Investments

Chapter 1: Investments - capital markets and products

BKM - Chapter 1 (1.1-1.5): The investment environment
BKM - Chapter 2 (2.1-2.3): Asset classes and financial instruments
BKM - Chapter 4 (4.1-4.4): Mutual funds and other investment companies

Financial Markets

Function

The focus of this course is on the investment in financial instruments
What is the rationale of this?

Financial markets allow to share or transfer risk: If you decide to buy stocks you take a risk you are willing to, but you want to be compensated for that

- Capital markets allow the risk that is inherent to all investments to be borne by the investors most willing to bear it.

Financial markets allow to separate the timing of income and consumption: store wealth to transfer to the future: The moment at which you store income and consume.

- Also, a saving account allows to transfer wealth over time but the fact that financial markets exist allow you to choose between more options to spread your consumption over time. Transferring wealth over time gives more opportunities and more wealth over time.

Financial markets allow to separate ownership and management in support of large-scale businesses (but be careful with associated agency problems): If a company is initiated by its founders, is limited in its growth because it doesn't have the necessary funding to grow, if it wants to grow beyond its part it needs extra funding.

A big firm has tens of thousands of stockholders with an ownership stake in the firm proportional to their holdings of shares. Such a large group of individuals cannot actively participate in the day-to-day management of the firm
They elect a board of directors that in turn hires and supervises the management of the firm.
= the owners and managers of the firm are different parties

But potential **agency problems**: managers engage in activities not in the best interest of shareholders. How can we mitigate this?

- Compensation plans tie the income of managers to the success of the firm.
- Board of directors can force out management teams that are underperforming
- outsiders such as security analysts and large institutional investors such as mutual funds or pension funds monitor the firm closely and make the life of poor performers at the least uncomfortable.
- Bad performers are subject to the threat of takeover

Financial markets allow to allocate resources efficiently i.e. to most productive real investments (= informational role of financial markets):

- This has to do with market efficiency: if its efficient instruments are priced based upon its fundamental value than money flows to productive opportunities otherwise it flows to unproductive opportunities. Efficient allocation to productive real investments.
- Stock prices reflect investors' collective assessment of a firm's current performance and future prospects. When the market is more optimistic about the firm, its share price will rise → easier for the firm to raise capital → more investment.
- In this manner, stock prices play a major role in the allocation of capital in market economies, directing capital to the firms and applications with the greatest perceived potential.
- The stock market encourages allocation of capital to those firms that appear *at the time* to have the best prospects.

- Markets stay open even during turbulent times, some countries closed their stock markets. If you open, you allow information to flow into the market and when you close it then not.

Real assets VS Financial Assets

An investment is the current commitment of money or other resources in the expectation of reaping future benefits.

Reals assets: the land, buildings, machines, …
Financial assets: claim to the income generated by real assets. Example: if we cannot own our own auto plant (a real asset), we can still buy shares in Ford or Toyota (financial assets) and thereby share in the income derived from the production of automobiles. So investors' returns ultimately come from the income produced by the real assets that were financed by the issuance of those securities.

Asset classes

The universe of financial instruments that we focus on is limited to traditional assets classes:

- **money market instruments**: short-term, low risk debt securities ("cash") → allow us "to save". Risk free assets, cash alternatives.
- **capital market instruments**: longer term, riskier and more diverse assets, namely stocks and bonds → allow us "to invest". Relatively low risk and extremely risky = wide range; 2 large categories stocks and bonds; also, combinations that allows different risk return profiles.

Each instrument is unique in terms of (un-)certainty of payments and timing of payments and has a distinct contribution to the investment portfolio

Money market instruments

Treasury bill: government debt obligation with a maturity < 1 year.

- This has the lowest risk.
- Most marketable of all money market instruments.
- The government raises money by selling bills to the public. Investors buy the bills at a discount from the stated maturity (equivalently, face) value. At maturity, the government pays the investor the face value of the bill. The difference between the purchase price and ultimate maturity value constitutes the investor's earnings.
- Are highly liquid; they are easily converted to cash and sold at low transaction cost and with not much price risk.
- Are exempt from state and local taxes

The **ask price** is the price you would have to pay to buy a T-bill from a securities dealer.
The **bid price** is the slightly lower price you would receive if you wanted to sell a bill to a dealer.
The **bid–ask spread** is the difference in these prices, which is the dealer's source of profit.

Certificate of deposit (CD): time deposit at a bank. Here you have bank risk.

- Time deposits may not be withdrawn on demand. The bank pays interest and principal to the depositor only at maturity.

Commercial paper (CP): short term unsecured debt issued by a large corporation. Here you have company risk.

Large companies often issue their own short-term unsecured debt notes rather than borrow directly from banks. Very often, commercial paper is backed by a bank line of credit, which gives the borrower access to cash that can be used to pay off the paper at maturity.

Commercial paper maturities range up to 270 days, but most often, it is issued with a maturity of less than 1 or 2 months. Commercial paper is considered to be a fairly safe asset because a firm's condition presumably can be monitored and predicted over a term as short as 1 month.

Bankers' acceptance: a bank promise to pay a prespecified amount.

This starts as an order to a bank by a bank's customer to pay a sum of money at a future date, typically within 6 months. Bankers' acceptances are considered very safe because traders can substitute the bank's credit standing for their own.

Repurchase agreement (repo or RP) **and reverse repo**: Dealers in government securities use this as a short-term loan (usually overnight) using other securities (usually government securities) as collateral. The dealer sells government securities to an investor on an overnight basis, with an agreement to buy back those securities the next day at a slightly higher price. The increase in the price is the overnight interest. The dealer thus takes out a 1-day loan from the investor, and the securities serve as collateral.

- Reverse repo is the one giving the loan and repo is the one taking the loan
- A *reverse repo* is the mirror image of a repo. Here, the dealer finds an investor holding government securities and buys them, agreeing to sell them back at a specified higher price on a future date.
- This is used to bring liquidity in the market.
- Repos are considered very safe in terms of credit risk because the loans are backed by the government securities.

Interbank loans: short term loans among banks (from which reference rates are being calculated e.g., eSTR, EURIBOR)

- Euribor has a longer maturity: rate, cost from borrowing from other banks.
- eSTR was EUNIA before and they would use the Libor. Euribor has another way to calculate, before then the rates were theoretical. Now there are real rates at which you will really borrow. Before it was possible to manipulate Libor rates, now with the new one it is not possible

Central bank deposits and loans: short term deposit/loan of a bank by its central bank (e.g. federal funds, MRO, deposit facilities, lending facilities, MRR).

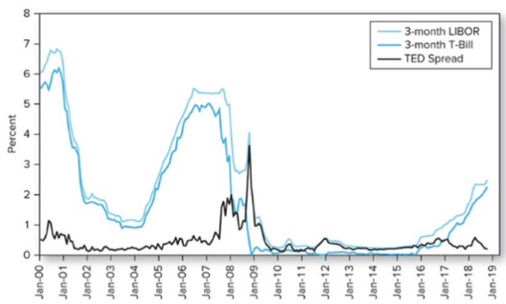
- MRO and MRR were at 0% or negative, now they are rising. This influences the liquidity in the market, they are doing this to calm down inflation

Call loans: loans that need to be repaid, on demand at any time; often made by banks to brokerage firms to fund individuals that buy on margin

Most money market instruments are low risk, but they are not risk-free! (In general T-bills have lowest risk, both in terms of credit risk, as well as liquidity risk.)
The securities of the money market promise yield greater than those on default-free T-bills, at least in part because of greater relative riskiness. In addition, many investors require more liquidity; thus, they will accept lower yields on securities such as T-bills that can be quickly and cheaply sold for cash.

In turbulent times rates diverge, when the economy is steady the rates are close to one another:

TED spread the difference between the LIBOR rate and Treasury bills, also peaked during periods of financial stress: both rates are close to one another in most periods of times. In 2008 libor rate spiked, because banks were in trouble, no confidence, so the rate at which you were able to trade increased because of the low confidence.



Source: BKM (2021) p.17
Figure 1.1 Short-term LIBOR and Treasury-bill rates and the TED spread

Capital market instruments: Bonds

Bonds are longer term debt instruments with wide range of maturities: you can choose between LT or ST e.g., market is turbulent, than there would be an inverse relation between the value of the bond and interest rates. If rates are going up the value of the bond is going down. This effect will be the biggest on longer term bonds, because you are stuck for a long time. With short term bonds you can roll over, the rates are expecting to rise over time.

- Various credit qualities: from high credit-quality instruments to (very) low credit-quality (junk) instruments
- Liquidity varies from very high to almost zero: Some bonds you can sell very quickly = very liquid. It's risky when you cannot easily find a counterparty, very illiquid, so you will have to lower the price.
- Often in smaller denominations such that they can be held by retail investors
- Often not traded on the exchange, but rather traded over the counter (OTC) via dealers
- For bonds you can observe a particular quote and when you go to the market you will see another price = stale market.
- Expressed as a % of the par value
- Typically pay coupon interest (annual or semi-annual)
- Performance measurement: hpr or yield with (including a risk premium for credit risk)

Examples:

- **Treasury notes**: government debt with 1 yr < maturity < 10 yrs
- **Treasury bonds**: government debt with maturity > 10 yrs

- **Treasury inflation-protected securities (TIPS):** government debt with principal and interest adjusted for inflation
    - Par value is adjusted to inflation, also the coupons = real risk-free instrument because it corrects for inflation. But it's very expensive, you can hedge for inflation, most investors will hedge with other types of bonds that correlates with inflation and you get a natural hedge.
- **Federal agency/municipal bonds**: debt issued or guaranteed by, respectively, a federal agency, state or local government
- **Corporate bonds**: debt issued by a large corporation; typically, larger default risk as compared to government debt; options included (callable, convertible)
    - *Callable bonds* give the firm the option to repurchase the bond from the holder at a stipulated call price.
    - *Convertible bonds* give the bondholder the option to convert each bond into a stipulated number of shares of stock.

Secured bonds: there is specific collateral backing them in the event of firm bankruptcy
Unsecured bonds (debentures): no collateral
Subordinated debentures: have a lower priority claim to the firm's assets in the event of bankruptcy.

- **Asset backed debt**: proportional ownership claims in an asset pool (cf securitization)

Capital market instruments: Stocks

Stocks (**common shares**) are issued by corporations and represent ownership in a firm

- share the distribution of profits
- have voting power at shareholders' meeting

- Equity represents a residual claim: you only receive payout if all other claims (e.g. salaries, debt, taxes) are met
- Equity has limited liability: minimum share price is zero (return is bounded at -100%). The most shareholders can lose in the event of failure of the corporation is their original investment
- Performance measurement: return (mostly reward for market risk)
- Normal distribution: it's not bounded, the lower bound should be bounded, but the normal distribution is not bounded → log normal would be an alternative
- Return: always account for the dividend

Note: **preferred shares** are related and possess common equity features, but also bond features:

- share of ownership
- no voting power
- promise of fixed dividend like a perpetuity *(often cumulative: The firm retains discretion to make the dividend payments to the preferred stockholders; it has no contractual obligation to pay those dividends. Instead, preferred dividends are usually cumulative; unpaid dividends cumulate and must be paid in full before any dividends may be paid to holders of common stock. In contrast, the firm does have a contractual obligation to make the interest payments on the debt.)*

The investment process

An investor’s **portfolio** is simply his collection of investment assets.
Investment assets can be categorized into broad asset classes, such as stocks, bonds, real estate, commodities, ...
Investors make two types of decisions in constructing their portfolios:

- The **asset allocation** decision: the choice among these broad asset classes
- The **security selection** decision: the choice of which particular securities to hold *within* each asset class.
    - Top-down portfolio construction starts with asset allocation.
    - Bottom-up portfolio is constructed from securities that seem attractively priced without as much concern for the resultant asset allocation.

Security analysis involves the valuation of particular securities that might be included in the portfolio.

Markets are competitive

Financial markets are highly competitive. This competition means that we should expect to find **few**, if any, “free lunches,” securities that are so underpriced that they represent obvious bargains = **no-free-lunch proposition**

Risk-return trade off

Actual or realized returns will almost always deviate from the expected return anticipated at the start of the investment period = risk
If you want higher expected returns, you will have to pay a price in terms of accepting higher investment risk. There is a **risk–return trade-off** in the securities markets, with higher-risk assets priced to offer higher expected returns than lower-risk assets.

Efficient markets

We should rarely expect to find bargains in the security markets. The security price usually reflects all the information available to investors concerning its value = **efficient market hypothesis**

According to this hypothesis, as new information about a security becomes available, its price quickly adjusts so that at any time, the security price equals the market consensus estimate of the value of the security. If this were so, there would be neither underpriced nor overpriced securities.

→ Choice between active or passive management:

- **Passive management** calls for holding highly diversified portfolios without spending effort or other resources attempting to improve investment performance through security analysis.

- **Active management** is the attempt to improve performance either by identifying mispriced securities or by timing the performance of broad asset classes.

Investment funds

Retail investors often do not trade individual securities directly but invest in a fund that allows them to pool assets in which they share ownership. Those funds are financial intermediaries that collect funds from individual investors and invest those funds in a potentially wide range of securities or other assets = **investment companies**
→ realization of economies of scale:

- **Diversification and divisibility**: By pooling their money, investment companies enable investors to hold fractional shares of many different securities. They can act as large investors even if any individual shareholder cannot. If you have only 5000 euros to invest in the portfolio you have will be concentrated but in an investment fund it is very wide
- **Lower transaction costs**: Because they trade large blocks of securities, investment companies can achieve substantial savings on trading costs. On a larger pool of money, you pay less transaction costs
- **Professional management:** Investment companies can support full-time staffs of security analysts and portfolio managers who attempt to achieve superior investment results for their investors. You will also be able to benefit of professional management because one person is specialized in managing the fund, as a retail investor you will have to do it on your own, you do not have the tools like investment funds have.
- **Record keeping and administration**: Investment companies issue periodic status reports, keeping track of capital gains distributions, dividends, investments, and redemptions, and they may reinvest dividend and interest income for shareholders.

Investors buy shares in investment companies, and ownership is proportional to the number of shares purchased. The value of a share in the investment company is the **net asset value (NAV)**:

$$\text{NAV} = \frac{\text{market value of assets - liabilities}}{\text{shares outstanding}}$$

Investment companies

Open-end fund ('mutual fund'): *managed* fund with specified investment policy that issues new shares when investors buy in and redeem shares when investors cash out; they are priced at NAV (traded at day-end only)

- Whenever you as an investor step in, it issues new shares and shares will be deleted when you exit.
- The number of shares outstanding depends on the flow of money, it's not fixed. When you enter or exit you do this at the NAV. They do not trade on organized exchanges. Instead, investors simply buy shares from and liquidate through the investment company at net asset value.
- The NAV is calculated at the end of each day. If you want to enter or exit you don't know at which price, but you know that you will always be able to enter or exit at the NAV. But you cannot do this whenever you want, there are specified times.

Closed-end fund: *managed* fund with specified investment policy and fixed number of shares that trade intra-day on an organized exchange at market determined prices (can be premium/discount to NAV)

- There are a number of fixed shares. On the moment these shares are created they are trading.
- You can buy intraday, and you can observe the prices at which you can enter or exit.
- These prices will deviate from the NAV. You will see that the prices will be lower, it's worth less.

- At the moment the fund is created the prices will mostly be at premium, you enter at a high price and you will only exit at a lower price. Because after the fund is created the prices are going downwards.

Unit trusts: *hybrid format with a maturity date, fixed portfolio ('units') of uniform assets, but redeemable via secondary market (popular diversification tool for buy-and-hold investor)*

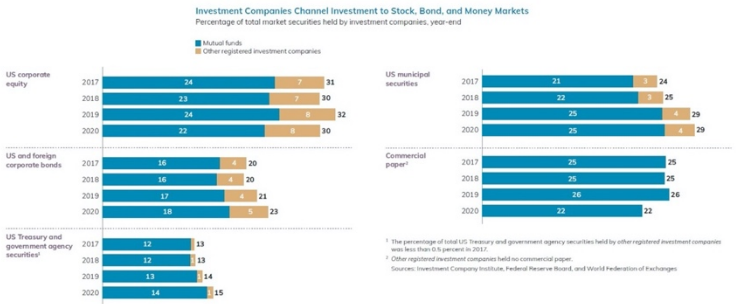
These are pools of money invested in a portfolio that is fixed for the life of the fund.
To form a unit investment trust, a sponsor buys a portfolio of securities that are deposited into a trust. It then sells shares, or "units," in the trust, called *redeemable trust certificates.* All income and payments of principal from the portfolio are paid out by the fund's trustees (a bank or trust company) to the shareholders.

- Little active management because once established, the portfolio composition is fixed = ***unmanaged***.
- Invest in relatively uniform types of assets; for example, one trust may invest in municipal bonds, another in corporate bonds.
  The uniformity of the portfolio is consistent with the lack of active management. The trusts provide investors a vehicle to purchase a pool of one particular type of asset that can be included in an overall portfolio as desired.
- Investors who wish to liquidate their holdings of a unit investment trust may sell the shares back to the trustee for net asset value.
  The trustees can either sell enough securities from the asset portfolio to obtain the cash necessary to pay the investor, or they may instead sell the shares to a new investor (again at a slight premium to net asset value).

Exchange traded fund (ETF): hybrid format, legally structured as open-end, but traded intra-day very close to NAV on an organized exchange
- This combines the best of 2 worlds; mutual and closed end.
- In term of legal structure its open end, the number of shares outstanding are the same.
- Difference is that you can trade intraday which is similar to the closed end.
- It seems like a simple product, it tracks, replicates an existing index. But the mechanism is quite complex, they have to make sure you can buy intraday at prices equal to the NAV.
- Behind the ETF you have a secondary market but also primary were dealers operates to make sure the prices equal the NAV.

Mutual funds are by far the biggest category but ETF's have indeed been more popular but are still less than the mutual funds.

Investment companies are important in the economy, you see here how they channel money to different markets.



Other investment organizations

Other intermediaries with similar purpose, but that are not formally organized or regulated as investment companies

Commingled funds: partnerships of investments, similar in form to open-end mutual funds for very large investors (e.g., retirement accounts). Commingled funds are similar in form to open-end mutual funds. Instead of shares, though, the fund offers *units,* which are bought and sold at net asset value.

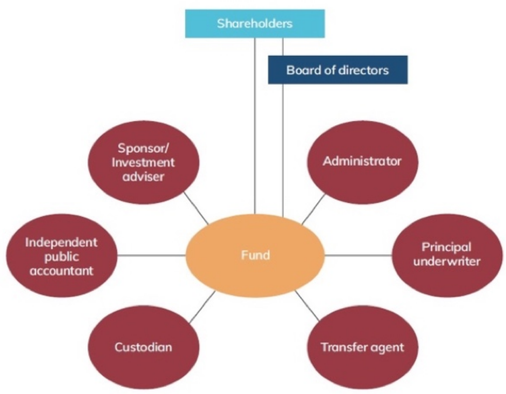
Real estate investment trust (REIT): an investment vehicle to invest in real estate or loans secured by real estate, similar in form to closed-end mutual funds (very regulated). Besides issuing shares, they raise capital by borrowing from banks and issuing bonds or mortgages.

There are two principal kinds of REITs. *Equity trusts* invest in real estate directly, whereas *mortgage trusts* invest primarily in mortgage and construction loans.

Hedge funds: private partnerships of investments with limited regulation (and exotic investment strategies because hedge funds are only lightly regulated)
- They are targeting professional investors; they follow very exotic policy.
- They cannot target retail investors because they are protected by law.
- The fees that you pay are extremely high. They also work with performance fee, if they can reach a particular target return, then they will do everything to reach that target return.
- They typically are open only to wealthy or institutional investors.
- Many require investors to agree to initial “lock-ups”: periods as long as several years in which investments cannot be withdrawn.
    - Allow hedge funds to invest in illiquid assets without worrying about meeting demands for redemption of funds.

Mutual fund

Organization



How is the "pooling of money" organized?

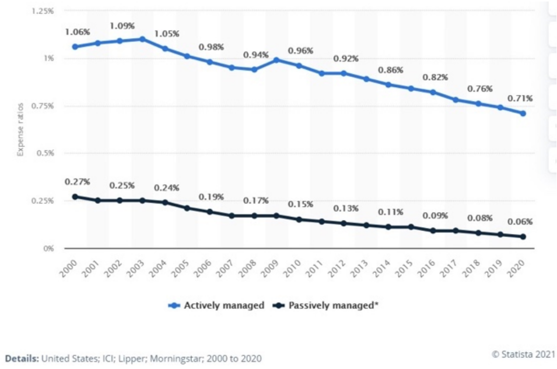
Parties involved:
- **Sponsors**: set up the fund (e.g. register the fund with the financial regulators, hire service providers); can be the advisor or distributor
    - Investment adviser: looks that what is in the fund is being managed, they manage assets in the fund
- **Fund**: manages the fund operations (e.g. compliance, risk, administration); can be outsourced to a management company
- **Board of directors** oversees the management of the fund in the interest of the shareholders
- **Shareholders**: are entitled to financial proceeds of the fund and have (specific) voting power
- **Advisors**: manage the fund’s portfolio in line with the investment policy and performs some administrative fund tasks

- **Administrator**: handles the back-off administration (e.g. fund accounting, data processing, bookkeeping, internal auditing)
- **Principal underwriter/distributor**: acts as sales agreement agent between fund and broker-dealers/distributes the fund
    - Buyers are reached via a distributor which is a very costly structure
- **Transfer agent**: maintains records of shareholders' accounts
- **Custodian**: safe keeps the assets and assesses the conduct of the fund and the quality of information it receives; needs to be independent
- **Auditor**: certifies the statements of the fund

Important: As many external parties are involved, this has implications for the costs of investing in a mutual fund e.g. operating expenses and management fees as reflected in the expense ratio. n individual investor choosing a mutual fund should consider not only the fund's stated investment policy and past performance but also its management fees and other expenses.

The investment policy and strategy are key determinants of the expense ratio, with actively managed funds having a higher expense ratio than passively managed funds

Average expense ratio; actively managed funds charges higher expense rates then passively managed funds
- For both types of the expense ratio have been declining over time
- The cost to enter a passively managed fund is very low = low margin business, you earn very low so you will have to do large values



Fee structure

Operating expenses are the costs incurred by the mutual fund in operating the portfolio, including administrative expenses and advisory fees paid to the investment manager. Shareholders do not receive an explicit bill for these operating expenses; instead they pay for these expenses through the reduced value of the portfolio.

In addition, there can be brokerage costs and front-end/back-end load:

Brokerage costs: many funds assess fees to pay for marketing and distribution costs. These charges are used primarily to pay the brokers or financial advisers who sell the funds to the public. One can avoid these expenses by buying shares directly from the fund sponsor, but many investors are willing to incur these distribution fees in return for the advice they may receive from their broker.

Front-end load: a commission or sales charge paid when you purchase the shares.
- *Low-load funds* have loads that range up to 3% of invested funds.
- *No-load funds* have no front-end sales charges.
- Loads effectively reduce the amount of money invested. *For example, each $1,000 paid for a fund with a 6% load incurs a sales charge of $60 and fund investment of only $940. You need cumulative returns of 6.4% of your net investment (60/940 = .064) just to break even.*

Back-end load: a redemption, or "exit," fee incurred when you sell your shares. Typically, funds that impose back-end loads reduce them by 1 percentage point for every year the funds are left invested. Thus, an exit fee that starts at 4% would fall to 2% by the start of your third year. These charges are known more formally as "contingent deferred sales loads.

Given this multitude of costs, transparency about costs and cost structure is crucial (cf moral hazard issues) and therefore regulated. For sure in low return environments the height of the fees is crucial!

Comparing these costs is important when entering a fund, it can decrease your return → cost structure has a huge impact on the cash flow you will earn.
(Cost are being regulated! Mifid regulation in EU: funds have to be transparent and open on the costs they charge)

Example:

Table 4.2 Impact of costs on investment performance

	Cumulative Proceeds (All Dividends Reinvested)		
	Fund A	Fund B	Fund C
Initial investment*	$10,000.00	$10,000.00	$ 9,400.00
5 years	15,922.92	15,210.60	14,596.24
10 years	25,353.93	23,136.23	22,664.92
15 years	40,370.85	35,191.60	35,193.90
20 years	64,282.18	53,528.53	54,648.80

*After front-end load, if any.
Notes:
1. Fund A is no-load with .25% expense ratio.
2. Fund B is no-load with 1.25% expense ratio.
3. Fund C has a 6% load on purchases and a .8% expense ratio.
4. Gross return on all funds is 10% per year before expenses.

Each investor must choose the best combination of fees. Obviously, pure no-load no- fee funds distributed directly by the mutual fund group are the cheapest alternative. However, many investors are willing to pay for financial advice, and the commissions paid to advisers who sell these funds are the most common form of payment.

In addition: mutual fund income is taxed.
While being geography-specific, there are two distributions to shareholders that can be taxed, often at a differential rate:

- ordinary dividends (taxed at 30% in BE)
- capital gains (not taxed in BE but we have an exchange tax TOB that does tax capital gain, when you enter or exit you have to pay a tax.)

A differential tax treatment also explains the existence of distribution funds vs accumulation funds:

- *Accumulation funds*: do not pay out dividends but reinvests all the income such that it only generates capital gains
- *Distribution funds*: pay out dividends

Taxation on capital gains is much lower as compared to dividends (in BE). This taxation schemes are very different:

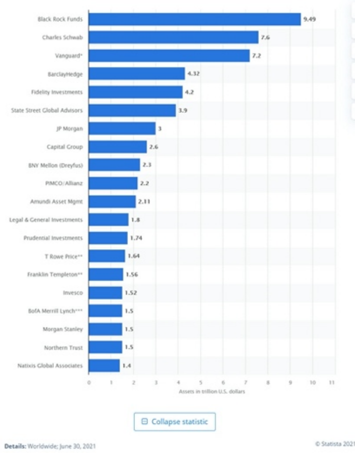
- Lower taxation schemes
- International taxation schemes: when you invest abroad this might be taxed differently.

→ As a retail investor it is difficult to keep track of all of this. This is where professional management comes in. They should be able to optimize this, because these matters allot. It's net returns that matter!

Leading companies?
It is hard to keep up with the biggest firms,we have 3 big firms:

- Black Rock
- Charles Schwab
- Vanguard

They are the leader in the market.



Investment policies

Mutual funds come in many different investment policies, described in the prospectus:

Equity funds primarily invest in stocks and commonly hold 4% to 5% of TA in money market instruments (needed to meet potential redemption of shares).

Different investment strategies, typically subdivided along size dimension (small-mid-large) and capital appreciation dimension (income/value-growth):

- *Income funds* tend to hold shares of firms with consistently high dividend yields.
- *Growth funds* are willing to forgo current income, focusing instead on prospects for capital gains. Growth stocks, and therefore growth funds, are typically riskier and respond more dramatically to changes in economic conditions than do income funds.

and with different sectoral focus (sector funds) and geographical focus (domestic, global, international, regional or emerging market funds)

Bond funds primarily invest in fixed income securities; different investment strategies, typically subdivided along issuer dimension (corporate, government, municipal bond), along maturity dimension (short, intermediate and long-term), credit risk dimension (investment grade and junk bonds) and by income-backing dimension (ABS, MBS, CDO)

Money market funds invests in short-term (often < 1 month) high-quality fixed income instruments ('cash'). These funds invest in money market securities such as Treasury bills, commercial paper, repurchase agreements, or certificates of deposit.

Money market funds are classified as prime versus government:

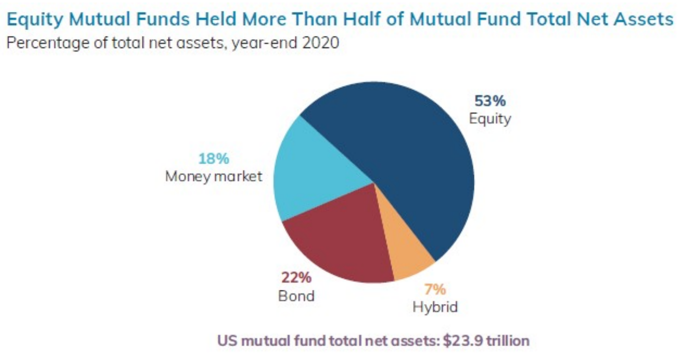
- *Government funds hold short-term U.S. Treasury or agency securities and repurchase agreements collateralized by such securities.*
- *Prime funds also hold other money market instruments such as commercial paper or bank CDs. While these assets are certainly at the very safe end of the credit-risk spectrum, they are riskier than government securities and are more prone to suffer reduced liquidity in times of market stress.*

Hybrid funds (asset allocation funds or balanced funds) invest in a mix of stocks and fixed income whereby capital appreciation of stocks can be combined and balanced with the stable income of bonds.

- *Balanced funds* hold a relatively stable mix of assets (often as funds of funds)
  These balanced funds hold both equities and fixed-income securities in relatively stable proportions.
  Many balanced funds are in fact **funds of funds.** These are mutual funds that primarily invest in shares of other mutual funds. Balanced funds of funds invest in equity and bond funds in proportions suited to their investment goals.

- *Asset allocation funds* vary the mix considerably (and thus engage in market timing and are not designed to be low-risk investment vehicles)
  These funds are similar to balanced funds in that they hold both stocks and bonds. However, asset allocation funds may dramatically vary the proportions allocated to each market in accord with the portfolio manager's forecast of the relative performance of each sector.

Index funds track the performance of a broad market index either by replicating the index, or selecting a representative sample of the index; they are passively managed funds and thus low cost. Investment in an index fund is a low-cost way for small investors to pursue a passive investment strategy—that is, to invest without engaging in security analysis.



Note: funds need to communicate about

- the fund's investment policies (objectives and strategies)
- risks of investing in the fund
- past performance
- distribution policy
- fees and expenses
- fund manager

via a **prospectus**, this is a legally binding document, but very lengthy...who reads the prospectus?
In EU, all UCITS investments funds need to publish a **Key Investor Information Document (KIID)** that provides retail investors with standardized and concise investment information that can easily be compared.

Next

Wide choice of investment funds...which to choose?

There are two considerations:

1. **Return**: How can I maximize the return (net of taxes and fees)
2. **Risk**: What level of risk is acceptable?

...and there is a trade-off between (expected) return and risk, so both dimensions are intertwined: we look for the investment that maximizes the expected return for a given level of risk
As will become clear, this discussion ultimately centers on **market efficiency** and **active versus passive management**

Basic principles of portfolio investment:

- **Capital allocation**: decision on how to divide one's wealth across a risk-free alternative and a risky alternative, driven by **risk aversion**
- **Asset allocation/security selection**: decision on how to select specific risky asset classes and securities within an asset class, driven by **diversification benefits** and **market (in-)efficiency**

Chapter 2: Investment concepts

BKM - Chapter 2 (2.4): Asset classes and financial instruments
BKM - Chapter 3 (3.8-3.9): How securities are traded
BKM - Chapter 5: Risk, return and the historical record
Oefeningen in handboek enkel degene waar zij een oplossing voor heeft voorzien!

Characterization of an investment

Investments are commonly characterized in **two dimensions**:

1. measure of **return**: EAR, APR, average return (arithmetic or geometric, IRR), HPR,...
2. measure of **risk**: volatility, VaR, ES, LPSD,... not each investment has the same amount of risk

→ Return and risk needs to be measured
→ When I want to evaluate to return, I will measure the actual return I made.
→ Risk is much more like a downside concept or a symmetric concept: it's the disposure around a mean. Having returns higher than what you expected is also volatility, but it is not intuitively risk. I am exposed to risk when I make return lager then expected, that's why we look at the downside of returns.

While no consensus over which measures are best, one should be **consistent** when comparing:

- across investments
- over time
  → you have to compute risk and return the similar way over time
  → different time horizons: 1 year vs 20 year → return concepts needs to be comparable
  →***Remark***: *if I do not add the investment horizon, this is annum basis = most common one (prof)*

Appendix

(als je dit niet goed begrijpt bekijk dan de paars gefluoriseerde stukjes in HB)

Measuring returns

- A return is a measure of how well a security performs as an investment over a **particular horizon**
- Return: $R_t = \frac{D_t + P_t - P_{t-1}}{P_{t-1}} = \underbrace{\frac{D_t}{P_{t-1}}}_{\text{income yield}} + \underbrace{\frac{P_t - P_{t-1}}{P_{t-1}}}_{\text{capital gain}}$

- This definition treats the dividend as paid at the end of the holding period
- If dividends are received earlier, this definition ignores reinvestment income

Different holding periods

Longer horizons provide greater returns...
But: for reasons of comparability, one should compute returns over the **same holding period** instead of total returns over the horizon

Example: total return on ZCB with differing maturities

Horizon T (in years)	P_0	P_T	$R_{0,T}$
0.5	97.36	100	0.0271
1	95.52	100	0.0469
25	23.30	100	3.2918

Source: BKM (2021) p. 120

Convention: compute the **effective annual rate (EAR)**

This is the percentage increase in funds invested over a 1-year horizon (computed as a compound return). *This measure allows to compare returns on investments with different horizons. We re-express each total return as a rate of return over a common period = the percentage increase in funds per year. These also account for compound interest.*

$$1 + \text{EAR} = (1 + R_{0,t})^{1/T}$$

With T = investment horizon in years

Example: EAR on ZCB with differing maturities

For the 0.5 year instrument: $1 + EAR = (1 + 0.0271)^2 = 1.0549$
For the 1 year instrument: $1 + EAR = (1 + 0.0469)^1 = 1.0469$
For the 25 year instrument: $1 + EAR = (1 + 3.2918)^{\frac{1}{25}} = 1.06$

Example: EAR on a 10-year investment
What annual return doubles your money in 10 years?

Horizon T (in years)	P_0	P_T	$R_{0,T}$
10	100	200	1.00

For the 10 year investment: $EAR = (1 + 1)^{\frac{1}{10}} - 1 = 7.18\%$

Annual percentage rates
For short-term instruments (T<1), annualized rates are often expressed using simple interest interest rates that ignores compounding = annual percentage rates (APR):

$$APR = \frac{1}{T}\; x\; R_{0,T} = n\; x\; R_{0,T}$$

with n = number of compounding periods per year

Example: For the 0.5 year instrument: $APR = 2 \times 0.0271 = 0.0542$

Relation between (multiperiod) returns, EAR and APR:

$$1 + \text{EAR} = (1 + R_{0,t})^{1/T} = (1 + T\; x\; APR)^{1/T} \quad \text{(eq 5.2)}$$

Annual percentage rates and effective annual rates

Compounding Period	T	EAR = 0.058 $r(T)$	EAR = 0.058 $APR = n[(1 + EAR)^{1/n} - 1]$	APR = 0.058 $r(T)$	APR = 0.058 $EAR = (1 + APR/n)^n - 1$
1 year	1.0000	0.0580	0.05800	0.0580	0.05800
6 months	0.5000	0.0286	0.05718	0.0290	0.05884
1 quarter	0.2500	0.0142	0.05678	0.0145	0.05927
1 month	0.0833	0.0047	0.05651	0.0048	0.05957
1 week	0.0192	0.0011	0.05641	0.0011	0.05968
1 day	0.0027	0.0002	0.05638	0.0002	0.05971
Continuous			$r_{cc} = \ln(1 + EAR) = 0.05638$		$EAR = \exp(r_{cc}) - 1 = 0.05971$

Table 5.2

Annual percentage rates (APR) and effective annual rates (EAR). In the first set of columns, we hold the equivalent annual rate (EAR) fixed at 5.8% and find APR for each holding period. In the second set of columns, we hold APR fixed at 5.8% and solve for EAR.

excel
Please visit us at
www.mhhe.com/Bodie12e

The difference between EAR and APR grows with the frequency of compounding. How far will these two rates diverge as the compounding frequency continues to grow? As n gets ever larger in Equation 5.2, we effectively approach *continuous compounding (CC),* and the relation of EAR to the annual percentage rate, denoted by r_{cc} for the continuously compounded case, is given by the exponential function:

Continuous compounded rates
What happens in the limit case?

- In the limit for $T \rightarrow 0$ (or $n \rightarrow \infty$) we approach continuous compounding:

$$1 + EAR = \lim_{T \rightarrow 0}(1 + T\ x\ APR)^{1/T} = \exp\left(r_{cc}\right)$$

- This continuous compounded rate can be computed as:

$$\ln(1+\text{EAR}) = r_{cc}$$

Logreturns
In finance we commonly use logreturns instead of returns where the lograte is a continuously compounded rate

A log return is defined as:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln\left(\frac{P_t - P_{t-1}}{P_{t-1}} + 1\right) \approx \frac{P_t - P_{t-1}}{P_{t-1}}$$

where the last proxy is valid for a return computed **over a short period**

There are some **advantages** to the use of logreturns

- It gives an additive relation for compound returns (instead of a multiplicative relation)
- It is consistent with limited liability

Statistical properties of returns

A major challenge in finance is to model the randomness/uncertainty in returns; to this end, make use of **probability theory**

One can proceed along two lines:

- Impose a parametric distribution
- Impose an empirical/historical distribution

There is a crucial **trade-off** in this choice: capture the stylized facts as good as possible versus distributional simplicity

Empirical key features
Based on historical (time-series) data, we can try to infer the probability distributions from which the returns are drawn

While fitting a full distribution is extremely hard, we start with expected returns (first moment) and standard deviations (second moment)

To obtain an estimate of expected returns we can compute the arithmetic average:

$$E(R) = \hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t$$

with an historical sample of T observations, where each observation is equally likely
If the time series of historical returns fairly represents the true underlying probability distribution, then the arithmetic average return from a historical period provides a reasonable forecast of the investment's expected future return.

Arithmetic mean vs geometric mean
While an arithmetic average is the unbiased forecast of the expected return, it does not capture the actual performance over a past period

Actual performance is measured by the geometric returns or time-weighted returns:

$$R_g = \left(\prod_{t=1}^{T} (1 + R_t) \right)^{\frac{1}{T}} - 1$$

The greater the volatility in rates of return, the greater the difference between arithmetic and geometric averages.

Volatility as a measure of risk
To obtain an unbiased estimate of the likelihood of deviations from the expected return we can compute the variance of returns:

$$Var\,(R) = \hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^{T} (R_t - \hat{\mu})^2$$

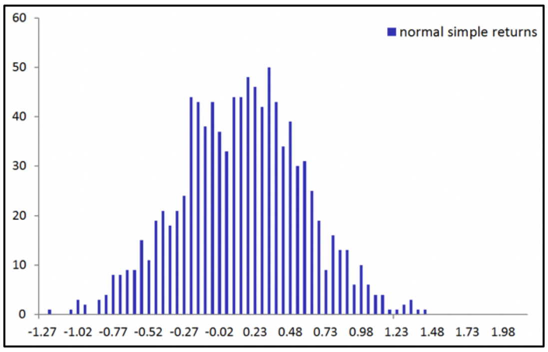
In finance, the standard deviation or volatility is a measure of risk: it is a measure of uncertainty of the outcome

While commonly used, standard deviation has been heavily criticized as a measure of risk
- risk is typically experienced as a loss (downside deviation)
- upside deviation will be interpreted as 'potential'
- only as long as the underlying distribution is more or less symmetric, standard deviation is a reasonable measure of risk

Note: Do more frequent observations lead to more accurate estimates? The answer to this question is surprising: Observation frequency has no impact on the accuracy ***of estimates of expected return****. It is the duration of a sample time series (as opposed to the number of observations) that improves accuracy.*
In contrast to the mean, the accuracy ***of estimates of the standard deviation*** *can be made more precise by increasing the number of observations. This is because the more frequent observations give us more information about the distribution of deviations from the average. Thus, we can improve the accuracy of estimates of SD by using more frequent observations.*

The most common assumption in finance: identically, independent, normal returns

$$R_t \sim IIDN(\mu, \sigma^2)$$



The simulated returns correspond to annual returns with an expected return of 10% and a volatility of 45%

Advantages

- It is easy to handle: two moments describe the full distribution of return (with volatility as a summary of 'risk')
- It is stable under addition: a portfolio of normally distributed components is also normally distributed → modelling co-movement can easily be done with a (Pearson) correlation
- Empirically, the normal distribution is a rough proxy of many financial variables
  - big frequency of small changes
  - small frequency of big changes

Disadvantages

- It is not stable under multiplication: if one period returns are normally distributed, multi-period compound returns are not normally distributed:

$$\text{For } R_t \sim N : R_{0,T} = \prod_{t=0}^{T} (1 + R_t) - 1 \not\sim N$$

- It violates limited liability: the range of the normal distribution is −∞, +∞, while returns are bounded at -100% (no negative prices)

  Note: we can overcome this in practice by simulating high frequency data such that the standard deviation is small
- Empirically: the normal distribution is not able to capture the observed skewness and kurtosis

Normal log returns zie hieronder

Deviations from normality

As we noted earlier normality of excess returns hugely simplifies portfolio selection. Normality assures us that standard deviation is a complete measure of risk and, hence, the Sharpe ratio is a complete measure of portfolio performance. Unfortunately, deviations from normality of asset returns are potentially significant and dangerous to ignore.

Normal and lognormal model fail at incorporating empirical observation of gain/loss asymmetry (skewness) and heavy tails (kurtosis)

Skewness (minor problem): third moment

What if large negative returns are more likely than large positive ones?

- negative skewness for stock indices (long left tail): negative values of skew indicate that extreme *bad* outcomes are more frequent than extreme positive ones
  - The SD will underestimate risk
- zero/positive skewness for individual stocks (long right tail): extreme positive outcomes dominate
  - The SD will overestimate risk

	μ	σ	S	excess K	Min	Max
			Panel A: Daily returns			
CRSP index (value weighted)	0.044	0.82	-1.33	34.92	-18.80	8.87
CRSP index (equal weighted)	0.073	0.76	-0.93	26.03	-14.19	9.83
IBM	0.039	1.42	-0.18	12.48	-22.96	11.72
General Signal corp.	0.054	1.66	0.01	3.35	-13.46	9.43
Wrigley Co.	0.072	1.45	-0.00	11.03	-18.67	11.89
Interlake Corp.	0.043	2.16	0.72	12.35	-17.24	23.08
Raytech Corp.	0.050	3.39	2.25	59.40	-57.90	75.00
Ampco-Pittsburgh Corp.	0.053	2.41	0.66	5.02	-19.05	19.18
Energen Corp.	0.054	1.41	0.27	5.91	-12.82	11.11
General Host Corp.	0.070	2.79	0.74	6.18	-23.53	22.92
Garan Inc.	0.079	2.35	0.72	7.13	-16.67	19.07
Continental Materials Corp.	0.143	5.24	0.93	6.49	-26.92	50.00

Source: Campbell, Lo, and MacKinlay (1998) p.21

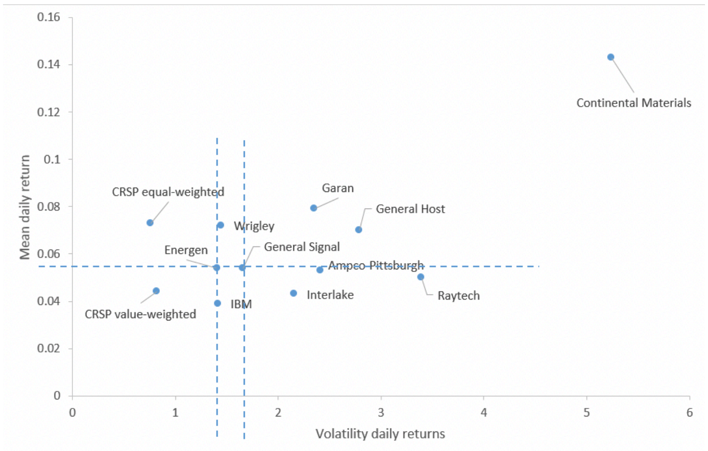
Kurtosis/peakedness (major problem): fourth moment

This concerns the likelihood of extreme values on *either* side of the mean at the expense of a smaller likelihood of moderate deviations.

