})}{V(r_{mt} - r_{ft})}$ is asset i's market beta

  What is the interpretation of intercept α_i?

The CAPM in data

- Take expectation in the market model equation

$$\underbrace{E(r_{it} - r_{ft})}_{\text{asset i's risk premium}} = \underbrace{\beta_i E(r_{mt} - r_{ft})}_{\text{compensation for market risk}} + \underbrace{\alpha_i}_{\text{extra compensation}}$$

- α_i represents "abnormal" compensation, on top of normal compensation for market risk

- If CAPM is true: $\alpha_i = 0$

The CAPM in data

- Take variance in the market model equation

$$\underbrace{V(r_{it} - r_{ft})}_{\text{asset i's total risk}} = \underbrace{V(\beta_i(r_{mt} - r_{ft}))}_{\text{market risk=systematic risk}} + \underbrace{V(\epsilon_i)}_{\text{idiosyncratic risk}}$$

The CAPM — Example

- Industry portfolio of US automobile stocks, monthly returns, 1926–2020 [data]

- Result of the market model regression:

  $\hat{\beta} = 1.2$ (std error = .05)

  > $\hat{\beta} > 1$ means the automobile industry is quite procyclical: when market goes up 1%, automobile stocks go up 1.2% on average

  $\hat{\alpha} = 0.04$ (std error = .12)

  > alpha is not statistically different from zero, as predicted by CAPM

Quiz

Q1. Automobile stocks have higher expected return than the market portfolio. True or false?

Q2. Automobile stocks have a more favorable risk-return tradeoff than the market portfolio. True or false?

Q3. Which industries have $\beta < 1$?

The CAPM and beyond

- The CAPM is a *model*: it says what asset prices and expected returns should be *under some assumptions*

  Under these assumption, we can prove that the CAPM is true

- Main assumption (not the only one): investors can perfectly diversify

  Perfect diversification implies that an asset's systemic risk is determined by its covariance with the market portfolio

- If assumption of perfect diversification does not hold

  CAPM does not hold: an asset's systemic risk depends on other factors (more on this later)

Road map

The CAPM

Market efficiency

Are markets efficient?

Issues for data scientists

Market efficiency — Definition

- The CAPM is a *normative* model: it says what asset prices and expected returns *should be*

- In **efficient markets**, prices and expected returns are equal to their normative value

  $\Rightarrow$ Every asset has alpha = 0

  $\Rightarrow$ It is impossible to find assets with superior expected return adjusted for the compensation for risk

- In **inefficient markets**, prices and expected returns sometimes differ from their normative value

  $\Rightarrow$ Assets can have alpha $\neq 0$

  $\Rightarrow$ Smart investors can find assets with superior expected return adjusted for the compensation for risk

Market efficiency — Example

- Yesterday, new results were released showing that company XYZ's revolutionary solar cells are 50% more efficient than current best technology. This is very good news for future earnings of XYZ

Q1. If markets are <u>efficient</u>, it is a good idea to buy company XYZ's stock today. True or false?

Q2. If markets are <u>inefficient</u>, it is a good idea to buy company XYZ's stock today. True or false?

Market efficiency — Example

- E-commerce company ABC announces that its sales are up 25% compared to last year. ABC stock price immediately loses 5%

Q3. In this particular instance, the market has been inefficient. True or false?

Hedge funds

- Hedge funds are in the business of searching for alpha
- Hedge funds invest their clients' money in capital markets
- Using data and/or economic analysis, they look for assets with nonzero alpha, buy undervalued assets (with $\alpha > 0$) and short-sell overvalued assets (with $\alpha < 0$)
- Hedge funds charge fees (which depend on the fund's performance) to their clients