High kurtosis means that there is more probability mass in the tails of the distribution than predicted by the normal distribution. That extra probability is taken at the expense of that there is less probability mass near the center of the distribution.

- very large positive kurtosis for stock indices
- large positive kurtosis for individual stocks (large variation)

Volatility is just one dimension of risk…



Volatility as a measure of risk: Energen seems less risky than General Signal Corp.

Excess skewness (likelihood of big losses) as measure of risk: General Signal Corps is less risky than Energen

Deviations from normality depend on the horizon over which returns are computed

	μ	σ	S	excess K	Min	Max
Panel B: Monthly returns						
CRSP index (value weighted)	0.96	4.33	-0.29	2.42	-21.81	16.51
CRSP index (equal weighted)	1.25	5.77	0.07	4.14	-26.80	33.17
IBM	0.81	6.18	-0.14	0.83	-26.19	18.95
General Signal Corp.	1.17	8.19	-0.02	1.87	-36.77	29.73
Wrigley Co.	1.51	6.68	0.30	1.31	-20.26	29.72
Interlake Corp.	0.86	9.38	0.67	4.09	-30.28	54.84
Raytech Corp.	0.83	14.88	2.73	22.70	-45.65	142.11
Ampco-Pittsburgh Corp.	1.06	10.64	0.77	2.04	-36.08	46.94
Energen Corp.	1.10	5.75	1.47	12.47	-24.61	48.36
General Host Corp.	1.33	11.67	0.35	1.11	-38.05	42.86
Garan Inc.	1.64	11.30	0.76	2.30	-35.48	51.60
Continental Materials Corp.	1.64	17.76	1.13	3.33	-58.09	84.78

Source: Campbell, Lo, and MacKinlay (1998) p.21

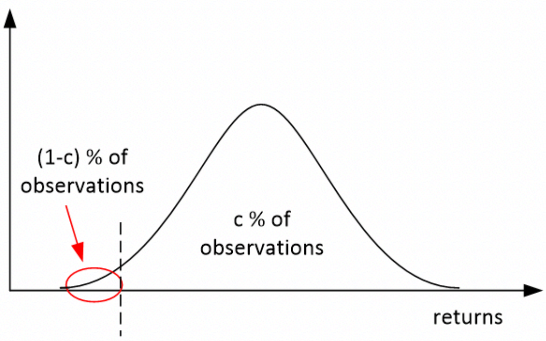
Alternative measures of risk

Risk as measured by standard deviation is only reasonable when returns are normally distributed

Alternative measures of risk exist that do capture extreme negative returns

- Value-at-risk (VaR)
- Expected shortfall (ES) also labelled conditional value-at-risk
- Lower partial standard deviation
- Frequency of extreme (3-sigma) events

Value-at-risk (VaR)
Value-at-risk is the loss corresponding to a very low percentile of the return (or value)) distribution Therefore, it is another name for the *quantile* of a distribution. The quantile, *q,* of a distribution is the value below which lie *q*% of the possible values. Thus the median is *q* = 50th quantile. Practitioners commonly estimate the 1% VaR, meaning that 99% of returns will exceed the VaR, and 1% of returns will be worse. Therefore, the 1% VaR may be viewed as the cut-off separating the 1% worst-case future scenarios from the rest of the distribution.



Expected shortfall

- VaR has the disadvantage of not being binding: what happens if VaR is breached? In other words, it tells you the investment loss at the first percentile of the return distribution, but it ignores the magnitudes of potential losses even further out in the tail.
- As an alternative to VaR: **expected shortfall (ES)** computes the expected loss to the left of the VaR (also labelled conditional value at risk) = the *expected* loss given that we find ourselves in one of the worst-case scenarios
- ES is much more conservative downside risk measure than VaR → it gives a better idea of **what is at stake**

Lower partial standard deviation (LPSD)

Standard deviation as a measure of risk suffers from two major problems

1. By focussing on both positive and negative returns it assumes symmetry
   → since return distributions are not symmetric, we should only focus on the downside returns to get a grasp of risk
2. By focussing on deviations from the sample average, the comparison is rather ad hoc
   → as an alternative we can use the risk-free rate as a benchmark (this focusses on negative excess returns)

Lower partial standard deviation is then calculated in the standard way by squaring the negative deviations (rather than negative deviations from the sample average) from the risk-free rate and then takes the square root to obtain a “left-tail standard deviation”.

The **disadvantage** of this measure is that we exclude positive excess returns: we thus throw away (valuable) information → the frequency of negative excess returns is ignored, that is, portfolios with the same average squared negative excess returns will yield the same LPSD regardless of the relative frequency of negative excess returns.

Practitioners who replace standard deviation with this LPSD typically also replace the Sharpe ratio (the ratio of average excess return to standard deviation) with the ratio of average excess returns to LPSD. This variant on the Sharpe ratio is called the **Sortino ratio.** *(zie hieronder)*

Relative frequency of large negative 3-sigma events

- Fraction of large negative 3-sigma events are compared with the fraction as observed in the normal distribution: Here we concentrate on the relative frequency of large, negative returns compared with those frequencies in a normal distribution with the same mean and standard deviation.
- Such big changes are often referred to as ’jumps’
- We compare the fraction of observations with returns 3 or more standard deviations below the mean to the relative frequency of negative 3-sigma returns in the corresponding normal distribution.
- While informative about downside risk, the measure has the drawback of statistical significance
    - the fraction of jumps is limited (under normality we can expect to observe a jump in 0.13%)
    - the use of this measure if therefore limited to large high-frequency samples (and large samples)

Risk-adjusted returns

The two-dimensional focus is crucial as there is a trade-off between (expected) returns (on average) and risk. On average higher return comes with greater risk. It’s not always the case on short periods but on average it is the case.

Off course, you want to assess exactly how much reward is expected and/or earned for the risk involved = **risk-adjusted returns**

You need to combine those 2 concepts, because returns come with risk. That means that you should try to come up with risk adjusted returns: you can combine the two dimensions in 1 concept.

Two common approaches:

- group investments into a comparison universe with similar risk characteristics: subclasses of risk with similar level of risk and for each of the subclasses you compare the returns
- compute a reward-to-risk measure: where you scale rewards with the level of risk you are exposed to, for example the sharp ratio

Note: since the first part of the course is on portfolio selection, and hence ex ante investment decision making, we now focus on ex ante measures. In a final chapter on performance evaluation, we introduce ex post measures.

Reward-to-risk

The most widespread reward-to-risk measure is the **Sharpe ratio**:

$$\mathrm{SR} = \frac{E(R) - R_F}{\sigma} = \frac{E(R^e)}{\sigma}$$

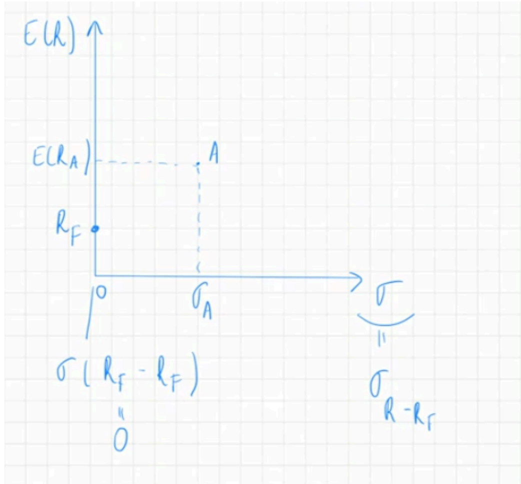
with σ the standard deviation of excess returns $R^e = R - R_F$ and $E\,(R^e)$ = **risk premium**
= ex ante concept
= the difference between the *expected excess* return on the asset and the **risk-free rate**

This Sharpe ratio is widely used in practice to evaluate the attractiveness of investments: the higher the Sharpe, the more reward per unit of risk.

The higher the sharp, the more attractive your investment will be because higher returns per unit of risk expected excess return

- Volatility (σ): volatility of **excess** returns (not just returns)
    - To be fully correct you should calculate this on excess return = the difference in any particular period between the *actual* rate of return on a risky asset and the risk-free rate.
    - You might think there is no risk in risk free rates, but there is because the type of risk you have in the Risk free rate depends on the time horizon and it depends on the type of instrument.
        - If I have to roll over my rf instrument. For example a T bill has a duration of 1 year. So, we have a roll over risk because we don't know what the interest rate will be over 1 year. There is some time variation on T bills.
        - Or imagine you have a T bill of 1 year but you have an investment horizon of 6 months. Then you have interest rate risk because I will have to sell my T bill after 6 months, today the rate is uncertain.
    - So: If your investment horizon is not aligned to your rf instrument you will have risk; **time variation.**
    - So: if you look at excess return you fully incorporate the possibility of time variation. This is more accurate.
    - If you plot 2-dimensional graph of return vs risk:

Therefore, the risk premium is the expected value of the excess return, and the standard deviation of the excess return is a measure of its risk.



Important: Sharpe ratios measured over different investment horizons cannot be compared!
If you scale investments over time, your SR will be mechanically impacted.

We can prove this for log returns IID normal; to work with log returns and impose a normal distribution. We impose identically and independently distributed. Then you can easily prove that you cannot compare this overtime:

So, a popular alternative is to assume IIDN logreturns:

$$r_t \sim IIDN\ (\mu, \sigma^2)$$

which implies that returns are IID lognormal

Advantages

- Limited liability is not violated as $Pt = P_{t-1}$ exp (r_t)
  Stability under addition makes that we can easily aggregate logreturns
- over time, preserving normality:

$$r_{0,T} = \ln(1 + R_{0,T}) = \ln\left(\prod_{t=0}^{T}(1 + R_t)\right) = \sum_{t=1}^{T}\ln(1 + R_t) = \sum_{t=1}^{T} r_t$$

$$\text{For } r_t \sim N : r_{0,T} = \sum_{t=1}^{T} r_t \sim N$$

- Log returns are easier to work with
- If you have a **log** return measured over multiple periods, then it is equal to executive of sums of single period log periods
- If returns are normally distributed, then multiple periods returns are NOT normally distributed
- If r_1 and r_2 are the returns in two periods, and each has the same normal distribution, then the *sum* of the returns, $r_1 + r_2$, would be normal. But the two-period compound return is not the sum of the two returns. Instead, invested funds would compound to $(1 + r_1)(1 + r_2)$, which is *not* normal.
- But if LOG returns are normally distributed then multiple periods LOG returns ARE normally distributed

Advantages - continued

Combining the stability property with the IID property also implies:

$$E\left(r_{0,T}\right) = T \times E\left(r_t\right)$$
$$Var\left(r_{0,T}\right) = T \times Var\left(r_t\right)$$
$$SD(r_{0,T}) = \sqrt{Var(r_{0,T})} = \sqrt{T} \times \sigma(r_t)$$

Proof:
Identical implies: $E(r_j) = E(r_i) = E(r_t)$ and $Var(r_j) = Var(r_i) = Var(r_t)$
Independence implies: $Cov(r_j, r_i) = 0$ (if we do not assume independence then you will have autocorrelation because of the covariance, correlation between periods)

$$E(r_{0,T}) = E\left(\sum_{t=1}^{T} r_t\right) = \sum_{t=1}^{T} E(r_t) \overset{identical}{=} T \times E(r_t)$$
$$Var(r_{0,T}) = Var\left(\sum_{t=1}^{T} r_t\right) \overset{independence}{=} \sum_{t=1}^{T} Var(r_t) \overset{identical}{=} T \times Var(r_t)$$

- $T/\sqrt{T} \times SR_t$ → SR of multiple periods will be larger than the SR of a single period
    - Log returns are IIDN; in context of SR you see that you cannot compare SR over time

Disadvantages

Empirically the observed skewness and kurtosis are still problematic
→ the shorter the horizon, the more problematic skewness (to a minor extent), and kurtosis (to a large extent) (see infra)

An alternative reward-to-risk ratio is the **Sortino ratio**

$$SR = \frac{E(R) - R_F}{LPSD}$$

with LPSD = lower partial standard deviation, i.e., standard deviation of returns lower than a chosen benchmark (e.g. risk-free rate). Volatility for a subset of returns, only those that are lower than a particular benchmark level. If you are interested in returns which are lower than 0 then you choose a benchmark. You can focus on the more intuitively version of risk.
Risk is captured in a different way: asymmetric version

It can be understood as the asymmetric version of the Sharpe ratio.

Portfolios and indices

A portfolio is an investment in multiple assets
- The return on a portfolio is a weighted average of the returns on the individual assets in the portfolio
- The weights w_n represent the fractions of the value of the portfolio p that is invested in each asset n:

$$R_{p,t} = \sum_{n=1}^{N} w_n R_{n,t}$$

Portfolio weights are most often positive but can also be negative.

Weighted average: weight represents the fractions of importance. Portfolio weights are most often positive but can also be negative. When do we observe negative weights?
- When you are allowed to short selling
- Leverage: we also take on the negative weight not only on the risky asset but also on the risk-free asset

Negative portfolio weights

Negative portfolio weights occur in two contexts:

1. **Leverage**: a negative weight on a money market instrument
    - Here you have a positive view
    - borrow money now, invest in a risky asset and commit to pay back the loan at a future moment = that is why it is a negative weight
    - purpose: leverage your position to profit more (from a larger investment) from an increase in the price of the risky asset. So, there is a particular risky asset from which I am positive about and I want to invest more than I have. I invest for example 10k + what I borrowed.
2. **Short selling**: a negative weight on a risky asset
    - Here you are very negative about an asset
    - borrow the asset now, sell the asset in the (spot) market, and commit to return the asset at a future moment (buy in future)
    - purpose: profit from a decrease in the price of the risky asset

Note: whether leverage and/or short selling is feasible, will impact the investment opportunity set!

→ The motive to take on the negative weight is very different.
→ These 2 expand your investment set of opportunities but in practice not everyone can do this. E.g., short selling is not for retail sellers.

Negative portfolio weights: leverage (buying on margin)

Leverage in practice:

Goal: borrow money from a broker to invest in a risky asset. Here you expand the cash you have to invest, so you have more money that you can invest. So, they buy on margin when they wish to invest an amount greater than their own money allows.

The loans are linked to call loans. These is what banks give to their clients to buy a margin. The brokers in turn borrow money from banks at the call money rate to finance these purchases; they then charge their clients that, plus a service charge.

- The portion of the purchased asset value contributed by the investor is the **margin**; the remainder is borrowed from the broker. *You are not allowed to borrow the whole amount; you have to invest a portion yourself.*
- The securities are collateral
- Lower bounds are set on **initial margin** e.g. 50% in the US as defined by Reg T. this means that minimum 50% needs to be invested by you.
- When the margin falls below a **maintenance margin** (e.g. 25% of market value of the assets), meaning the value of the stock is no longer sufficient collateral to cover the loan from the broker. Then additional margin needs to be posted, the broker will issue a *margin call,* which requires the investor to add new cash or securities to the margin account.
    - In case the investor does not act, the broker reduces the position to restore the margin, the broker may sell securities from the account to pay off enough of the loan to restore the percentage margin to an acceptable level.

  So, when the asset value of your portfolio drops, the maintenance margin changes. When it is too low you will be asked to pass additional margin to the broker.
- The broker charges a rate on the borrowed money equal to the call money rate plus a service charge for the loan

Such investment strategy is known as **buying on margin** (investors have easy access to a source of debt financing called *broker's call loans).* Buying on margin amplifies your exposure to fluctuations in risky asset returns (both gains and losses).

Example:
Suppose that an investor wants to buy 100 shares of $100 each. He pays for $6,000 in cash and borrows the remaining $4,000 from his broker. The maintenance margin equals 30%. What is the balance sheet of this transaction and what is the initial margin?

Assets		Liabilities and owner's equity	
Share value	$10,000	Loan from broker	$4,000
		Equity	$6,000

The initial margin (in %) equals:

$$\frac{\$6,000}{\$10,000} = 60\%$$

Assume now that the share price increases to $140. What is the new balance sheet and margin? Is the maintenance margin satisfied? What is the return earned? How does this differ from the return earned when the purchase was fully funded by the investor?

Assets		Liabilities and owner's equity	
Share value	$14,000	Loan from broker	$4,000
		Equity	$10,000

The margin (in %) equals:

$$\frac{\$10,000}{\$14,000} = 71\% > 30\%$$

The return equals:

$$R = \frac{\$14,000 - \$4,000}{\$6,000} - 1 = 67\%$$

Had the purchase been fully funded by the investor, the return equalled:

$$R = \frac{\$14,000}{\$10,000} - 1 = 40\%$$

You made a profit, that's why you leveraged

You see that you can increase the investment value, if you did not borrow the money your return would be only 40%. Thus, they can achieve greater upside potential, but they also expose themselves to greater downside risk. There is also a downturn! This indicates the riskiness of such a leveraged position you will always be stuck with the loan position you took.

Assume now that the share price declines to $70. What is the new balance sheet and margin? Is the maintenance margin satisfied? What is the return earned? How does this differ from the return earned when the purchase was fully funded by the investor?

Assets		Liabilities and owner's equity	
Share value	$7,000	Loan from broker	$4,000
		Equity	$3,000

The margin (in %) equals:

$$\frac{\$3,000}{\$7,000} = 43\% > 30\%$$

The return equals:

$$R = \frac{\$7,000 - \$4,000}{\$6,000} - 1 = -50\%$$

Had the purchase been fully funded by the investor, the return equalled:

$$R = \frac{\$7,000}{\$10,000} - 1 = -30\%$$

How far could the stock price fall before getting a margin call?

The maintenance margin (in %) should be at least 30%:

$$30\% = \frac{\text{Equity}}{\text{Share value}} = \frac{\text{Share value-Loan}}{\text{Share value}}$$

$$30\% = \frac{\text{Share value-\$4,000}}{\text{Share value}}$$

$$\rightarrow \quad \text{Share value} = \$5,714.29$$

$$\rightarrow \quad \text{Stock price} = \frac{\$5,714.29}{100} = \$57.14$$

57 dollar is really the lower bound and you want to avoid that. It comes in handy to first calculate this before you take up a leveraged position.

Negative portfolio weights: short selling

Short selling in practice:

- Goal: borrow risky assets from a broker to sell on the spot market (can be called at any time!).

A **short sale** allows investors to profit from a decline in a security's price. An investor borrows a share of stock from a broker and sells it. Later, the short-seller must purchase a share of the same stock in order to replace the one that was borrowed. This is called *covering the short position*

- The short-seller anticipates the stock price will fall, so that the share can be purchased later at a lower price than it initially sold for; if so, the short-seller will reap a profit.
- Deposit the proceeds of the short sale in a margin account (The short-seller cannot invest these funds to generate income)
- Post **margin** (cash or collateral) to cover for losses should the risky asset price increase.
- Later: buy back the risky asset and return to the broker, along with interim income (any proceeds that that asset had during the period, like dividends…)
- Profit in case the risky asset's price has decreased while one is short

Order of buying and selling is reversed in a **short sale**

Like investors who purchase stock on margin, a short-seller must be concerned about margin calls. If the stock price rises, the margin in the account will fall; if margin falls to the maintenance level, the short-seller will receive a margin call.

What is the cost of short selling transaction? You have to pay for this, your broker will not give it for free. You have to deposit the proceeds of the short sales on a margin account. Not all the interest on a margin account is for you, so the net return will be lower than on a margin account.

Example

Suppose dot Bomb shares are currently priced at $100. As you are pessimistic on them, you tell your broker to short 1,000 shares. Assume that the broker applies a margin requirement of 50%, and that you own $50,000 in Treasury bills that you can use as collateral. How does your account with the broker looks like after this short sale?

Assets		Liabilities and owner's equity	
Cash	$100,000	Short position in dot Bomb	$100,000
Treasury bills	$50,000	Equity	$50,000

Assume you are right about Dot Bomb and that the share price drops to $70. What is the profit of this short sale if you close out your position at this share price?

Assets		Liabilities and owner's equity	
Cash	$100,000	Short position in dot Bomb	$70,000
Treasury bills	$50,000	Equity	$80,000

You can now close the position at a profit. You buy 1,000 shares at a total price of $70,000, to return the ones you have borrowed. Because your account was credited for $100,000 when the shares were borrowed and sold, your equity increases to $80,000, implying a profit of $30,000

So, in general portfolio rates are positive but in 2 cases it can be negative.

Short selling has a **negative connotation** and is occasionally banned by regulators:

- Following financial crisis: SEC restricted short sales in stocks that experienced large price drops in a single day
  - They restrict that in very turbulent times: as a regulator you don't want to intervene into the financial market.
    - Information will not pass: you prohibit the market to taking out its role on information processing
  - But on the other hand why they intervene is because short selling could really disrupt the market, and have an effect on market stability.
- Following sovereign debt crisis: France, Italy, Spain and Belgium banned short sales on number of financial stocks. Here the regulator intervened to prevent a downhill. We had a financial crisis, and than the failure of financial institutions is very dangerous.
- Following COVID-19 pandemic: Belgium banned short sales on instruments trading on Belgian trading venues

Motivation to ban short selling: protect against destructive speculation that destabilizes the market..
BUT: What about the informational role of financial markets?

There is no unique advice you can give to regulators when to intervene or not. The common ground is that regulators try to avoid bans, but when there is panic in the market, they hope to stabilize it by intervening.

Market indices

Indices are portfolios of instruments that are representative for a particular market. It could be limited to a country, an instrument, …

Such indices play a major role in financial markets:

- give insights into performance of a (sub-)market
- act as benchmark to evaluate investments (cf. KIID, in EU, all mutual funds have to publish a KID, and they will have to list a benchmark)
- can be bought as an investment product as a passive instrument (cf. index funds and ETFs; they mimic an index)

Traditional weighting schemes are: *(if you do not understand read p47 pages in HB)*

- price-weighted: asset with highest price carries most weight
- market-value-weighted: asset with highest market cap carries most weight

Market indices: price-weighted index

A price-weighted index corresponds to a portfolio in which each component of the index is weighted according to its current price (as if you hold a single asset of each component)

For example: DJIA (= 30 US blue-chips)

The **weight** of a component n in an index with N components is calculated as:

$$w_n = \frac{P_n}{P_1 + \cdots + P_N}$$

The **value** can, in theory, be calculated as an arithmetic average of the prices:

$$V = \frac{1}{N}\sum_{n=1}^{N} P_N$$

However, in practice one adjusts the value for events like stock-splits, dividends, mergers, index adjustments. The value of the index does not equal the arithmetic average. Where does the difference come from?

This is done by using a 'divisor' to compute an average such that the event leaves the index unaffected**:**

$$V = \frac{1}{d}\sum_{n=1}^{N} P_N$$

with d = the divisor

We do not calculate the arithmetic average by dividing N but we divide by a divisor.

- What is the purpose of the divisor?
    - Some corporate events will influence the price of your index.
    - Divisors are calculated so that the corporate events do not affect the price.

For example: the DJIA divisor is 0.1492 (October 2022) *implying that a change in price by $1 in one of the index components corresponds to a 6.70 point movement in the index*.

Market indices: market-value-weighted index

A market-value-weighted index corresponds to a portfolio in which each component of the index is weighted according to its market capitalization
For example: S&P500 (= 500 largest US companies)

The **weight** of a component n in an index with N components is calculated as:

$$w_n = \frac{MC_n}{MC_1 + \cdots + MC_N}$$

with MC_n = P_n × stocks outstanding

We are going to weight components according to their size. The index is, in practice, corrected for the free-float, by computing the market value of freely tradable shares. SO, we only calculate to the stocks that are freely traded in the market.

Contrary to the price-weighted index, a market-value-weighted index is unaffected by stock splits. (A corporate event will have no impact on the index with this calculation)

While the above weighting schemes can be easily criticized for overweighting particular firms (there will always be some companies that are favored; you give e.g., more weights to bigger companies), one major advantage is that they reflect the returns of obvious portfolio strategies, i.e. you can easily mimic such indices. The indices are important to have a replica strategy without too many transaction costs, you don't have to balance allot your portfolio.

That is also why existing index funds and ETFs often track such price- or market-value weighted index

You don't have the same with an equally weighted portfolio. Alternative weighting schemes, such as equal-weighting, do not immediately translate into a buy-and-hold strategy, but need constant rebalancing.

Finally note the index construction business is big business due to:

- legal requirement of mutual funds to use/publish benchmarks (If a mutual fund wants to list a particular index into its KID, they will have to pay. Its really big business and its expensive to list these. As a mutual fund its important.)
- popularity of index funds and ETFs tracking standard, but also tailor-made indices: Most of the ETFs are replicating tailor made indices and not standard indices. It's very expensive.

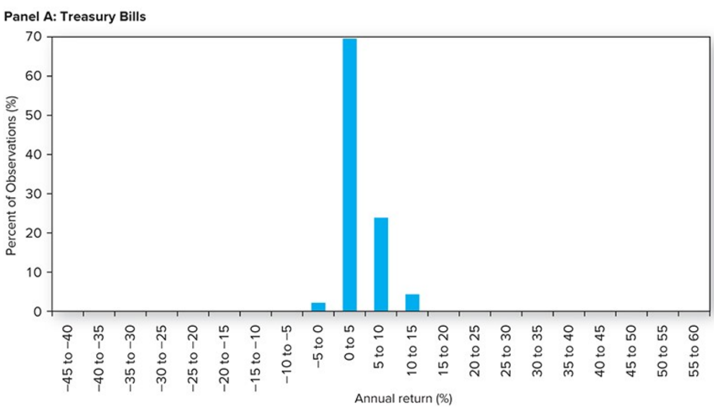
Finally, there are also bond market indices: several bond market indicators measure the performance of various categories of bonds. The major problem with bond market indexes is that rates of return on many bonds are difficult to compute because the infrequency with which the bonds trade makes reliable up-to-date prices difficult to obtain. In practice, some prices must be estimated from bond-valuation models. These "matrix" prices may differ from true market values.

An historic perspective on portfolio returns

US T-bills

If you think about investment, you should have numbers in your head as benchmark.
That will allow you if the numbers you observe are reasonable.
The reference will typically be on an annual basis

T-bills are widely considered the least risky of all assets. There is essentially no risk that the U.S. government will fail to honor its commitments to these investors, and their short maturities mean that their prices are relatively stable.

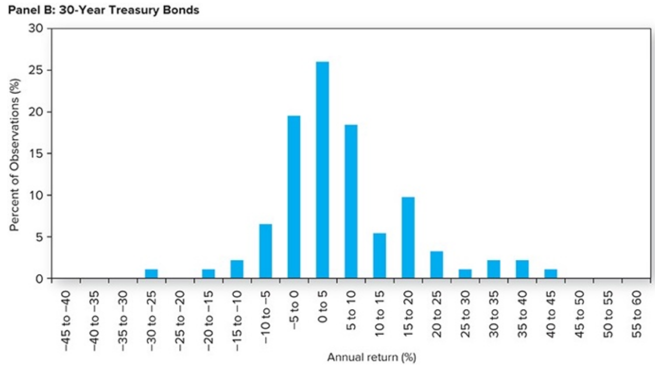


Here annual returns are very limited
The spread on returns over time is very low
- Returns are low
- Volatility is low

US T-bonds

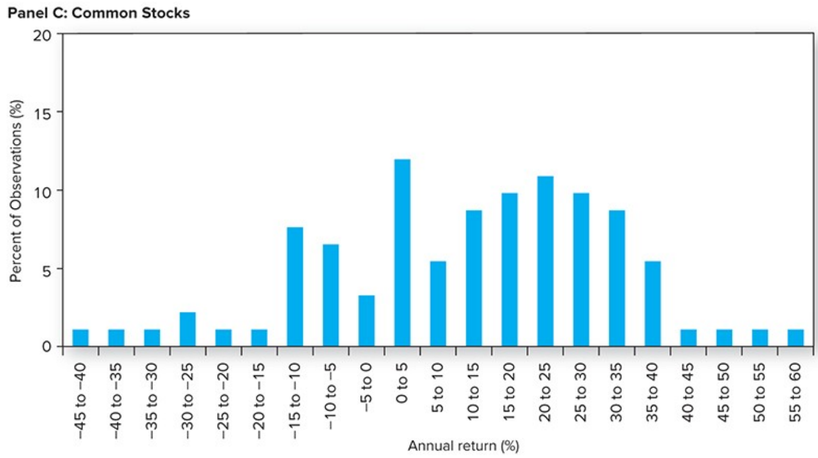
Treasury bonds are also certain to be repaid, but the prices of these bonds fluctuate as interest rates vary, so they impose meaningful risk

- Range is increased
- Returns earned could be higher or lower
- Bigger spread
- Higher returns and higher volatility



US common stocks

Finally, common stocks are the riskiest of the three groups of securities. As a part-owner of the corporation, your return will depend on the success or failure of the firm.



Similar analysis

- range is very wide
- volatility is large and returns are larger

An historic perspective on portfolio returns: comparison

Table 5.4

Risk and return of investments in major asset classes; estimates from annual data, 1927–2018

	T-Bills	T-Bonds	Stocks
Average	3.38%	5.58%	11.72%
Risk premium	N/A	2.45	8.34
Standard deviation	3.12	11.59	20.05
Max	14.71	41.68	57.35
Min	−0.02	−25.96	−44.04

Here you can see of the number you have are reasonable or not.
This table shows that the standard deviation of the return on stocks over this period, 20.05%, was nearly double that of T-bonds, 11.59%, and more than 6 times that of T-bills. Of course, that greater risk brought with it greater reward. The excess return on stocks (i.e., the return in excess of the T-bill rate) averaged 8.34% per year, providing a generous risk premium to equity investors.

	Market Index	Big/Growth	Big/Value	Small/Growth	Small/Value
A. 1927–2018					
Mean excess return (annualized)	8.29	8.07	11.69	8.99	15.38
Standard deviation (annualized)	18.52	18.35	24.70	26.06	28.21
Sharpe ratio	0.45	0.44	0.47	0.34	0.55
Lower partial SD (annualized)	21.68	21.10	25.44	28.95	26.18
Skew	0.19	−0.11	1.63	0.68	2.18
Kurtosis	7.85	5.63	18.43	7.85	22.32
VaR (1%) actual (monthly) returns	−13.61	−14.48	−19.40	−20.48	−20.57
VaR (1%) normal distribution	−11.79	−11.69	−15.69	−16.80	−17.78
% of monthly returns more than 3 SD below mean	0.94%	0.75%	0.85%	0.85%	0.57%
Expected shortfall (monthly)	−19.60	−19.80	−23.87	−24.67	−25.33
B. 1952–2018					
Mean excess return (annualized)	7.60	7.46	10.04	7.17	13.16
Standard deviation (annualized)	14.76	15.37	16.42	22.13	18.41
Sharpe ratio	0.52	0.49	0.61	0.32	0.71
Lower partial SD (annualized)	17.25	17.14	17.60	23.81	18.26
Skew	−0.54	−0.38	−0.32	−0.41	−0.34
Kurtosis	1.95	1.84	2.25	2.11	3.44
VaR (1%) actual (monthly) returns	−10.71	−10.94	−12.26	−16.96	−14.97
VaR (1%) normal distribution	−9.28	−9.70	−10.19	−14.26	−11.27
% of monthly returns more than 3 SD below mean	0.62%	0.66%	1.06%	0.93%	1.19%
Expected shortfall (monthly)	−18.85	−17.78	−21.16	−24.11	−24.45

Table 5.5

Statistics for monthly excess returns on the market index and four "style" portfolios

Source: Authors' calculations using data from Prof. Kenneth French's Web site: **http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.**

Here you can look at subsegments in styles:
Small or Large caps: average return is different between those two.

- Small caps have higher returns and risk will also be larger.

Value or Growth companies: average return is different between those two.

- Value companies will have higher returns then growth companies

There is no stability over time, return earned change over time. For example: the **first** full period, small value earned the highest excess return on average. This is also what we expected because small and value companies earned the highest. But they also have the highest volatility. But the question is also what does this mean in term of risk adjusted return? We can look at the sharp ratio. We see here that the highest SR is generated by those companies.

If we go to the **second** period things have changed. Still or small value companies earn highest average return but in terms of SD no longer have the largest volatility but the small growth have the largest volatility.

Chapter 3: Capital allocation to risky assets

BKM - Chapter 6: Capital allocation to risky assets

Excel assignment on portfolio selection covers lecture material in:

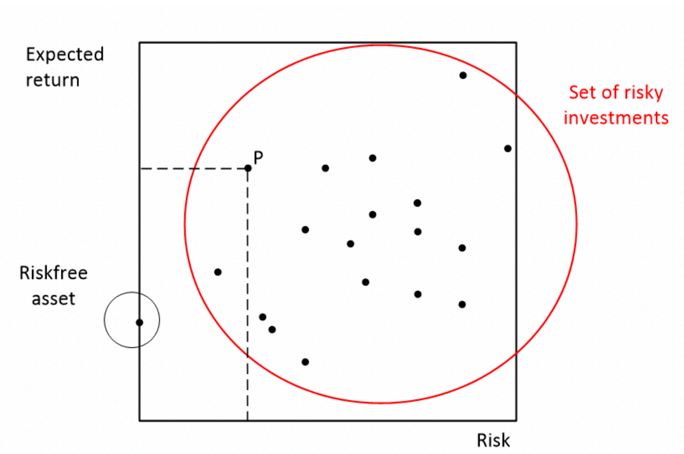
- *Capital allocation to risky assets*
- *Optimal risky portfolios*
- *Index models*

Practice your understanding and Excel skills by reviewing the practice videos on Toledo, and solving the Excel exercises

Process of portfolio construction involves two main steps:

1. **Choice between all risky investment opportunities**: in which instruments do you invest, and in which proportions? (Assume portfolio P); this is a technical aspect of asset allocation, driven by diversification benefits, and in an inefficient market, also by selection skills.
2. **Choice between this risky portfolio P and the risk-free investment**: which proportion of your wealth is invested in the risk-free asset? This choice is called the **capital allocation to risky assets.** This choice is driven by the investor's aversion to risk.
    - Outperformance: detect under or overpriced assets in the market

The drivers of these 2 decisions are different but we can do it in two steps.
We focus on the second choice this chapter.



We can just think about risk as volatility of **returns** (in contrary of what we did before): meaning that we assume that the risk-free asset has no volatility (ik denk dus dat we niet er van uit gaan dat de riskfree asset ook risico heeft)

Each dot is a potential risky opportunity with each his level of return and risk.

How will we combine P with risky assets and how much wealth to invest in P and how much in the risk-free asset?
→ this decision is the first control of risk: how much of your wealth to invest in the risky asset, its only in the second order that we decide in WHICH asset to invest.

First investment opportunity set then we can think about risks and risk aversion = optimal allocation

The optimal capital allocation will depend in part on the risk–return trade-off offered by the risky portfolio. But it will also depend on the investor's attitude toward risk, so we need a way to measure and describe risk aversion. Therefore, we will show how risk aversion can be characterized by a "utility function" that investors can use to rank portfolios with different expected returns and levels of risk. By choosing the overall portfolio with the highest utility score, investors optimize their trade-off between risk and return; that is, they achieve the optimal allocation of capital to risky versus risk-free assets.

Investment opportunity set

- Most straightforward way to **control risk** in a portfolio is with the **capital allocation**: reduce risk by increasing the allocation to the risk-free asset and investing less in the risky asset
  - *When we shift wealth from the risky portfolio to the risk-free asset, we do not change the relative proportions of the various risky assets within the risky portfolio. Rather, we reduce the relative weight of the risky portfolio as a whole in favor of risk-free assets.*
- Common risk-free asset: **government T-bills**
  - government can issue default-free bonds by its power to control money
  - supply
    short term nature makes them rather insensitive to interest rate changes or inflation
- In practice, investors use a broad range of money market instruments as risk-free asset (e.g. CD and CP) even though they yield higher returns compared to T-bills in stress periods (see earlier)

Portfolios of one risky and one risk-free asset

- Suppose there is a limited choice of instruments to invest in
  - a risk-free asset denoted ***F*** (e.g., T-bill)
  - a risky asset denoted ***P*** (e.g., portfolio of stocks and bonds)
- Let y denote the fraction of the investment budget to be allocated to the risky portfolio P
- The remainder $(1 - y)$ is invested in the T-bill F
- The T-bill yields a sure return of R_F, while the risky portfolio return equals R_P (known ex post), with an **expected rate of return** of $E(R_P)$ and a **level of risk** σ_P
- We want to look for the optimal Y.
- We have to make a decision today given the risk; we will form a best estimate of the riskiness of the instrument.

The complete portfolio, denoted C, has a return R_C as follows:

$$R_C = yR_P + (1-y)R_F$$
$$= R_F + y\,\underbrace{(R_P - R_F)}_{\text{excess return}}$$

What are the risk-expected return characteristics of this portfolio?

The expected portfolio return of this portfolio C is characterized as

$$E(R_C) = R_F + y\,(E(R_P) - R_F)$$

- the base rate of return is the risk-free rate
- the portfolio is expected to earn a proportion y of the risk premium of the risky portfolio:

$$\underbrace{E(R_C)-R_F}_{\text{risk premium on C}} = \underbrace{y(E(R_P)-R_F)}_{\text{risk premium on P}} \qquad (1)$$

The risk of the complete portfolio is characterized as:

$$\sigma_C = y\sigma_P \qquad (2)$$

which shows that the risk of the portfolio is proportional to the risk of the risky portfolio and the proportion invested in it.

Remarks:

- Volatility = measure of risk

- Volatility of a sum = variance of that sum then square root of that result (otherwise computational error)
  - Vol(A+B) → $\sqrt{[\mathrm{Var}(\mathrm{A}+\mathrm{B})]} = \sqrt{[\mathrm{Var}(\mathrm{A}) + \mathrm{var}(\mathrm{B}) + 2\mathrm{cov}(\mathrm{A},\mathrm{B})]}$
- R_f has no volatility it's not stochastic: only $y(R_p\text{-}R_f)$ is stochastic so we will only look at the volatility of this and that is why we have equation (2).

We want one formula so we can apply this into the two-dimensional space. The above equations can be rewritten into a linear relation between the expected portfolio return $E(R_C)$, and the standard deviation σ_C

from (2) we know that:

$$y = \sigma_C/\sigma_P$$

substituting this into (1) gives following linear relationship:

$$E(R_C) = R_F + \frac{(E(R_P)-R_F)}{\sigma_P}\sigma_C$$

This linear relation in the expected return-risk plane is called the **capital allocation line (CAL)**: it summarizes the feasible set of portfolios combining the two instruments F and P.

The *investment opportunity set,* is on this CAL, it is the set of feasible expected return and standard deviation pairs of all portfolios resulting from different values of *y.*

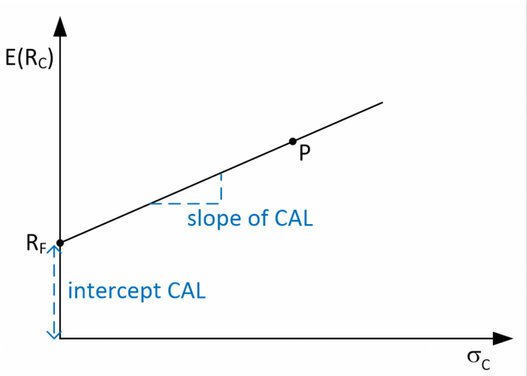
Excel exercise

Capital allocation line

You earn 1.5% p.a. on a 3 month T-bill. You expect to earn 4% p.a. on a portfolio of stocks, with a volatility of 10%.

What is your investment opportunity set?

In general: the opportunity set of one risky portfolio and a T-bill



This is what the CAL will look like if you graph it → important for the excel assignment!

On this line you have different combination of risk-free and risky asset.

We can ask ourself if it is reasonable that we can go beyond P? That is when you take a leveraged position *(here you take negative weight on the risky asset; borrowed money …see earlier)*

Once you have this line you can also describe this in term of intercept and slope

Intercept = R_F *and* *Slope* = $\frac{(E(R_P)-R_F)}{\sigma_P}$

The slope of this CAL(P)

$$S_P = \frac{E(R_P) - R_F}{\sigma_P} = \frac{risk\ premium}{SD\ of\ returns}$$

This is the **Sharpe ratio** of the risky asset P: this ratio measures the increase in expected return per unit of additional risk

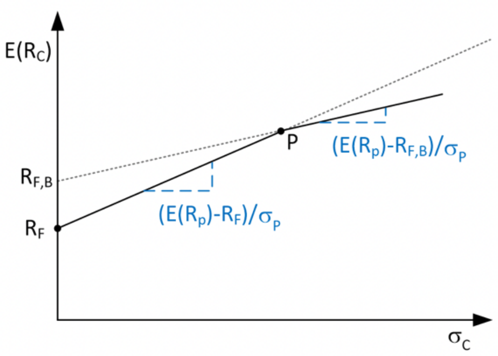
- Sharpe ratios are used in **performance measurement**: they allow to rank investments according to their Sharpe ratio
- The portfolio with the highest Sharpe ratio yields the highest expected return per risk and is therefore most attractive

Sharp ratio is not only a theoretical concept but a widely used metric if you want to compare investments; it's like a risk adjusted return. Important remarks:

- If you have just a single risky asset there is only one sharp ratio you can achieve
- All portfolios that combine the T-bill F and the risky portfolio P lie on the same CAL and have the same Sharpe ratio → we can only change the Sharpe ratio by introducing additional risky assets (see chapter 4)
    - *Every dot on the line has the same Sharpe ratio: you cannot change the Sharpe ratio. You can only increase this by increasing the amount of risky asset → if you only have one you have to be satisfied with the same sharp ratio.*

- Also **leveraged positions** (points on the CAL to the right of P) only yield an identical Sharpe ratio
    - borrow money at the risk-free rate to additionally invest in the risky asset
    - this increases the expected return and the risk stepwise
    - That's because you will have higher return BUT also higher risk: so, in terms of risk/return the ratio will stay the same. It is even worse because as a retail investor you will not be able to borrow at the Rf rate of return, if you want to borrow money you will have to pay a higher rate than the T bill:
        - In reality, only the government can borrow at the risk-free rate
            - non-government investors borrow at a rate that exceeds the lending rate of the risk-free asset $R_{F,B} > R_F$
            - borrowing at a higher rate will change the Sharpe ratio of a leveraged position: that is why in the graph here under, the CAL is kinked after P:

The opportunity set with differential borrowing and lending rates



Points beyond P:
you will have to borrow at $R_{F,B}$; so you cannot go beyond and the slope of CAL will then go down. So, if you take a leveraged position, you will have a LOWER sharp ratio.

So, when you take a leveraged position, after point P as a nongovernment investor you borrow at a rate higher than lending rate of the risk-free asset = kinked curve. To the left of P, the investor is borrowing ate the risk-free rate R_F.

It all boils down to preferences, if you want a higher expected return, you can only do this with a lower SR. The points on the CAL will depend on risk preferences →

→ This will determine which point on the CAL will be chosen.

- The investor will need to decide on one position on this CAL; this choice will be determined by his **attitude towards risk**
- The majority of investors is risk-averse: they are only willing to take on risk if they are rewarded for this risk (in expected terms): People need higher returns to be compensated for the risk.
  → this corresponds to the empirical observation that risky assets earn, on average, a higher return than the riskfree rate of return (over the period 1926-2012, the average return of S&P500 exceeded the T-bill return by +/- 8% per year)

Risk, Speculation, and Gambling

***Speculation** is the assumption of considerable investment risk to obtain commensurate gain.*

- *"Considerable risk" the risk is sufficient to affect the decision*
- *"Commensurate gain" a positive risk premium, that is, an expected return greater than the risk-free alternative.*

*To **gamble** is to bet or wager on an uncertain outcome.*

The central difference between gambling and speculation is the lack of "commensurate gain."

A gamble is the assumption of risk for enjoyment of the risk itself, whereas speculation is undertaken in spite of the risk involved because one perceives a favorable risk–return trade-off.
*Notice that a risky investment with a risk premium of zero, sometimes called a **fair game,** amounts to a gamble because there is no expected gain to compensate for the risk entailed.*

A risk-averse investor will reject gambles, but not necessarily speculative positions.
The important point is that even highly risky positions may be willingly assumed by risk-averse investors if they believe they are adequately compensated by the risk premium. We therefore should expect higher risk premiums to be associated with greater risk.

St. Peterburg Paradox

This recognition of risk aversion dates back to 1738 with Daniel Bernoulli in his famous St. Peterburg Paradox

St. Peterburg Paradox

To enter a coin-toss game, you need to pay an entry fee. Afterwards, a coin is tossed, until the first head shows up. The number of tails (n) that appears until the first head is tossed, is used to compute the payoff W as $W = 2^n$.

How much are you willing to pay to enter this game?

It depends on the payoff what you willing to pay for the fee:

Tails n	Proba P(n)	Payoff W(n)	P(n) x W(n)
0	1/2	\$1	\$1/2
1	1/4	\$2	\$1/2
2	1/8	\$4	\$1/2
3	1/16	\$8	\$1/2
⋮			
n	$\left(\frac{1}{2}\right)^{n+1}$	$\$2^n$	\$1/2

Table: Payoffs and probabilities of various outcomes

The expected payoff is thus: $E(W) = \sum_{n=0}^{\infty} P(n)\, W(n) = \infty$, but most participants will only be willing to pay a finite and modest fee.

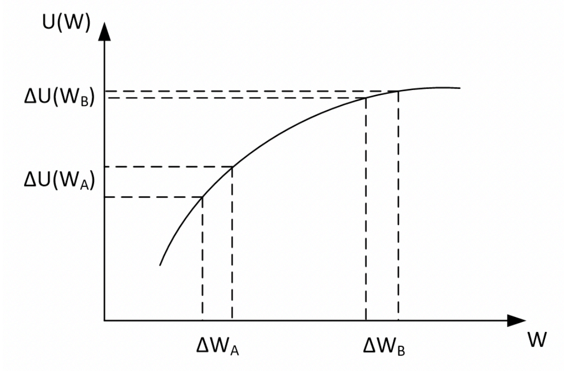
How to solve this paradox?

Compute expected payoff of the game; you list a number of scenarios
Infinity is the answer: normally people would be willing to pay a high price. How come that people are only willing to pay a low fee while the payoff of the game is infinite?

- Instead of optimizing expected payoffs, people tend to **optimize expected utility** that they derive from the payoffs
  → in doing so, people do not value each payoff in an identical way
  → the greater the wealth, the lower the appreciation for an incremental dollar; this is the idea of decreasing marginal utility
- These insights can be translated into a utility function that increases with the payoff, but at a decreasing rate

A common example of such utility function is the **log utility function**: $U = ln\,(W)$

This translates into risk aversion



Certainty equivalent rate

The **certainty equivalent** is the rate that a risk-free investment would need to offer to provide the same utility as the risky portfolio. In other words, it is the rate that, if earned with certainty, would provide a utility score equal to that of the portfolio in question.

- Applied to the above game, this gives a subjective expected utility value which is finite and modest:

$$E(U(W)) = \sum P(n) \ln(W(n)) = \sum (1/2)^{n+1} \ln(2^n) = 0.693$$

- To determine the fee that one is willing to pay to enter this game, we can compute the certainty equivalent value (CE) of the game as

$$\ln(CE) = 0.693 \Longrightarrow CE = 2$$

  - this CE is the amount that yields a utility identical to the expected utility of the game
  - as both yield the same utility, one would be indifferent between the game or receiving the CE for sure → this CE is then the maximum amount that an investor will pay to play the game

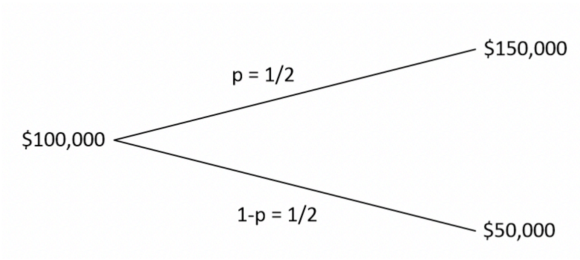
A portfolio is desirable only if its certainty equivalent return exceeds that of the risk-free alternative. If risk aversion is high enough, any risky portfolio, even one with a positive risk premium, will be assigned a certainty equivalent below the risk-free rate and will be rejected by the investor. At the same time, a less risk-averse investor may assign the same portfolio a certainty equivalent rate greater than the risk-free rate and thus will prefer it to the risk-free alternative.

CE=2: indifferent between playing the game or receiving 2 dollars directly = maximum value that you willing to pay to enter the game
- If the fee is lower you enter the game, if the fee is higher you prefer receiving 2 dollars

Risk aversion and fair games

Will someone with logutility accept a fair game?

Assume you have logutility over your wealth, and you are offered the following game:



Would you take on the gamble, do you prefer the sure amount, or are you indifferent between both?

If you are risk averse should you accept a fair game? It depends on your utility function

To appreciate this fair game, we need to look at the utility of the different outcomes

Choice	E(W)	E(U(W))
No gamble	100, 000	11.51
Gamble	100, 000 $= \frac{1}{2} \times 150,000 + \frac{1}{2} \times 50,000$	11.37 $= \frac{1}{2} \times 10.82 + \frac{1}{2} \times 11.92$

The fair gamble has lower expected utility than the sure bet, and thus the gamble will not be preferred $\rightarrow$ the loss in expected utility from losing $50,000 is larger than the gain in expected utility from winning $50,000

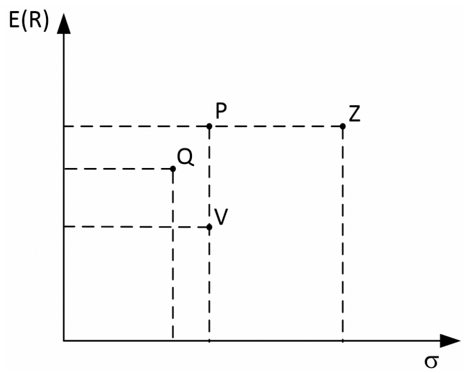
This is a general result: risk averse investors will reject fair games or worse
They consider only risk-free or speculative prospects with positive risk premiums.

- Identical E(W) want fair game
- Depending on your utility function; in this case log utility = risk averse
- The risk averse person will always choose to not gamble and reject fair game
- Because you already have 100k you are already rich so you thing that if you loose 50k it is worser than winning 50k

Choosing among portfolios

Investors choose among portfolios by assigning welfare or utility to competing portfolios: more attractive portfolios receive higher utility

Investors like **high expected returns** and/or **low risk**



- P ≻ V: higher expected returns, for a same level of risk
- P ≻ Z: lower risk, for a same level of expected returns

  What about P and Q?

  - Less risk is preferred to more risk
  - Risk loving people will always take V over P because you don't like expected returns, you like risk
  - What about risk neutral investors?
    Indifferent between P and Z then you value P and Z equally.
  - P or Q? This is a bit more individual specific. We cannot uniquely rank P and Q.

- While some portfolios can be ranked unambiguously, this does not hold for all of them
- In those cases, it will depend on the degree of risk aversion of the individual investor: **how severely is risk penalized**?
- This trade-off between risk and expected returns can be represented in a utility function:

$$U = E(R) - \frac{1}{2} A\sigma^2$$

  - higher expected returns yield higher utility
  - higher variance of returns yields lower utility
- Degree of penalization (The extent to which the variance of risky portfolios lowers utility) depends on the risk aversion of the investor, as captured by A
  - *A* high: high penalty = high aversion to risk
  - *A* low: low penalty = low aversion to risk

Notice that risk-free portfolios (with variance = 0) receive a utility score equal to their (known) rate of return because they receive no penalty for risk.

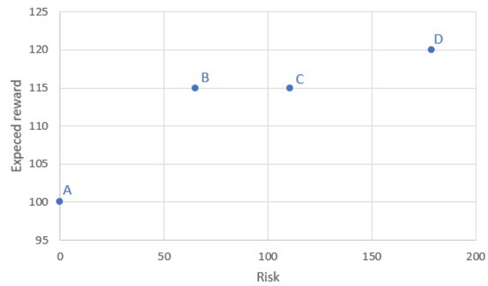
Risk-neutral investors (with *A* = 0) judge risky prospects solely by their expected rates of return. The level of risk is irrelevant to the risk-neutral investor, so there is no penalty for risk. In this case, a portfolio's certainty equivalent rate is simply its expected rate of return.

A **risk lover** (for whom *A* < 0) is happy to engage in fair games and gambles; this investor adjusts the expected return *upward* to take into account the "fun" of confronting the prospect's risk. Risk lovers will always take a fair game because their upward adjustment of utility for risk gives the fair game a certainty equivalent that exceeds the alternative of the risk-free investment.

In context of investment analysis, we work with **mean variances utility and not log utility.** Because we have a preference over those 2 dimensions: risky asset an Rf.
- Variance go up → U goes down

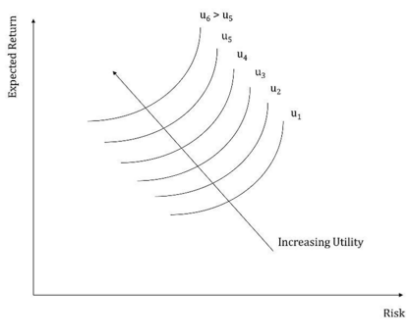
A = sensitivity of my risk.

Choice	E(W)	σ(W)
Get €100 for sure	€100	€0
50% chance to either €50 or €180	€115	€65
66.67% to €45 and 33.33% to €255	€115	€110.68
75% to €40 and 25% to €360	€120	€178.88
Indifferent between (at least) 2 choices		

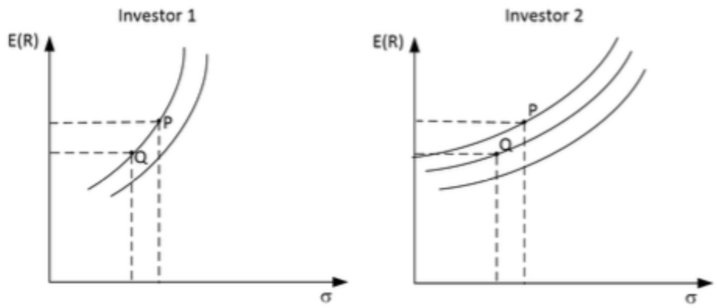


Indifferent is also reasonable between two choices: both opportunities give you the same amount of U = indifference curve → all solutions on his curve yields the same amount of U.

Which indifference curve if you like returns but dislike risk? You try to go as much upward (northwest) because in this direction we simultaneously increase the expected return *and* decrease the volatility of the rate of return. The slope of the ICdepends on the risk aversion →



The investor's utility function that trades off risk and expected return can also be summarized by its **indifference curves:** this gives combinations of risk and expected return that yield a constant utility



Which investor is most risk averse?

- **Investor 1**: steep → most risk averse → the steeper, the more risk averse you are. The increase in expected return is sufficient for Q en P so he is equally happy: starting at *P,* an increase in standard deviation lowers utility; to compensate, expected return must be higher. Thus, point *Q*, with higher risk but higher expected return, is equally desirable to this investor as *P.*
    - Steeper curves mean that investors require a greater increase in expected return to compensate for an increase in portfolio risk.

- **Investor 2**: flat → he is also confronted with the same increase in risk and expected return but here the increase in return is much more then he expected. It makes him happier and ends up in a higher U = P to be indifferent between Q and P you see that we would end up at a point lower then P, to be indifferent the increase in risk and return would be marginal. This investor value risk differently.

Excel exercise

Risk aversion

Assume again the investment opportunity set determined by a 3 month T-bill on which you earn 1.5% p.a., and a portfolio of stocks on which you expect to earn 4% p.a., with a volatility of 10%. What is the utility you derive from an equally weighted portfolio, assuming mean-variance utility with a risk aversion parameter of 6? List/draw a set of portfolios for which you are indifferent. How does it change your results for a risk aversion parameter of 3?

Estimating risk aversion

While theoretically clear and appealing: how to estimate risk aversion?
Multiple approaches exist, but it is not easy:

- use of questionnaires
- analysis of portfolio holdings over time, together with estimates of expected returns/variance of returns
- purchase of insurance contracts: prices of insurance contract correspond to your risk aversion: the more you willing to pay to more risk averse you are

Economists generally agree that 1 < risk aversion < 10 is a reasonable range

Capital allocation decision

How much risk should you take?

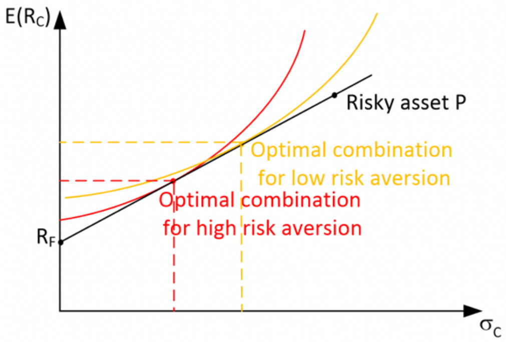
We have shown how to find the CAL, the graph of all feasible risk–return combinations available for capital allocation. The investor confronting the CAL now must choose one optimal complete portfolio, C, from the set of feasible choices. This choice entails a trade- off between risk and return. Individual differences in risk aversion lead to different capital allocation choices even when facing an identical opportunity set. In particular, more risk-averse investors will choose to hold less of the risky asset and more of the risk-free asset.

To pinpoint the capital allocation decision, we bring together:

1. The investment opportunity set as summarized in the CAL:
$$E(R_C) = R_F + \frac{(E(R_P) - R_F)}{\sigma_P}\sigma_C$$

2. Investor mean-variance preferences as summarized in the utility function:
$$U = E(R) - \frac{1}{2}A\sigma^2$$

Graphical solution



Higher indifference curves correspond to higher levels of utility. The investor thus attempts to find the complete portfolio on the highest possible indifference curve. When we plot indifference curves on the investment opportunity set represented by the capital allocation line, we can identify the *highest possible* indifference curve that still touches the CAL. That curve is tangent to the CAL, and the tangency point corresponds to the standard deviation and expected return of the **optimal complete portfolio.**

The point on the call depends on the mean variance utility.
I move the IC until it is tangent with the CAL. If it crosses the CAL it's not an optimal solution because I can increase my utility by using one that is tangent. I can choose maybe one that is way above the CAL but it is not tangent with the CAL so not optimal. So, this is the optimal solution.

Analytical solution

The investor chooses the risky portfolio weight *y* to maximize his utility *U*:
$$\max_y U = E(R) - \frac{1}{2}A\sigma_C^2$$

Substituting for equation (1) and (2), this becomes:
$$\max_y U = R_F + y(E(R_P) - R_F) - \frac{1}{2}Ay^2\sigma_P^2$$
The first-order condition of this problem determines the optimal portfolio weight y*:
$$y = \frac{E(R_P) - R_F}{A\sigma_P^2}$$

- $y*$ is *directly* proportional to the risk premium offered by the asset: you invest more in the risky asset if the risk premium is larger on P
- $y*$ is *inversely* proportional to the level of risk aversion and the level of risk (as measured by variance): more risk averse, the higher A, the lower allocation to risky asset

This equation van be rewritten towards A:

$$A = \frac{E(R_P) - R_F}{y\sigma_P^2}$$

Estimate risk premium and riskiness; you will know how much wealth risky people on average invest in risky assets and you can estimate the risk aversion of people

> **Excel exercise**
> **Capital allocation decision**
> Assume again the investment opportunity set determined by a 3 month T-bill on which you earn 1.5% p.a., and a portfolio of stocks on which you expect to earn 4% p.a., with a volatility of 10%. What is the optimal capital allocation for a mean-variance utility investor with a risk aversion parameter of 6 and 3, respectively?

Next

- The current chapter takes portfolio P as given, but the question raises how the composition of P is determined
- Mean-variance portfolio theory provides a framework to determine the 'best' P on the basis of **diversification benefits**.
- Depending upon additional assumptions this leads to two main approaches in practice:
  - MV + market efficiency + CAPM: market-capitalization weighted portfolio is optimal (CAL = CML) → passive strategy to profit from diversification benefits
  - MV + market inefficiency: smart selection of instruments → active strategy to profit from under/overpriced assets, typically in addition to diversification benefits

  What to prefer ultimately depends on your skill and cost-structure

Chapter 4: Optimal risky portfolios

BKM - Chapter 7: Optimal risky portfolios

How to choose the risky portfolio P among various risky assets?
There is no such thing as A risky asset. There are many risky assets to choose from. This is what we will focus on in this chapter.
In principle, asset allocation and security selection are technically identical; both aim at identifying the optimal risky portfolio, the combination of risky assets that provides the best risk–return trade-off or the highest Sharpe ratio. In practice, asset allocation and security selection are typically separated into two steps, in which the broad outlines of the portfolio are established first (asset allocation), while details concerning specific securities are filled in later (security selection).

Diversification

A simple numerical example

Suppose you have been investing in portfolio P; you have extra cash and want to make an extra investment → P, A or B?

Table: Returns (p.a.)

	Portfolio P	Stock A	Stock B
2007	0%	-1%	7%
2008	4%	-1%	5%
2009	8%	-4%	4%
2010	2%	-7%	8%
2011	-1%	-2%	8%
2012	-9%	-3%	7%
2013	-7%	-6%	9%
2014	0%	1%	7%
2015	2%	3%	3%
2016	1%	3%	-2%
2017	9%	8%	-4%
2018	6%	8%	-8%
2019	7%	9%	-9%
2020	6%	8%	-6%
2021	5%	7%	-4%

Based upon these historical returns I am calculating my investment parameters.

Table: Investment parameters

	Portfolio P	Stock A	Stock B
$E(R_i)$	2.20%	1.53%	1.67%
σ_i	5.17%	5.49%	6.44%

The decision to invest will always be based upon the expected return and the risk because those are my two decision variables. This is what I take into account.

Suppose I do this in a naïve way and I just look at the expected return and the volatility.

P is the most attractive one because the risk is lower, and the return is higher.
Suppose this is the data you are provided with. What would be your conclusion, which kind of investment would you choose?
Most likely you will go for portfolio P. because this has the highest expected return with the lowest level of risk. But this is not necessary the whole picture. We are not explicit about one important piece of information:

Table: Investment parameters

	Portfolio P	Stock A	Portfolio (95%-5%)
$E(R_i)$	2.20%	1.53%	2.17%
σ_i	5.17%	5.49%	5.09%

	Portfolio P	Stock B	Portfolio (95%-5%)
$E(R_i)$	2.20%	1.67%	2.17%
σ_i	5.17%	6.48%	4.70%

There is a missing link.
Let's continue as if we had invested in A or in B.

The top panel: investment parameters of 95% of initial portfolio and 5% of stock A
The lower panel: 95% of initial portfolio and 5% of stock B

→ are these new portfolios attractive this attractive vis a vis the initial portfolio

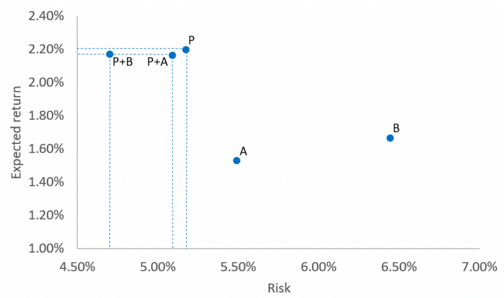
Stock A
- The expected return is lower than P
- But the risk is lower than the initial portfolio
- Even tough stock A is not attractive, adding it to the initial portfolio will make it attractive

Stock B
- Expected return is identical as A
- But here the risk is significantly lower as compared to the initial portfolio

  → Comparing these different investment opportunities: we can rule out A because B will be more attractive.

What about choosing between P and B? That is a harder decision.
- Level of risk is lower with B but also expected return will be lower; so your decision depends on how much risk you are willing to take.

If you just link these 3 investment opportunities as individual opportunities, you want to come to a conclusion whether you will invest in A or B. Well it seems A and B individually would NOT be attractive. But once you add them to your existing portfolio, the dynamics of those new portfolios are attractive.
Most people would choose P+B: I only need to give up a small amount of expected returns, but my risk is significantly reduced. The reduction in risk is so big for the small cost in terms of expected return.



How can we explain this?
We can explain this by diversification. It is a wakeup call to never focus on risk and expected return of individual investments. We need to look at diversification benefits.
You could also see this on the table; when portfolio P is doing great, stock B is not doing good → good diversification: they are negatively correlated.

A stylized diversification strategy with two risky assets
We introduce a simple model to illustrate the benefits of diversification

- Assume two risky assets *e* and *d* both with identical expected returns:

$$R_d = \mu + \varepsilon_d$$
$$R_e = \mu + \varepsilon_e$$

  - Where they both exposed to a single risk factor: ε_d and ε_e, modelled as random variables with mean zero *E* (ε_i) = 0 *(expected value of risk factor is zero: I do not expect this risk, but it can occur)* and identical variance σ^2_{idio} *(they have the same size)*
  - where the risks are uncorrelated, i.e., *cov* (ε_d , ε_e) = 0

- Intuitively, this set-up implies that the only risk in the stocks is **idiosyncratic risk** or so-called **firm-specific risk**: this means that the risk is unique to asset d or unique to asset e

Compare the investment parameters of following two strategies:
1. Invest all your wealth in a single risky asset
2. Invest half of your wealth in each of 2 risky assets

Strategy 1 yields following variance-expected return trade-off:

$$E(R_1) = \mu$$
$$Var(R_1) = \sigma^2_{idio}$$

→ This is a summary of investing everything in one risky asset:

- So, you expect to earn a return μ
- In terms of risk same amount of variance.

Strategy 2 yields following variance-expected return trade-off:

$$E(R_2) = \frac{1}{2}\mu + \frac{1}{2}\mu = \mu$$

$$Var(R_2) = Var\left(\frac{1}{2}\varepsilon_d + \frac{1}{2}\varepsilon_e\right) = \frac{1}{4}\left(Var(\varepsilon_d) + Var(\varepsilon_e) + 2cov(\varepsilon_d, \varepsilon_e)\right)$$
$$= \frac{\sigma^2_{idio}}{2}$$

Basic diversification reduces the risk

- Here we still expect to earn μ. Half of it coming from e and half of it coming from d.
- But in terms of risk, we do see an important modification as compared to the first strategy.
- So, if you take the variance of the two risk factors (variance of sum = sum of variances + cov): there is no covariance because they are not correlated, so the covariance term drops out (=0).
- Then we are left with: (var(ε_d)+ Var(ε_e)), but they are identical so = $2\sigma^2_{idio}$. And we divide by 4 → $= \sigma^2_{idio}/2$

This simple strategy halves the amount of variance I am exposed too and it does not cost me in terms of expected returns. This is the power of diversification.

Assume now that there are N of these risky assets (all with identical expected returns):

$$R_i = \mu + \varepsilon_i \quad i = 1, \ldots, N$$

Strategy 3, in which an equally weighted portfolio is constructed, yields following variance-expected return trade-off:

$$E(R_3) = \frac{1}{N}\sum_{i=1}^{N}\mu = \mu$$
$$Var(R_3) = \frac{1}{N^2}Var\left(\sum_{i=1}^{N}R_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\left(\sigma^2_{idio} + \underbrace{crossterms}_{0}\right) = \frac{\sigma^2_{idio}}{N}$$

For many risky assets N, diversification eliminates idiosyncratic risks

- We still expect to earn μ.
- The larger N, the smaller the variance you are exposed to. In the limit this fraction will go to zero. So, if we diversify, we can eliminate ALL idiosyncratic risk.

The previous model is stylized in that it assumes only the presence of idiosyncratic risks and that there was only a single risk factor → in reality, there is also **systematic risk**: risk that is shared among assets.

Assume there are N risky assets, all with identical expected returns:

$$R_i = \mu + \varepsilon_{sys} + \varepsilon_i \qquad i = 1, \ldots, N$$

Where:

- ε_{sys} is a systematic risk factor with (we assume identical over all assets)

- Expectation E(ε_{sys}) ≠ 0
    - Variance σ_{sys}^2
- The systematic and idiosyncratic risk factors are uncorrelated, i.e. $Cov(\varepsilon_{sys}, \varepsilon_i) = 0, \forall i$
- The idiosyncratic risk factors are uncorrelated, i.e. $Cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

If you have this set of returns, with those two types of risk, what is than the performance of such a portfolio?

Strategy 4, in which an equally weighted portfolio is constructed, yields following variance-expected return trade-off:

$$E(R_4) = \mu + E(\varepsilon_{sys})$$
$$Var(R_4) = Var\left(\varepsilon_{sys} + \sum_{i=1}^{N} \frac{\varepsilon_i}{N}\right) = \sigma_{sys}^2 + \frac{\sigma_{idio}^2}{N}$$

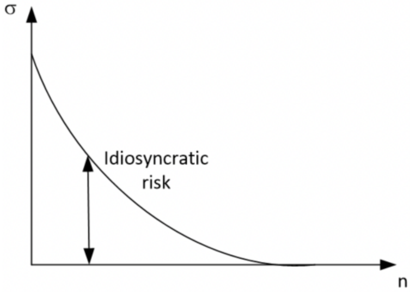
where for many risky assets N, the idiosyncratic risk component becomes negligible

For many risky assets N, diversification eliminates idiosyncratic risks, but not systematic risk

- The expected value of the systematic value is not zero. This means that you are rewarded for the systematic risk you took. So, you get a return to be exposed to systematic risk.
- What happens in term of variance? All the cross terms are zero; we will be left with the sum of variances. If you calculate this variance, you have the systematic and the idiosyncratic risk.
- What is different from other strategy? There is no factor N where systematic risk will diminish. I will always be exposed to the same amount of systematic risk. When I diversify my investments, I can't get rid of my systematic risk, but I can get rid of idiosyncratic risk.
    - There is no way for you to get rid of this risk; the only way you are able to is to accept this risk by being rewarded for it. So when we talk about diversification benefits we only talk about idiosyncratic risk.
    - The market will only reward you for the systematic risk and that is why it is not incorporated in the expected return formula.

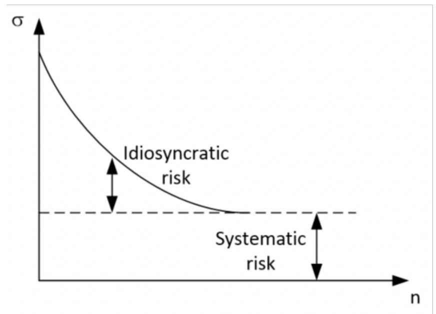
Graphical representation of diversification benefits

Case 1: all risk is idiosyncratic



The more assets we add the more our volatility is reduced, and we could go to zero.

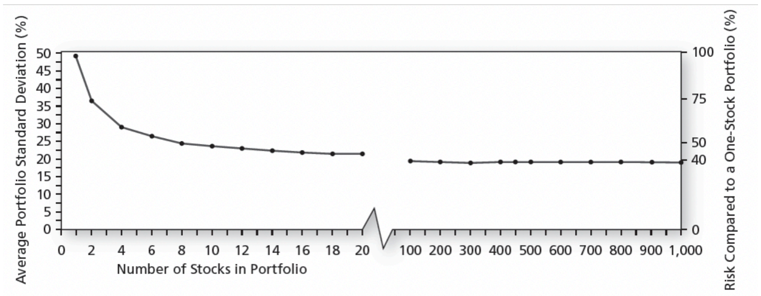
Case 2: some risk is idiosyncratic ; some risk is systematic



We cannot reduce systematic risk.

Empirical evidence of diversification benefits

- Average portfolio risk decreases rapidly with increasing number of stocks
- In the limit: portfolio risk is reduced to +/- 19%



If you want to take this advice in practice how many stocks on average do you need to eliminate this idiosyncratic risk? If you randomly select stocks, you should be okay with 15-20 stocks. This is what we see on this graph. This is important when we talk about elimination of idiosyncratic risk you can achieve this. Diversifying pays of immediately.

Mean-Variance analysis

The simple models we have discussed so far are useful to introduce the concept of diversification, but they are unrealistic

- We have imposed very particular return dynamics
- We have assumed naive diversification by adding more and more risky assets in equal proportions

In a next step, we elaborate on this to obtain a more realistic result:

- We introduce more general return dynamics: we will not impose structure on returns (for now)
- We look into **optimizing the portfolio composition** as to **minimize variance for a given expected return** (so not just randomly add more and more assets)

→ This is **efficient diversification**

The idea of diversification in a portfolio selection MV model was first formalized by Harry Markowitz in 1952

Basic idea:

- Combining risky assets diversifies risk: it pays off because combining risky assets will help you achieve different risk profiles
- Identification of the efficient frontier: portfolios with highest expected return, for any level of risk (or portfolios with lowest risk, for any level of expected return) so not in a naïve way!

Such formal approach to investments was groundbreaking and still is today the basis of the asset management industry. This pushed the industry to think in an efficient way about portfolio composition. Because even before this paper, the industry knew that the idea of diversification was the good way. But the public could not quantify the risk, risk reduction and the returns. With this model we know the optimal portfolio we can achieve.
Markowitz, H.(1952). Portfolio Selection. Journal of Finance, 7(1): 77-91
Markowitz received the 1990 Nobel Prize in Economics (together with William Sharpe and Merton Miller)

We think in a consistent way about portfolio composition. With this mean variance model, we know what the best combination is.

Portfolio with two risky assets

Assume 2 risky assets *d (debt/bonds)* and *e (equity) but* can also be considered as 2 asset classes with following rate of return distribution characteristics (who are known ex post):

- returns: R_d and R_e
- expected returns: $E\ (R_d)$ and $E\ (R_e)$
- risk: σ_d and σ_e
- correlation: $\rho(R_d,R_e)=\rho_d$
- covariance: $cov\ (R_d\ , R_e) = cov_{de} = \rho_{de}\, \sigma_d\, \sigma_e$

We look at a portfolio of these two assets, where a fraction w_d and w_e is invested in *d* and *e*, respectively

Portfolio with two risky assets: portfolio structure

- The portfolio return R_P is (**ex post**):

$$R_P = w_d R_d + w_e R_e$$

But **ex ante** when we have to decide upon w_d and w_e we will do this based upon our best estimates of expected returns and risk:

- The expected return of the portfolio is a weighted average of the expected returns on the risky assets:

$$E\ (R_P) = w_d E\ (R_d) + w_e E\ (R_e)$$

→ the expected return is **additive in the components**

- The variance of the portfolio is:

$$\sigma_P^{\ 2} = w_d^{\ 2} \sigma_d^{\ 2} + w_e^{\ 2} \sigma_e^{\ 2} + 2 w_d w_e \rho_{de} \sigma_d \sigma_e$$

→ the variance of the return is **not a weighted average of the variances (not additive in the components); it is a weighted average of the covariances**

This means that portfolio variance will always be lower than the weighted average of the variance:

Portfolio with two risky assets: power of diversification

- When the portfolio weights are non-negative, it holds that:

$$\sigma_P^{\ 2} \leq (\ w_d\, \sigma_d + w_e\, \sigma_e\ )^2$$

  - with an equality only if $\rho_{de} = 1$
    - In this special case, with perfect **positive** correlation (and only in this case), the standard deviation of the portfolio is just the weighted average of the component standard deviations.

        - In all other cases, the correlation coefficient is less than 1, making the portfolio standard deviation *less* than that weighted average.
    - the portfolio variance is lower than the weighted average variance
    - This inequality holds because if you look at the right side of the inequality, this is the square of a sum: $(w_d \sigma_d)^2 + (w_e \sigma_e)^2 + 2 w_d \sigma_d w_e \sigma_e$
        - The one difference between this equation and the one above: ρ
        - The value of this correlation is max 1; so in general this value would be smaller than 1. Which means that the sum above will be smaller (or equal) to this equation.
- This summarizes the **power of diversification**: less than perfectly correlated stocks **always offer diversification benefits**; the lower ρ, the greater the diversification benefit
- So, we look for attractive ρ, attractive correlations

How low can the risk of a portfolio be?

Portfolio with two risky assets: Global minimum variance portfolio

The global minimum variance portfolio is the portfolio combination for which the variance is minimal; the weights w_d and w_e for which we minimize the risk, will depend on the correlation between *d* and *e*

- Special case $\rho_{de} = 1$

$$\sigma_P^2 = (w_d \sigma_d + (1 - w_d)\, \sigma_e)^2 = 0$$

$$w_d = \frac{-\sigma_e}{\sigma_d - \sigma_e}$$

- Special case $\rho_{de} = -1$

$$\sigma_P^2 = (w_d \sigma_d - (1 - w_d)\, \sigma_e)^2 = 0$$

$$w_d = \frac{\sigma_e}{\sigma_d + \sigma_e}$$

For perfectly correlated assets (+1 or -1), **all portfolio risk can be eliminated** by holding the right combination of the two risky assets

If these two assets are perfectly correlated (positive or negative), then there is a particular mix that can eliminate risk al together: I have to minimize portfolio risk (portfolio variance), I calculate the derivative and I find the fraction (w_d).

Which of the two are the least attractive?

- The return will be the highest with -1 = risk free rate (both weights will be positive, you will invest a positive fraction in d and e)
- The one with +1 is the least attractive = risky asset (if you look at the weight, if one weight is positive the other will be negative) → very weird strategy. If you have two assets that are perfectly correlated than in fact you have a single asset. Because this is an identical asset, in equilibrium. Because it acts in the same way. So if you go long in one asset you will have to go really short in the other asset, that is the only way to get to zero.

Our correlation is between 1 and -1 in normal times between 0 and 1. (but not many asset have a negative correlation)

In turbulent times all correlations tend to go to 1 and then there is less diversification benefits. But there is one asset class that typically go down during turbulent times: government bonds. During crisis periods there is 1 government bonds that shows negative correlation with stock and it's the German bond. Prices of German bonds are driven up during crisis periods, because everyone wants to fly into the safe heavens.

In the more general case where $-1 < \rho_{de} < 1$ (not perfectly correlated): portfolio risk cannot be eliminated altogether, the best we can do is minimizing the risk, this is called the **global minimum variance portfolio**:

$$Min\ \sigma_P^2 = w_d^2\sigma_d^2 + (1-w_d)^2\sigma_e^2 + 2w_d(1-w_d)\,cov_{de}$$

$$\frac{\partial \sigma_P^2}{\partial w_d} = 2w_d\left(\sigma_d^2 + \sigma_e^2 - 2cov_{de}\right) - 2\left(\sigma_e^2 - cov_{de}\right) = 0$$

Following portfolio weights thus give minimal risk:

$$w_d = \frac{\sigma_e^2 - cov_{de}}{\sigma_d^2 + \sigma_e^2 - 2cov_{de}}$$

$$w_e = 1 - w_d = \frac{\sigma_d^2 - cov_{de}}{\sigma_d^2 + \sigma_e^2 - 2cov_{de}}$$

This is the lowest you can go. Apart from that there are many other combinations: Because going for the lowest amount of variance might not be the best option. You can look and analyze different combinations and always calculate the corresponding general return.

Portfolio with two risky assets: Portfolio characteristics as a function of portfolio weights

Apart from looking to the global minimum variance portfolio you can look more general at other options: different portfolio weights correspond to different portfolio expected returns and risk

How are the portfolio's expected return and risk as a function of the portfolio weights?

Two risky assets (BKM p. 196)

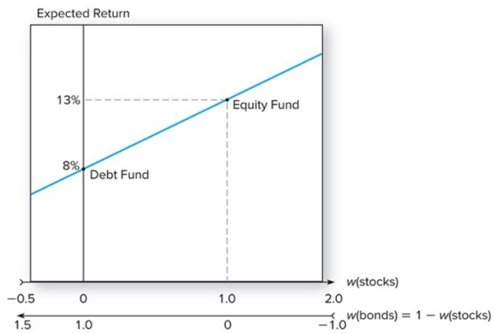
Assume two assets d and e with following return dynamics:

	Asset d	Asset e
$E(R)$	8%	13%
σ	12%	20%

What is the portfolio's expected return and risk for given correlations $\rho = -1, 0, 0.30, 1$? Also derive the corresponding minimum variance portfolio.

First case: expected returns in terms of portfolio weights

This is a simple case, because when we analyze expected return correlation does not enter. We can ignore different scenarios of correlations. The relationship between expected return and portfolio weights is here independent from correlations.

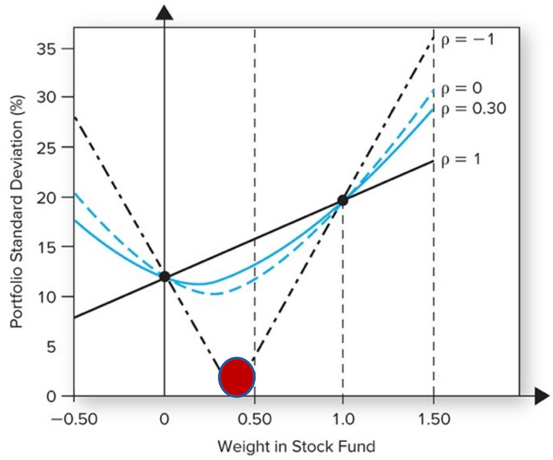


We have a linear relationship (this is also what we expected): expected return of portfolio is linear in the weights. If all is invested in equity; what you earn is the expected return on the equity instrument.

If all is invested in debt instrument; then you expect to earn the rate of return on your debt instrument.

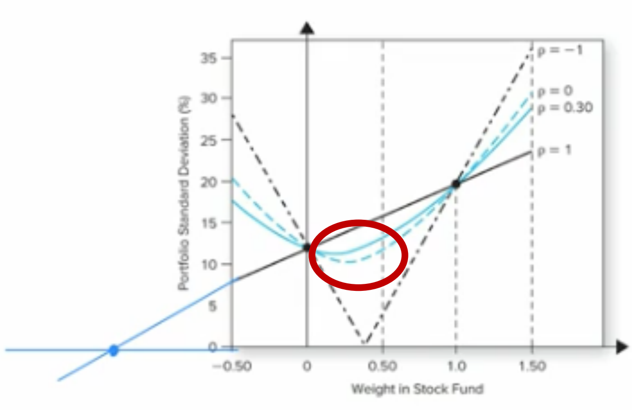
- Area where you have zero weight of the assets = between 'Debt Fund' and 'Equity Fund'
- If you go beyond those points, you will go short; go long 100% in one and short (negative weights) in other asset.

Case 2: Correlation
When we talk about volatility correlation enters.



Here we have the different scenarios meaning that we have different relationships for different correlations.

- p=1: linear relationship between portfolio SD and weight invested in stocks.
    - Here you will find also the combination that will allow you to eliminate the risk all together. Where? The blue point here under.
    - That is more than -100% invested in stocks, so that is not very standard
- p = -1: here also which combination allows us to eliminate all the risk? (the red dot here above)
    - you concentrate more your portfolio into the stock investment (so to the right) your volatility increases
    - the same goes when you hold for a lower weight in stocks but bigger concentration in bonds, your risk increases
    - so when you deviate from the red point, you increase your risk in any direction



Slightly more general scenarios:

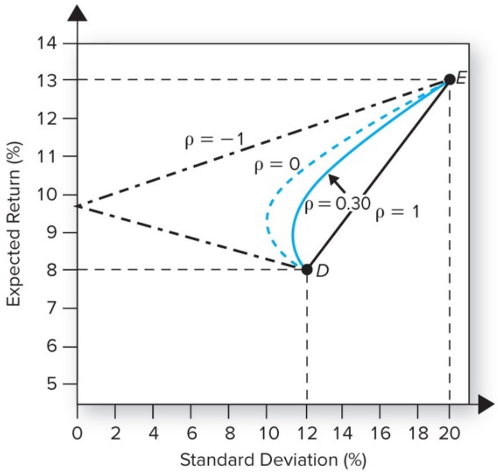
- you will not be able to eliminate all the risk in your portfolio
- you can see that on the graph because those two blue lines to not reach zero
- the minimum points in the red circle are the **minimum variance portfolios**
    - **=** the minimum level to which portfolio standard deviation can be held
    - also here when you start deviating from that you see that your risk increases
    - but taking up more risk is a cost but you will be rewarded
- ideally we combine this graph with the first case, then we get a full picture of how expected returns and volatilities can be combined for varying portfolio weights.