Hedge funds

- Hedge funds hire finance & data experts

Rank	Firm	Headquarters	AUM as of second quarter 2021 (millions of USD)
1	Bridgewater Associates	Westport, CT	$105,700
2	Man Group	London, UK	$76,800
3	Renaissance Technologies	East Setauket, NY	$58,000
4	Millennium Management	New York City, NY	$52,314
5	The Children's Investment Fund Management	London, UK	$40,000
5	Elliott Management	New York City, NY	$42,000
6	D.E. Shaw & Co.	New York City, NY	$39,738
7	Two Sigma Investments	New York City, NY	$39,550
8	Farallon Capital	San Francisco, CA	$38,100
9	Citadel LLC	Miami, FL	$37,630
10	Davidson Kempner Capital Management	New York City, NY	$37,350

- Hedge funds have a raison d'être only if markets are inefficient

If markets are efficient, there is no alpha

Road map

The CAPM

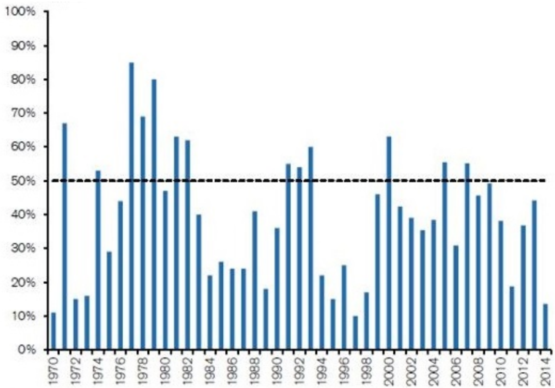
Market efficiency

Are markets efficient?

Issues for data scientists

Performance of mutual funds

- Fraction of mutual funds with $\alpha > 0$



Source: John C. Bogle (1970-1999) and "SPIVA® U.S. Scorecard: Year End 2014," S&P Dow Jones Indices Research (2000-2014).
Note: U.S. general equity funds (1970-1999) and U.S. large capitalization equity funds (2000-2014).

- It does not seem so easy to generate alpha

Are markets efficient?

- Whether markets are efficient or not is a hotly debated issue

- Nobel laureates rewarded for research on market efficiency disagree with each other!

Eugene F. Fama (2013)

Lars Peter Hansen (2013)

Robert J. Shiller (2013)

Are markets efficient?

- Academics and investors have found many "market anomalies", that is, portfolios with nonzero alpha

- Example: the "profitability anomaly." Companies with high profits have higher expected returns than predicted by the CAPM ($\alpha > 0$)

Example: the profitability anomaly

- US stock return data, monthly, 1963–2020 [data]
- At the beginning of each month, construct a portfolio that contains all stocks of companies in the top tercile of previous year's company earnings. The return of this portfolio is

$$r_t^{top} = \sum_{i \in TopTercile_t} w_i r_{i,t} \qquad \text{with portfolio weights } w_i = \frac{1}{\# TopTercile_t}$$

- Estimate the CAPM regression

$$r_t^{top} - r_{ft} = \alpha + \beta \left(r_{mt} - r_{ft} \right) + \epsilon_t$$

$\beta = 0.97$ (s.e.=0.01)

$\alpha = 0.13$ (s.e.=0.04) statistically different from zero

$\Rightarrow$ Profitable companies have abnormally high expected return

Example: the profitability anomaly

- Similarly, construct the portfolio that contains all stocks in the <u>bottom</u> tercile of previous year's company earnings

$$r_t^{bottom} = \sum_{i \in BottomTercile_t} w_i r_{i,t} \qquad \text{with } w_i = \frac{1}{\#BottomTercile_t}$$

- Estimate the CAPM regression

$$r_t^{bottom} - r_{ft} = \alpha + \beta \left(r_{mt} - r_{ft} \right) + \epsilon_t$$

$\beta = 1.13$ (s.e.=0.02)

$\alpha = -0.20$ (s.e.=0.06) statistically different from zero

$\Rightarrow$ Unprofitable companies have abnormally low expected return

Example: the profitability anomaly

- Hedge funds exploit such anomalies by constructing long-short portfolios

  NB: taking a long position means buying
  taking a short position means short-selling