- Combining the portfolio risk and expected returns gives the well-known risk-expected return diagram

  Remember that expected returns are a linear function of the portfolio weights, while variances of returns are quadratic in the weights;
  *Plotting expected returns vs variance is a parabola, and thus plotting expected returns vs standard deviation is a **hyperbola**.*
- All combinations of expected returns and risk that can be constructed from the two available assets is called the **investment opportunity set**

Case 3: combination

Here on this graph, we have plotted the **investment opportunity set** for those different scenarios.



- p = 1: if we combine D and E all those combinations will lie on the black straight line (linear relationship)
  - intuitively means this that there is no diversification benefits to realize
- p = -1: the investment set is the whole black striped line
  - D: not attractive because it will be dominated by the dotted line above; more expected return for the same amount of risk
  - So not the whole set is equally attractive
  - But here we can see the extreme diversification benefits that we can realize:
    - If you start at the right (p = 1) you see that we will move to the left direction, the lower your correlation the higher the diversification benefits
  - You can go to the no risk scenario (point under 10 by expected return):
- The point when having extremely positive correlation where you will have no risk is →

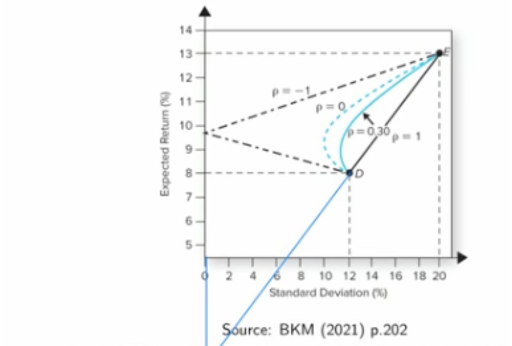


Figure 7.5 Portfolio expected return as a function of standard de

But this is not attractive as you will have a negative return.

More realistic cases:

- Here you see the diversification benefits that you realize
- The investment opportunity set is pushed toward lower level of volatility
- The lower the correlation the more the set is pushed to the left
- The degree to which you have a curved relationship as compared to the standard case of no diversification benefits captures your degree of diversification

Conclusion: Potential benefits from diversification arise when correlation is less than perfectly positive. The lower the correlation, the greater the potential benefit from diversification. In the extreme case of perfect negative correlation, we have a perfect hedging opportunity and can construct a zero-variance portfolio

UOVT

Excel exercise

The investment opportunity set of two stocks

What is your investment opportunity set when following 2 stocks are available to invest in (in % p.a.):

	Stock 1	Stock 2
$E(R)$	18.11%	9.66%
σ	32.64%	24.32%

The correlation between both is 28%. Determine also the minimum variance portfolio.

How to determine the investment opportunity set with many risky assets?

- When N risky assets are available, it is impossible to trace all feasible portfolio combinations by hand
- We therefore focus on a subset of portfolio combinations: portfolios that yield a particular target return, for a minimal amount of risk (only the attractive ones)
- *Suppose that I have 10 assets to choose from and I want to earn a return of 5%, then I will look for that combination that allows me to reach 5% for minimum risk.*
- To compute this set of minimal variance portfolios we solve:

$$\min_{w} \sigma_P^2$$

$$E(R_P) = R_{\text{target}}$$
$$\sum_i w_i = 1 \text{ and } w_i \geq 0 \quad \forall i$$

where we repeat this optimization for different levels of a target return (for example start with 5%, than 6%, … that allows me to trace all the attractive portfolios)

Constraints:

- weights all need to adopt to 1, (100%).
- All weights should be positive or zero. you can drop this constraint but in practice we often impose this contraint. Because this is a short selling constraint. If you omit this constraint, it will impact the possible solutions that you end up with.

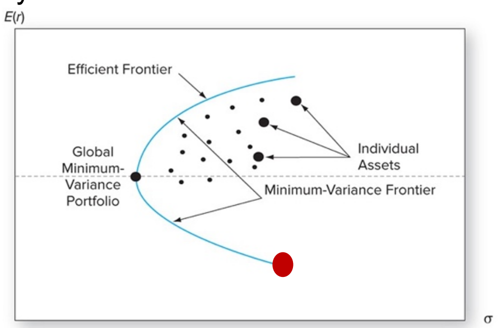
How would the solution look like?

↓

Portfolio with two risky assets: Minimum variance frontier

The portfolios obtained this way delineate the investment opportunity set

They form the **minimum variance frontier**



Here we look at the solution in a graphical way. Now we want to know how we can combine all those assets (black dots) into attractive portfolios.

For example, the red dot:

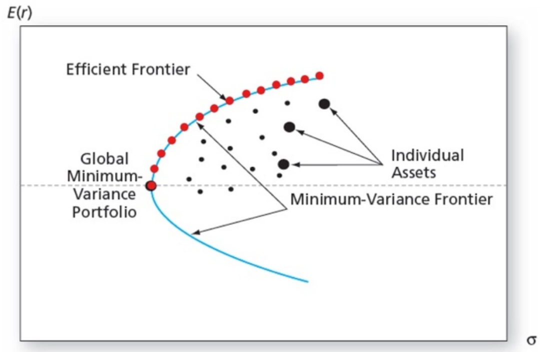
- It allows me to attain an expected return with an amount of risk.
- All the sets on the blue line are the minimum variance portfolio set

The black dot to the most left = Global minimum variance portfolio: the set that allows us to achieve the lowest level of risk. We cannot go lower than this amount of risk.

But not all the portfolios on the blue line are attractive. The lower arm of the curve, have lower target returns with the same amount of risk, they are always dominated by some portfolio in the upper arm.

Portfolio with two risky assets: Efficient frontier

- Clearly, not all combinations on the minimum variance frontier are equally attractive
- Portfolios on the lower arm of the frontier are dominated by the portfolios on the upper arm of the frontier: for each portfolio on the lower arm, there is a portfolio on the upper arm with superior risk-expected return trade-off
- The upper arm frontier is therefore called the **efficient frontier**
- You will never want to go for sets on the inner side (or on the lower arm)
- The variance minimization problem is to define the efficient frontier



Based on this information you should be able to find the efficient part of the frontier (UOVT):

Excel exercise

The investment opportunity set of five stocks

What is your investment opportunity set when following 5 stocks are available to invest in (in % p.a.):

	Stock 1	Stock 2	Stock 3	Stock 4	Stock 5
$E(R)$	11%	16.2%	13.5%	7%	11.2%

Σ	Stock 1	Stock 2	Stock 3	Stock 4	Stock 5
Stock 1	0.046	0.025	0.021	0.013	0.020
Stock 2	0.025	0.054	0.016	0.013	0.018
Stock 3	0.021	0.016	0.049	0.011	0.014
Stock 4	0.013	0.013	0.011	0.020	0.011
Stock 5	0.020	0.018	0.014	0.011	0.042

The optimal portfolio

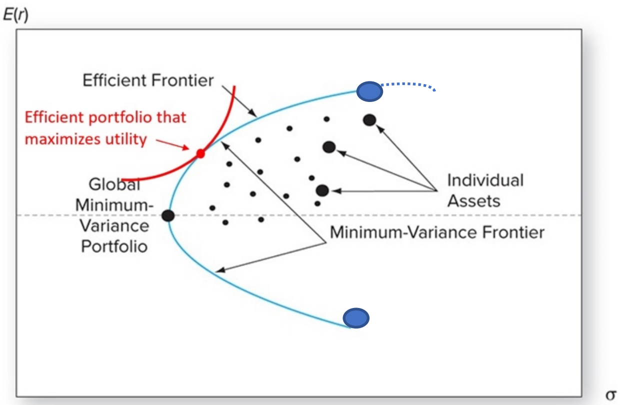
Which portfolio on the efficient frontier should we choose?

- The ultimate portfolio that one chooses on the efficient frontier is determined by the risk preferences: the investor chooses the portfolio *P* that maximizes utility
- To determine this portfolio that maximizes utility, we compute:

$$\max_{w} U = E(R_P) - \frac{1}{2} A \sigma_P^2$$

$$\sum_i w_i = 1 \text{ and } w_i \geq 0 \quad \forall i$$

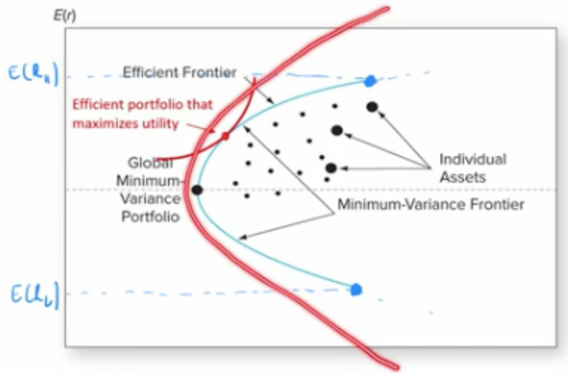
So, you will maximize your utility over those portfolio weights. You will look for the weights that allow you to maximize utility.



Graphically: if you maximize your utility over a set of assets, it boils down to looking for that portfolio where your indifference curve is still tangent to your opportunity set.
For someone who is less risk averse the utility curve might be tangent elsewhere.

We talked about a short selling constraint: what or how would the set look like if I would not impose a short selling constraint?

- If there is one, the set is bounded. What bounds this? (Blue dots)
  - These are not combination of assets but unique assets:
    - The lower bound is that individual asset that has the lowest expected return from all the assets that you choose
    - The upper point will be that asset with the highest return
  - This also means that not all points on the efficient frontier are combinations of stocks. This upper point is a single asset. So, it steps aways from the idea that you are diversifying risk by making a portfolio of assets.
    - It could be that other points on the efficient frontier consists of one or maybe two assets. They do not necessarily consist of many stocks. Because the efficient frontier is a mechanic result of the original minimization problem. So, nothing prevents that the solutions come out very concentrated. (You might come across that with the assignments)
    - This is also one of the critiques of the MV model. This is not desirable, because you know that the input parameters you set are estimated. And you might be wrong about your estimates. So, the solution will be extremely sensitive to your investment parameters. But we will come back to that, there are solutions.
- If there is no one, then it is not bounded if continues (see dotter line here above)
  - Then you can take on negative weights: they can extremely large meaning that the other one can also by extremely positive. There is no bound.
  - You will not be limited to the set on the efficient frontier, it will be wider:



  - So this sets also that the more constraints you impose the tighter your investment set (efficient frontier) will be.

Excel exercise

The optimal portfolio based on five stocks

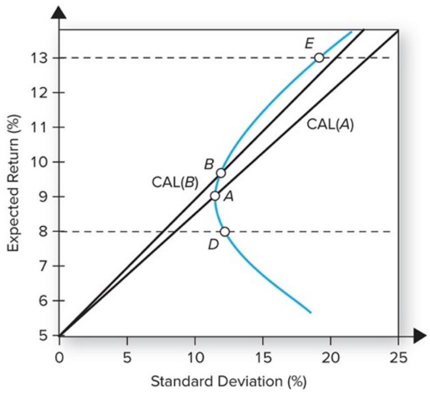
What is the optimal portfolio of the set of 5 stocks introduced earlier, for an investor with mean-variance utility function and a risk aversion parameter of 3 and 6 respectively?

What now if we include a risk free asset?

Mean-Variance analysis with a T-bill

Investment opportunity set

- Expanding the risky asset investment set with a risk-free asset expands the investment opportunity set (0,0 point)
- Each combination of the risky asset investment set with the risk-free asset is now available (cf CAL)



Has the efficient frontier changed?

Suppose you did your first exercise and you have determined your minimum variance frontier
Combining risk-free asset with A:

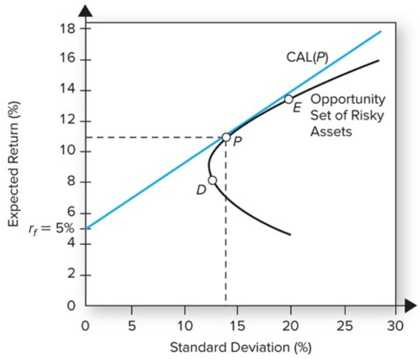
- All the investment opportunities will lie on a straight line that goes to A (CAL(A))
- So, combining these two allows me to achieve the whole line
- Is this strategy attractive?
- You have more opportunities; all these points before (left) A were previously not feasible because the lowest amount of risk was A you could not go lower than that.
- Your investment set is widened (here no frontier)
- What about the upper part on the CAL(A)?
    - Not interesting because those options are dominated by something on the frontier (blue line). All the points within the frontier will be dominated by something on the frontier.
    - So, adding the risk free asset to A is interesting but only for the lower part of the CAL and not the upper part.

If you review this graph, you see that you can do better than choosing A, you choose B:

- You combine the risk free asset with B.
- B is better than CAL(A) because the slope is steeper and the slope of this CAL is the sharp ratio and we said the higher the sharp the more attractive this investment is. But we can do better again:

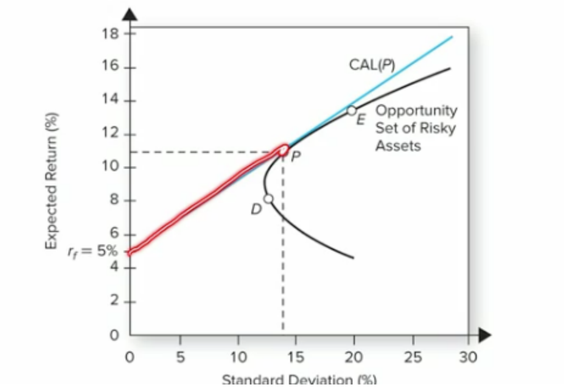
Efficient frontier by the tangency portfolio

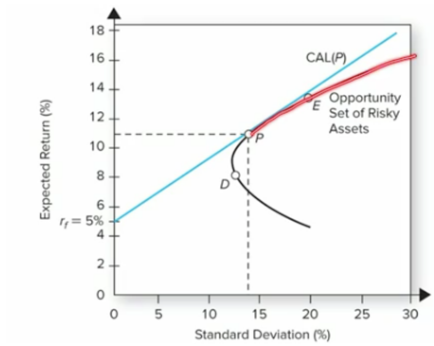
- Based on the Sharpe ratios (slope of the CAL) the combinations on the CAL of B dominate any combinations on the CAL of A
- Continuing this logic eliminates all risky portfolios but the **tangency portfolio P** → none of the combinations on this CAL(P) are dominated: this is the new efficient frontier
- Choose a portfolio on this CAL(P) in line with risk aversion



You push your sharp until you are tangent to your risky set and that is in portfolio P. I cannot push it more because then I will not be tangent anymore. P is our tangency portfolio, the one with the highest SR. All the other combinations have lower SR.
If you compare this scenario with extending with a risk free asset VS just having a risky frontier. Is it worthwhile doing that?

- The efficient frontier is no longer the upper arm of the black curve but the CAL(P) because all of these portfolios are dominant, you cannot find a single portfolio that dominates.
- How does this efficient set changes if you are not allowed to leverage (not short selling we already talked about that)? If you have a leverage constraint how will your efficient frontier look like? Then it would stop at P and continues on the back curve:





  - The upper arm of the black curve become efficient if we have a leveraged constraint because I could not reach the blue upper line. Then I could invest all of my wealth in E.
  - **So, your efficient set will depend on wheter you have a short selling constraint or a leverage constraint.**
  - If leverage can be done your efficient set is your full call, if you cannot leverage your efficient set is the blue line to the P and after the P it's the black curve.

Which portfolio will you choose? The one that maximizes utility. That portfolio is the portfolio y = risk premium/(risk aversion A x variance of P)

Portfolio optimization

The optimal portfolio combines the risk-free asset and the risky asset that has the **highest Sharpe ratio** (i.e. the tangency portfolio) in such proportions that investor **utility is maximized**.

Formally, this maximization problem can be solved in **two steps**:

1. Find the combination of stocks P with the highest Sharpe ratio (find the tangency portfolio). (technical aspect, which is the same for everyone). *Once find, you will combine P with risky asset than:*
2. Choose a portfolio on the CAL(P) in line with investor preferences (personal aspect)

In a first step, we calculate the tangency portfolio.

This tangency portfolio has portfolio weights that result in the highest Sharpe ratio

- We thus need to solve the following maximization problem:

$$\max_{w_i} SR = \frac{E(R_P) - R_F}{\sigma_P} \qquad s.t. \sum_i w_i = 1$$

  where R_P and σ_P are the portfolio return and volatility, respectively

- In addition to this tangency portfolio *P*, we can also define the **complete risky efficiency frontier** (see above)
  - Why not only P? because we know that depending on constraints we impose, part of our risky frontier might be relevant. (leveraged vs short position)

(You can easily do this by excel)
For the case of two risky assets *d* and *e* and a T-bill, the tangency portfolio is obtained with following weights:

$$w_d = \frac{E(R_d^e)\,\sigma_e^2 - E(R_e^e)\,cov(R_d^e, R_e^e)}{E(R_d^e)\,\sigma_e^2 + E(R_e^e)\,\sigma_d^2 - [E(R_d^e) + E(R_e^e)]\,cov(R_d^e, R_e^e)}$$

$$w_e = 1 - w_d$$

Here the superscript *e* indicates returns in excess of the risk-free T-bill return, i.e. $R_i^e = R_i - R_F$

In a second step, we choose a specific portfolio on the CAL(P)

This portfolio combination of the tangency portfolio and the risk-free asset maximizes investor utility:

- We thus solve following maximization problem:

$$\max_{y} U = E(R_C) - \frac{1}{2} A \sigma_C^2$$

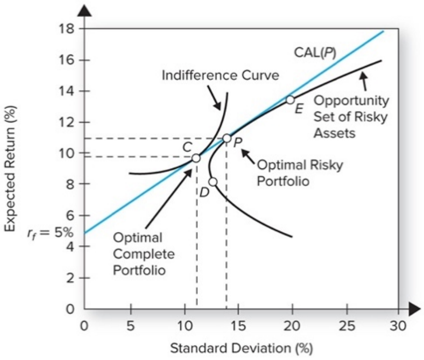
  where ***y*** *(not w)* is the weight allocated to the tangency portfolio; R_C and σ_C are the return and volatility of the complete portfolio, respectively. We want to find our complete portfolio.

- From the previous discussion, we know that this optimal portfolio equals:

$$y^* = \frac{E(R_P) - R_F}{A\sigma_P^2}$$

This is what you need to do for the assignment!

The optimal allocation on the CAL(P) is determined by the indifference curve that is tangent to the efficient frontier



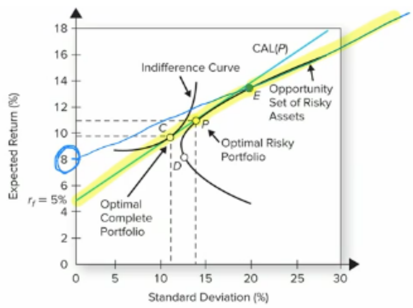
Here a graphical representation. You need to determine opportunity set, global minimum variance portfolio and you can determine portfolio P, the CAL(P). And on this CAL you define the optimal complete portfolio that maximizes utility. This determines the full portfolio selection.

If you can leverage at the same rate at which you invest than the blue curve will be your efficient set.

If you cannot leverage your CAL will go up to P and then we continue on our curve set (black curve).

Imagine leverage is feasible but not at the same rate at which you can invest. You have to pay a higher borrowing rate. Suppose the rate is 8%. Then how does my efficient frontier look like?

- Non leveraged positions are the under yellow curve until P
- Leveraged positions are the above yellow curve from E
- What about in between P and E? I will be located on the risky set, the curved part.
- It could be that depending on your risk preferences the portfolio you hold on to will be different from others: where you are located is in line with your risk preferences (your IC)



Excel exercise

The optimal portfolio based on five stocks and a T-bill

What is the optimal portfolio of the set of 5 stocks and T-bill introduced earlier, for an investor with mean-variance utility function and a risk aversion parameter of 3?

Portfolio management in practice

Input

- To define the optimal portfolio we need a large input list of expected returns and covariances
    - using accurate and consistent estimates is indispensable, but also a great challenge (GIGO)
    - You need to estimate your input parameters: exp returns, correlations. That is not easy because you need. To be consistent. The solution can be very different, it is sensitive to your input parameters. A crucial step is taking the time choosing your input parameters. If you don't do that it's not worth to try to continue with the second step (looking for highest SR…)

- GIGO: garbage in garbage out. If your input parameters don't make sense, your output will also not make sense. The fact that different input parameters can lead to different output is a problem, because those are estimates so we can make forecast errors. This is why you need to do **robustness analysis**.
  - You start off with a set of input parameters, you calculate your efficient frontier for this input parameters. Then you change your efficient frontier, you re do your analysis, many times for different sets of reasonable parameters, … that will give you multiple frontiers, e.g., 200 frontiers. Then you take an average of those frontiers. The portfolio that you end up with will be much more balanced than the portfolios based on a single frontier cause they can be extremely concentrated on the frontier (assignment: very often if you are given a choice of 10 assets you will choose for 3, but when you start playing with the input parameters you will see that not only you want to invest in those 3 but also in other 3). This will smoothe out your portfolio weights over the different assets and you will end up with much more balanced portfolios. The good news is that if you have a much more balanced portfolio, that over time the rebalancing that you need to do will also be limited. Suppose that after 3 months you want to re do your analysis. If you did this in a robust manner the recalibration that is needed to your portfolio will be limited as compared to the situation where your portfolio was based upon a single frontier.

- We can easily adapt the portfolio selection problem to account for different constraints, e.g.
  - on short-sales/leverage
  - on the assets in which can be invested (e.g. only SRI, or a particular regional focus)
  - You can easily change a portfolio selection model by imposing additional constraints. It could be that you are offered a particular investment assets but you have a a particular ESG profile and only a subset of assets will be viable. **The only thing you have to remember** is that each constraint that you impose will limit your investment set. No constraints will give you the widest opportunity set, and each additional constraint will squeeze your opportunity set. So, each constraint comes at a cost.
- Note: each constraint comes at a cost, i.e. the efficient frontier will be restricted and the tangency portfolio will have a lower Sharpe ratio

Separation property

- When starting from the same input list, investors will end up with the same tangency portfolio, and thus identical CAL.
- Degree of risk aversion only comes into play in the capital allocation, i.e. when choosing the optimal point along the CAL(P)
- This result is the **mutual fund separation property** of modern portfolio theory: similar investors choose the same optimal mutual fund (=tangency portfolio); depending on risk preferences, they combine this mutual fund with the risk-free investment in different proportions

  This theorem is from James Tobin (1958) which (amongst others) earned him the 1981 Nobel prize in Economics
  Tobin, J. (1958). Liquidity preference as behaviour towards risk. Review of Economic Studies, 25: 65-86

Own notes:
If you apply the Markovitz model to the asset management industry, we can see that this industry is extremely efficient industry. It is an industry that can easily be scaled to a larger customer base. Why? This tangency portfolio that we are looking for is the portfolio that you will advise to all of your clients. Let's say you are in a scenario where you can leverage and no short selling constraint. The optimal portfolio from my input list is P. All of my clients should be advised P in combination with risky asset. = mutual fund separation property. If you start off from the same input list and you all have the same set of constraint I will all end up advising you some combination of P and a risky asset. That means that this is a business that you can easily scale to a larger client base. Because the only thing that I need to ask you is what is your risk aversion. This determines where on the CAL you will be allocated.

The conclusion is that a portfolio manager will offer the same risky portfolio, P, to all clients regardless of their degree of risk aversion. The client's risk aversion comes into play only in capital allocation, the selection of the desired point along the CAL. Thus, the only difference between clients' choices is that the more risk-averse client will invest more in the risk-free asset and less in the optimal risky portfolio. This result is called a ***separation property;*** *it tells us that the portfolio choice problem may be separated into two independent tasks.*

The first task, determination of the optimal risky portfolio, is purely technical. Given the manager's input list, the best risky portfolio is the same for all clients, regardless of risk aversion. However, the second task, capital allocation, depends on personal preference. Here the client is the decision maker.

Consistency with theory

This theorem was radical for investment advisors:

- The composition of the risky portfolio should be identical for all investors, no matter what degree of risk aversion
- Investment portfolios among different investors should only differ in the proportions invested in the optimal risky portfolio and the risk-free asset

This mutual fund theorem makes professional asset management rather efficient and thus a low-cost business

Is this consistent with conventional investment advice?

Is this in line with the typical investment advice that companies give? Do we see that the same set of clients with same preferences end up being advised the same two assets?

If you go to an investment manager the first question that they ask is not what the risk tolerance is, they immediately start talking about the composition of the risky portfolio. Never about the composition of the complete portfolio. That means that they tie the risk tolerance to the composition of the risky portfolio. And that is in contradiction with what this theory says.

Because my risk preferences do not impact my choice of P. It only impact what fraction of P I hold on to vis a vis my risk free asset. So if you for example are risk tolerant that does not means that you will prefer P over another portfolio. I still prefer P. This is exactly what we see with all asset managers.

Asset allocation puzzle

We see here different advisors and the average of amount of money invested in cash, bonds and stocks. If the mutual fund separation property holds, the amount of bonds and stocks, should be constant in function of different degree of risk aversion. There is no reason to give more bonds to someone who is maybe more conservative vs a moderate investor; the ratio should be identical.
What matters is how do I allocate my wealth of cash VS portfolio P? What we see in reality is that the ratio of bonds to stocks this goes down with risk appetite. The higher the risk appetite the lower the proportion of bonds vs stocks. This means that depending upon the risk aversion, they offer to all clients, different portfolios. To some extent that must be suboptimal. Because some get P others may be getting another portfolio E, … that is counter the portfolio selection model.

	Percent of portfolio			
Advisor & investor type	Cash	Bonds	Stocks	Bonds/stocks
A. Fidelity				
Conservative	50	30	20	1.50
Moderate	20	40	40	1.00
Aggressive	5	30	65	0.46
B. Merrill Lynch				
Conservative	20	35	45	0.78
Moderate	5	40	55	0.73
Aggressive	5	20	75	0.27
C. Jane Bryant Quinn				
Conservative	50	30	20	1.50
Moderate	10	40	50	0.80
Aggressive	0	0	100	0.00
D. The New York Times				
Conservative	20	40	40	1.00
Moderate	10	30	60	0.50
Aggressive	0	20	80	0.25

- CMW find that investment advisors recommend different risky portfolios to investors with a different degree of risk aversion
- The ratio of bonds to stocks is much higher for conservative investors as compared to moderate and aggressive investors

This is inconsistent with the mutual fund separation theorem!
Are investment professionals wrong?

We can come up with some reasons why the portfolio selection model is not in line with reality. Ultimately, we impose allot of assumptions and the solution, the mutual fund separation property, is a consequence of the assumptions we have imposed and maybe they are not realistic. So are there arguments we can put forward about how indeed it might be wise to have varying ratios of bond versus stocks or why this model does not hold?

Horizon effects of risk

- Risk-free investing is not always risk-free: this is only the case when the maturity of the instrument matches the investment horizon
- When maturity and investment horizon differ, one is exposed to:
    - **interest rate risk**: maturity > investment horizon
    - **reinvestment risk**: maturity < investment horizon
- In addition, the majority of risk-free investments are fixed in nominal terms, not in real terms
    - this exposes you to **inflation risk**
    - to protect against such inflation risk, you could buy inflation-protected securities

This could explain the advice that conservative LT investors should hold on to more bonds as they become the risk-free investment

Own notes:
What is the risk free alternative? When is risk free investing really risk free? There are so called horizon effects of risk. What risk free is depends on your investment horizon.
If I have an investment horizon of 1 year than your risk free investment should be a 1 year T bond because than the maturity of my instrument is perfectly aligned with my investment.
If I have a longer investment horizon than this 1-year T bill is no longer risk free because after 1 year I will have to roll over my investment and today I do not know what the future rate will be.

So, it's only when investment horizon is in line with the horizon of instrument, that this T bond is the risk free alternative. So, this means that whenever the maturity of your instrument is longer than your investment horizon you will be exposed to interest rate risk; because you have to sell off your risk free investment before maturity. The price you get is dependent on the than valid interest rate. Whenever the maturity of your instrument is shorter than your investment horizon you will be exposed to roll over risk, you will have to buy a new investment in the future and today you don't know what the rate will be. So, this means that if you have investors with the same preferences but with a different investment horizons that allocations can be indeed different.

How is the risk of stocks over the long run?

There are also horizon effects to riskfree assets. Mainly the risk of stocks is known to vary with the investment horizon. With the current analysis we were not very explicit about the investment horizons nor about the distribution of returns. We go back to the discussion if Sharpe ratios are comparable over time. (we showed that they are not comparable with investment with different investment horizons)

- Also the risk of stocks varies with the investment horizon
  Assume an asset with log-returns r_t whose returns are identically and
- independently distributed (**IID**) with mean μ and variance σ^2
- The mean and variance over 2 periods:

$$E\left(r_t + r_{t+1}\right) = E\left(r_t\right) + E\left(r_{t+1}\right) = 2E\left(r_t\right) = 2\mu$$
$$var\left(r_t + r_{t+1}\right) = var\left(r_t\right) + var\left(r_{t+1}\right) = 2var\left(r_t\right) = 2\sigma^2$$

- Expected log-returns and the variance of log-returns grow linearly over time
- In such cases the MV trade-off is **independent of the horizon**

This result is only valid when you assume IID! If that is not the case this result does no longer hold. This means that as long as my log returns are IID normal I do not really have horizon effects in terms of expected returns and variances. But when my log returns are no longer IID normal I MIGHT have some horizon effects of risk here. This is what we see in reality. We see horizon effects in risk.

- Assume now that logreturns r_t are serially correlated
- The mean and variance over 2 periods:

$$\begin{aligned} E\left(r_t + r_{t+1}\right) &= E\left(r_t\right) + E\left(r_{t+1}\right) = 2\mu \\ var\left(r_t + r_{t+1}\right) &= var\left(r_t\right) + var\left(r_{t+1}\right) + 2cov\left(r_t, r_{t+1}\right) \\ &= 2\sigma^2 + 2cov\left(r_t, r_{t+1}\right) \end{aligned}$$

- Expected log-returns grow linearly over time, but the variance of logreturns no longer grows linearly over time
- In such cases the MV trade-off is dependent of the horizon: investment horizon matters for risk in the presence of serial correlation

"Given this striking results, it might seem puzzling why the holding period has never been considered in portfolio theory. This is because modern portfolio theory was established when the academic profession believed in the random walk theory of security prices. As noted earlier, under a random walk, the relative risk of securities does not change for different time frames, so portfolio allocation does not depend on the holding period. The holding period becomes crucial when data reveal the mean reversion of stock returns."
Source: Siegel (2008), p. 35-36

Does serial correlation increase or decrease risk?

- The impact of serial correlation on risk depends on the sign (and size) of the covariance
- When the covariance is positive, the log-returns are characterized by momentum:
$$var(r_t + r_{t+1}) > 2\sigma^2$$
  → momentum increases long-run risk
- When the covariance is negative, the log-returns are characterized by mean reversion:
$$var(r_t + r_{t+1}) < 2\sigma^2$$
  → mean reversion decreases long-run risk

Do we observe serial correlation that increases or decreases risk?

- Numerous studies have documented either momentum or mean reversion in financial markets, depending on the horizon
- Stocks show momentum over the short horizons of up to 9 months, but mean reversion over longer horizons of 3-5 years (see later)

This could explain the advice that LT investors can hold on to risk (= more stocks), as the stocks are in fact less risky over the longer horizon

Remark: not everyone is convinced about these results; when accounting for factors such as uncertainty and estimation risk, it is less clear that stocks are less risky over the long horizon (see Pastor and Stambaugh (2012), JF)

Own notes:

- *So, when we have horizon effects in risk that means that my covariances are not zero. Otherwise, this independent assumption means that this would be zero. so over the long period of time the variance is not just twice the single period variance, but we have an additional covariance term here. Depending on this covariance term is negative or positive over the longer end is either higher or lower than the shorter end.*
- *What does it mean if the covariance is positive? Then we have momentum. Higher returns yesterday are likely to be followed by higher returns today.*
  If we have momentum than over the longer run equity investing is more risky than on the shorter run. Thus, stocks become more risky on the longer end than shorter end.
- *What does it mean if the covariance is negative? Then low returns yesterday are most likely to be followed by high returns today. This is mean reversion.*
  Implying that investing in the long end reduces the risk.
  → depending upon you have momentum or mean reversion, investing in stocks is more or less risky.
- *Do we observe momentum or mean reversion? Yes! We both observe them. It is the horizon that matters. If we are on the short end up until 6 or 9 months we overserve some momentum. But over longer periods of time, 3-5 years we observe mean reversion. So for very long investors 5 years, you can assume that stock returns show mean reversion. And investing in stocks become than less risky as compared to a long term investor.*
  So you will have different types of portfolio depending on the investment horizon.

Life-cycle approach

- When taking a LT view, the portfolio decision is part of decisions taken over aggregate wealth over the lifetime: this is the life-cycle approach of LT investing in which portfolio allocation depends on the age of the investor
- Crucial in this approach is that overall wealth not only depends on the investment portfolio, but also on **human capital** (and thus labour income)
- In this view, younger investors can take on more risk (invest in more stocks)
    - they can compensate realizations of high risk by working harder (see Bodie et al. (1992), JEDC)
    - human wealth (risk-free or uncorrelated with stocks) is a large fraction of younger investors' wealth portfolio (Campbell and Viceira (2002), Jagannathan and Kocherlakota (1992), JF)

If you think about this approach than you think about your investment in stocks and bonds as part of your overall wealth. Important part of your wealth is not only the money you have at the bank but also your human capital. You want to maximize your capital over your full life cycle. When you are young you have allot of human capital and limited amount of capital. When you get older your human capital goes down.

Very often human capital is assumed to be risk free or at least uncorrelated with risky assets. That means that if you would plot your overall portfolio, for someone who is young will have allot of HC and limited amount of C. This capital you can invest in risky assets. For the same amount of risk aversion, it means that you can be tilted to very risky assets as compared to the older person. Because you hold allot of your wealth in risk free and the older person hold only a limited amount in the risk free asset. That is a life cycle approach of investing. Your human capital could be correlated to your capital if you invest al of your wealth into the company you work for. If the company is not going good, you might be fired and your stocks will go down.

"Financial planners typically advise people to shift investments away from stocks and toward bonds as they age. The planners commonly justify this advice in three ways. They argue that stocks are less risky over a young person's long investment horizon, that stocks are often necessary for young people to meet large financial obligations (like college tuition for their children), and that younger people have more years of labour income ahead with which to recover from the potential losses associated with stock ownership. This article uses economic reasoning to evaluate these three different justifications. It finds that the first two arguments do not make economic sense. The last argument is valid—but only for people with labour income that is relatively uncorrelated with stock returns. If a person's labour income is highly correlated with stock returns, then that investor is better off shifting investments toward stocks over time."

Source: Jagannathan and Kocherlakota (1992), JF p.11)

Chapter 5: Index models

Readings
BKM – Chapter 8: Index models

Can mean-variance optimization be used in practice?
→ active or passive management? This model is active management because you have to come up with input parameters and based upon that data you maximize your SR and that determines the different weights you allocate to your portfolio.

Can MV optimization be used in practice?

We should be critical about this optimization. Even though many MV statistical software exists, serious questions can be raised as to whether they work well:

1. **Obtaining reliable and accurate estimates of returns is hard** (cf GIGO)
   Inconsistent/noisy inputs lead to unreasonable solutions or no solution
   → sensitivity of the model to the inputs you give: This means you have to be careful in estimating and putting your input parameters.
   → Consistent estimates: if you have a large set of assets and you define expected returns, to some extend these expected returns will be linked to risk. If you have two companies with very similar expected returns that says something about their risk profile. In the same way if you have correlation dynamics between two assets that also says something about the correlation dynamics.
   → You have to think careful about the input parameters
2. **Estimating the covariance matrix requires many estimates** (for a portfolio of N assets, the covariance matrix has $N(N+1)/2$ inputs)
   → this is about the practical implementation. The number of parameters that you have to estimate explodes when the number of assets in your portfolio increases. This makes the mean variance optimization not so practical to implement.
   Suppose you have N assets in your portfolio; than your covariance matrix has $N(N+1)/2$ inputs. Example: a portfolio of 200 assets has a covariance matrix with 20,010 (20k) inputs
3. **When the number of assets to include is large, while the historical data availability is limited ($N > T$), the covariance matrix is singular,** the MV optimization then detects arbitrage opportunities, with unreasonably concentrated portfolios as a consequence even when the number of assets to include is large relative to the historical sample length (still $N < T$), the MV optimization will perceive near-arbitrage opportunities
   → so, when the number of assets is large relative to your time series dimension that you have available than you might also have estimation issues. The var-cov matrix is typically estimated based upon historical data and when the cross section is larger or very large vis a vis the time series dimension you might end up in a covariance matrix at a singular. Then the outcome that you will have: either you don't find an outcome or your outcome would be very unrealistic.

Such difficulties and limitations have initiated a search for more **practical solutions** to find optimal portfolios

- The baseline of these solutions is to **introduce more structure** on the return dynamics in order to shorten the input list
- Different approaches can be distinguished in the literature
  - statistical-based index models (single-index models and multi-index models)
    - here you will attain the most structure
    - more general approach
  - equilibrium pricing models (CAPM and APT)

        - this is only valid in equilibrium
- Both approaches share the idea that risk can be decomposed in **systematic risk** and **idiosyncratic risk**
- They are not identical because in the CAPM we are in equilibrium and all assets are priced with fundamental value. This is not the case in a statistical based single index model; here you can allow for under or over pricing.
- Even though the models resemble one another the fact that you allow under or overpricing in the single index model, makes that the implications on the portfolio selections angle very different.

In the current course, we focus on the statistical models

Single-index model

A statistical model

Assume we have an asset i with returns R_i

$$R_i = \beta_i m + \varepsilon_i + \mu_i$$

with
- R_i are joint normally distributed
- m is a common risk factor → **systematic risk**
- β_i is the response of asset i to the common factor/index
    - *each asset has its own sensitivity to the systematic risk factor*
- ε_i is a random variable with mean 0 and variance σ^2 and uncorrelated with the common factor → **idiosyncratic risk**
- μ_i is the baseline return = deterministic
    - *the risk free rate at least should be included here*

To make this model operational, we impose **intuitive economic structure**: the common risk factor m is proxied by a broad market index (e.g. S&P500 for US assets; or Euronext100 for Belgian assets). Then we can translate this model into a statistical model.
Because the systematic factor affects the rate of return on all stocks, the rate of return on a broad market index can plausibly proxy for that common factor. This approach leads to a model which is called the **single-index model** because it uses the market index to stand in for the common factor.

The single-index model can then be rewritten as a regression equation:

$$R_{i,t}^e = \beta_i R_{m,t}^e + \varepsilon_{i,t} + \alpha_i$$

With

- $R_{i,t}^e$ excess returns for asset i over time t
- $R_{m,t}^e$ excess returns on a broad market index over time t (generating systematic risk)
- β_i the response of asset i to the systematic risk factor
    - Beta is the amount by which the security return tends to increase or decrease for every 1% increase or decrease in the return on the index, and therefore measures the security's sensitivity to the market index
- $\varepsilon_{i,t}$ a residual with mean 0 and variance $\sigma_{\varepsilon i}^2$ and cov(ε_i, R_m^e) = 0, and cov$(\varepsilon_i, \varepsilon_j) = 0$ (capturing idiosyncratic risk)
- $\alpha_i = E\left(R_{i,t}^e \middle| R_{m,t}^e = 0\right)$ the expected excess return when the market excess return is zero (=constant)

- This is part of our baseline return in the previous model but here we see that the baseline return of the previous model is entangled into two components: R_f+α_i. So the μ_i consists of R_f +α_i. (see here under)
- How does this relate to the CAPM? $\alpha_i = 0$ in CAPM.
- Here in this model, I allow for some intercept. So, I allow for over or underpricing. Whilst in equilibrium (CAPM) α_i should be zero because you know that in equilibrium if you are not exposed to risk, you should only earn the risk-free return. Meaning that the excess return should be zero. But we allow for the excess rate of return to be equal to α_i.
- If α_i > 0 than this asset is underpriced.
- If $\alpha_i < 0$ it means that this stock is overpriced.

(You can estimate this model with OLS.)

So how does this model put structure?

We look at the expectation of an excess return. (when you take the expected value than $\varepsilon_{i,t} = 0$ because the $E(\varepsilon_{i,t}) = 0$).
→ This equals the risk premium. The security's risk premium then equals:

$$E(R_{i,t}^e) = \beta_i E(R_{m,t}^e) + \alpha_i$$

This risk premium can be decomposed into two parts

1. The systematic risk premium $\beta_i E(R_{m,t}^e)$: you expect to earn a return in excess of the risk free rate because you are exposed to systematic risk and you need to be rewarded for this exposure
2. A non-market premium α_i or **alpha**: it cannot persist in equilibrium as it is arbitraged away (superior security analysis lies in the identification of α_i investments): you expect to earn an excess return in function of the under or overpricing α_i.

→ This means that the security risk premium is decomposed into a systematic risk premium and into a non-market risk premium.

What about risk and covariance?

The security's variance then equals:

$$\sigma^2(R_{i,t}^e) = \sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{\varepsilon i}^2$$

Here we take the variance of the excess returns. An asset's risk can be decomposed into two parts

1. Systematic variance $\beta_i^2 \sigma_m^2$.
2. Idiosyncratic variance $\sigma_{\varepsilon i}^2$.

The fact that there is no covariance term is due to the fact that the idiosyncratic risk component $\sigma_{\varepsilon i}^2$ has zero covariance with the market risk component $\sigma_{\varepsilon i}^2$.
The approach is similar to the previous chapters but different in the sense that we can now compute the systematic and idiosyncratic risk.

Now we see that the index model vastly reduces the number of parameters that must be estimated.

Characteristics of the index portfolio
What is the β, firm-specific risk and α of the index portfolio m?

This means you have to define

$$R_m^e = \alpha_m + \beta_m R_m^e + \varepsilon_m$$

Not just intuitively but you also need to derive this!
If you have your OLS model $R_m^e = \alpha_m + \beta_m R_m^e + \varepsilon_m$

1) $\beta_m = \frac{cov(independent\ var, dependent\ var)}{var(independent\ var)} = \frac{cov(R_m^e, R_m^e)}{var(R_m^e)} = 1\ (for\ this\ model)$
2) $E(R_m^e) = E(\alpha_m + \beta_m R_m^e + \varepsilon_m) = \alpha_m + 1E(R_m^e) + 0$

   $E(\alpha_m) = \alpha_m$ because this is deterministic
   $E(\beta_m) = 1$ we have proven this above
   $E(\varepsilon_m) = 0$ because we do not expect this risk

   So $E(R_m^e) = \alpha_m + E(R_m^e)$ and this equation can only hold if $\alpha_m = 0$. Otherwise, this equality can never hold.

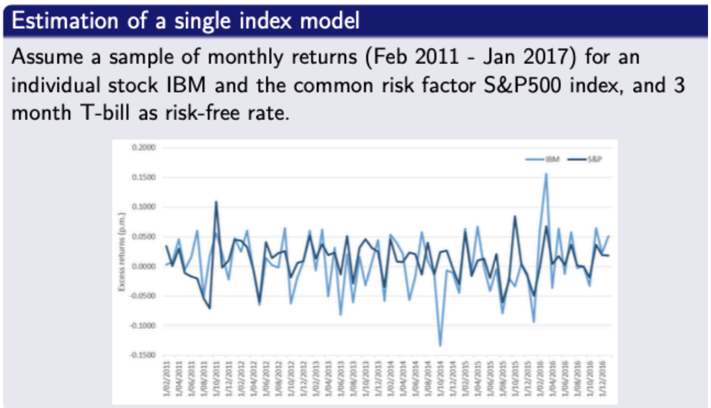
3) What is ε_m?

$$R_m^e = \alpha_m + \beta_m R_m^e + \varepsilon_m$$

   We know that $R_m^e = 0 + 1R_m^e + \varepsilon_m$
   Again, this equation can only hold if $\varepsilon_m = 0$.

Estimation

Before starting estimation your model, you should **visualize your data**. We construct a regression and get the **security characteristic line (SCL).** This line describes the (linear) dependance of IBM's excess returns on the excess return of the market index portfolio. The intercept is alpha, and the slope is beta (measure of **systematic risk**).
The vertical distance of each point from the regression line is the value of IBM's residual, this is a measure **of firm-specific risk**.



So, we have plot it in the graph here. This is an idea of the data you have, and this allows to look at the outliers or potential errors in your data.
If you look at both excess returns:

- IBM mimic a bit the S&P returns. But the variability in the IBM seems to be larger.

In a next **phase we look at the** descriptive statistics. Observations:

- S&P returns are higher than IBM returns
- The total risk in terms of volatility is higher on IBM stock than on the market
- The higher volatility can also be observed in the minima and maxima
    - IBM stock have lower minima and higher maxima = spread is large
- Correlation: they both behave alike but by far the match is not 100% so the IBM do not fully track the S&P
- The standard error (not in this table) of the regression is the SD of the residual. High standard errors imply greater impact of firm-specific event from one time period to the next.

Estimation of a single index model

- IBM excess returns follow S&P500 excess returns, but there is important firm-specific movement in returns → the correlation is 51% implying that IBM does not track changes in the returns of the S&P500 very closely
- Volatility of IBM excess returns seems slightly larger → see the sample volatility and min/max observations

	R^e_{IBM}	$R^e_{S\&P500}$
Mean	0.0041	0.0102
St.Dev	0.0480	0.0329
Min	-0.1340	-0.0703
Max	0.1557	0.1093
Correl	51.24%	

Here we do the **OLS regression**:

- Beta is smaller than 1 implying that the movements in the S&P are not transferred to the IBM stock to the fullest, so the effect is dampened a bit
- Alpha is very small, negative; but we cannot say anything about it because it is not statistically significant
    - Not unexpected result: if you run such regressions many individual stocks, alpha will typically not be significant.
    - Alpha will not be determined using OLS in practice
    - Beta will be determined using historical analysis but alphas not. Alpha is a result of a thorough security analysis.
- Explained variances (**adjusted R²**): this gives an idea of the amount of variance that is explained by the model. So, in this case we have 25%. What about the rest? The rest is idiosyncratic. So, the R^2 gives you allot of information. **It allows you to disentangle systematic and idiosyncratic risk**. R^2 = (explained var / total var).
    - You can characterize your explained variance and total variance in function of the model parameters.
    - Explained variance: $\beta_i^2\sigma_m^2$
    - Total variance: $\beta_i^2\sigma_m^2 + \sigma_{\varepsilon i}^2$
    - $\frac{\beta_i^2\sigma_m^2}{\beta_i^2\sigma_m^2+\sigma_{\varepsilon i}^2}$
- So, once you have estimated the single index model you have an idea of what the Beta is. You know what your variance is on the market. Then you can calculate this idiosyncratic variance as left-over term. You can play around with this term.

Estimation of a single index model

	Estimate	T-stat
α	-0.0035	-0.6864
β	0.7482***	4.9917
$\overline{R}^2$	25.20%	
Observations	72	

- β is smaller than 1 indicating a relatively low market sensitivity
- α is negative, though not statistically significant
- $\overline{R}^2$ is only 25% which implies that the variation in the index returns explains about 25% of the variation in the IBM returns; the remainder is idiosyncratic risk

Advantages

This single-index model greatly **shortens the input list** in a MV optimization: for a portfolio of N assets, only (3*N* + 2) estimates are needed:

- N non-market premia (α)
- N betas
- N firm-specific variances $\sigma_{\varepsilon i}^2$
- 1 market risk premium $E(R_{m,t}^e)$
- 1 systematic risk variance σ_m^2

This is clearly a great reduction in the number of parameters to estimate. You have a gain in number of parameters compared to our previous MV method. Of course, if you only have 2-3 assets in your portfolio than the single index model is not really a gain. It's only when the number of assets to include becomes larger, than the bigger the gain will be.

What about the covariances of returns?

Here covariances are not included: your covariances can be derived from this set of parameters. So no extra need to estimate your covariances.

No additional parameters are needed to obtain estimates of the covariances. We can easily prove this, by deriving expressions for the different dimensions of risk

Variance of asset i:

$$\sigma^2\left(R^e_{i,t}\right) = \sigma_i^2 = \beta_i^2\sigma_m^2 + \sigma_{\varepsilon_i}^2$$

Covariance between asset i and j:

$$cov\left(R^e_{i,t}, R^e_{j,t}\right) = cov_{ij} = \beta_i\beta_j\sigma_m^2$$

Correlation between asset i and j:

$$\rho\left(R^e_{i,t}, R^e_{j,t}\right) = \rho_{ij} = \frac{\beta_i\beta_j\sigma_m^2}{\sigma_i\sigma_j} = \frac{\beta_i\beta_j\sigma_m^2\sigma_m^2}{\sigma_i\sigma_j\sigma_m^2} = \rho_{im}\rho_{jm}$$

You can rewrite this expression of the correlation for between two correlations. So to characterize the correlation between i and j, I need the correlation of i and the correlation between j and the market. And the product between both describes the correlation between both stocks.

But what is so attractive about this expression? Note that these risk expressions clearly indicate that **specialization of security analysis** within e.g. industries in such index-model is possible. It allows for specialization in this asset management industry/sector.

Example: suppose I am financial analyst of the automobile sector. I might know my sector but I do not necessary know the food sector. How do I come up with a correlation of a stock of the automobile sector and the food sector. Without structure you need to have someone that is specialized in both sectors. Know that we have imposed a single index model you do not need to be an expert in both sectors. You can have one sector for the automobile sector and one for the food sector. Because each individual analyst will come up with his expertise advise. This automobile stock is correlated with the market and the other analyst will come up with a correlation of the food stock with the market. And combining those two estimates allows you to come up with a correlation between those both stocks.

Side note: you have to be able to compute that cov_{ij} is equal to those Beta's.

The single-index model also offers important insights into **the benefits of diversification**

Assume the single-index model for a portfolio of N assets:

$$R_i^e = \alpha_i + \beta_i R_m^e + \varepsilon_i$$

For an equally weighted portfolio of N assets, it then holds that:

$$R_p^e = \alpha_p + \beta_p R_m^e + \varepsilon_p$$

with

$$\alpha_p = \frac{1}{N}\sum_{i=1}^{N}\alpha_i \quad \beta_p = \frac{1}{N}\sum_{i=1}^{N}\beta_i \quad \varepsilon_p = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i$$

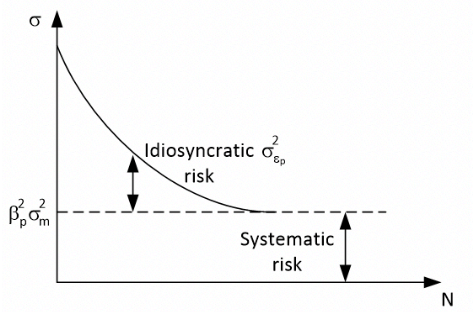
To understand the benefits of diversification, we need to analyze the portfolio variance:

$$\sigma_p^2 = \beta_p^2\sigma_m^2 + \sigma_{\varepsilon p}^2$$

1. The systematic risk component ($\beta_p^2\sigma_m^2$) depends on the individual sensitivities and the variance of the common factor; this part will be there no matter how many or which assets are added. That is why you get a return for this risk.
2. The idiosyncratic component ($\sigma_{\varepsilon p}^2$) depends on the firm-specific risks; as they are independent (all with zero mean), the law of averages implies that for increasing N these components cancel out:

$$\sigma_{\varepsilon_p}^2 = \frac{1}{N}\left(\frac{1}{N}\sum_{i=1}^{N}\sigma_{\varepsilon_i}^2\right) = \frac{1}{N}\overline{\sigma}^2 \overset{N\to\infty}{=} 0$$

The conclusion is the same as before: you will always be stuck with systematic risk. You can characterize this type of risk in terms of the Beta's and the market risk. Idiosyncratic risk can be diversified.



As a function of N (number of stocks in portfolio), idiosyncratic risk goes to zero, but systematic risk remains there. The difference with this graph in the previous chapters is that you can know characterize that systematic risk.

Can we construct a portfolio with zero systematic risks?
In a long portfolio this would be very hard because the betas are positive. To have systematic risk = 0 you need to have a beta on the portfolio that is zero. But with almost all beta's being positive, that is very hard. The only way to do this is to go for a long-short position. That is a typical trading strategy that some hedge fund managers use: two similar companies, the exposure to the market in terms of the beta's is almost identical, and I go long in one company, and I go short in the other company. What is than the Beta of your portfolio? 0.

Why would a hedge fund manager go for such a position? If beta is zero, you know that the return associated with the systematic risk is also zero because they cancel out. They are not interested in systematic risk, so they are not interested in the reward they get for the systematic risk. This is exactly why they go for such a long-short position. They are interested in the respective Alpha's. Because what is the expected (excess) returns? It's the following

+ $E(R_i^e) = \alpha_i + \beta_i E(R_i^e)$
- $\underline{E(R_j^e) = \alpha_j + \beta_j E(R_j^e)}$

α_i

If I go long in one and short in the other than of course the beta's drops out. What I am exposed to is than the difference between those two Alpha's. So, if I think that stock i is underpriced, I go long in that stock. But I am only interested in Alpha, so I look for a stock that is very similar such that systematic risk is canceled out, so I go short in the other stock. That allows me to remove that exposure and just be stuck to the Alpha's. Suppose α_j is zero than you are left with α_i.

Portfolio construction in a single-index world

The main advantage of the single-index model is that less parameters need to be estimated:
1. It allows for a specialization and organizational decentralization, without loss of consistency
2. The estimation of the market risk premium is dis-entangled from the estimation of the non-market risk premia
3. Common risk is dis-entangled from idiosyncratic risk

The input list of a single-index model in a portfolio allocation model consists of
1. Macro-economic analysis to estimate the risk and risk premium of the market index
2. Statistical analysis to estimate expected returns and risk of individual securities, absent any security analysis
   - estimates of the β sensitivities of the individual portfolio constituents
     - can be estimated by an OLS regression
   - estimates of the residual variances $\sigma_{\varepsilon i}^2$
     - can be estimated by an OLS regression
3. Security analysis to generate incremental (alpha) risk premia or non-market returns
   - These are estimated by security analysts who perform thorough analysis

Assume an investment set consisting of **N individual risky securities** The optimal risky portfolio is found as:

$$\max_{w_i} S_p = \frac{E\left(R_p^e\right)}{\sigma_p} \qquad s.t. \sum_{i=1}^{N} w_i = 1$$

with:

$$E\left(R_p^e\right) = \alpha_p + E\left(R_m^e\right)\beta_p = \sum_{i=1}^{N} w_i \alpha_i + E\left(R_m^e\right)\sum_{i=1}^{N} w_i \beta_i$$

$$\sigma_p = \left(\beta_p^2 \sigma_m^2 + \sigma_{\varepsilon_p}^2\right)^{\frac{1}{2}} = \left(\left(\sum_{i=1}^{N} w_i \beta\right)^2 \sigma_m^2 + \sum_{i=1}^{N} w_i^2 \sigma_{e_i}^2\right)^{\frac{1}{2}}$$

Once you have those parameters you continue like we did in the previous chapters. We look for the tangency portfolio, the one who has the highest SR. Allocating those weights to the different stocks such that the SR is the steepest. The only difference is that there is structure behind these expected returns and the volatility.

This is still active management. Suppose that instead being in a single index model, we are in a CAPM world:

Portfolio construction in a CAPM world

Mean-variance optimization, CAPM and passive investment
Assume a world in which:
1. Investors are alike: they are mean-variance optimizers with common time-horizon and common set of information as reflected in input estimates
2. Markets are well-functioning: there are no frictions and impediments to trading (borrow at the same Rf for example)

→ This leads to a very different tangency portfolio.

It then follows that:
- All investors calculate an identical efficient frontier
- All investors identify an identical CAL
- All investors arrive at an identical tangency portfolio

The tangency portfolio is the (market-weighted) **market portfolio** and CAPM holds such that there is no α in equilibrium → If I know we will all end up in the market portfolio than there is no need for me to do active strategy → investors can skip the trouble of MV and security analysis and obtain an efficient portfolio by just holding the market portfolio.

The passive strategy is efficient!

The CAL of the market portfolio is labeled **CML = Capital Market Line**

So you see that imposing the difference between a single index model and CAPM is the choice of choosing between active or passive management.

- If I impose single index model, I allow active management
- If I impose an equilibrium model, there cannot be Alpha; it is not as if there is an asset which we all believe the Alpha would be positive or for which we would all believe the Alpha is negative.

That does not mean that at one point in time you are smarter than all others and you know that a particular asset has an Alpha, if that's the case you want to seize that Alpha and engage in active management. If you believe that we are in an equilibrium world and that there are no Alpha's to seize than you just hold on to the market portfolio. That means that you buy a market-based portfolio: S&P 500, …
So the passive investing comes from a belief of **market efficiency**, if you believe in that than holding on to the market is the best you can do. So it has a theoretical foundation. But in practice the market is not fully efficient but however still this passive investing is popular and attractive. Because the effort, the costs, all play in the advantage of passive investing = it's low effort and costs. This compensates for the fact that even in reality passive investing might not be as optimal as CAPM explains, still it might be attractive because of the low cost and the low effort.

Portfolio construction in a pragmatic way

Can you not combine the best of the two worlds? Because active management is to some extent risky. You will only engage in active management if you believe you are better informed than others. But you should not forget about diversification benefits!

So on the one hand try to exploit some market inefficiencies but on the other hand also profit from diversification. This is why in practice you do not see investors going for a full 100% portfolio in a single asset, they will be more careful.

Does the CAPM hold, at all times? NO!
Is the CAPM a reasonable benchmark model? YES!

A pragmatic approach therefore consists of combining:

- the **diversification benefits** and cost-efficiency of a passive portfolio
- security analysis targeted at **outperformance** of an active portfolio

Assume an investment set consisting of **N+1 securities**:

- 1 market portfolio → the market portfolio can be thought of as the **passive portfolio**
- N individual risky assets → these individual securities can be combined into an **active portfolio**

For example: your portfolio consists of a tracker of the S&P 500 and a number of companies of the S&P 500. The investment in the S&P500 allows you to profit from diversification = passive part. You might have superior information about a number of companies in the S&P500, and in addition to this tracker you will also hold on to an active portfolio in which you try to exploit your Alpha's.

The portfolio problem then reduces to solving for:

1. the optimal weights of this passive and active portfolio (→ e.g., 80%/20%)
2. the optimal weights within the active portfolio

As before...

The optimal risky portfolio is found as:

$$\max_{w_i} S_p = \frac{E\left(R_p^e\right)}{\sigma_p} \qquad s.t. \sum_{i=1}^{N+1} w_i = 1$$

$$E\left(R_p^e\right) = \alpha_p + E\left(R_m^e\right)\beta_p = \sum_{i=1}^{N+1} w_i \alpha_i + E\left(R_m^e\right) \sum_{i=1}^{N+1} w_i \beta_i$$

$$\sigma_p = \left(\beta_p^2 \sigma_m^2 + \sigma_{\varepsilon_p}^2\right)^{\frac{1}{2}} = \left(\left(\sum_{i=1}^{N+1} w_i \beta\right)^2 \sigma_m^2 + \sum_{i=1}^{N+1} w_i^2 \sigma_{e_i}^2\right)^{\frac{1}{2}}$$

Maximize SR over N+1 assets (instead of N). that will give you the optimal weights and to different individual assets that you can combine in an active portfolio.

Unlike before...

Here you are not obliged to do this optimization. There exists closed formed solution to this: **analytical expressions** can be derived for the optimal risky portfolio, making the traditional MV optimization superfluous.

It can be shown that the optimal risky portfolio in the single-index model is a combination of

- The **active portfolio** (denoted a)
- The **passive market portfolio** (denoted m)

See BKM for a step-by-step procedure of the analytical optimization procedure: (**Treynor black model**)

Once security analysis is complete, the optimal risky portfolio is formed from the index-model estimates of security and market index parameters using these steps:

1. Compute the initial position of each security in the active portfolio as $w_i^0 = \alpha_i/\sigma^2(e_i)$.
2. Scale those initial positions to force portfolio weights to sum to 1 by dividing by their sum, that is, $w_i = \dfrac{w_i^0}{\sum_{i=1}^{n} w_i^0}$.
3. Compute the alpha of the active portfolio: $\alpha_A = \sum_{i=1}^{n} w_i \alpha_i$.
4. Compute the residual variance of the active portfolio: $\sigma^2(e_A) = \sum_{i=1}^{n} w_i^2 \sigma^2(e_i)$.
5. Compute the initial position in the active portfolio: $w_A^0 = \left[\dfrac{\alpha_A/\sigma^2(e_A)}{E(R_M)/\sigma_M^2}\right]$.
6. Compute the beta of the active portfolio: $\beta_A = \sum_{i=1}^{n} w_i \beta_i$.
7. Adjust the initial position in the active portfolio: $w_A^* = \dfrac{w_A^0}{1 + (1-\beta_A) w_A^0}$.
8. Note: The optimal risky portfolio now has weights: $w_M^* = 1 - w_A^*; w_i^* = w_A^* w_i$.
9. Calculate the risk premium of the optimal risky portfolio from the risk premium of the index portfolio and the alpha of the active portfolio: $E(R_P) = (w_M^* + w_A^* \beta_A) E(R_M) + w_A^* \alpha_A$. Notice that the beta of the risky portfolio is $w_M^* + w_A^* \beta_A$ because the beta of the index portfolio is 1.
10. Compute the variance of the optimal risky portfolio from the variance of the index portfolio and the residual variance of the active portfolio: $\sigma_P^2 = (w_M^* + w_A^* \beta_A)^2 \sigma_M^2 + [w_A^* \sigma(e_A)]^2$.

Extra explanation:

Alpha: A positive-alpha security is a bargain and therefore should be overweighted in the overall portfolio. Conversely, a negative-alpha security is overpriced and, other things equal, its portfolio weight should be reduced.

Assume first that the active portfolio has a beta of 1. In that case, the optimal weight in the active portfolio would be proportional to the ratio $\alpha_A/\sigma^2(e_A)$. This ratio balances the contribution of the active portfolio (its alpha) against its contribution to the portfolio variance (via residual variance). The analogous ratio for the index portfolio is $E(R_M)/\sigma^2{}_M$, and hence the initial position in the active portfolio is:

$$w_A^0 = \frac{\alpha_A/\sigma_A^2}{E(R_M)/\sigma_M^2}$$

Next, we adjust this position to account for the actual beta of the active portfolio. For any level of $\sigma^2{}_A$, the correlation between the active and passive portfolios is greater when the beta of the active portfolio is higher. This implies less diversification benefit from the passive portfolio and a lower position in it. Correspondingly, the position in the active portfolio increases. The precise modification for the position in the active portfolio is:

$$w_A^* = \frac{w_A^0}{1 + (1 - \beta_A)w_A^0}$$

Notice that when $\beta_A = 1, w_A^* = w_A^0$.

The Sharpe ratio of a risky portfolio with weights w_a and $w_m = 1 - w_a$ exceeds the Sharpe ratio of the passive-only portfolio as follows:

$$S_p^2 = S_m^2 + (\frac{\alpha_a}{\sigma_{\varepsilon_a}})^2$$

- The active portfolio contributes to the Sharpe ratio by the ratio of its alpha to its residual standard deviation, known as the **information ratio: active return over tracking error**
- It measures the **extra return** from security analysis, compared to its **idiosyncratic risk** when we over-or underweight securities relative to the passive portfolio → alpha comes at a cost of increased variance!
- So, if you have an asset that has an Alpha you want to seize that Alpha. However, seizing that Alpha means you need to allocate wealth to that Alpha. By increasing the portfolio weight to that asset, you deviate from the optimal passive weights, and optimal diversification. You expose yourself to additional idiosyncratic risk.
- You want **that portfolio** for which this tradeoff is optimized
- So seizing Alpha comes at a cost. This explains why you will not go 100% in a single asset with positive Alpha.

To achieve the highest Sharpe ratio possible, this information ratio needs to be maximized

This result has a compelling interpretation: It says that the position in each security will be proportional to its ratio of alpha (which investors seek) to diversifiable risk (which they wish to avoid). The higher the ratio, the more of the security they will hold in the active portfolio. Scaling this ratio so that the total position in the active portfolio adds up to w^*_A, the weight in each security is

$$w_i^* = w_A^* \frac{\alpha_i / \sigma^2(e_i)}{\sum_{i=1}^{n} \frac{\alpha_i}{\sigma^2(e_i)}}$$

The model thus reveals the central role of the information ratio in efficiently taking advantage of security analysis.

- The positive contribution of a security to the active portfolio is made by its addition to the non-market risk premium (its alpha).
- Its negative impact is to increase the portfolio variance through its firm-specific risk (residual variance).

In contrast to alpha, the market (systematic) component of the risk premium, $\beta_i E(R_M)$, is offset by the security's nondiversifiable (market) risk, $\beta^2_i \sigma_M^2$, and both are driven by the same beta. This trade-off is not unique to any security, as any security with the same beta makes the same balanced contribution to both risk and return. Put differently, the beta is a property that simultaneously affects the risk and risk premium of a security.

We see from the equation above that if a security's alpha is negative, the security will assume a short position in the optimal risky portfolio. If short positions are prohibited, a negative- alpha security would simply be taken out of the optimization program and assigned a portfolio weight of zero. As the number of securities with nonzero alpha values increases, the active portfolio will itself be better diversified and its weight in the overall risky portfolio will increase at the expense of the passive index portfolio.

Finally, we note that the index portfolio is an efficient portfolio only if all alpha values are zero. Unless security analysis reveals that a security has a nonzero alpha, including it in the active portfolio would make the portfolio less attractive. In addition to the security's systematic risk, which is compensated for by the market risk premium (through beta), the security would add its firm-specific risk to portfolio variance.
With a zero alpha, however, there is no compensation for bearing that firm-specific risk. Hence, if all securities have zero alphas, the optimal weight in the active portfolio will be zero, and the weight in the index portfolio will be 1. However, when security analysis uncovers securities with non-market risk premiums (nonzero alphas), the index portfolio is no longer efficient.

Characteristics of the optimal risky portfolio
What are the allocation implications for a security with negative *α*?

Suppose a stock with negative Alpha: you want to short sell because you want to decrease the weight in those stocks. Go long in a positive alpha stock and go short in a negative alpha stock. This goes also back to the information ratio. Even for a negative Alpha that can increase the Sharp. Those negative alpha stocks are overpriced stocks so you will go short in those ones.

When is the market portfolio m efficient?

If there is no Alpha. Then holding on to the market is the best you can do. If the market is priced efficiently and thus there is no alpha than you should hold on to the market:

$$S_p^2 = S_m^2$$

The mean-variance model: a critical note

We impose allot of assumptions and they are not always realistic…

	μ	σ	S	excess K	Min	Max
			Panel A: Daily returns			
CRSP index (value weighted)	0.044	**0.82**	**-1.33**	**34.92**	-18.80	8.87
CRSP index (equal weighted)	0.073	**0.76**	**-0.93**	**26.03**	-14.19	9.83
IBM	0.039	**1.42**	**-0.18**	**12.48**	-22.96	11.72
General Signal corp.	0.054	**1.66**	**0.01**	**3.35**	-13.46	9.43
Wrigley Co.	0.072	**1.45**	**-0.00**	**11.03**	-18.67	11.89
Interlake Corp.	0.043	**2.16**	**0.72**	**12.35**	-17.24	23.08
Raytech Corp.	0.050	**3.39**	**2.25**	**59.40**	-57.90	75.00
Ampco-Pittsburgh Corp.	0.053	**2.41**	**0.66**	**5.02**	-19.05	19.18
Energen Corp.	0.054	**1.41**	**0.27**	**5.91**	-12.82	11.11
General Host Corp.	0.070	**2.79**	**0.74**	**6.18**	-23.53	22.92
Garan Inc.	0.079	**2.35**	**0.72**	**7.13**	-16.67	19.07
Continental Materials Corp.	0.143	**5.24**	**0.93**	**6.49**	-26.92	50.00

MV optimization is focused on a view of risk as variance, volatility. Is that reasonable?
Look at the table: it plots the descriptive statistics. We ignore in the MV setting the skewness and kurtosis. Which imply that returns are distributed normally. Is a normal distribution realistic?

- Variance as a proxy of risk is reasonable for a normal distributed variable, but in reality, returns show evidence of **skewness and kurtosis.** We observe large positive excess kurtosis and large excess skewness. Which we ignore in the MV setting.
- A portfolio model that neglects these dimensions is then likely not optimal.

"In particular, skewness and kurtosis are generally ignored as criteria for evaluating portfolio decisions, even though studies generally admit that asset returns are positively skewed and leptokurtic." Source: You and Daigler (2010) p. 164

Maybe we do not care…

"In measuring these gains from diversifcation, previous studies assumed that returns are normally distributed. However, this assumption is contradicted by numerous empirical findings (e.g., Fama, 1965; SImkowitz and Beedles, 1978; Singleton and Wingender, 1986). These studies imply that the two-parameter CAPM, which assumes either quadratic utilty functions for investors or normally distributed stock returns, is insufficient in making optimal investment decisions." Tang and Choi (1998) p. 119

How is the investor's appetite for skewness and kurtosis?
Skewness has to do with symmetry or asymmetry. With no skewness you will have a symmetric distribution.
An investor will be averse to left-skewness: extremely negative returns can happen.
But will be happy with right-skewness.
→ the consequences of skewness is not so easy as the consequences of variance: it depends on the type of skewness.

Also, an investor will be averse towards kurtosis: kurtosis has also a symmetric characteristic. This has to do with the tails of your distribution; fatter tail as compared to the normal distribution. So, you will have more observations on the extremes.

Suppose you do a MV analysis; we add a number of stocks in the portfolio, and we realize diversification benefits because of the lower variance. Can we use the same reasoning for skewness and kurtosis?

What are the benefits of diversification for skewness and kurtosis?
A number of studies find that the idea that a portfolio diversifies volatility cannot just be extended to the risk as captured in skewness and kurtosis.

1. Skewness: adding more assets seems to introduce undesired negative skewness as compared to desired positive skewness
2. Kurtosis: adding more assets increases the kurtosis

"Diversification then is a two-edged sword: it eliminates undesired variance in return distributions, but also **eliminates desired skewness**."

"It is found that diversification does not reduce either skewness or kurtosis. As the portfolio size increases, portfolio returns become **more negatively skewed** and **more leptokurtic**."

"We find that **skewness risk and kurtosis may even increase** (i.e. skewness decrease) by building diversified portfolios. Hence, a general assumption of risk mitigation by diversification does not hold if higher moments are taken into account." Source: Walther (2014)

- Also, diversification benefits also seem to vary over time due to **time-varying correlations.** It hinges on correlations. When the market is volatile those correlations increases, than we value those correlations more.
- In particular, correlations seem to increase in a more volatile market, implying that **benefits of diversification disappear exactly when needed most.**

"Correlations are **higher during bear markets** and tend to fall during periods of recovery."

"In our study, most of the correlations are larger during the bear market relative to the bull market, confirming previous findings that markets possess higher correlation during more volatile times." and "Thus, one needs to take the variability of correlations into account in order to evaluate the effectiveness and stability of diversification."

- MV optimization is mechanical: small differences in estimates can have large effects on estimated allocations
- This results in concentrated portfolios, that require lots of rebalancing

To overcome this a number of solutions exist:

- Opt for the pragmatic approach in which passive and active management are combined: outperform in the margin
- Apply a robust resampling approach to estimate the efficient frontier: monte-carlo simulation where you preform many tangency portfolios. The ultimate portfolio that you will retain will be an average of these that you did. This is a much more diversified portfolio.

If you rebalance your portfolio than with this robust resembling approach, the trades that you need to do will be smaller than the trades in a standard MV optimization.

Chapter 6: Market efficiency

BKM - Chapter 11: Efficient market hypothesis
Stock returns seem to move, to a large extent, randomly. Is such behavior in line with economic theory?

This is linked to the previous lecture: realizing outperformance → searching for alpha:
- Positive alpha = underpriced → go long in that asset
- Negative alpha = overpriced → go short in that asset

The degree to which you can detect potential overpricing is linked to the degree of market efficiency. Is the market efficient yes, or no? and if we detect market efficiencies how large are those?
- The market is not at all times efficient
- It is not because you observe particular market efficiencies that you can profit from that
    - There are some mechanisms that these mispricing stay in equilibrium

Efficient market hypothesis

The concept of **market efficiency** was introduced by Fama in 1970:

"In general terms, the ideal is a market in which prices provide **accurate signals for resource allocation**: that is, a market in which firms can make production-investment decisions, and investors can choose among the securities that represent the ownership of firms' activities under the assumption that security prices at any time 'fully reflect' all available information. **A market in which prices 'fully reflect' available information is called efficient.**"
Source: Fama (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, JF 25: 383
→ This is the famous efficient market hypothesis

- Having markets that are efficient is the ideal scenario
- This has to do with the fact that in an efficient market, money flows to the best investment opportunities
- Having an efficient market is not whether you have a positive alpha or not, it has to do with an accurate allocation of capital in an economy
- But this definition is quite broad, we cannot implement this definition so:

This concept of market efficiency was formalized by Malkiel in 1989:

"Formally, the market is said to be **efficient with respect to some information set, φ, if security prices would be unaffected by revealing that information to all participants.** Moreover, efficiency with respect to an information set, φ, **implies that it is impossible to make economic profits by trading on the basis φ.**"
Source: Malkiel (1989). Efficient Market Hypothesis, The New Palgrave of Finance: 127

This definition can you use in practice; it boils down to look at what is contained within φ

Intuition

What is the mechanism of this market efficiency?

Implications:
EMH implies that news about future performance will quickly be incorporated into current stock prices
As new information is, by definition, unpredictable, stock prices also move unpredictably.

Stock price changes follow a random walk, that is, that price changes should be random and unpredictable.

All the information within φ is already reflected in current prices. So current prices are a reflection of what the market expects for the future.
When do prices change? When new information comes into the market → φ set is extended, prices will move. New information is unpredictable otherwise it would not be new. If new information comes into the market in an unpredictable manner than the stock prices will also change in an unpredictable manner.

Don't confuse randomness in price *changes* with irrationality in the *level* of prices. If prices are determined rationally, then only new information will cause them to change. Therefore, a random walk would be the natural result of prices that always reflect all current knowledge. Indeed, if stock price movements were predictable, that would be damning evidence of stock market **inefficiency** because the ability to predict prices would indicate that all available information was not already reflected in stock prices. Therefore, the notion that stocks already reflect all available information is referred to as the **efficient market hypothesis (EMH)**

Which information set?

Different versions of EMH exist, based on the type of information that is available

1. **Weak-form efficiency** relates to past market trading data (narrow information set)
   - Past prices, past volumes, …
   - Really narrowed down to price information

   Implication: technical analysis is useless
   - **Technical analysis** is essentially the search for recurrent and predictable patterns in stock prices.
   - This is because whatever the fundamental reason for a change in stock price, if the price responds slowly enough, the analyst will be able to identify a trend that can be exploited during the adjustment period.
   - The implication is that you cannot use past returns for future returns, because past returns are already incorporated in the current prices. This is what technical analysis does and that is why it is useless.

   *One of the most commonly heard components of technical analysis is the notion of **resistance levels** or **support levels.** These values are said to be price levels above which it is difficult for stock prices to rise, or below which it is unlikely for them to fall, and they are believed to be levels determined by market psychology.*

2. **Semi-strong-form efficiency** relates to all publicly available information w.r.t. the prospects of a firm (broad information set)
   - Maro-economic news, fundamental company information, … will be reflected in prices

   Implication: fundamental analysis is useless
   - **Fundamental analysis** uses earnings and dividend prospects of the firm, expectations of future interest rates, and risk evaluation of the firm to determine proper stock prices. Ultimately, it represents an attempt to determine the present value of all the payments a stockholder will receive

from each share of stock. If that "intrinsic value" exceeds the stock price, the fundamental analyst would recommend purchasing the stock.

   - Fundamental analysis uses a broader set of company information to predict future returns. This will not pay off, because all that information is already reflected in current prices.
      - The trick is not to identify firms that are good, but to find firms that are *better* than everyone else's estimate. Similarly, troubled firms can be great bargains if their prospects are not quite as bad as their stock prices suggest.
      - This is why fundamental analysis is difficult. It is not enough to do a good analysis of a firm; you can make money only if your analysis is *better* than that of your competitors because the market price will already reflect all commonly recognized information.

3. **Strong-form efficiency** relates to all information, including private information, w.r.t. the prospects of a firm (extremily broad information set)
   - Everything is included in the current prices

   Implication: in the long run, all strategies are useless
   - Here you cannot outperform the market

Notice one thing that all versions of the EMH have in common: They all assert that prices should reflect *available* information. The EMH asserts only that at the given time, using current information, we cannot be sure if today's prices will ultimately prove themselves to have been too high or too low. If markets are rational, however, we can expect them to be correct on average.

Empirical evidence on market efficiency

Most analysis is done on the first two efficiencies. Not on the last one because that is hard to analyze, and trading on private information is in most cases illegal. So studies restrict to the first two one.

A never ending debate

The idea of market efficiency has never been welcomed in Wall Street (as they have active strategies presuming that the market is inefficient → academics assume that the market is efficient and practitioners don't ➔ debate) as it goes against a large part of their business model, and the debate will most likely never be settled because of

- **Magnitude issue**: mispricings are relatively small, so only managers of large portfolios can earn enough trading profits allowing us to observe statistically significant profits
   - If a stock is mispriced, it will be small deviations. So, if you want to outperform you need a firm that reaps many of these inefficiencies. According to this view, the actions of intelligent investment managers are the driving force behind the constant evolution of market prices to fair levels. Rather than ask the qualitative question, Are markets efficient?, we ought instead to ask a more quantitative question: How efficient are markets?

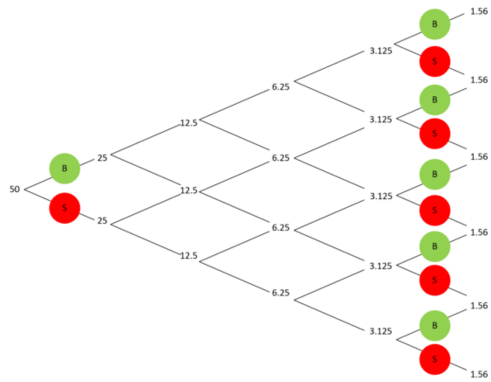
- **Selection bias**: successful strategies are kept secret
    - Once the market is aware of a particular firm that generates outperformance than everyone will look at all of the trades that the firm is doing so they can replicate that.
    - Only investors who find that an investment scheme cannot generate abnormal returns will be willing to report their findings to the whole world. Hence, opponents of the efficient markets' view of the world always can disregard evidence that various techniques do not provide investment rewards and argue that the techniques that *do* work simply are not being reported to the public. This is a problem in *selection bias;* the outcomes we are able to observe have been preselected in favor of failed attempts. Therefore, we cannot fairly evaluate the true ability of portfolio managers to generate winning stock market strategies.

- **Survivorship bias**: unsuccessful strategies are stopped
    - If an asset manager starts in a particular fund in mind that they will have arbitrage opportunities and it is not effective than they will stop the fund.

⇨ These two together means that you do not have the lower end or the upper end of the spectrum. What we observe, what we have data for is everything in the middle.

- **Lucky event**: after the facts, analyses will always point to at least one outperforming strategy
    - There will always be one person who will be better ex post (do not replicate this person). That is just statistically the case. You need to pinpoint ex ante who will outperform. Look at the example here under:

In a large population of unskilled asset managers, there are always a few that perform particularly well **by luck**. Do not be impressed by evidence that the best few have statistically significant outperformance. Are you able to indicate whom of them will be the one that will be correct five times in a row = skill! So do not be impressed who is outperforming ex post.



→ If any stock is fairly priced given all available information, any bet on a stock is simply a coin toss with equal likelihood of winning or losing the bet. However, if many investors using a variety of schemes make fair bets, statistically speaking, *some* of those investors will be lucky and win a great majority of the bets. For every big winner, there may be many big losers, but we never hear of them. The winners, though, turn up in *The Wall Street Journal* as the latest stock market gurus; then they can make a fortune publishing market newsletters.
→ Our point is that after the fact there will have been at least one successful investment scheme. A doubter will call the results luck; the successful investor will call it skill. The proper test would be to see whether the successful investors can repeat their performance in another period, yet this approach is rarely taken.

- **Joint hypothesis**: how to interpret any anomaly? What is the benchmark model from which to measure anomalies? Is it a lacuna in risk measurement, or is it inefficiency?
    - An incorrect price, you make a statement on a particular asset pricing model about the inefficiency. Because if I believe that a particular asset is wrongly priced, I must have an idea of what asset is correctly priced. So, what is the correct price? There is no unique model where there is consensus that, that model values fundamental prices.
    - Making a statement of a mispricing will always be conditional on a particular model. If you believe that a particular asset is mispriced, is it really mispriced or is it you just using a different asset pricing model?
    - It is always vis a vis a benchmark model that you conclude if something is over – or underpriced.

As we have no clear/unique idea on what the 'correct' price of a stock is, it is hard to evaluate an 'incorrect' price. This joint **hypothesis issue l**eads to two opposing camps in finance. Central to their discussion is the question of which benchmark model we should use to describe the trade-off between risk-return

- **Rational camp** with Fama and French (1993), JFE 33: 3-56
    - CAPM is not the correct benchmark; there are other risk factors such as momentum and value.
    - This model misses out on important risk factors; if you then impose CAPM than it is as if there is inefficiency.
    - EHM is alive - the CAPM is dead
- **Behavioral camp** with Lakonishok, Shleifer and Visnhey (1995), JF 50: 541-578
    - the CAPM is the correct model to adjust for risk; you can add factors, but they are not true for economic underpinning. So, they believe that the CAPM is the correct model.
    - EMH is dead - CAPM is alive

So what does the EMH tells us?

Weak-form efficiency

- Early tests of EMH were tests of serial correlation (do returns have a history?) Remember: serial correlation refers to correlation between a stock return today, and a stock return on a previous day
    - If past returns are correlated with current returns, then past returns have predictive power, and the market is efficient. This boils down to autocorrelation.
- Positive serial correlation: positive returns tend to be followed by more positive returns → this is called **momentum**
- Negative serial correlation: positive returns tend to be followed by negative returns → this is called **reversal**
- In the market we do observe momentum, on a rather short term horizon while reversal is observed on longer periods of time.

- Prediction of EMH: we should not observe any serial correlation (on the condition that risk is constant over time)
- However, quite some empirical evidence shows the opposite: the degree and the sign of the serial correlation depends on the **investment horizon** that is taken into account
    - for the short to medium horizon (until 12 months): evidence of momentum
    - for longer horizons (3-5 years): evidence of reversal

- ⇨ this is a violation of market efficiency, because we should not observe negative or positive correlation. Returns should not be autocorrelated. This means that the market is not efficient.
    - This is not what the rational camp says: they say momentum indeed predicts returns but this is a risk factor. You should include momentum as a risk factor in your model. Once you add that, alpha will disappear. In the Fama French world, momentum is just the risk factor and when you add that we live in an efficient market.
    - What about reversal? Rational camp explains this by variation as a risk aversion. They argue that even though on the short term, risk aversion should be constant, over the long run it is not unlikely that risk aversion changes. There might be episodes when investors are more or less risk averse. For example during crisis people are more risk averse. Thus we might observe reversals. If there is a shock in mean risk aversion than our market risk premium will go up, than prices will go down. If after a period of time market risk premium goes down, prices goes up again. = reversal
    - Momentum and reversal can thus also be explained by the rational camp.
- This evidence has led to the **fads hypothesis:** episodes of overshooting are followed by correction
    - over the short horizon: overreaction leads to momentum
    - over a longer horizon: subsequent correction leads to reversal

Can we profit from the observed momentum?

It is not because you detect potential anomalies that you can profit from them. This has to do with limits to arbitrage.

- A momentum strategy consists of buying recent winners and selling recent losers
    - This is profitable on the ST but you need to make allot of trades and taking into account trading costs, this will not be so profitable.
    - this generates alpha in a CAPM and even Fama-French 3-factor model
    - this has led to an extended factor model, including momentum as the 4th factor → known as the Carhart model (1997), JF 52: 57-82
- While momentum profits are positive on average, they are occasionally large and negative (cf end of technology bubble)
- Further, to exploit a momentum strategy, frequent trading is necessary: such strategy is thus only profitable when transaction costs are low otherwise you won't pursue this strategy and then these anomalies stay in equilibrium.

- Also, the long-run reversal can be framed in the rational camp by introducing time-varying risk aversion
- Over longer horizons, it is reasonable to assume non-constant compensation for risk
- Shocks in the market risk premium (or the discount rate) can then generate the observed reversal
    - an increase in the market risk premium translates into decreasing stock prices
    - when risk premia revert to their initial level, stock prices increase again
    - so this generates a pattern of reversal

Semi-strong-form efficiency

Tests of semi-strong-form efficiency are tests of the predictive power of a broader information set, i.e. so-called fundamental analysis. Findings such as these, are difficult to reconcile with the efficient market hypothesis and therefore are often referred to as efficient market **anomalies.**

Tests of risk-adjusted returns are *joint tests* of the efficient market hypothesis *and* the risk adjustment procedure. If it appears that a portfolio strategy can generate superior returns, we must then choose between rejecting the EMH and rejecting the risk adjustment technique. Usually, the risk adjustment technique is based on more-questionable assumptions than is the EMH; by opting to reject the procedure, we are left with no conclusion about market efficiency.