- Short-selling

  at t: borrow stock i from a broker (investment bank) and sell the stock in the market $\Rightarrow$ cash flow $= P_{i,t}$

  at $t+1$: buy back the stock and return it to the broker $\Rightarrow$ cash flow $= -P_{i,t+1}$

  $\Rightarrow$ profit $= P_{i,t} - P_{i,t+1}$

- Profit on a short position $(P_{i,t} - P_{i,t+1})$ = minus the profit on a long position $(P_{i,t+1} - P_{i,t})$

Example: the profitability anomaly

- Long-short portfolio to exploit the profitability anomaly
  - buy the top tercile portfolio ("long leg" of the long-short portolio)
  - short-sell same amount in the bottom tercile portfolio ("short leg")
- Profit per €1 in the long and short legs = $r_t^{top} - r_t^{bottom}$

$$\alpha = \alpha^{top} - \alpha^{bottom} = 0.13 - (-0.20) = 0.33 \text{ (s.e.=0.09)}$$

$\Rightarrow$ the long-short portfolio earns $0.33 \times 12 = 4$ pp per year higher return than predicted by CAPM

Many other "market anomalies"

- Portfolios of stocks with the following characteristics have been found to have positive alpha
  - High ratio book value/market value of equity (value effect)
  - Low market capitalization (small size effect)
  - High return in previous year (momentum effect)
  - Low return in previous five years (long-term reversal)
  - Low return in previous week (short-term reversal)
  - Low investment
  - Low idiosyncratic volatility
  - etc.

Road map

The CAPM

Market efficiency

Are markets efficient?

Issues for data scientists

Issues for data scientists in finance

- Two issues when you find an investment strategy with positive alpha

  1. Risk adjustment: have we correctly adjusted for systematic risk?

  2. p-hacking: out-of-sample validity?

Risk adjustment

- Ex.: Companies with high profits have higher expected return than predicted by CAPM

- Interpretation 1: the market is inefficient (alpha is truly > 0)

  Companies with high profits are undervalued, perhaps because investors under-react to good news about companies' profits[2]

- Interpretation 2: we did not adjust for risk correctly (alpha is not correctly measured)

  Risk adjustment based on CAPM (systematic risk = covariance with market risk) may not be correct (systematic risk may depends on covariance with other risk factors and on higher order moments[3])

---

[2] Augustin Landier and coauthors, 2019, "Sticky Expectations and the Profitability Anomaly," *Journal of Finance* [pdf]

[3] Hugues Langlois, 2020, "Measuring skewness premia," *Journal of Financial Economics* [pdf]

Risk adjustment

- Which interpretation is correct?
- We cannot know for sure!
  - Known as the "joint hypothesis problem": cannot know if we reject the market efficiency hypothesis or the risk adjustment
  - We can use other models than the CAPM to include additional risk adjustment, but we can never eliminate the joint hypothesis problem
  - Practical solution: ask yourself (using data and common sense) if the portfolio is exposed to a risk factor that could explain the higher expected return

Risk adjustment

- Ex.: Hedge funds performance, 1996–2012[4]

	Average hedge funds returns (per year)		
	mean return	alpha	beta
Before fees	9.3%	7.3%	0.37
After fees	5.5%	3.6%	0.26

  - Positive alpha
  - Hedge funds exposed to all sorts of exotic risks: is risk adjustment for covariance with market enough?

[4] Jurek and Stafford, 2015, "The Cost of Capital for Alternative Investments," *Journal of Finance* [pdf]

p-hacking (a.k.a. data snooping)

- Researchers have looked for and found many strategies with positive alpha. However, after researchers publish their results, the alpha tends to disappear[5]

Q. How do you interpret this fact?

[5] McLean and Pontiff, 2016 "Does Academic Research Destroy Stock Return Predictability?" *Journal of Finance*

p-hacking

1. Type I errors (false positive): the strategy has positive alpha in-sample by luck, but zero alpha out-of-sample
   - Familiar to data scientists
2. Market efficiency: once the alpha is discovered, hedge funds trade on it, which drives alpha to zero
   - Hedge funds buy assets with positive alpha $\Rightarrow$ price $\uparrow$ $\Rightarrow$ future return $\downarrow$ $\Rightarrow$ α vanishes
   - Alpha is not exogenously determined, it depends on the behavior of the people who try to predict alpha!