A number of so-called anomalies has been detected in the literature:

- dividend yield: stocks with high dividend yield, provide higher returns
- P/E effect: stocks with low P/E ratio provide higher returns
- neglected firm effect: less covered firms earn higher returns
  - Because small firms tend to be neglected by large institutional traders, information about smaller firms is less available. This information deficiency makes smaller firms riskier investments that command higher returns.
  - In this sense the neglected-firm premium is not strictly a market inefficiency, but is in fact a type of risk premium.
- liquidity effect: less liquid stocks provide higher returns
  - The dramatic dependence of returns on book-to-market ratio is independent of beta, suggesting either that high book-to-market ratio firms are relatively underpriced, or that the book-to-market ratio is serving as a proxy for a risk factor that affects equilibrium expected returns.
- B/M effect: higher B/M firms provide higher returns
- small firm effect (size effect): stocks of small firms provide higher returns (but small firm effect is concentrated in January)

While some anomalies can be linked to risk, this is not true for all factors; certainly not for the B/M effect of FF

We do have a number of predictive variables and some of them do have an intuitive meaning. Some of these factors can be used and represents risk. For other factors it is not clear if it is a risk factor, and then it could be a market inefficiency.

For example the small firm effect: is it reasonable to ask for a higher compensation for small firms vs large firms. Because you could think about small firms being less diversified than large firms. So you could think about small firms being more risky than a large company.

What about market effect? Value VS growth companies. Whitin the Fama French model, value companies have higher expected returns, and is more risky compared to a growth company. Is that reasonable? It goes counter your intuition. Typically, practitioners would say you will expect higher returns on the growth companies, but this is on average not what we observe. It is the value companies that yield higher returns. So, in a rational context you can only explain this by that value companies are more risky. Why would a value company be more risky?

Mutual fund and analyst performance

A large spectrum of funds are being marketed from active management to passive management

How does this align with market efficiency?

Active management is still more popular than passive management. That means that people still believe that there is someone out there that can outperform the market. Because if you invested in an active fund, you should have the belief that that manager has the skills to outperform the market → your choice for active vs passive is directly linked to your belief of market efficiency.

Active funds are actively run by fund managers who reject EMH

- they try to outperform the market by collecting for information that is not yet reflected in prices

- they take positions on individual stocks, macroeconomic themes or investment styles: costly research is needed
- they trade frequently (often with turnover rates of 100% per year): this is because these market inefficiencies are very small and so they have to trade allot so transaction costs are incurred

⇨ active funds are costly in terms of transaction costs but also in terms of input

Passive funds hold mechanically a portfolio of stocks, consistent with an index and thus accepting EMH

- they do not take any positions: no research input is needed
- they do not trade frequently; only need to rebalance to align with the index they track → by-and-hold strategy

⇨ low-cost strategy

Do active mutual fund managers beat the market?

This depends on:

- Height of transaction costs
- Quality of the information they get

Equity research analysis

Extent to which active managers can beat the market depends on the quality of the investment research, which is performed by **financial analysts**

1. Sell-side analyst: works for a brokerage house and issues recommendations
    - they follow up on a limited set of firms and sell their research to (institutional) investors
    - were paid by soft dollars in the past (suppose I am an asset manager and there is a sell side analyst and is giving information and I am happy about the quality about these forecasts. I will not pay them directly but redirect my trades to these brokerage houses and the fee that I am paying to these analysts is incorporated in the bid-ask price), but now the regulator asks for transparency because ultimately it is the investor who pays for that advise in an indirect manner and so nowadays this is no longer allowed, and these analysts are paid by hard dollars.
2. Buy-side analyst: works for a mutual fund
    - they advise a fund manager based on all sell-side analyst information in combination with the fund's strategy

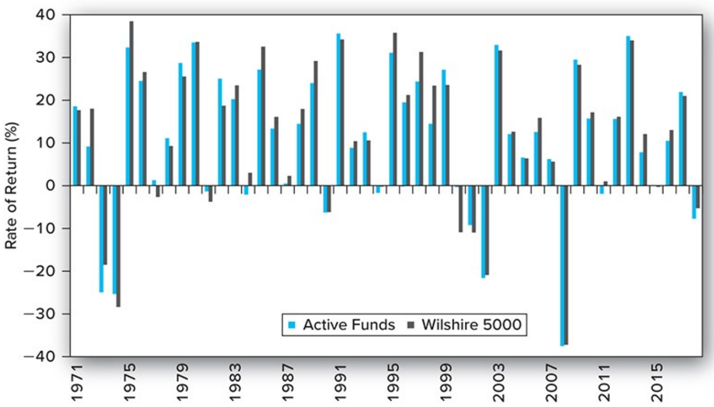
Are these analysts good? Because if they are visionary than they will be able to generate Alpha. But **on average** they are not, but that does not mean that they have not an important role to play. If it were not for these analysts, many pieces of information would stay there in the market and would never be incorporated in the prices. They make sure that all of that information is being processed. So, **on average they have an extremely important role**. However, on an **individual basis** are they any good? Can **this analyst** make consistent forecast compared to the other analyst? That is much harder.

- Financial analysts are important in assimilating and processing available information and disseminating it: this increases market efficiency
- But: it is well documented that financial analysts are systematically biased
    - strategic biases
        - e.g. please management because if I am a financial analyst I would want to keep the management of the firm I have in portfolio happy, because as long as they are happy with me than I get information. But if that is the role than you would not be neutral to the information you get, because you **might be overly optimistic.**

        - This you typically see in the recommendations. If you plot the distribution you expect to find a bell shaped distribution but in reality we see a really skewed distribution.
        - This also means that if you see someone issuing a hold, that is not good. Because often a hold is the lower end of the issue. This boils down to how to read the recommendation.
        - see Ljungqvist et al. (2007), JFE 85: 420 - 456)
    - behavioral biases
        - e.g. overconfidence; they overestimate the precision of their forecast. Once they do a good job they become overly optimistic, overly confident and they become wrong.
        - see Hilary and Menzly (2006), MS 52: 489-500
- Conclusion: financial analysts play an important role and thus have value added, but ambiguity about the accuracy of their advice remains

Active VS passive funds

The following graph plots over time the rates of returns of active vs passive funds. You see that we would expect given the prevalence of active funds, that they outperform passive funds. Is this the case? No, there is no clear winner. Some years active funds are outperforming, some years passive funds are.



At first glance, the performance of active fund managers seems disappointing (net returns, including trading costs and expenses)

- Average annual index returns is 1% higher than the average fund return
- The fund beats the index in only 19/48 years (= 40%)

Two important considerations remain:

1. **Do active mutual fund managers beat the market, excluding costs?**
    - This is not good publicity for active funds. But these are averages, meaning that there could be hidden profits. Active managers do beat the market, but it is because of the transaction costs that they do not outperform. Maybe if we exclude transaction costs, these managers will outperform the market? So managers are efficient, but it is the market (inefficiencies) that make that they cannot outperform the market.

2. **On average active managers underperform, but are there good managers that consistently outperform?**
    - It is not because on average they do not outperform, that there is not a particular subset of managers that do a good job.

Outperformance and costs

There is a huge difference between net and gross returns earned by active fund managers:

"We find that **funds hold stocks that outperform the market by 1.3 percent per year**, but their **net returns underperform by 1 percent**. Of the 2.3 percent difference between these results, 0.7 percent is due to the underperformance of non-stock holdings, whereas 1.6 percent is due to expenses and transactions costs."
Source: Wermers (2000), JF 55: 1655

Fund managers seems to be skilled, but once you account for transaction costs, and active funds depend on a liquidity of cash. Net returns are just as low or as high as the market.

Individual outperformance

What about persistence in mutual fund performance? (are there any good managers)

Such persistence can be driven by

- superior fund manager skills
- trend in investment style/costs

"Using a sample free of survivor bias, I demonstrate that **common factors in stock returns and investment expenses almost completely explain persistence in equity mutual funds' mean and risk-adjusted returns.** Hendricks, Patel and Zeckhauser's (1993) "hot hands" result is mostly driven by the one-year momentum effect of Jegadeesh and Titman (1993), but individual funds do not earn higher returns from following a momentum strategy in stocks. **The only significant persistence not explained is concentrated in strong underperformance by the worst-return mutual funds**. The results do not support the existence of skilled or informed mutual fund portfolio managers."
Source: Carhart (1997), JF 52: 57

We cannot detect a subset of managers that consistently performs well. The ranking is changed on a yearly basis. There is however once consistency: the bottom, there is persistence in the underperforming managers. So, try to avoid that subset!

Role of asset management in a EMH world

Even in an efficient market, there is an important role for asset management:

1. Chasing outperformance, they make sure all information is picked up and they contribute to market efficiency
2. They offer diversified portfolios (at low cost) with optimal risk-return trade-off
    - Even if all stocks are priced fairly, each still poses firm-specific risk that can be eliminated through diversification. Therefore, rational security selection, even in an efficient market, calls for the construction of an efficiently diversified portfolio providing the systematic risk level that the investor wants.
3. They can offer investors an investment portfolio, matching their individual preferences/risk aversion
    - They can make you aware of your risk profile. As an investor you do not always know that.
    - Investors of varying ages also might prefer different portfolio policies with regard to risk bearing.
4. They allow you to invest in a broader set of assets
    - A number of assets are not available to you as a retail investors but only to institutional investors.
5. There could be tax considerations
    - If you think about the ETF sector in EU, they do not offer US ETF's. Why not? Because the taxation is completely different.
    - High-tax-bracket investors generally will not want the same securities as low-bracket ones.

In conclusion, there is a role for portfolio management even in an efficient market. Investors' optimal positions will vary according to factors such as age, tax bracket, risk aversion, and employment. The role of the portfolio manager in an efficient market is to tailor the portfolio to these needs, rather than to beat the market.

Chapter 7: Behavioral finance and limits to arbitrage

BKM - Chapter 12 (12.1): Behavioral finance and technical analysis
Many of the quiz questions come from Kahneman (2011) - Thinking fast and slow.
An excellent (and amusing) overview of the history of behavioral finance is given by Thaler (2016) - Misbehaving: The making of behavioral economics.

This is closely linked to market efficiency. Because if the market is not efficient, how come? We might not be as rational as we hope we are and even if you are rational there might be some mechanisms that prevent you to take advantage of potential mispricings (limits to arbitrage).

The behavioral critique

Conventional finance

- (some) investors behave rationally
- prices are correct and equal to intrinsic value

Behavioral finance

- (not all) investors do not behave rationally

Are investors rational?
If not, what are the consequences on the market?
Is this a sufficient condition to observe anomalies?

The behavioral critique rests on two categories of anomalies when investors (humans) take decisions under uncertainty

1. Investors do not always process information correctly because they use heuristics
    - These heuristics lead to irrational decisions
2. Investors take decisions that are behaviorally biased, even when resolving uncertainty
    - This is about framing: the way in which a problem is framed has an impact on the decision that you will be taking

In both cases this leads to consistent and predictable errors

Information processing

- When making decisions under uncertainty, people tend to rely on heuristics or simple rules of thumb instead of using statistics (because of limited time and attention)
- While sometimes useful, many of these heuristics do often lead to poor decisions
- The reason is that heuristics can be misleading because they **ignore basic rules of statistics and probability theory**
- Investors' limited analytic processing capacity may also cause them to overreact to salient or attention-grabbing news and underreact to less salient information.

Representativeness bias

Steve is very shy and withdrawn, invariably helpful but with little interest in people or in the world of reality. A meek and tidy soul, he has need for order and structure, and a passion for detail.
Is Steve more likely to be a librarian or a farmer?

How many librarians and farmers are there in the world? This is what you should ask. We tend to give librarian as an answer, based on representativeness.

- Representativeness bias makes that people tend to neglect the size of their sample
  - We used a small description to conclude that he is a farmer and you neglect all the information that you have
  - a small sample is considered representative
  - one relies on resemblance and relevant statistical facts are ignored
- In a finance context, this translates into inferring patterns too quickly and extrapolating apparent trends in the future
- This could explain the observed anomalies of overreaction and subsequent correction ("**fads hypothesis**") where short-run news is extrapolated too far into the future: on the short-term we have momentum and in the long run we have reversal
  - This can be explained by the representativeness bias: on the short term people observe particular trend of phenomenon and they extrapolate that into the future and after a while they realize that they are overshooting and correct downwards again = reversal
  - This fads hypothesis can be explained by the representativeness bias

For empirical evidence in line with overreaction and correction anomalies see Chopra, Lakonishok and Ritter (1992), JFE, p.235:

"A highly controversial issue in financial economics is whether stocks over- react. In this paper we find an **economically-important overreaction effect** even after adjusting for size and beta. In portfolios formed on the basis of prior five-year returns, extreme prior losers outperform extreme prior winners by 5–10% per year during the subsequent five years. Although we find a pronounced January seasonal effect, our evidence suggests that the overreaction effect is distinct from tax-loss selling effects. Interestingly, the overreaction effect is substantially stronger for smaller firms than for larger firms. Returns consistent with the overreaction hypothesis are also observed for short windows around quarterly earnings announcements."

Similar mistakes have also been observed in other domains e.g. sports
A famous example is baseball: scouts using traditional 'representativeness' methods underperform in picking baseball players as compared to statisticians using track records.
Source: Lewis (2004), Moneyball: The art of winning an unfair game

Representativeness and conjunction fallacy

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Which alternative is more probable?

1. Linda is a bank teller
2. Linda is a bank teller and active in the feminist movement

Most people tend to answer number 2.

- Basic probability theory learns us that the likelihood of a conjunction of events can never be larger than the likelihood of an individual event
- this is the conjunction fallacy, the likelihood of two events happening at the same time is always smaller than the likelihood of a single event
- However, one may perceive the conjunction as more plausible, if it is more representative of how one characterizes the event/people
- Again we ignore statistical law because Lina is representative for being socially engaged
- In the current example: Linda is perceived to be representative for being active in the feminist movement, and thus we assume this to be very likely

Overconfidence

Relative to the other students, how would you classify your driving ability?
Poor - below average - average - above average - excellent

The distribution of the answers should be bell shaped because it is relative to other students. But in reality, it is not bell shaped → it is not by accident, it always looks like this, and this is proof of overconfidence. You believe too much in your driving abilities. And typically, the males believe that they are excellent and above average and that has to do with testosterone. Females with high levels of testosterone also turn out to be more overconfident.
People overestimate the precision of their abilities and forecasts, labelled **overconfidence**:

1. The prevalence of active investing is hard to reconcile with the difficulty of outperforming
   For each investor who outperforms, there will be someone that is underperforming. So the prevalence, all the asset managers that believes they can outperform, that seems to be a case of overconfidence.
   The popularity of active investment is inconsistent with its poor performance, but consistent with overestimation of ability of the fund managers
2. Male investors trade more frequently than female investors
   Ex post, the performance of males and females are identical. Males trade much but the fact that they trade so much eats up their outperformance.
   This is consistent with greater overconfidence of men, linked to higher levels of testosterone (Barber and Odean (2001), QJE)
3. CEOs overpaying corporate acquisitions
   Many M&A activity is not value adding but value destroying. This is counter to the CEO who thinks that with a particular acquisition he can add value and generate return.
   CEOs overestimate their ability to generate returns, overpay for target companies and undertake value-destroying mergers (Malmedier and Tate (2008), JFE)
4. Financial analysts make predictable errors in their earnings forecasts
   We can link this with overconfidence, they overweight the accuracy of their own signals.
   Overconfidence about the precision of one's value-relevant information would be consistent with value-versus-growth (e.g., book-to-market) anomalies. If investors respond too strongly to signals about the fundamental value of a stock, then those signals will cause stock prices to overshoot their intrinsic values. Stocks with high prices relative to proxies for intrinsic value will be more prone to be overvalued and therefore poor investments.
   Similarly, low-priced stocks would be more apt to be undervalued. These errors could lead to value anomalies such as the lower average returns of stocks with high ratios of market- to-book value compared to low market-to-book stocks.
   They overweigh private information yielding biased earnings forecasts (Bosquet, De Goeij and Smedts (2015), ABR)

Anchoring

Assume a class of 30 students. Is the likelihood that (at least) 2 students share their birthday (day/month combination) larger than 35%?

What is this likelihood?

Did the global worldwide avocado production over the period 2015-2020 more than double (growth of more than 100%)?

What is your estimate of the growth in global worldwide avocado production over the period 2015-2020?

Many students answered 28% on the first question and 126% on the second question.
Birthday problem: people typically underestimate this likelihood. This problem is typically used to show that we are bad at statistics.

- Anchoring refers to the situation where one makes a forecast based from some initial value and adjusts from there; when the anchor is relevant, this can be a useful tool to come to a more accurate forecast
- There is, however, evidence that people also use anchors that are not related to the variable that needs to be forecasted (e.g. the last number that one has observed) → if such non-relevant anchor is used, one's forecast is biased by it
- A famous experiment has been done by Kahneman and Tversky (1974) who used a spinning wheel as an anchor when asking about the number of countries in UN
    - Even if the anchor is completely irrelevant and as a participant you know that still people are biased by this anchor.
    - The number on which the spinning wheel landed had an impact on the number that the students answered even though they know that the spinning wheel has nothing to do with the question.
    - The consequence is that if you use a nonrelevant anchor, your forecast will be worse than using no anchor at all. If the anchor is relevant it can improve your forecast.

Behavioral biases

We make decisions that are not consistent, those decisions are influenced by the way you are presented to a particular problem. This has to do with the concept of **nudging**: presenting information in such a way that you steer the decisions of consumers, investors, …

Even if the processing of information is perfect, individuals tend to make decisions that are *less than rational* using this information
Such behavioral biases are often the result of the framing of a particular problem

- Loss versus cost
- Gain versus loss
- Action versus non-action
- Overall wealth versus segregated wealth
- Level of wealth versus changes in wealth

Framing

Individuals may act risk averse in terms of gains but risk seeking in terms of losses. But in many cases, the choice of how to frame a risky venture—as involving gains or losses—can be arbitrary.

Regret avoidance

Mr. Brown almost never picks up hitchhikers. Yesterday he gave a man a ride and was robbed.
Mr. Smith frequently picks up hitchhikers. Yesterday he gave a man a ride and was robbed.

Who of the two will experience greater regret over the episode?

Mr. Brown - Mr. Smith - both equally

The majority answered Mr. Brown. In general it is true that whatever the decision we take, when it turns bad we feel bad about it. So both will feel bad. But it also true that Mr. Brown will feels worse. Because this action was unconventional, and this had a negative consequence.
When a decision turns bad, one often feels regret over that decision

- The more unconventional the decision, the more regret one feels
- Regret comes with a guilty feeling *"I should have known better"*

Similar behavior is also observed in finance contexts

- Investors require **increased rates of return for out-of-favor firms**
  An out-of-favor firm is an unconventional investment decision, and you are only inclined to do that decision when you are rewarded for that. In this context the high rates of returns for book-to-market companies is explained.
  → This is consistent with the size and B/M effect of Fama-French: small companies, or high book-to-market companies are considered less conventional investments, which leads to depressed prices and thus high expected returns (see De Bondt and Thaler (1987), JF)
- Financial analysts' **herd** with earnings forecasts that are overly optimistic
  Financial analysts' herd: suppose I am an analyst and I see my colleagues come with strong buy decisions; will I just go with the flow? I will do that because:
  Being the only one who is wrong, hurts much more than being wrong in the crowd. The explanation is that if I think that a company is doing bad and I would like to issue a sell but everyone is issuing a buy, if I also issue a buy than no one will blame me for that. You can hide within the crowd. Suppose you do stand out and you were wrong about the sell, than you might be fired. The cost is too high to stand out of the crow.

Mental accounting

You have bought two $80 tickets to the theater. When you arrive at the theater, you open your wallet and discover that the tickets are missing.

Will you buy two more tickets to see the play?

You go to the theater, intending to buy two tickets that cost $80 each. You arrive at the theater, open your wallet and discover that the cash with which you were going to make the purchase is missing. You could use your credit card.

Will you buy the tickets to see the play?

These set of questions are the same. But the answer on those questions differs:
First question the answers were mostly no and second questions answers were mostly yes.
→ how come we decide so differently? This has to do with mental accounting.

Different frames can evoke different accounts, and the significance or experience of a loss depends on the account to which it is posted

- In the case of lost tickets: you are likely posting the loss to the 'evening-out-account'
- In the case of lost cash: you are more likely posting the loss to the 'general-expenses-account'
  In relative terms the lost cash is a smaller part of the overall budget as compared to the lost tickets in relation to the evening-out budget

Rationally it will be better to view different accounts as part of your overall wealth portfolio

- it allows you to integrate the risk-return profiles of all different accounts in a unified and aggregate framework

Mental accounting has been observed in several investment contexts

- Investors segregate accounts or monies and **take risks with their gains that they would not take with their principal** (Statman (2008)): this is the so called **house money effect** which creates momentum in the market (Thaler and Johnson (1990), MS)
    - When people go to the casino and gain allot of money they feel like that gained money is not theirs and they typically take allot of risk with those gains.
- Investors will **not dip into their capital** with a tendency to keep on to losers and sell winners too quickly (Shefrin and Statman (1985), JF); this is the **disposition effect**: this has to do with the fact that investors do not want to dip into their capital.
    - Behavioral motives are consistent with some investors' irrational preference for stocks with high cash dividends (they feel free to spend dividend income but would not "dip into capital" by selling a few shares of another stock with the same total rate of return) and with a tendency to ride losing stock positions for too long (because "behavioral investors" are reluctant to realize losses). In fact, as an empirical rule, investors are more likely to sell stocks with gains than those with losses.

Risk aversion

You are offered a gamble on the toss of a coin. If the coin shows tails, you lose €1000. If the coin shows heads, you win €1500. Will you accept this gamble?
Majority of the answers do not take on the gamble.

To choose between the two options, you need to balance the two outcomes of the game

- the benefit of winning €1500
- the cost of losing €1000

Although the expected value of the gamble is positive (€250), many people will dislike this gamble → this can be explained by aversion to risk: the fear of losing €1000 is more intense than the hope of gaining €1500.

This is **risk aversion**, which is captured in a concave utility function. But the conventional finance literature assumes asymmetry that we have the same attitude for gains as well as for losses. But the behavioral finance literature doubts this, they introduced loss aversion.

Is risk aversion synonym for loss aversion?

Loss aversion

Upon graduation, your parents give you $1,000. Afterwards you are offered the following two choices. Which of the two do you prefer?
50% chance to win $1,000 get $500 for sure

In addition to what you own, you are given $2,000. Afterwards you are offered the following two choices. Which of the two do you prefer?
50% chance to lose $1,000 lose $500 for sure

If you put those two options next to each other, you will see that they are identical in terms of outcome.
But the answers: the majority goes for the gamble. While in the winning scenario people tend to have the $500 for sure.
→ this has to do with loss aversion: people act differently towards gains and losses

- Conventional finance assumes symmetry in attitude to uncertainty: people have the same attitude to risk for gains as to risk for losses
- Behavioral finance states that people's attitude is different for favorable outcomes than for unfavorable outcomes: losses loom larger than gains
- This latter view implies that people are willing to take more risks to avoid a loss, than to make a gain: they tend to be risk-averse with respect to gains, but risk-loving with respect to losses
- The outcome is thus the same but the reference point influences the choices

It is not the utility of wealth that matters, it is the **reference point** (from which the options are evaluated) that matters

- In the first problem: we have a reference point of €1,000
  In the second problem: we have a reference point of €2,000

Being richer by €1,500 is then a gain wrt the reference point of €1,000; but a loss wrt the reference point of €2,000
This explains why people tend to choose for the safe amount in the first problem, but for the gamble in the second problem

There is quite some empirical evidence documenting **an asymmetry in people's risk attitude**
For example, it turns out that the most recent situation is often used as a benchmark, and that this recent experience impacts our behavior towards risk
Coval and Shumway (2005) find that CBOT traders are highly loss-averse/risk-seeking depending on the reference point

- they are more likely to take on more risk in an afternoon following a morning with losses (risk-seeking)
- their behavior also has important short-term consequences for prices, but not for the long term (reversal)

The weakness of behavioral finance

While the behavioral critique of full rationality is well taken, there is considerable debate among financial economists about the strength of the behavioral critique.
The main critique (weakness) is that the behavioral approach is **too unstructured**, and **lacks coherence**

- You cannot compare behavioral finance with conventional finance (very structured and coherent)
- A combination of irrationalities can always be found that could explain an anomaly: a unified approach that can explain a range of behavioral anomalies would be more interesting and convincing
- Some behavioral explanations contradict one another
  e.g. the asymmetric risk attitude of investors around a reference points hints at investors that are risk averse with respect to gains, while the house-money effect implies risk-seeking behavior towards gains

Limits to arbitrage

Relevance of the behavioral critique

Is this critique relevant for the aggregate market (market as a whole)?

- Behavioral finance would not matter if the marginal trader:
  - A single investor who;
  - is rational; and
  - is able to exploit the mistakes made by behavioral investors (if you are all irrational and I am rational I can exploit your mistakes and make a profit from that and then I will push prices to equilibrium)
- Such marginal investor would correct any mis-pricings by arbitrage: he would profit from buying and selling mispriced assets
- While textbook arbitrage requires no capital and entails no risk, in reality it is (almost always) risky and costly: these costs and risks of real arbitrage trading are even more relevant/urgent when arbitrageurs manage other people's money
  - So there are a number of factors in practice which prohibit you to take advantage of the mispricings you observe in the market
  - Having a single rational investor is not sufficient for the behavior critique not to matter

The behavioral critique states that, in practice, limits to arbitrage activities exist

Fundamental risk

This boils down that setting an arbitrage strategy is not without risk. When you set up an arbitrage strategy you believe that prices will converge to their original price but nothing prevents the price to move even further from the intrinsic value before going to the intrinsic value. So when you set up this kind of strategy you will incur losses. You will have to deal with these losses and if you invest money for other investors than it is hard to explain to them why you are losing money.

- Exploiting a market mispricing is not risk-free: the presumed market mispricing can get worse, before disappearing
- Convergence between intrinsic value and market value can take too long, possibly longer than the arbitrageur's investment horizon
  Assume a fund manager who takes an arbitrage trade hoping that particular spreads would decrease; if the spreads do not convergence immediately, and even get worse initially, the fund will realize losses in the short-run. This puts that fund at risk of losing clients and running out of capital
- This risk is called **fundamental risk** of arbitrage or **noise trader risk** and will limit arbitrage activities
  - You will need sufficient liquidity to overcome periods of losses
- This fundamental risk is nicely summarized in the famous quote by J.M. Keynes:
  "Markets can remain irrational longer than you can remain solvent"
  → incurring these losses is costly. If you have set up an arbitrage strategy and you have losses in the meantime, you may be asked to post margin so you need liquidity to post his margin so you need deep pockets = allot of cash.

- A prominent example of noise trader risk is Long-Term Capital Management (LTCM) which was a hedge fund with Merton and Scholes, two Nobel-prize winners, as members of the board
- This hedge fund's strategy was arbitrage or convergence trading on different asset classes (equity, bonds, derivatives)
  - due to the small profits per arbitrage trade, you need to take on a large amount of trades to make a sizeable profit (to do so, they were highly leveraged)
  - for example, they had 2.3 billion pairs of trades on Royal Dutch - Shell (Siamese twin stocks)

- When the market panicked in 1997/1998 (Asian and Russian crisis), the spreads widened (diverged) and LTCM went into distress (being highly leveraged made them extremely vulnerable)
  Lowenstein, R. (2000). When genius failed: the rise and fall of LTCM
  - When LTCM was bankrupt, the SEC intervened and they looked into other market participants to help them out. So LTCM had to show their books, they gave allot of information about the strategies that they were pursuing. So everyone in the market knew what the positions were that they were sitting on. So the one with the deepest pockets bought the portfolio and waited till the spread converged.

Implementation costs

Exploiting a market mispricing is also not costless (very costly)
- You need capital to take out the necessary trades
- With each trade you make, you incur transaction costs
- If the cost of trading is larger than the gains you will realize you will not set up the trade.
- So if you are in an environment where trading costs are high you will not make profits.

Depending on the necessary trades, these implementation costs can be significant
- This is for sure true when you try to exploit an overpricing as this requires short-selling: short selling is costly as you need to compensate the short-seller for lending you the asset
- Moreover, not only the horizon of the arbitrage convergence is uncertain, also the horizon of the short sale is uncertain as the borrowed security can be reclaimed on short notice; when the asset is reclaimed before convergence took place, you will need to take out a new short sale
- Other investors, such as many pension or mutual fund managers, face strict limits on their discretion to short securities.

Such **implementation costs** thus limit arbitrage activity to push prices to intrinsic values

Model risk

If you believe there is a mispricing in the market; then the market thinks otherwise. You believe that the market price you observe is not the correct one. The question is who is the smart one in this room? You will always doubt your own belief, only when you are very sure on your belief you will trade on that belief. Because if you trade on it, you will incur costs, so when having doubts you will not engage in those trades.

- Finally, a judgement on the existence of a discrepancy between fair and market value is subjective as it is based on one's own analysis
  - we do not observe intrinsic values, we can only estimate them
  - then there is always a risk that one's own model is faulty
- The arbitrage opportunity can thus be more apparent than real
  - arbitrageurs are aware of this and will therefore be prudent and conservative in setting up arbitrage transactions
  - they will act only when they are really convinced about the mispricing
- This **model risk** will also limit arbitrage activity

The law of one price

- Empirical evidence on the limits to arbitrage focus on violations of the law of one price → in a rational and competitive market we should see that identical assets trade at identical prices. If that is not the case than limits to arbitrage are at play.
- In practice, however, the market is not perfectly competitive, there are transaction costs and barriers to trade → such market imperfections will limit arbitrage forces

- Several examples have been observed pointing to violations of the LOP
    - closed-end funds do not trade at their NAV
    - Siamese twin companies trade at different prices
    - dual share classes trade at different prices
    - equity carve-outs display negative market values for the mother company

Lamont and Thaler (2003). JEP 17(4): 191-202

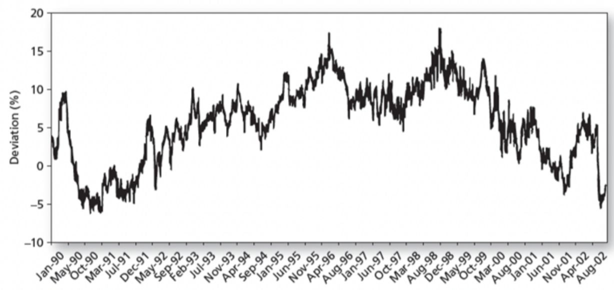
Closed-end funds

Number of shares is fixed. What we observe is that the price that you pay for these closed end funds can deviate from the NAV. Small deviations can be explained by costs and expenses but the deviations we see are too large. This paper seems to see that there is a correlation on the one hand between the firms that we observe having large deviations from the NAV and funds that are hard to arbitrage. So limits to arbitrage are one of the reasons why we observe huge differences from the NAV.

- It is observed that fund prices and NAV
    - differ, with the fund being valued, on average, less than NAV
    - vary considerably over time (funds go from premia wrt NAV to discounts of up to 30%)
- Part of the divergence can be explained by the fact that this is not a pure LOP-case (costs, expenses)
- But observed divergences are just too large..
  Interestingly, Pontiff (1996) shows that the divergence is highest for hard-to-arbitrage funds, which suggests that **limits to arbitrage** are responsible for the mispricings

Siamese twin companies

- Siamese twins are firms that have two types of shares with fixed claims on cash-flows and assets of the firm
    - such parity should be reflected in the ratio of the market prices
- A famous example is the Royal Dutch/Shell company (Shell was delisted in 2005)
    - Royal Dutch received 60% of cash-flows
    - Shell received 40% of cash-flows
    - this should translate in a ratio of the Royal Dutch to Shell stock price of 1.5
    - but looking at the graph here you see big deviations



- The substantial deviations observed between the Royal Dutch/Shell ratio are surprising
    - the companies were identical
    - the shares traded in a liquid Dutch and UK market (even with easy access to US investors via ADRs)
- A successful arbitrage strategy would therefore be to buy the cheap stock and to sell the expensive stock
- Normal market behavior should put prices to equilibrium cause everyone who wants to buy would buy the cheap one. But we do not observe this.
- But price differentials seemed to persist
- Note that even 'normal' market trading (demand factors) did not push prices to its intrinsic ratio. These deviations can be explained by limits to arbitrage
    - whoever wanted to buy the shares should buy the cheapest version
    - this was counter to the fact that many investors did buy the expensive shares: a possible explanation could be found in the investment restrictions of large institutional investors (e.g. index effect of S&P500)

    - Many investors have investor restriction; some can invest on Royal Dutch but not in Shell. E.g., index effect; many investors have in their prospectus a restriction for example you can only invest in stocks that are part of the S&P500. If you then have an arbitrage strategy where one is part of the S&P500 and the other one is not than you cannot set up this strategy. The same goes with bonds, most investors are restricted to invest in bonds with a typical credit rating. We really see this index effects and credit rating effect.

Dual share classes

- Dual share classes exist when a firm has two sorts of stock outstanding, with identical cash-flow rights, but different voting rights
- In normal circumstances, both classes trade at a very similar price
    - except at times of battle for corporate control (voting rights become crucial)
    - except when liquidity between both classes differs a lot
- In practice, we see that there are times in which the shares with the more voting rights trade at unreasonable high premia (e.g. 15%-20%) so even when there is no battle for corporate control
- There are even cases where shares with the more voting rights trade at a discount (e.g. McData shares in 2001)
  this can only partly be explained by a difference in liquidity
    - This can be explained by limits to arbitrage: in this case it is explained in terms of liquidity issues. Very often these shares classes have different liquidity (risk) and some investors might be prohibited to invest in low liquid stocks.

Equity carve-outs

- In case of an equity carve-out, a company creates a subsidiary and IPOs this subsidiary, while retaining a large equity stake
    - the shareholders of the parent firm are thus entitled to the remaining shares in the subsidiary
    - the share price of the parent firm should reflect this stake in the subsidiary company
- When doing such equity-carve out, a relation between the stock price of the parent firm and the stock price of the subsidiary is thus established
- In practice we do not observe such relation: there are a number of cases where we have seen that these carve-outs displayed important violations of the law of one price:
- Lamont and Thaler (2003) showed that several tech stock carve-outs displayed LOP violations
    - they showed 'negative stub values': the implied stand-alone value of the parent company's assets <u>without the subsidiary</u> was **negative**
    - The parent company should be worth at least the subsidiary
- A prominent example is the 3Com/Palm equity carve-out

Story of Com/Palm

"On March 2, 2000, 3Com sold a fraction of its stake in Palm to the general public via an initial public offering (IPO) for Palm. In this transaction, called an equity carve-out, 3Com retained ownership of 95 percent of the shares. 3Com announced that, pending an expected approval by the Internal Revenue Service (IRS), it would eventually spin off its remaining shares of Palm to 3Com's shareholders before the end of the year. 3Com shareholders would receive about 1.5 shares of Palm for every share of 3Com that they owned."

Source: Lamont and Thaler (2003), JPE, p.230

Story of Com/Palm

- In theory, we should have seen that the share price of 3Com should have been at least 1.5 times the share price of Palm
- In reality, however, Palm shares sold for more than the 3Com shares
    - the stub-value of 3Com was negative, while it was a profitable company, even with cash assets worth $10 per share 3Com
- This was a potential arbitrage strategy, but limits to arbitrage made that the negative stub values persisted for 2 months

- Clearly, such negative stub values, cannot be aligned with the existence of limited liability
- No one seized that arbitrage opportunity:
    - Thus, there must be something else at work that could lie at the basis of such anomaly:
        - **strange preferences**: investors are, for unclear reasons, buying expensive Palm shares instead of cheaper 3Com; this points to irrational markets
        - **limits to arbitrage**: rational arbitrageurs are limited in their activity to eliminate the mispricing; this points to rational, but imperfect markets

Why limits to arbitrage?

- This has to do with rational investors but imperfect markets
- Some stocks are hard to short, making an arbitrage strategy expensive or even impossible
- Moreover, the arbitrage strategy is risky (see Mitchell, Pulvino and Stafford (2002), JF)
- The link between the parent and subsidiary turns out to disappear without convergence in 30% of cases: What happened in these equity carve-outs, in normal circumstances you should see a convergence of these prices. However in allot of equity carve-out the link between the two companies will disappear before convergence.
    - for example, the parent goes bankrupt after using the subsidiaries' stake as collateral when issuing debt. So, no convergence will ever happen and the market knew about this.
- The time to termination varies significantly, making the arbitrage horizon risky (and potentially costly)
- The limit to arbitrage in this case was the inability of investors to sell Palm short. Virtually all available shares in Palm were already borrowed and sold short, and the negative stub values persisted for more than two months.

Impossibility of strong EMH

- What are the implications of limits to arbitrage for EMH?
- One of the main determinants of limits to arbitrage are the costs that come with investing: when you trade it is costly
    - The consequence is that you will only collect information if it gives you a benefit
- Clearly, when collecting and processing information, for sure hard to get information, this is also costly: the only motivation to make such costs is then that the benefits outweigh these costs
- This leads to a **paradox**:
    - If the market is strong form EMH (all information reflected in prices), no one will make costs to gather information, as there are no benefits associated to this
    - **but** if no one gathers private information, then the market cannot reflect this private information!
    - So, once you introduce limits to arbitrage in perfect markets transaction costs, the market actually cannot be strong form EMH

Strong form EMH is theoretically implausible when arbitrage is costly

Grossman and Stiglitz (1980) therefore redefine EMH:

Partially strong form EMH

Some, but not all private information is reflected in prices

- Investors make an effort to collect and process new private information
- Market prices will reflect this information
- In equilibrium, the marginal benefit to collect new private information equals its marginal costs
  - so, all private information for which the marginal benefit to collect new private information is smaller than its marginal costs, will not be collected and will not be reflected in the price

- This modified version of EMH has important implications for active versus passive management:
  - investors that have a cost of information-collection which is below the cost of other investors, can outperform the market and engage in active portfolio management
  - investors with a high cost of information-collection should engage in passive portfolio management
    - these investors free-ride on the efforts of the informed investors
- This modified version of EMH also implies that some active investors **underperform** the market
  - **the market adds up**: the market portfolio reflects the average of all trades
  - if some investors outperform the market, some others must be underperforming
  - but who outperforms and who underperforms?

Who wins, who loses?

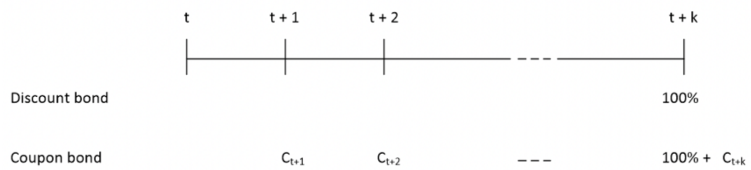
- The winners are active managers that are:
  - cost-efficient; ánd
  - skilled
- The losers are active managers that are
  - non-skilled (e.g. distorted beliefs leading to suboptimal strategies)
  - in need of liquidity (making that they have to abort their investment strategy)
- For non-skilled investors, or investors with a high cost to collect information it is optimal to stick to passively managed highly diversified funds ("hold on to the market")
  - this allows them to maximally profit from the winning active managers that push prices to intrinsic values

Chapter 8: Fixed income instruments

BKM - Chapter 14: Bond prices and yields

Fixed-income basics

- Fixed-income securities promise to pay specific payments on specific dates (this can be determined in nominal or in real terms): not necessarily fixed it can be variable amounts, but you know beforehand what the maturity and the computation is
- We can distinguish between two broad categories of fixed-income instruments
    - **zero-coupon bonds (ZCB)**, sometimes labelled discount bond or deep discount bond, make no intermediate payments, but only a single payment at the maturity date
        - the return you make on ZCB is the price you paid – par value you are paid back at maturity
    - **coupon bonds (CB)** make intermediate payments of $C up to and including the maturity date
        - return is the price you paid – par value you are paid back at maturity and the intermediate coupons
- The final payment of both ZCB and CB can be normalized to 100% (or 1); expressing the price as a % of the par value
    - when the price of the bond is larger than 100%, it sells at a premium
    - when the price of the bond is smaller than 100%, it sells at a discount



- The maturity of the bond determines the end of the contract
    - for most bonds this maturity is fixed
    - for some bonds the maturity is variable (according to a pre-specified formula)
        - callable or puttable bonds and also convertible bonds
- The coupon rate determines the interest payment (coupon payments) paid annually or semi-annually
    - for treasury bonds this coupon is typically fixed
    - for some corporate bonds this coupon is time-varying (according to a pre-specified formula)
    - you also know beforehand on which formula this coupon is being computed
    - so there might be uncertainty about the level of the coupon but you know a priori how the coupon is computed

Variable coupon bonds

- **Floating rate bonds** they have a base rate and then some extra rate based upon macro-economic conditions, have coupons with adjustable coupon rate, where the coupon payments are tied to some measure of current market conditions
  e.g a coupon rate of 2% + current T-bill rate, adjusted on an annual basis
    - Beforehand I do not know what the T-bill rate will be but I know how the coupon is computed
    - The introduction of adjustable coupons dampens the interest rate effect; because suppose you have a bond with a basis of 2% + T-bill; what happens if the T-bill rate increases?

- ⇨ Bond prices go down
- ⇨ So an increase in the level of the Tbill rate means that you will loose on your capital depreciation and thus decrease the value of your bond
  - But that loss is dampened by the fact that you get a higher coupon

The major risk in floaters has to do with changes in the firm's financial strength. The yield spread is fixed over the life of the security, which may be many years. If the financial health of the firm deteriorates, then investors will demand a greater yield premium than is offered by the security. In this case, the price of the bond will fall. Although the coupon rate on floaters adjusts to changes in the general level of market interest rates, it does not adjust to changes in financial condition.

- Alternative variable coupon bonds are **inverse floaters** that have coupon payments that are inversely tied to some measure of current market prices → this amplifies interest rate changes (through coupon and price effects)
  - Here the coupon is also tied to market conditions but inversely related!
  - Example: 2% **-** Tbill rate; here interest rate effect might be amplified instead of dampened
    - If the Tbill rate increases, your coupon will be reduced
    - You incur a capital loss because the value of your bond will go down
      - Double negative effect
- **Asset-backed bonds** make coupon payments that are (variable or fixed) tied and backed by the income of a pre-specified group of assets e.g. MBS, credit backed securities
- **Indexed bonds** have coupon payments that are tied to some price index → hereby both coupon payments and final payment are variable e.g. TIPS
  - Par value is adjusted to some index
  - Coupon value will then also be adjusted (% of par value)
  - Inflation products: they allow you to protect (hedge) you against inflation and adjust to constant **real returns**

TIPS

- TIPS (Treasury Inflation Protected Securities) are US T-bonds whose principal, and thus coupon payments are corrected for inflation
  - TIPS payouts are fixed in real terms, not just in nominal terms
  - TIPS have a lower bound on the final repayment value equal to the initial face value, in case of deflation
- TIPS play an important role in the economy as they represent the riskless investment for the long run

TIPS payout

Consider a newly issued TIPS with 3-years to maturity, a par value of $1000 and a coupon rate of 4%. Assume that inflation over the next years turn out to be 2%, 3% and 1%, respectively. As the face value is adjusted on the basis of actual inflation, so are the coupon payments.

Time	Inflation in Year Just Ended	Par Value	Coupon Payment +	Principal Repayment =	Total Payment
0		$1,000.00			
1	2%	1,020.00	$40.80	$ 0	$ 40.80
2	3	1,050.60	42.02	0	42.02
3	1	1,061.11	42.44	1,061.11	1,103.55

Source: BKM (2021) p.432

Table 14.1 Principle and interest payments for a Treasury Inflation Protected Security.

TIPS payout

The nominal return of this TIPS after the 1st year:

$$R_{0,1} = \frac{40.80 + 20}{1000} = 6.08\%$$

The real return if this TIPS after the 1st year:

$$R_{1,r} = \frac{1 + 0.0608}{1 + 0.02} = 4\%$$

The *nominal* rate of return on the bond in the first year is

$$\text{Nominal return} = \frac{\text{Interest} + \text{Price appreciation}}{\text{Initial price}} = \frac{40.80 + 20}{1,000} = 6.80\%$$

The real rate of return is precisely the 4% real yield on the bond:

$$\text{Real return} = \frac{1 + \text{Nominal return}}{1 + \text{Inflation}} - 1 = \frac{1.0608}{1.02} - 1 = .04, \text{ or } 4\%$$

→ This TIPS mechanism assures that you get a constant real return and not just nominal returns

Variable maturity bonds

- **Callable bonds** can be repurchased by the issuer at a pre-specified call price before maturity date (often with period of call protection)
  - corporates do this to refund at more favorable rates: when the outstanding bond is an expensive debt instrument and can fund itself cheaper in the market (when interest rates have declined)
  - the bond holder is compensated for the risk of calling by higher coupons
- **Puttable bonds** can be retired early by the bondholder
  - investors do this to switch to a more favorable investment: you use that option when interest rates have increased and have thus a more attractive bond that you can buy in the market
  - the bond holder is penalized for the option of retiring by lower coupons
- **Convertible bonds** can be exchanged for shares of the firm's common stock
  - conversion ratio: number of shares for which the bond can be exchanged
    Example: you have a bond with par value 1000 and it entitles you to 10 shares. Then based upon this conversion ratio you will look at the market conversion value.
  - market conversion value: value of shares for which the bond can be exchanged
    suppose the bond has a par value of 1000, you can convert this bond in 10 shares and the stock is traded at 80. Thus, the value of the stock is 800 and this would be the market conversion value. In this case you will not want to convert.
  - conversion premium: excess bond value over conversion value
    - when this premium is positive you will not want to convert because the bond value should be compared to the conversion value, and the higher of the two is the one you will prefer.
      - ⇨ Bond value means you will be repaid at par
      - ⇨ Conversion value means you will be repaid in stocks: this number depends on the value of the portfolio of stocks, how many stocks you get. So you will need to look up the conversion rates or the market conversion value.
  - the bond holder is penalized for the option of conversion by lower coupons
    - so a normal bond is higher in value than a convertible bond because here you will have to pay for the potential extra you can make = no free lunch

Present value relation

Bond pricing is a present value calculation that depends on

1. coupon payments (C_t)
2. principal repayment (*par value*)
3. market (spot) interest rates ($R_{0,t}$)
   - ⇨ return that investors ask to hold on to these cashflows

$$P_{0,T} = \sum_{t=1}^{T} \frac{C_t}{(1+R_{0,t})^t} + \frac{par\ value}{(1+R_{0,T})^T} \quad (1)$$

Remarks:

- Each cash flow is discounted by its (time-) corresponding market interest rate (cf term structure of interest rates)
- If coupons are paid annually, the market interest rate is an annual rate; if coupons are paid semi-annually, the market interest rate is semi-annual

The summation sign in Equation (1) directs us to add the present value of each coupon payment; each coupon is discounted based on the time until it will be paid. The first term on the right-hand side of Equation (1) is the present value of an annuity. The second term is the present value of a single amount, the final payment of the bond's par value.

You may recall from an introductory finance class that the present value of a \$1 annuity that lasts for T periods when the interest rate equals r is $\frac{1}{r}\left[1 - \frac{1}{(1+r)^T}\right]$. We call this expression the T-period *annuity factor* for an interest rate of r.[5] Similarly, we call $\frac{1}{(1+r)^T}$ the *PV factor,* that is, the present value of a single payment of \$1 to be received in T periods. Therefore, we can write the price of the bond as

$$\begin{aligned} \text{Price} &= \text{Coupon} \times \frac{1}{r}\left[1 - \frac{1}{(1+r)^T}\right] + \text{Par value} \times \frac{1}{(1+r)^T} \quad \textbf{(14.2)} \\ &= \text{Coupon} \times \text{Annuity factor}(r, T) + \text{Par value} \times \text{PV factor}(r, T) \end{aligned}$$

Example of bond price

Price of a coupon bond

Assume a 4% coupon bond with coupons paid semi-annually, a time to maturity of 10 years and a par value of €1000. Market interest rates are flat at 2% for the first 5 years, but then at 2.5% for the last 5 years (expressed on a semi-annual basis). At what price does this bond sell?

This bond has 20 semiannual coupon payments:

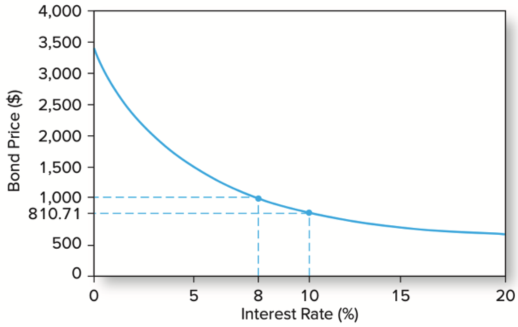
$$P_{0,T} = \sum_{t=1}^{10} \frac{20}{(1.02)^t} + \sum_{11}^{20} \frac{20}{(1.025)^t} + \frac{1000}{(1.025)^{20}} = €926.66$$

Why is this bond issued at a discount?

Remark: if the professor is not explicit about the frequency at which the bond is valid then always per annual basis.

Why is this bond issued at a discount?
A coupon of 2 is not as attractive as compared to a discount rate of 2.5, so this pushes down the price of the bond below its par value.

The shape of the curve in this figure implies that an increase in the interest rate results in a price decline that is smaller than the price gain resulting from a rate decrease of equal magnitude. This property of bond prices is called convexity because of the convex shape of the bond price curve. This curvature reflects the fact that progressive increases in the interest rate result in progressively smaller reductions in the bond price. Therefore, the price curve becomes flatter at higher interest rates.



See page 436 HB for extra information

Accrued interest

- The pricing equation (1) computes the **flat price** referred to as the **clean price**
  - it assumes that the next coupon is in exactly 1 period
  - this is also the price that is quoted in the market
  - but this is not the price you pay, you pay the dirty or invoice price
- To compute the **invoice price** or **dirty price**, the flat price needs to be adjusted with any accrued interest (the period you hold on to the bond before selling it)

  invoice price = flat price + accrued interest
- **Accrued interest** will be positive when one is between coupon dates:

$$\text{coupon payment} \times \frac{\text{days since last coupon}}{\text{days separating coupons}}$$

Example of accrued interest

Invoice price of a coupon bond

Retake the 4% coupon bond with coupons paid semi-annually, a time to maturity of 10 years and a par value of €1000. Market interest rates are flat at 2% for the first 5 years, but then at 2.5% for the last 5 years. If the bond has been issued 30 days ago, at what price does this bond sell?

The clean price of the bond (€926.66) needs to be corrected for the accrued interest over the past 30 days:

$$20 \times \frac{30}{182} = 20 \times 16.48\% = 3.30$$

This gives a dirty price of:

$$€926.66 + €3.30 = €929.96$$

Such computations can easily be done in financial calculators, e.g. in Excel (see BKM and Toledo for an example)

182: it is typically done on the actual number of days

Yield to maturity

- To evaluate and summarize the performance of a bond, we use the **yield to maturity (YTM)**
  - It is not easy to do it on the basis of the market interest rates: they vary allot
- YTM is the interest rate *Y (that do not depend on the moment of the cashflow small t)* that makes the present value of the bond's payments equal to its market price:

$$P_{0,T} = \sum_{t=1}^{T} \frac{C_t}{(1+Y_{0,T})^t} + \frac{par\ value}{(1+Y_{0,T})^T} \qquad (2)$$

- YTM can be interpreted as the internal rate of return (IRR); it assumes that all coupons can be reinvested at the YTM
- Two bonds which are identical but with different maturities will be different because of the different YTM

Yield to maturity of a ZCB

- The price of a ZCB is the present value of the bond's payoff where the relevant discount factor is the YTM

$$P_{0,T} = \left(\frac{1}{1+Y_{0,T}}\right)^T$$

$$Y_{0,T} = \left(\frac{1}{P_{0,T}}\right)^{\frac{1}{T}} - 1$$

- The YTM of a ZCB is equal to the per-period average return if the bond is held until maturity.
- ZCB yields on Treasury issued bonds play a major role as reference rates summarized in the **yield curve** or **term structure of interest rates** (this is the set of YTM, at a given moment, on ZCBs of varying maturities (see infra))

Yield to maturity of CB

For a coupon bond, there is no analytical expression for the YTM: it needs to be computed by trial & error or a financial calculator

Price of a coupon bond

Retake the 4% coupon bond with coupons paid semi-annually, a time to maturity of 10 years and a par value of 1000. Market interest rates are flat at 2% for the first 5 years, but then at 2.5% for the last 5 years (expressed on a semi-annual basis). What is the YTM of this bond?

The YTM is the rate of return that satisfies:

$$926.66 = \sum_{t=1}^{20} \frac{20}{(1+Y_{0,20})^t} + \frac{1000}{(1+Y_{0,20})^{20}} \rightarrow Y_{0,20} = 2.47\% \text{ (per 0.5 year)}$$

This YTM can be annualized using simple interest techniques, resulting in a **bond equivalent yield** of 4.94%. Compound interest techniques result in an **effective yield** of 5.00%

→ this is a semi-annual yield → needs to be annualized

Yield to maturity

*A bond will sell at par value when its coupon rate equals the market interest rate. In this case, the coupon payments are sufficient to provide fair compensation for the time value of money.

*When the coupon rate is lower than the market interest rate, the coupon payments alone will not provide bond investors as high a return as they could earn elsewhere. To receive a competitive return, they also need some price appreciation. The bonds, therefore, must sell below par value to provide a "built-in" capital gain on the investment.

When bond prices are set according to the present value formula, any discount from par value provides an anticipated capital gain that will augment a below-market coupon rate by just enough to provide a fair total rate of return. Conversely, if the coupon rate exceeds the market interest rate, the interest income by itself is greater than that available elsewhere in the market. Investors will bid up the bond price above par value. As the bond approaches maturity, its price will fall because fewer of these above-market coupon payments remain. The resulting capital loss offsets the large coupon payments so that the bondholder again receives only a competitive rate of return.

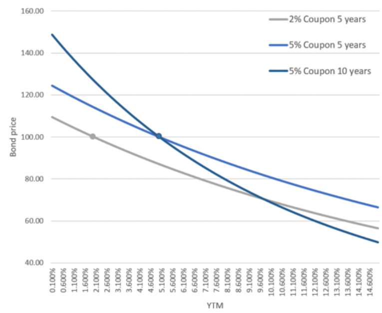
- YTM and price are inversely related, in a non-linear way
  Example: zero bond of $100/(1+y)^T$ → price and yield are inversely related
- Impact of coupon: higher coupon bonds sell at higher prices
  For yield = coupon: bond is priced at par
- Impact of maturity: longer maturity increases the steepness and the curvature

  - for yield < coupon: longer maturity bond sells at higher price
  - for yield > coupon: longer maturity bond sell at a lower price

Graph: if you plot bond prices you see indeed the inverse nonlinear relationship

→ Compare **grey VS light blue**

- You see the impact on the price of the bonds of the level of the coupon
- They are identical bonds apart from the coupon that is paid out
- When plotting the pricing function: you see that for each level of YTM the higher coupon bond (light blue) is always more expensive than the lower coupon bond (grey curve)
  - Why? Because it has higher coupons and is thus more attractive.



→ Compare **light blue VS dark blue**

- Looking at the impact of maturity:
  - Identical bonds, identical coupons but different maturities
  - The relationship is a bit more different
  - We see that the two pricing functions cross one another
  - For low levels of the yield, it is the long maturity bond that is more expensive
  - For high yield of maturity, it is the short yield that is more expensive
  - They cross one another: **the bond sells at par when its coupon = YTM**
  - *(For the grey curve this was also the case: the bond sells at par when its coupon = YTM)*

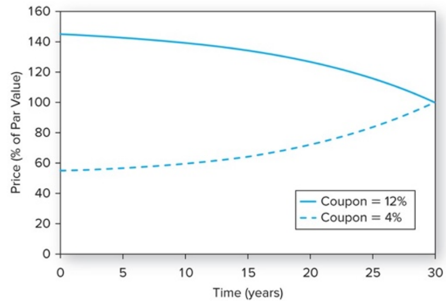
➔ if the YTM is lower than the coupon, the longer maturity bond is more expensive (and vice versa):

- The bond with an attractive coupon will trade above par; it pushes up the price of the bond but more for the longer-term bond. This is because this bond is more attractive, it yields higher returns, and you will hold on to it as long as possible. This is translated as the price being pushed above par
- The opposite occurs when YTM are very high in the market: the coupon will then be quite low and this pushed the price of the bond down, that is why the bond is sold below par. Then you will prefer the bond with the shortest maturity, and the prices of the longest bond maturity will be pushed up even more than the shorter one.

Not only coupons matter, but total returns also. This is what you want to look at the concept of YTM. This allows you to look at the performance of a bond: accounting for coupons and **capital appreciation/depreciation.**

- As YTM is the average return that is guaranteed when the bond is held until maturity (and with reinvestment at YTM), only YTM of bonds with similar maturity can be compared
- Yields of similar maturity/risk are similar (have the same YTM) as prices of such bonds adjust accordingly: what differs is the proportions earned by capital gains/losses and coupons (investors care about total returns!)
  - If coupon is low for one bond and the other bond has a high coupon, everyone will want to buy the bond with the higher coupon, and this pushes up these prices making the bonds more expensive. The total return of the coupon effect and the price effect will be identical. Because to you as an investor it does not matter if you get a coupon or capital appreciation, but **the whole return matters at the end.**

These two bonds have the same return at the end, even though they have different coupon rates. This is because the value at which you can buy these bonds are different:



- The 12% coupon bond seems very attractive because everyone seems to want to have a 12% coupon bond, and this pushes up the price: you will pay a very high premium
- The other with the low coupon, it will be sold at a discount.
  → they both have a repayment value of 100.
  - So, the low coupon bond will raise capital appreciation, and together with the coupon payments they both add up to the YTM of 8%
  - For the high coupon bond: this will have a lower value at the end, you pay a high price and over time the value of this bond declines. Meaning that you realize a capital loss/depreciation, in combination with the coupon payments, you realize a YTM of 8%
    ➔ this shows that it does not matter how you earn income, it matters how much income you gain at the end

Current yield

The **current yield** is the ratio of the bond's annual coupon to its price

- For a bond selling at a premium: coupon rate > current yield (coupon value as a function of the price) > YTM (IRR)
  - Coupon rate > current yield: C/100 > C/P
    The P is higher than par value so the ratio C/P will be smaller
  - Current yield > YTM: (coupons VS current prices) > (IRR that reflects coupons, and reflect repayment values)
    If you have a bond selling at 140 while in the future it will be repaid at par. What does that mean? You will lose capital on that bond, it declines from 140→100. This means that the current yield that reflects C/P will be above the YTM, because this YTM is a yield concept that also reflects capital depreciation/appreciation.
- For a bond selling at par: coupon rate (C/100) = current yield (C/P) = YTM
  - Example: 2-year coupon bond, 4% coupon bond. Then the yield is also 4%. Exam: I might give you just the information that the bond is selling at par with the coupon rate, than you should also know what the yield is of this bond.
- For a bond selling at a discount: coupon rate < current yield < YTM
  - Coupon rate < current yield: C/100 < C/P
    The P is lower than par value so the ratio C/P will be larger
  - A bond selling at 60 and will be paid back for 100. Here you will have a capital appreciation. So your YTM > current yield.

Coupon rate, current yield and YTM

Current yield vs coupon rate and YTM

Retake the 4% coupon bond with coupons paid semi-annually, a time to maturity of 10 years and a par value of 1000. With a price of 926.66, the bond's effective annual yield was computed at 5%. The current yield then equals:

$$\frac{40}{926.66} = 4.32\%$$

For this discount bond we thus have

$$\textit{coupon rate } (4\%) < \textit{current yield } (4.32\%) < \textit{YTM } (5\%)$$

The coupon rate is lower than the current yield as the former divides the annual coupon payment over the higher par value instead of the price. In addition, the YTM is higher than the current yield as the latter incorporates the future capital gain on the bond when converging to face value.

Realized returns versus YTM

YTM does say something only when you are holding your bond till maturity. Yield to maturity will equal the rate of return realized over the life of the bond if all coupons are reinvested to earn the bond's yield to maturity.

But suppose I do not hold it till maturity? Or what if the reinvestment rate does not equal the YTM?

- Realized returns over the maturity will differ from YTM when reinvestment of coupons cannot be done at the YTM, or when the bond is not held until maturity
    - Even at the start when I buy the bond it might be that I do not want to hold it till maturity; I will still want to come up with a return. So, I have to do my own analysis
- At maturity (or when the bond is sold) performance can be computed as the realized compound return, accounting for the actual reinvestment rates
- But the problem is that before maturity, with future interest rate uncertainty, the rates at which interim coupons will be reinvested are not yet known. Therefore, although realized compound return can be computed *after* the investment period, it cannot be computed in advance without a forecast of future reinvestment rates.
- Before maturity, investment rates can only be forecasted based upon
    - forecasts of future reinvestment rates
    - forecast of future bond price
- Forecasting the compound returns over various holding periods is called **horizon analysis**
    - I take into account the current price that I paid, the expected value of the bond in the future when I sell the bond and also reinvestment rates (YTM assumes that you can reinvest at the YTM)
    - The forecast of total return depends on your forecasts of *both* the price of the bond when you sell it at the end of your horizon *and* the rate at which you are able to reinvest coupon income. The sales price depends in turn on the yield to maturity at the horizon date.

*This demonstrate that as interest rates change, bond investors are subject to two offsetting sources of risk. On the one hand, when rates rise, bond prices fall, which reduces the value of the portfolio. On the other hand, reinvested coupon income will compound more rapidly at those higher rates. This **reinvestment rate risk** offsets price risk.*

Forecasting bond returns

Horizon analysis

Suppose you buy a 30-year 7.5% coupon bond with annual coupon payments for $980. You plan to hold this bond for 20 years. Your forecast is that the bond's YTM will be 8% when the bond is sold and that the reinvestment rate on the coupons is 6%. What is the YTM of this bond?

What is your best forecast of the compound return (p.a.)?

(If you do not understand this look at the video 1:35.)

Horizon analysis

maturity	30
face value	1000
coupon rate	0.075
coupon/year	1
price	980
Horizon	20
YTM (at t = 20)	0.08
Reinvestment rate	0.06

At horizon 0	
YTM	7.67%

At horizon 20	
E(Price)	966.45
E(Coupons)	2758.92
E(value)	3725.37
E(compound HPR)	6.90%

Source: BKM (2021) p.443
Example 14.9 Horizon analysis.

Realized return over a single period

Apart from this long horizon analysis you could also limit yourself to 1 period returns.

- To compare the performance of bonds of different maturities we can compute **1-period realized returns**
- We limit ourselves to **zero coupon bonds** because this really reflects what determines the rate of return and what the link is with the observed yields in the market
- Assume we buy a ZCB at t with a maturity of T (at a price P_{t}, T), and sell this bond at t+1 when it has a maturity of T-1 (at a price $P_{t+1}, T-1$)
- Its 1-period realized return is computed as:

$$1 + R_{t,1} = \frac{P_{t+1,T-1}}{P_{t,T}}$$

For a ZCB this 1-period realized return can be written as (the ratio of the two prices):

$$1 + R_{t,1} = \frac{P_{t+1,T-1}}{P_{t,T}} = \frac{(1 + Y_{t,T})^T}{(1 + Y_{t+1,T-1})^{T-1}}$$

The log version of this realized return is even more intuitive:

$$\begin{aligned} r_{t,1} &= p_{t+1,T-1} - p_{t,T} \\ &= Ty_{t,T} - (T-1)\, y_{t+1,T-1} \\ &= y_{t,T} - (T-1)\, (y_{t+1,T-1} - y_{t,T}) \end{aligned}$$

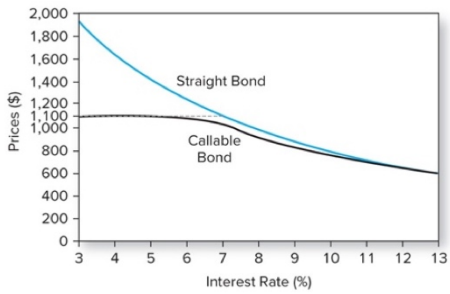
- You can rearrange this by plugging in the functions; you express prices in function of its yields
- The log returns are just the difference in price: the price I observe today – the price of yesterday
- $\ln(1+Y_{t,T})^T = T\ln(1+Y_{t,T}) = Ty_{t,T}$
- the realized return equals = the YTM at the moment I bought the bond – any changes in the YTM
- a decrease in YTM drives up the realized return (and price)
- when yields go up, return goes down because when yield go up, prices go down and this means that I will only be able to sell this bond at a lower price
- in context of the ZCB you easily see how realized returns are linked to YTM
- the realized return equals the initial YTM only if the YTM is unchanged over the period

Yield to call

YTM calculates the average return when the bond is held to maturity, but for callable bonds such yield calculation is not informative when the risk of being called is high

YTC is similar to YTM with the only difference that you assume that the bond will terminate and reach maturity at the moment that it is called, and that the repayment value is the call value.

Whether a bond is likely to be called, depends on the current level of interest rates:



- For high levels of interest rates: call risk is small, and the callable bond behaves very much like a straight bond
- For low levels of interest rates: call risk is high, and the value of a callable bond **diverges** from the value of a straight bond
    - the divergence reflects the firm's option to redeem the bond early
    - for very low levels of interest rates the bond's value converges to the call value
- For a bond that is likely to be redeemed early it is more intuitive to compute the **yield to call**
- This is the average return of the bond over the period until the bond is likely called at its call price (e.g. yield to first call)
- There are periods where these bonds can be called (yield to first call) and when these bonds cannot be called:
    - This analysis suggests that investors might be more interested in a bond's yield to call than its yield to maturity, especially if the bond is likely to be called. The yield to call is calculated just like the yield to maturity except that the time until call replaces time until maturity, and the call price replaces par value. This computation is sometimes called "yield to first call," as it assumes the issuer will call the bond as soon as it may do so.

Yield to call

Retake the 30-year 7.5% coupon bond with annual coupon payments that quotes at a price of $980. Assume this bond can be called in 10 years at a call price of $1100.

What would be its yield to first call?

Based on current conditions, is a call likely?

maturity	30
face value	1000
coupon rate	0.075
coupon/year	1
price	980
Horizon	20
YTM (at t = 20)	0.08
Reinvestment rate	0.06
first call date maturity	10
call price	1100

yield to call	8.48%
YTM	7.67%

Default risk

There is always a risk that the bond issuer will default on its obligation, and this risk will be reflected on the market interest rates. That is because the market interest rates will be like a fair compensation, as agreed in the market to take on those cash flows with that particular default risk.

Yields spreads

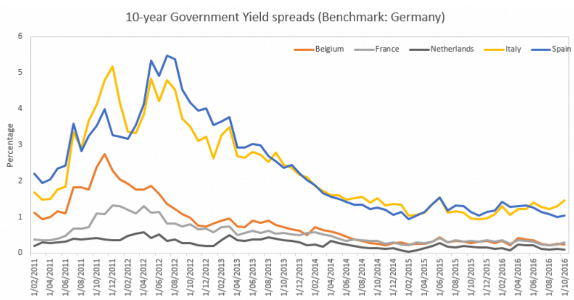
Yields on comparable bonds of different issuers will be different for two reasons:

1. Differences in credit risk or default risk: there is always a likelihood that the issuer will default on its obligations
2. Differences in liquidity risk: there is always a risk that selling the bond at prevailing market prices is not easy

Such differences in risk are reflected in the **yield spread** (most likely the yield spread reflects credit risk).

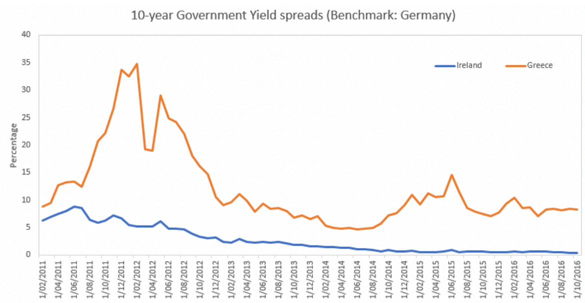
Yield spreads in EMU

Here to give you an idea of credit risk:



Here you see the yield spread for those countries, from 2011 until 2016. You indeed see that these spreads have been considerable.
Italy and Spain have been traded at high spread, implying that these governments were paying much more than the German government (who are considered safe heaven).
At very low spreads we see a convergence.

These two countries suffered during the crisis, and you can see that on this graph. Greece has defaulted on a number of bonds.
This makes clear that credit risk is real. Even within EU, government bonds have credit risk.



Credit ratings

How can we get an idea of credit risk? Look at the credit rating by a credit rating agency.

- The default risk of a bond is typically captured in its credit rating issued by a credit rating agency e.g. S&P, Moody's and Fitch
- These rating agencies issue short term and long term ratings (of which the LT are best known)
    - ratings are forward looking opinions of the relative credit risks
- The rating system across the different credit rating agencies are similar (though not identical).
    - Investment grade
    - Speculative grade or junk bonds (below BBB or Baa)
- The opinion as reflected in this credit rating differ between the different credit rating agencies. They all use other definitions of the ratings.
- Whenever you look at a particular rating agency, it is important to understand what triple A means or B, … **for that particular agency**.

Example: Credit ratings by Moodys

- Moodys ratings reflect
    - likelihood of a default on contractual payments
    - expected financial loss suffered in the case of default
- Long-term ratings are assigned to issuers/issues with an original maturity of ≥ 1 year
- Expected credit loss: determined by likelihood of default, loss given default and exposure at default = normal calculation but for Moody they only use likelihood of default and loss given defaul

- Short-term ratings are assigned to obligations with an original maturity ≤ 13 months

Global Long-Term Rating Scale	
Aaa	Obligations rated Aaa are judged to be of the highest quality, subject to the lowest level of credit risk.
Aa	Obligations rated Aa are judged to be of high quality and are subject to very low credit risk.
A	Obligations rated A are judged to be upper-medium grade and are subject to low credit risk.
Baa	Obligations rated Baa are judged to be medium-grade and subject to moderate credit risk and as such may possess certain speculative characteristics.
Ba	Obligations rated Ba are judged to be speculative and are subject to substantial credit risk.
B	Obligations rated B are considered speculative and are subject to high credit risk.
Caa	Obligations rated Caa are judged to be speculative of poor standing and are subject to very high credit risk.
Ca	Obligations rated Ca are highly speculative and are likely in, or very near, default, with some prospect of recovery of principal and interest.
C	Obligations rated C are the lowest rated and are typically in default, with little prospect for recovery of principal or interest.

Note: Moody's appends numerical modifiers 1, 2, and 3 to each generic rating classification from Aa through Caa. The modifier 1 indicates that the obligation ranks in the higher end of its generic rating category; the modifier 2 indicates a mid-range ranking; and the modifier 3 indicates a ranking in the lower end of that generic rating category. Additionally, a "(hyb)" indicator is appended to all ratings of hybrid securities issued by banks, insurers, finance companies, and securities firms.*

Note: For more information on long-term ratings assigned to obligations in default, please see the definition "Long-Term Credit Ratings for Defaulted or Impaired Securities" in the Other Definitions section of this publication.

* *By their terms, hybrid securities allow for the omission of scheduled dividends, interest, or principal payments, which can potentially result in impairment if such an omission occurs. Hybrid securities may also be subject to contractually allowable write-downs of principal that could result in impairment. Together with the hybrid indicator, the long-term obligation rating assigned to a hybrid security is an expression of the relative credit risk associated with that security.*

Global Short-Term Rating Scale	
P-1	Issuers (or supporting institutions) rated Prime-1 have a superior ability to repay short-term debt obligations.
P-2	Issuers (or supporting institutions) rated Prime-2 have a strong ability to repay short-term debt obligations.
P-3	Issuers (or supporting institutions) rated Prime-3 have an acceptable ability to repay short-term obligations.
NP	Issuers (or supporting institutions) rated Not Prime do not fall within any of the Prime rating categories.

Credit risk determinants
The definition that different credit agencies use are different but also the methodologies. Credit rating agencies use proprietary methodologies, there is no transparency about this.

Common determinants of credit risk include typical financial ratios
- Coverage ratio: Ratios of company earnings to fixed costs. A decreasing measure of earnings to fixed costs signals distress
- Leverage ratio, debt-to-equity ratio: high indebtedness signals distress
- Liquidity ratio: The two most common liquidity ratios are the *current ratio* (current assets/current liabilities) and the *quick ratio* (current assets excluding inventories/current liabilities). A low ability to pay bills coming due with liquid assets signals distress
- Profitability ratio: Measures of rates of return on assets or equity. The *return on assets* (earn- ings before interest and taxes divided by total assets) or *return on equity* (net income/ equity) are the most popular of these measures. A low return on equity or return on assets signals distress
- Cash flow-to-debt ratio: This is the ratio of total cash flow to outstanding debt. A low ability to cover total debt with its yearly operational cash flow signals distress

	Aaa	Aa	A	Baa	Ba	B	C
EBITA/Assets (%)	12.3%	10.2%	10.8%	8.7%	8.5%	6.7%	4.1%
Operating profit margin (%)	25.4%	17.4%	14.9%	12.0%	11.5%	9.0%	4.6%
EBITA to interest coverage (multiple)	11.5	13.9	10.7	6.3	3.7	1.9	0.7
Debt/EBITDA (multiple)	1.9	1.8	2.3	2.9	3.7	5.2	8.1
Debt/(Debt + Equity) (%)	35.1%	31.0%	40.7%	46.4%	55.7%	65.8%	89.3%
Funds from operations/Total debt (%)	45.5%	43.4%	34.1%	27.1%	19.9%	11.7%	4.6%
Retained Cash Flow/Net Debt (%)	31.4%	30.1%	27.3%	25.3%	19.7%	11.5%	5.1%

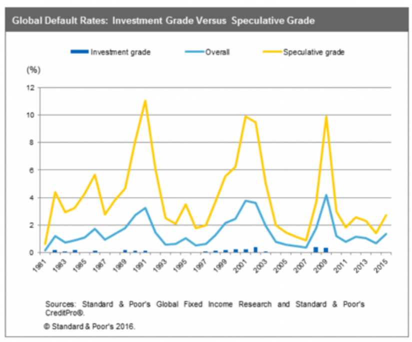
Table 14.3

Financial ratios by rating class

Note: EBITA is earnings before interest, taxes, and amortization. EBITDA is earnings before interest, taxes, depreciation, and amortization.
Source: Moody's Financial Metrics, *Key Ratios by Rating and Industry for Global Non-Financial Corporations*, December 2016.

There is thus a correlation between the financial ratios of a company that reflects the health of a company and the credit ratings.
Ratios tend to improve along with firms' rating classes.
- We see for example in the table that the higher the credit ratings the higher the operating profit margin.
- The higher the debt of the company, the lower the credit rating on average.

Default rates
Are these credit ratings any good? Do companies with high rating have a lower chance of default?
- Defaults on investment grade are rare (dark blue bars) which shows that indeed companies with high credit rating have a low risk of defaulting
- But much more common for speculative grade (yellow line)
- There is time-variation in default rates; actual default rates vary over time with a clear link to economic growth
  - When we enter a recession, default rate spikes: it is the companies that are already in difficulties that will go bust quickly



Bond indentures

A bond is issued with an **indenture** (=contract) which stipulates a set of restrictions to protect the rights of the bondholders, it stipulates also your right as a bondholder or bond issuer

- **Collateral provisions**: some bonds are backed by specific collateral which can be seized by the bondholders in case of default (←→ debenture bonds)
  e.g.
  - mortgage bond: if the collateral is property
  - collateral trust bond: if the collateral takes the form of other securities held by the firm
  - equipment obligation bond: in the case of equipment

  Collateralized bonds generally are considered safer than general **debenture bonds,** which are *unsecured,* meaning they do not provide for specific collateral.

- **Sinking fund** specifications: when the issuer needs to repay to bond at the end of maturity it is a big burden so to spread the repayment burden of a bond issue over several years, a sinking fund allows to call bonds early
  - Here the bonds will be called **in the context of a sinking fund specification.** There are two ways:
    1. repurchase fraction of outstanding debt in the open market
    2. repurchase fraction of outstanding debt at special call price

  The sinking fund call differs from a conventional bond call in two important ways. First, the firm can repurchase only a limited fraction of the bond issue at the sinking fund call price. At most, some indentures allow firms to use a doubling option, which allows repurchase of double the required number of bonds at the sinking fund call price. Second, while callable bonds generally have call prices above par value, the sinking fund call price usually is set at the bond's par value.

- **Dividend policy**: restrictions on dividend payments force the firm to retain assets rather than pay them out to shareholders, as not to jeopardize the payout of the bondholders
  e.g. dividends are not allowed if the cumulative dividend paid since inception > cumulative retained earnings + stock sale proceeds

- **Future borrowing**: subordination clauses restrict the amount of additional borrowing and priority ranking
  - *If you bought a bond today, you would be understandably distressed to see the firm tripling its outstanding debt tomorrow. The bond would be riskier than it appeared when you bought it. To prevent firms from harming bondholders in this manner,* ***subordination clauses*** *restrict the amount of additional borrowing. Additional debt might be required to be subordinated in priority to existing debt; that is, in the event of bankruptcy, subordinated or junior debtholders will not be paid unless and until the senior debt is fully paid off.*

Stated versus expected YTM

A difference exists between the **stated YTM** and the **expected YTM** (assuming reinvestment at YTM)
If the market is in doubt that the cashflows will be paid out as promised than there will be a difference.

- The stated YTM is the maximum possible YTM that is only achieved in case the bond does not default, this assumes that the cashflows will all be paid.
- The expected YTM is the YTM that corrects for expected default losses. If you expect that the bond will not be repaid at full than you will discount that amount.

Expected yields

Suppose a firm has issued a 9% coupon bond, 20 years ago. The bond has 10 years left to maturity and trades at \$750, but the firm has financial difficulties. Investors believe that the semi-annual coupon payments will still be paid, but that the final payment will be \$700 instead of \$1000 par value. What is the stated and the expected yield?

maturity (annual)	10
face value	1000
expected repayment	700
coupon rate (annual)	0.09
coupon/year	2
price	750

Promised YTM	6.83%
Promised bond equivalent yield	13.66%
Promised bond effective yield	14.12%
Expected YTM	5.82%
Expected bond equivalent yield	11.63%
Expected bond effective yield	11.97%

Yield spreads

To compensate for the possibility of default, corporate bonds must offer a **default premium.** The default premium, also called a credit spread, is the difference between the promised yield on a corporate bond and the yield of an otherwise-identical government bond that is riskless in terms of default.

- If the firm remains solvent and actually pays the investor all of the promised cash flows, the investor will realize a higher yield to maturity than would be realized from the government bond.
- If, however, the firm goes bankrupt, the corporate bond is likely to provide a lower return than the government bond. The corporate bond has the potential for both better and worse performance than the default-free Treasury bond. In other words, it is riskier.

The pattern of default premiums offered on risky bonds is sometimes called the *risk structure of interest rates.*

Credit derivatives

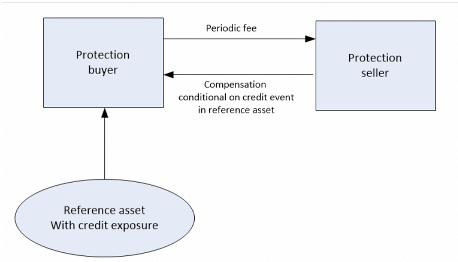
- Credit derivatives are instruments to **transfer** credit risk to a party better placed/more willing to take the risk → to this end, the credit risk must be **isolated** and **priced**
- You must make sure that these are correctly priced (this is what went wrong during the credit crisis)
- Parties involved in a credit risk transfer are
  - the protection buyer who transfers the credit risk (the one who want to get rid of the risk)
  - the protection seller who takes on the credit risk
  - (sometimes) the special purpose vehicle who act as intermediary in this transfer

Credit default swaps (CDS)

- Credit default swaps are largest segment of market of credit derivatives
- It is an **insurance policy** on a 'credit event' (e.g. default, yield spread widening, rating downgrade)
  - the protection buyer receives a compensation for a credit event loss
  - the protection seller receives a fee payment to take on this credit risk
- Different CDS products exist
  - standard CDS
  - digital CDS
  - basket CDS
  - portfolio CDS

Standard CDS
Two involved parties make following payments

- the protection buyer receives a compensation for a credit event
- the protection seller receives a periodical fee for taking on credit risk



- The reference asset within the CDS can be any asset with default risk (OTC contract/tailor-made) agreed between the seller and buyer
- The CDS premium is defined as a fraction of the notional amount of the reference asset
- The CDS contract expires in case of
    - maturity of the CDS is reached, without a credit event happening
    - credit event occurs and triggers the conditional payment

CDS cashflows

Assume 2 parties entering a 5 year CDS with a $1,000 6% coupon bond as a reference asset. The CDS premium is 1% p.a., paid on a semi-annual basis. The credit event is defined as default on the coupon payments. Assume default occurs after 3 years, and that recovery is estimated at $400. What are the CDS cash flows?

Time (in months)	Periodic fee	Conditional payment	Net CF to protection buyer
6	$5	$0	-$5
12	$5	$0	-$5
18	$5	$0	-$5
24	$5	$0	-$5
30	$5	$0	-$5
36	$5	$600	$595
Total			$570

Digital CDS

- A digital CDS is very similar to a standard CDS with the only difference related to the size of the conditional payment
    - in a standard CDS the conditional payment equals **the actual loss** as defined by the recovery rate
    - in a digital CDS the conditional payment is **pre-agreed**, irrespective of the recovery rate → 'fixed recovery CDS'
- The predefined recovery value is generally larger than the actual recovery value in case of default → the CDS premium on a digital CDS is typically larger than on a standard CDS

Basket CDS

- A basket CDS is a CDS on a pool of reference assets (=multi-name CDS) as compared to a single reference asset (=single-name CDS)
- The CDS insures against a credit event on a number of components of the pool
    - first-to default basket CDS: only a single event can occur, and than the CDS is terminated.
    - second-to-default basket CDS: you have example 20 instruments in the basket, if the first default event occurs your default will be compensated. If a second credit event occurs, you still be compensated and then the basket CDS is terminated.
    - n^{th}-to-default basket CDS
- The basket CDS is terminated when the default number is reached
- The CDS premium on the basket is lower than the sum of CDS premia on the individual components

Portfolio CDS

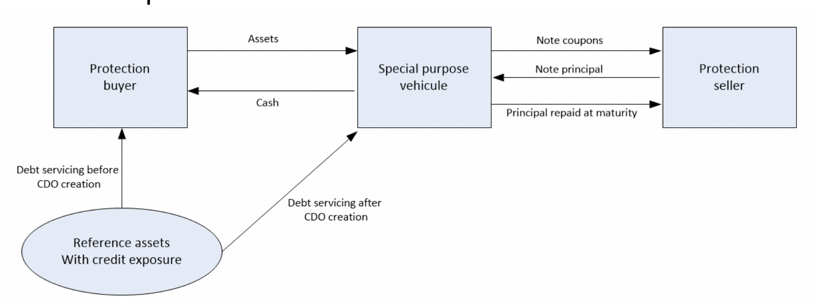
- A portfolio CDS is similar to a basket CDS (a number of assets), but insures against a credit event on the portfolio as measured by **an amount**
- The portfolio CDS is terminated when the credit amount has been reached
- Smaller defaults are covered as compared to a basket default

Collateralized debt obligation (CDO)

- A CDO is a credit instrument backed by a pool of assets, where the untradeable debt (like a mortgage, loan, …) is transformed into tradeable securities (= securitization)
  - CLO: collateralized loan obligation
  - CBO: collateralized bond obligation
  - CMO: collateralized mortgage obligation
- The transformation of the debt is possible because of the pooling → this allows for a more accurate evaluation of the credit risk (cf insurance business)
  - It is more difficult to calculate the risk of an individual car accident than to calculate the risk of a pool of car accidents
  - The likelihood that one particular debt issuer will default might not be accurate to estimate, but in the pool of assets we can more accurately calculate the likelihood of default and losses

Traditional CDO

- Assets with credit risk are pooled and transformed into tradeable assets
- The CDO involves 3 parties
  - the protection buyer transfers assets to a special purpose vehicle in exchange for the principal value
  - the special purpose vehicle transforms the assets into tradeable notes
  - the protection seller invests in CDO notes, where the coupon payments reflect the degree of credit risk
    - they are the ultimate note owners, they have bought instruments, bonds, notes where these reference assets are backing these notes. This means that if there are losses in the credit instruments, these losses will be taken by these bondholders. Not each bondholder takes the same amount of risk, you have different categories.
    - If you are a seller and buy a note with low credit worthiness than that means you have high credit default risk, than you need to be compensated for that



Standard CDO example

Protection seller types:
Senior: all holders take the smallest risk
Mezzanine: all holders take smaller risk
Equity: all holders take the largest risk → if there is a credit loss this group will fist take the loss and when they all incurred it than we go to the mezzanine, …

Principle of a CDO

Assume a CDO is written on a portfolio of loans with principal $100mio. The loans earn a spread of 400bp. The CDO consists of 3 tranches

CDO tranche	Credit rating	Notional (in mio $)	Spread (in bp)
Senior	AAA	80	70
Mezzanine	A	15	200
Equity	unrated	5	1300
Total portfolio		100	151

Do you see within this example the motivation to initiate this structure?