More on data in finance

Two courses in the Spring, 2023

- Data Analysis in Finance, Augustin Landier & Hugues Langlois
- Economic Value of Data, Jean-Edouard Colliard & Thierry Foucault

Until tomorrow

- Practice problems on Slack
- Email me the group composition for the group work
- See you tomorrow morning ☺

Bonus slides

Bonus slides: Heuristic derivation of the CAPM

- **CAPM:** Under the following assumptions

  (i) investors have mean-variance preferences and

  (ii) investors can diversify perfectly,

  the risk premium on asset i's is

$$E(r_i) - r_f = \beta_i \times (E(r_m) - r_f)$$

  where $\beta_i = \dfrac{Cov(r_i, r_m)}{V(r_m)}$ and

  - r_m: return on the market portfolio = portfolio of all risky assets $i = 1, \ldots, n$ traded in capital markets

    $r_m = \sum_i w_i^m r_i$ where $w_i^m = \frac{\text{Asset } i\text{'s market capitalization}}{\text{Sum of all assets' market capitalizations}}$

  - r_f: risk-free rate = return on risk-free government bonds

Bonus slides: Heuristic derivation of the CAPM

- Assumption (i): investors have mean-variance preferences, i.e., they assign subjective value $E(r) - \frac{1}{2}\gamma V(r)$ to risky return r

  $\gamma \geq 0$ is the coefficient of risk aversion: investors with higher γ are more risk averse, investors with $\gamma = 0$ are risk neutral

- Determine the optimal portfolio for a mean-variance investor
  - Denote x_i the % of the portfolio invested in each risky asset $i = 1, \ldots n$, the rest $\left(1 - \sum_i x_i\right)$ % is invested in the risk-free asset
  - Portfolio return: $r_P = \sum_i x_i r_i + \left(1 - \sum_i x_i\right) r_f$
  - Optimal portfolio solves $\max_{x_1,\ldots,x_n} E(r_P) - \frac{\gamma}{2} V(r_P)$

Bonus slides: Heuristic derivation of the CAPM

$E(r_P) = r_f + \sum_i x_i(E(r_i) - r_f)$

$V(r_P) = \sum_{i,j} x_i x_j Cov(r_i, r_j)$

– First order condition w.r.t. x_i:

$\frac{dE(r_P)}{dx_i} = E(r_i) - r_f$

$\frac{dV(r_P)}{dx_i} = 2\sum_j x_j Cov(r_i, r_j) = 2Cov(r_i, r_P)$

$$\Rightarrow \quad \underbrace{E(r_i) - r_f}_{\substack{\text{Asset i's marginal} \\ \text{contribution to} \\ \text{portfolio} \\ \text{expected return}}} - \frac{\gamma}{2} \underbrace{2Cov(r_i, r_P)}_{\substack{\text{Asset i's marginal} \\ \text{contribution to} \\ \text{portfolio risk}}} = 0 \qquad (1)$$

Bonus slides: Heuristic derivation of the CAPM

- Assumption (ii): investors can perfectly diversify (i.e., can choose $x_i \neq 0$ for all i) and they hold the market portfolio (i.e., x_i proportional to market portfolio weight w_i^m)

(NB: The second part of the assumption can be derived as an implication of the first part.)

$\Rightarrow$ $r_P = r_f + s\, r_m$ where $s = \sum_i x_i$

(1) $\Rightarrow$ $E(r_i) - r_f = \gamma\, s\, Cov(r_i, r_m)$

This also holds for the market portfolio (proof: sum over i with weights w_i) $\Rightarrow$ $E(r_M) - r_f = \gamma\, s\, V(r_M)$

Therefore $E(r_i) = r_f + \beta_i \times (E(r_m) - r_f)$ where $\beta_i = \frac{Cov(r_i, r_m)}{V(r_m)}$