- What is the cost to the SPV? Return it needs to pay to the ultimate noteholders. In this example it is 151 basis points.
- What is the reward to the SPV? That is the return on the loans, the credit sensitive instrument they earn a spread in this example of 400bp.
- The difference in spread is clearly a gain: it earns 400 basis points and needs to pay 151
- You can call this an arbitrage strategy as these assets are the same and just repacked (see here under)

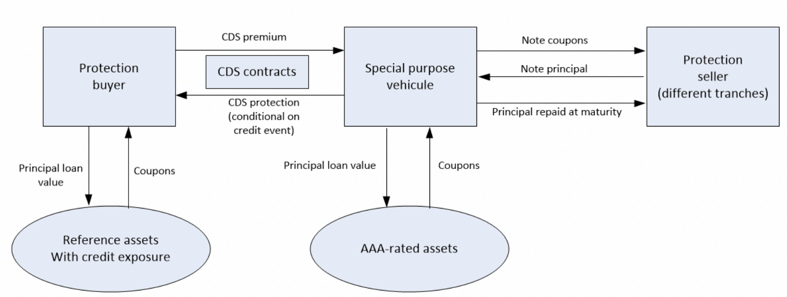
- Depending upon the motivation to securitize we distinguish between
  - a balance sheet CDO
  - an arbitrage CDO
- A balance sheet CDO sells off assets to the SPV to remove the credit risks from the balance sheet
  - creation of additional lending capacity
  - reduction of risk capital leading to improved ROE
- An arbitrage CDO aims to create value by repacking the debt into tradeable securities
  - as the value of the underlying collateral equals the value of the notes, the notes' return should equal the debt's return
  - in practice: the return on the notes is lower than the return on the debt → arbitrage profits (see example above)

Synthetic CDO

This is different from a standard CDO in the way credit risk is transferred.

- Characteristic of a standard CDO is the transfer of the underlying debt pool to the SPV i.e. not only the credit risk is transferred, also the assets and all their risk (and their market risks) are transferred
- In a synthetic CDO a CDO structure is combined with CDS: here they set up a structure such that it is just the credit risk that is transferred and not the interest rate risk → much more efficient because a credit derivative in general is used to transfer credit risk and a standard CDO does this but the side effect is that you transfer market/interest rate risk
  - the CDS allows to transfer credit risk
  - the CDO structure allows to issue CDO notes based on the CDS as collateral
  - the SPV does not take ownership of the reference asset, but takes on the credit exposure by being the bank's counterparty in the CDS

As the SPV does not buy the assets, it does not need funding → it can use the principal of CDO notes to invest in high grade low risk assets to generate additional returns



Chapter 9: Managing fixed income portfolios

BKM - Chapter 15: The term structure of interest rates
BKM - Chapter 16: Managing bond portfolios

→ how to manage interest rate risk in fixed income instruments? The interest rate is the discount factor. Whenever this changes, the value of your fixed income instrument will change. So having an understanding of current and future interest rates is crucial for managing your interest rate risk.

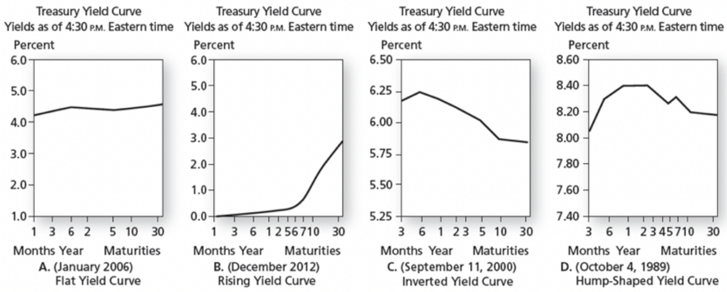
The yield curve

- The **yield curve** or **term structure of interest rates** is the set of YTMs, at a given moment, on (Z)CB issued by the government of varying maturities
  - It is a summary of base rates in the economy
  - It gives you the different levels of interest rates that the government needs to pay on debt (most of the time on ZCB)
  - It gives you info about current interest rates and depending on the shape of the yield curve you know something about future interest rates.
  - This yield curve is a key concern for a fixed income investor
  - it is central to bond pricing as it gives the discount rates for various future cash flows
  - it allows the investors to assess their expectations about future rates as compared to the market
- There exist two types of yield curves
  - the **pure** yield curve (yield on zero coupons) uses stripped (a strip is a coupon that is being sold as an instrument separately which than can be viewed as a ZC) or ZC treasuries
  - example of a strip:
    - suppose you have a coupon with 2 years till maturity

$$P = \left(\frac{C}{1+R_{0,1}} + \frac{C+Par\ value}{(1+R_{0,2})^2}\right)$$

$$\rightarrow P = \frac{C}{1+Y} + \frac{C+Par\ value}{(1+Y)^2}$$

    - on that instrument you can calculate the yield: it is that yield that makes sure that if we discount future CF by this yield, we get the observed market price
    - $R_{0,1}$ and $R_{0,2}$ are the market interest rates (these are pure yields)
    - Y are the YTM and for a coupon bond these YTMs are some kind of average of the underlying zero yields $R_{0,1}$ and $R_{0,2}$ e
    - When I talk about pure yields I thus talk about the R and this instrument can be stripped into two separate instruments:
      - I can trade the first CF as a separate instrument $P_1 = \frac{C}{1+R_{0,1}}$
      - I can trade the second CF as a separate instrument $P_2 = \frac{C+Par\ value}{(1+R_{0,2})^2}$
      - And the rate of returns at which I discount these strips are the zero yields
      - But if you would compute the yield on the combined product that is not the zero yield but the YTM
      - So the yields you observe are either the pure yields or the on-the-run yields;
  - the **on-the-run yield curve** uses recently-issued coupon bonds selling at or near *par*
    - this means that you know what the coupon of that bond is; if the $y_{0,2}$ = 5% than the coupon is 5%.

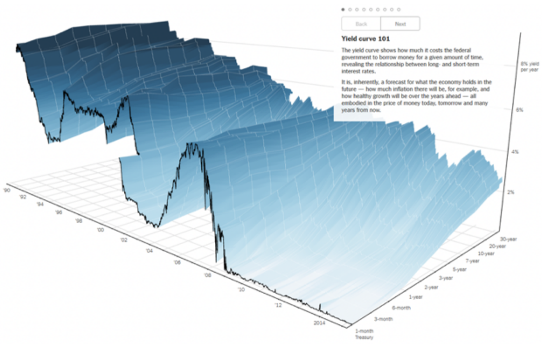


To give you an idea of the variability of the yield curve and how they change you see here these graphs:

- the first one is a flat curve: the yield on short maturities is more or less equal to the yield on longer maturities
- the second one is an upward sloping yield curve
- the third one is an inverted yield curve implying that yields on short term bonds were higher as compared to yields on long term bonds (this is typical)
- the last one is a hump shaped: upward sloping on the short-end and downward sloping on the long end

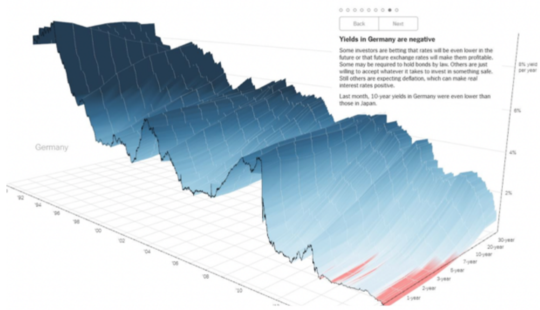
what you need to remember is that both the location of the yield curve as the shape of the yield curve can vary significantly over time. This is always a snap shot each day that there will be a new yield curve.

The US yield curve over time



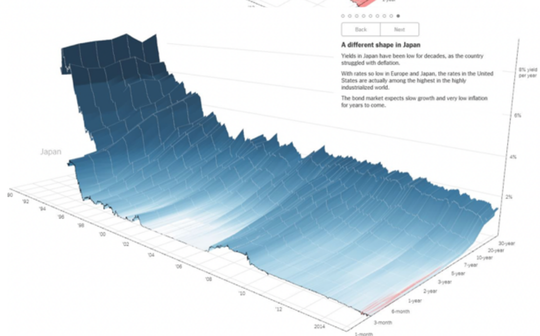
This is a 3D graph from the NYT. Here you see the evolution of the yield curve over time. So the view of the previous section where we always looked at the yield curve at one particular moment in time is actually the view from this angle. You see indeed that over time this yield curve changes dramatically.

The German yield curve over time



Here you see a red area for Germany, this implies negative yields. This corresponds with the fact that yields on the short end for some European countries were negative.

The Japanese yield curve over time



Japanese yield curves has been quite stable for a long period of time.

Explaining the yield curve

How can we explain the yield curve's behaviour?

To study the behaviour of the yield curve we analyze:
1. The yield curve under certainty
2. The yield curve under uncertainty

Explaining under certainty

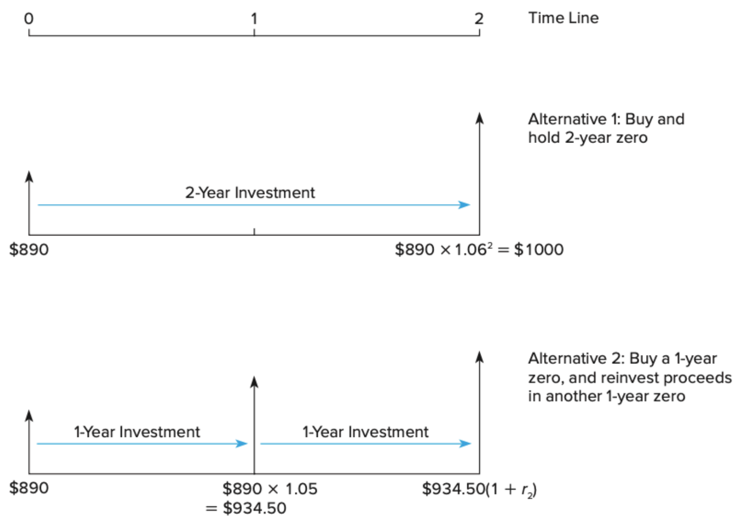
Here in this example, we talk about zero coupons so the yield = return.
No uncertainty about future interest rates = you know the rate at which you can invest in the future. Suppose you have the two following investment options.

Assume the 1-year yield $Y_{0,1}$ ($= R_{0,1}$) is 5% and the 2-year yield $Y_{0,2}$ is 6%. There is no uncertainty about future interest rates. You have the choice between following 2-year bond investments:

1. Today you invest in a 2-year bond that you keep until maturity
2. Today you invest in a 1 year bond; in one year you invest in a 1-year bond

What will be the future 1-year rate $R_{1,1}$?

If there is no uncertainty the two investment options are identical. Suppose I go for the first alternative. I know this will give me a yield of 6%.
The alternative 2 is that I invest 890 in a one-year bond which means I get 934.50 after one year. I can reinvest this amount in again a one-year bond. This will give me r_2 or example $R_{0,2}$ with 0 the moment when you start and thus from which your rate is valid and 2 means for how long the rate is valid.
$R_{1,1} = r_2$; R that starts in one year and that last one year. The handbook labels this r_2 which is the same.



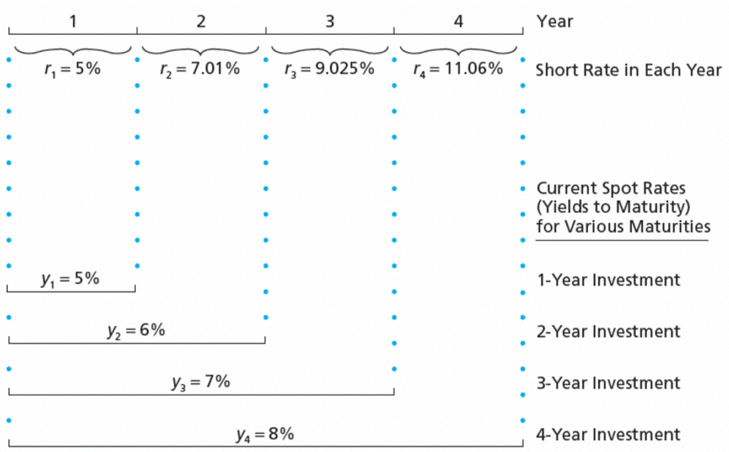
In a world without risk, **bond investments with identical overall maturity, must offer identical returns**. If not, there are arbitrage opportunities:

$$(1 + Y_{0,2})^2 = (1 + R_{0,1})(1 + R_{1,1})$$
$$R_{1,1} = 7.01\%$$

Next year's 1-year rate will be just enough to make rolling over a series of 1-year bonds equal to investing in the 2-year bond.

So while the 1-year bond offers a lower yield to maturity than the 2-year bond (5% versus 6%), we see that it has a compensating advantage: It allows you to roll over your funds into another short-term bond next year when rates will be higher. Next year's interest rate is higher than today's by just enough to make rolling over 1-year bonds equally attractive as investing in the 2-year bond.

- The **short rate** is the rate for a given short maturity at different points in time
    - $R_{0,1}$ and $R_{1,1}$ are short rates because they are valid for a short period of time
- The **spot rate** is the rate the prevails today for a given maturity → a spot rate is the geometric average of its component short rates *(Hb: They call the yield to maturity on zero-coupon bonds the **spot rate,** meaning the rate that prevails today for a time period corresponding to the zero's maturity. → is da nu voor ZCB of ??)*
    - A spot rate is always a rate which the first subscript is a zero: $Y_{0,2}$ it is valid on the spot as of now.
    - $R_{1,1}$ is a future rate and no spot rate



Here we have on the bottom spot rates: *$Y_{0,1}$= 5%, $Y_{0,2}$ = 6%, $Y_{0,3}$ = 7% and $Y_{0,4}$ = 8%.* On top we have the future rates: *$R_{0,1}$ = 5%, $R_{1,1}$ = 7.01%, $R_{2,1}$ = 9.025%, $R_{3,1}$ = 11.06%.*

You can easily see that the longer spot rates are a geometric average of the short rates. You can play around with them. For example the yield on a 4 year bond is a geometric average of all the short returns: $(1+ Y_{0,4}) = (1+Y_{0,1})(1+ R_{1,1})(1+ R_{2,1})(1+ R_{3,1})$ (this is investing in a 4 year bond and rolling over 4 times) or

$= (1+ Y_{0,3})^3 (1+ R_{3,1})$ (investing in a 3 year bond and rolling over in the last year)

- When the term structure is **upward sloping**
    - *$Y_{0,1} < Y_{0,2}$ → $R_{1,1}$ < or > $Y_{0,1}$?*
        - *$(1+Y_{0,2})^2 = (1+Y_{0,1})(1+R_{1,1})$*
        - *You know that $Y_{0,1} < Y_{0,2}$ this means then that $R_{1,1} > Y_{0,1}$*

        - *This means that market expect future ST interest rates to be higher than current interest rates and this say something about how the market thinks about future market interest rates; they expect the CB to increase interest rates and the economy is then doing well*
    - long spot rates are higher than short spot rates
    - it therefore implies that future short rates are expected to rise
    - *(When next year's short rate, $R_{1,1}$, is greater than this year's short rate, $Y_{0,1}$, the average of the two rates is higher than today's rate, so $Y_{0,2} > Y_{0,1}$ and the yield curve slope upward.)*

When the term structure is **downward sloping**

-
    - *$Y_{0,1} > Y_{0,2} \rightarrow R_{1,1}$ < or > $Y_{0,1}$?*
        - *$(1+Y_{0,2})^2 = (1+Y_{0,1})(1+R_{1,1})$*
        - *You know that $Y_{0,1} > Y_{0,2}$ this means then that $R_{1,1} < Y_{0,1}$*
        - *This means that market expect future ST interest rates to be lower than current interest rates and they expect the CB to decrease interest rates and the economy is then doing not so well, the economy then needs extra stimulus that is why interest rates decrease*
    - long spot rates are lower than short spot rates
    - it therefore implies that future short rates are expected to decrease
- *(If next year's short rate were less than $Y_{0,1}$, the yield curve would slope downward.)*

From this term structure of interest rates, we can extract information about future interest rates.

More generally, we can write:

$$(1+Y_{0,T+1})^{T+1} = (1+Y_{0,T})^{T}(1+R_{T,1})$$

- with $Y_{0,T}$ the YTM of a ZCB with a maturity T
- with $R_{T,1}$ the short rate earned on an investment starting at T

This allows us to solve for the short rate $R_{T,1}$ in the future period starting at T, based on the observed yield curve:

$$(1+R_{T,1}) = \frac{(1+Y_{0,T+1})^{T+1}}{(1+Y_{0,T})^{T}}$$

- the short rate $R_{T,1}$ makes up for the difference between the total return on a (T+1)-period ZCB and the total return on a T-period ZCB
- The numerator on the right-hand side is the cumulative growth of an investment in an *(T+1)*-year zero held until maturity. Similarly, the denominator is the growth of an investment in an T-year zero. Because the former investment lasts for one more year than the latter, the difference in these growth multiples must be the gross rate of return available in year *n* when the T-year zero can be rolled over into a 1-year investment.
- the short rate is therefore a **break-even interest rate** that makes sure that there is no arbitrage opportunity between investing for T+1 periods and investing for T periods and thus that the return between those two investments is identical

Forward rates as forward contracts

What is the meaning of this future rate <u>under certainty</u>?

In general, we do not know what the future rates will, be as **the world is uncertain.**

Of course, if we contractually agree today about a future interest rate, we take away uncertainty (= forward rate)

To account for this (recognizing that future interest rates are uncertain), we call the future interest rate, in the absence of uncertainty, a **forward rate**

$$(1+F_{T,1}) = \frac{(1+Y_{0,T+1})^{T+1}}{(1+Y_{0,T})^{T}}$$

Example: I sign a contract today to invest in one year time in a ZCB at a particular rate of return.

A forward rate is thus a guaranteed rate, that allows to lock in a forward loan → **Proof:**
Today at time 0 I buy a ZCB with total maturity T+1; at T+1 I get $1. Today I pay the price $-P_{0,T+1}$
In addition to this transaction, I do another transaction: I sell a number of ZCB with maturity T. How many? I sell ($P_{0,T+1}/P_{0,T}$) amount and for each of these bonds I get $P_{0,T+1}$.
If you combine the CFs of these transactions, you have 0 at time 0 and at T+1 you have a CF of 1. What is the CF at time T? That is this ratio $-(P_{0,T+1}/P_{0,T})$.

	0	T	T+1
Buy a ZCB with maturity (T+1)	$-P_{0,T+1}$		1
Sell ($P_{0,T+1}/P_{0,T}$) ZCB with maturity T	($P_{0,T+1}/P_{0,T}$) x $P_{0,T}$	$-(P_{0,T+1}/P_{0,T})$	
Net cash flows	0	$-(P_{0,T+1}/P_{0,T})$	1

This strategy effectively engineers a **synthetic forward loan:** because I set up a transaction today and I know for sure I lend out a particular amount of money at time T and I get 1 for sure at T+1 (= forward loan) → A loan in the future that is determined at a rate today. Now I can easily determine the rate of return that I earn on this forward loan. This should be the forward rate that you earn on such a forward loan.

To determine the corresponding rate of return on this loan:
We rewrite instead of prices, in terms of yields and we know that the forward rate is kind of a break-even yield.

$$\frac{P_{0,T+1}}{P_{0,T}} = \frac{(1+Y_{0,T})^{T}}{(1+Y_{0,T+1})^{T+1}} = \frac{1}{(1+F_{T,1})}$$

The loan granted at T $\left(-\frac{P_{0,T}}{P_{0,T+1}}\right)$, which is repaid at T+1 (1), thus yields a return of $F_{T,1}$!

This forward rate that we derive is a guaranteed rate that you can lock in, in some kind of synthetic forward loan.

Important: Note that these forward rates <u>are not</u> (necessarily) the rates that will prevail in the future (they will much likely deviate because of the economy etc.…)

Even more: Forward rates are not (necessarily) the expected value of the future rates!

$$F_{1,1} \neq E_0(R_{1,1})$$

*What can we say about futures rates ($F_{1,1}$) in the presence of **uncertainty**?*
And can these rates say something about the expectations ($E_0(R_{1,1})$)?

Forward rates are easily calculated but the expectations not. This is because it depends on how the investors behave:

The expectations hypothesis

- Assume that investors only care about the expected value of the interest rate (simple scenario)
  - You are indifferent between investing in the future at a certain rate or an uncertain rate
- Forward rates will then be **equal** to expected rates:

$$F_{T,1} = E\left(R_{T,1}\right)$$

$$\left(1 + E\left(R_{T,1}\right)\right) = \frac{\left(1 + Y_{0,T+1}\right)^{T+1}}{\left(1 + Y_{0,T}\right)^{T}}$$

Then:

$$(1+Y_{0,1}) = (1+F_{1,1})$$

$$(1+Y_{0,1}) = (1+E(R_{1,1}))$$

- This is the key result of the **expectations hypothesis**
  - yields on long-term bonds depend only on expectations of future short rates.
- Then looking at the yield curve says something about future and current interest rates and locked in forward rates
- In this view, an upward sloping curve is evidence of expected increases in interest rates (that investors expect an increase in interest rates)

The liquidity preference theory

*However, **in reality** investors do care for risk*

- Assume that investors not only care about expectations, but also about risk
  - For a long-term investor: a long-term bond that is held until maturity is *risk free*, while rolling over short term bond is *risky* because I might know today's rate of return, but I do not know what next years rate of return of the bond will be
  - For a short-term investor: a short-term bond held until maturity is *risk free*, while selling a long-term bond before maturity is not. That is risky because I do not know today at which rate I will sell the two-year bond in 1 year time. So, when we account for risk, short-term investors will shy away from the long-term bond unless its expected return is higher.

- Given the different risk-profiles of the different investments the forward rate differs from the expected future rate by a **risk premium**:

$$F_{T,1} = E\left(R_{T,1}\right) + risk\ premium$$

$$\left(1 + E\left(R_{T,1}\right) + risk\ premium\right) = \frac{\left(1 + Y_{0,T+1}\right)^{T+1}}{\left(1 + Y_{0,T}\right)^{T}}$$

I am not indifferent; I care for risk. I will only take on risky options if I am rewarded for that. Most investors prefer liquidity so short-term investors dominate the market.

Is this risk premium positive or negative?

⇨ We can derive this:

ST investor prefers ST investments → two options

1) Invest in 1-year ZCB
2) Invest in 2-year ZCB that you sell after 1 year

Which of the two will this investor (who is risk averse) prefer more? The first one because there you know for sure what your rate of return will be. The second one is a risky option. If you prefer the first one this means that you are only willing to go for the second option if you get some extra return. This means that the return you can earn on a one-year ZCB will be smaller than the rate of return on a 2-year ZCB that you sell off after one year which today can only give you the best estimate (your expectation):

$$(1 + R_{0,1}) < \frac{(1 + R_{0,2})^2}{1 + E(R_{1,1})} \quad (1)$$

These guaranteed rates are embedded in the yield curve. From the yield curve we know that:

$$(1 + R_{0,1}) = \frac{(1 + R_{0,2})^2}{1 + F_{1,1}} \quad (2)$$

If we combine these two equations:

$$\frac{(1 + R_{0,2})^2}{1 + F_{1,1}} < \frac{(1 + R_{0,2})^2}{1 + E(R_{1,1})}$$

This can only be valid if:

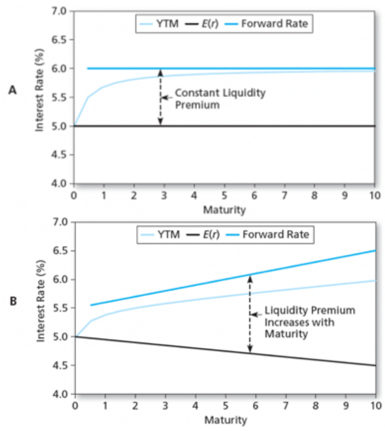
$$F_{1,1} > E(R_{1,1})$$

So, the sign of the risk premium: $F_{1,1} = E(R_{1,1}) + R\pi$ which is > 0

- The sign of the risk premium, for the market as a whole, depends on the proportions of short-term versus long term investors.
- Both short term and long-term investors demand a risk premium
  - for a short-term investor $F_{T,1} > E(R_{T,1})$: the forward rate embodies a positive risk premium as compared to the future short rate. Short-term investors will be unwilling to hold long-term bonds unless the forward rate exceeds the expected short interest rate.
  - for a long-term investor $F_{T,1} < E(R_{T,1})$: the forward rate embodies a negative risk premium as compared to the future short rate. Long-term investors will be unwilling to hold short bonds unless $F_{T,1} < E(R_{T,1})$.
- According to the **liquidity preference theory**, **short term investors dominate the market**: the liquidity premium is positive and thus $F_{T,1} > E(R_{T,1})$. This is why we call this risk premium a liquidity premium because it is the short-term investors that dominate the market.

The term structure thus reflects expected future rates, but the forecasts are clouded by a risk premium

Figure 15.4 Yield curves. **Panel A**: Constant expected short rates. Liquidity premium of 1%. Result: a rising yield curve. **Panel B**: Declining expected short rates. Increasing liquidity premiums. Result: a rising yield curve despite falling expected interest rates.



Whenever you have this liquidity risk premium, extracting information from the yield curve about expected future rates becomes a bit more tricky. Because we observe a yield curve which means we observe spot rates, but we do not know what kind of expectations and risk premia are behind this. These graphs show you that it is not obvious to extract information from the yield curve alone and draw conclusions from the expected future interest rate.
For example, in **panel A** we see an upward sloping yield curve. But we see a constant E(r) and constant risk premium.

In **panel B** you also have an upward sloping yield curve but, in this scenario, you have downward sloping E(r). So, it is not because you have an upward sloping yield curve that the E(r) is also upward sloping.

Interpreting the shape of the yield curve is not as easy in terms of what this says about future interest rates. It says something about forward rates but it is harder to say something about **expected** future rate.

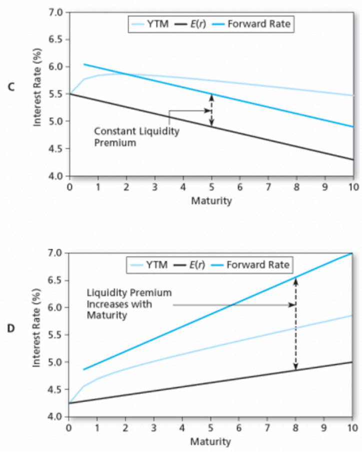


Figure 15.4 (concluded) Panel C: Declining expected short rates. Constant liquidity premiums. Result: a hump-shaped yield curve. **Panel D**: Increasing expected short rates. Increasing liquidity premiums. Result: a sharply rising yield curve.

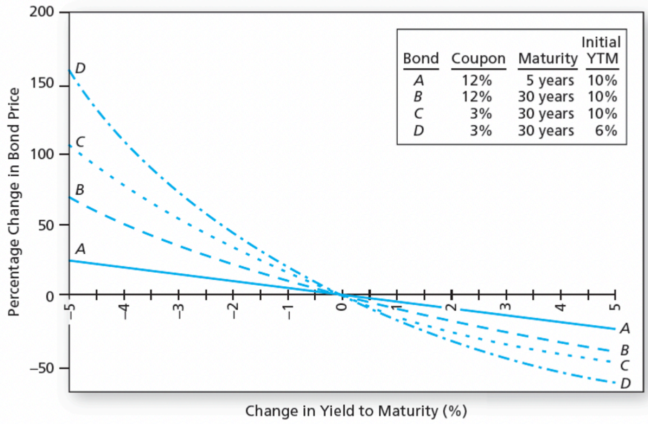
Interest rate risk

Interest rate on ZCB or treasury instruments vary over time. Each day we have a new yield curve. If you look over time it varies allot. This is important because fixed income instruments depend on the general level of market interest rates. When we value a fixed income instrument this is a nonlinear function of interest rates. And these changes will impact the value of outstanding instruments. This will create **interest rate risk**.

- TS graphs have shown that interest rates fluctuate substantially over time
- The basic bond pricing equation shows that bond prices depend on current market rates
  - whenever market interest rates change, the price of an outstanding bond will adjust
  - the price adjustment sets off any arbitrage opportunities between bonds outstanding and newly issued bonds
- Given the **inverse relation** between bond prices and interest rates
  - an increase in interest rates decreases bond prices and results in a capital loss
  - a decrease in interest rates increases bond prices and results in a capital gain

Can we say anything about the size of these bond price changes?
Do all bonds change in the same way, or do we see so called different interest rate sensitivity?

Bond pricing relationships
The starting point is this graph with 4 bonds.
Here we plot the change in YTM on the X-axis = Δy
On the Y-axis we have $\Delta P/P$.
- we see the inverse relationship
- the impact is nonlinear; the curves are curved
- it means that a particular change in yields (decline) will not have the same impact on the percentage change in bond prices as an increase in yields



Bond A vs B:
- B reacts much more than bon A
  - The longer the time to maturity, the more impact on bond prices
  - Long maturity bonds are more sensitive to changes in yields than short term bonds
  - The impact is at a decreasing rate: if you increase the maturity, the bond price sensitivity will increase at **a decreasing rate**. So, changing from 1 to 2 years and from 2 to 3 years, the change from 2 to 3 will have a bigger sensitivity.

Bond B vs C:
- A low-coupon bond (C) has a higher sensitivity to interest rates than bond B. This is because the reaction for a given change in yield is bigger for bond B as compared to bond C. So, high coupons reduce interest rate sensitivity.

Bond C vs D:
- The higher the discount rate, the lower the present value will be (c)
- The more distanced cashflows are more impacted than the more near ones
- This means that bond D, with a lower YTM, is more sensitive to changes in interest rates as compared to bond C

Summary
1. Bond prices and yields are inversely related: as yields increase, bond prices fall
2. An increase in a bond's YTM results in a smaller price change than a decrease of equal magnitude
3. Long term bonds tend to be more price sensitive as compared to short term bonds
4. As the maturity increases, the price sensitivity increases at a decreasing rate
5. Interest rate risk and coupon rates are inversely related: as coupons increase, interest rate sensitivity decreases
6. Price sensitivity and YTM are inversely related: a bond selling at a higher YTM, is less sensitive to changes in yields

Can we formalize this?
- Maturity seems to be a **major determinant** of bond price sensitivity → we will focus on this characteristic
- To see how the maturity impacts bond price sensitivity to interest changes, we start with a numerical example of **ZCBs**

Yield to Maturity (APR)	T = 1 Year	T = 10 Years	T = 20 Years
8%	924.56	456.39	208.29
9%	915.73	414.64	171.93
Fall in price (%)*	0.96%	9.15%	17.46%

*Equals value of bond at a 9% yield to maturity divided by value of bond at (the original) 8% yield, minus 1.

- We see here that a change in 1% in YTM for T=1 changes the bond price by 1%
- For T = 10; a 1% increase in YTM% results in a fall in price by 9.15% (almost 10%)
- For T = 20; a 1% increase in YTM% results in a fall in price by 17.46% (almost 20%)

- General observation: For a ZCB the bond price changes almost 1 to 1 with maturity: a 1% change in YTM changes the bond price by T%

We will formalize this:

Bond price sensitivity to interest rates

- To formalize this relation between maturity and bond price sensitivity we take the first derivative of the ZCB price wrt the YTM (for ease of notation we leave out time and maturity subscripts):

$$P = \frac{1}{(1+Y)^T} \rightarrow \frac{dP}{dY} = -T(1+Y)^{-T-1}$$

$$\frac{dP}{P} = -T\left(\frac{dY}{1+Y}\right) \qquad (1)$$

The equation (1) corresponds to the graph we saw in the beginning

- For a ZCB, the sensitivity of price to a small change in the YTM is proportional to maturity

Important: by taking the first derivative we are focusing on the linear effect of a change in interest rates!

Can we generalize this sensitivity analysis to **coupon bonds**?

Yield to Maturity (APR)	T = 1 Year	T = 10 Years	T = 20 Years
8%	1,000.00	1,000.00	1,000.00
9%	990.64	934.96	907.99
Fall in price (%)*	0.94%	6.50%	9.20%

*Equals value of bond at a 9% yield to maturity divided by value of bond at (the original) 8% yield, minus 1.

The coupon is 8% because we know that when a bond is priced at par that the coupons equals YTM.

- For a 1-year maturity it is also close to 1%
- For a 10-year maturity, the price falls by 6.50% which is smaller than what we observe for ZCB
- For a 20-year maturity, the price falls by 9.20% which is smaller than what we observe for ZCB
  → just the introduction of coupons reduces the maturity

For a CB the sensitivity of price to a 1% change in the YTM is smaller than its maturity

Do coupons reduce maturity?

The question than is: what is a maturity in the context of a coupon bond?

- Characteristic to a coupon bond is that cash flows arrive at many different times: each of the cash flows has its own 'maturity'
- To deal with the ambiguity of a bond making many payments, we need to measure the **effective maturity**
- Such effective maturity should be an <u>average</u> of the maturity of the cash flows of a coupon bond; this takes into account that CFs are received in different moments in time

Macaulay's duration (D) is a measure of effective maturity of a bond. It is the weighted average of the times to payment, where the weights are the present value of each payment divided by the price of the bond

Macaulay's duration
Formally:

- Consider a coupon bond with a maturity of T
- Coupon payments are made in periods $t = \{1, ..., T\}$
- We will weight them as the relative importance of the CF received at t, this means relative vis a vis the total CFs received
- The weight w_t associated with the payment C_t made at t equals:

$$w_t = \frac{\overbrace{C_t / (1+Y)^t}^{present\ value\ of\ C_t}}{P}$$

The duration of the bond is then:

$$D = \sum_{t=1}^{T} t \times w_t \qquad (2)$$

→ this means that the duration of **a coupon bond** will always be **smaller** than the **ZCB**. Because you indeed receive one payment at the end (T), but the weight of that payment will not be 1 because you also receive intermediate payments.

Consider a 4% coupon bond with a face value of 1000 and a maturity of 9 years (coupons paid on an annual basis). The YTM on this bond equals 5%. What is its duration?
vvConsider also a zero-coupon bond with a face value of 1000 and a maturity of 9 years. The YTM on this bond equals 5%. What is its duration?

Assume now that the YTM declines to 4%. What happens to the price of these bonds? What happens to its durations?

You see also the last cashflow has the biggest weight.
You also see what happens when the YTM declines to 4%. The first CF has a higher PV than the other, and all of these CF will be higher as compared to the 5% YTM. You also see that the impact on the nearby CFs is minor as compared to the impact on the more distant CFs. This means that the weights change (relative importance). More weight is given to the more distant CFs. This impact our duration in the sense that our duration increases.

Face value		1000				
Coupon		4%				
Maturity		9 years				
	YTM = 5%				YTM = 4%	
t	CF	PV(CF)	Weight	CF	PV(CF)	Weight
1	40	38.095	0.041	40	38.462	0.038
2	40	36.281	0.039	40	36.982	0.037
3	40	34.554	0.037	40	35.560	0.036
4	40	32.908	0.035	40	34.192	0.034
5	40	31.341	0.034	40	32.877	0.033
6	40	29.849	0.032	40	31.613	0.032
7	40	28.427	0.031	40	30.397	0.030
8	40	27.074	0.029	40	29.228	0.029
9	1040	670.393	0.722	1040	730.690	0.731
Price		928.92			1000	
ΔP/P (exact)		7.65%				
Duration		7.676			7.733	

You see for a ZCB that the duration equals the maturity of the bond. This is because you don't receive any CF.

Face value		1000				
Coupon		0%				
Maturity		9 years				
	YTM = 5%			YTM = 4%		
t	CF	PV(CF)	Weight	CF	PV(CF)	Weight
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	1000	644.609	1	1000	702.587	1
Price		644.609			702.587	
ΔP/P (exact)		8.99%				
Duration		9			9	

Properties of Macaulay's duration

- The duration of a ZCB equals it's time to maturity
- Holding maturity constant, a bond's duration is lower when the coupon rate is higher
    - when coupon rate is higher relatively larger payments are received **early**
    - this means that the weights of this early moments will be slightly higher and so the bonds duration will be lower.
    - In other words, a higher fraction of the total value of the bond is tied up in the (earlier) coupon payments, whose values are relatively insensitive to yields, rather than the (later and more yield-sensitive) repayment of par value.
- Holding the coupon rate constant, a bond's duration generally increases with its time to maturity. Duration always increases with maturity for bonds selling **at par/at a premium**.
    - since duration is a weighted average of the maturity of all payments, longer term bonds will most often have longer durations
- Holding other factors constant, the duration of a coupon bond is higher when the bond's yield to maturity is lower
    - a lower YTM increases the present value of all payments, but more so of the more distant ones.
    - at lower yields a larger fraction of the bond's payments is thus later
- The duration of a perpetuity can be explicitly solved as $(1 + Y)/Y$
    - makes it obvious that maturity and duration can differ substantially. The maturity of the perpetuity is infinite, whereas the duration of the instrument at a 10% yield is only 11 years. The present-value-weighted cash flows early on in the life of the perpetuity dominate the computation of duration.

Macaulay's duration as interest rate sensitivity

Remember the price-sensitivity equation for a ZCB, this equation now generalizes to coupon bonds, once we replace maturity with Macaulay's duration:

$$\frac{dP}{P} = -D \times \left(\frac{dY}{1+Y}\right) \qquad (3)$$

Duration (D) thus measures the sensitivity of the bond price to changes in the YTM. It is important as it captures the bond's risk associated with changes in interest rates and is therefore a measure of **interest rate risk**.
Remark: this assumes a **parallel shift** in the term structure!

Proof:
Consider a C% coupon bond with a maturity of T years, a yield Y and a par value of M:

$$P = \sum_{t=1}^{T} \frac{C}{(1+Y)^t} + \frac{M}{(1+Y)^T}$$

To determine the approximate change in the bond price for a small change in the YTM, we compute the first derivative wrt YTM:

$$\frac{dP}{dY} = \frac{(-1)\,C}{(1+Y)^2} + \frac{(-2)\,C}{(1+Y)^3} + \ldots + \frac{(-T)\,(C+M)}{(1+Y)^{T+1}}$$

$$= -\left(\frac{1}{1+Y}\right)\left(\left(\sum_{t=1}^{T}\frac{t \times C}{(1+Y)^t}\right) + \frac{TM}{(1+Y)^T}\right)$$

Dividing both sides of the above equation by the bond's price gives:

$$\frac{dP}{dY}\frac{1}{P} = -\left(\frac{1}{1+Y}\right)\left(\left(\sum_{t=1}^{T}\frac{t \times C}{(1+Y)^t}\right) + \frac{TM}{(1+Y)^T}\right)\frac{1}{P}$$

$$= -\left(\frac{1}{1+Y}\right)\underbrace{\left(\sum_{t=1}^{T} t \times w_t\right)}_{\text{Macaulay duration}}$$

Rearranging gives us:

$$\frac{dP}{P} = -D\left(\frac{dY}{1+Y}\right)$$

which completes our proof.

We impose **two important assumptions**:

- The yield curve is flat, but we know in reality that the yield curve is not flat. (It is typically upward sloping.) This result is only correct when the yield curve is flat otherwise, we need to discount for every CFs with its corresponding interest rate.
- We also assume that when interest rate changes that the changes are parallel.

A number of **modifications** exists that incorporates this idea that the yield curve is not flat and that changes are not always parallel. But the complexity that you introduce does not outweigh using a simple formula like this. So, this formula will not be exact, but it is so simple to use which is better.

Modified duration as interest rate sensitivity

Modified duration is often seen as the true measure of interest rate sensitivity. This is because the relationship between bond prices and yields is much more nether (hechter) when we introduce modified duration.
Practitioners often find it useful to work with modified duration:

$$D^* = \frac{D}{1+Y}$$

In this case, equation (3) is further simplified to:

$$\frac{dP}{P} = -D^* dY$$

Modified duration (D*) is a measure of the proportional change in the bond price caused by a given small change in the YTM.

Modified duration
If $D*$ = 10, then a 1% increase in the yield (e.g. from 3% to 4%) causes the bond price to drop by 10%

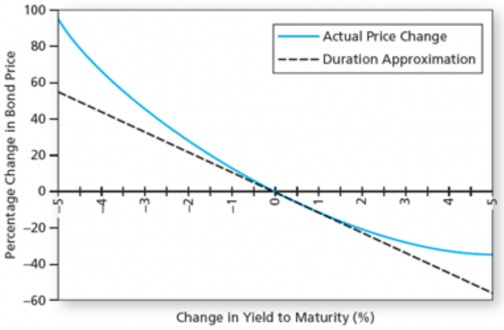
Calculating the impact of an interest rate change

Macaulay duration
Retake the 4% coupon bond with a face value of 1000 and a maturity of 9 years (coupons paid on an annual basis). The initial YTM on this bond equals 5%. What is the impact on the price of the bond when the YTM decreases to 4%?

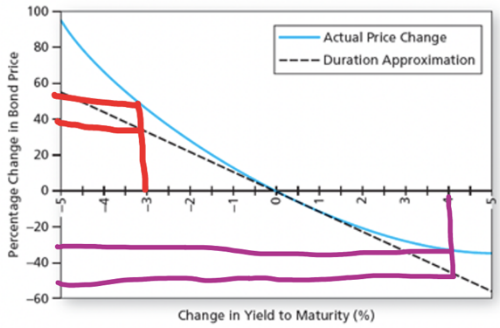
Face value	1000	
Coupon	4%	
Maturity	9 years	
	YTM = 5%	YTM = 4%
Price	928.92	1000
ΔP/P (exact)	7.65%	
Duration	7.676	7.733
ΔP/P (D)	7.31%	

Graphical representation of modified duration

Suppose we plot this:



- You see the actual price change on the blue **curved line**
- You see the duration on the dotted **straight line**:
- You take the first derivative on the initial yield; this means that you plot a tangent line at the initial yield at point 0. And you read off the effect on bond prices on your tangent line as compared to the curved actual price line. Then you see indeed that you make a small mistake by assuming a linear relationship.
    - The straight line is the % price change predicted by the duration rule
    - The slope of the straight line is the modified duration of the bond at its initial YTM



- Suppose (red example) you have a large change in YTM, then you know that the price will increase till almost 50; when you use duration the price change will be much lower. So, you make a mistake, and the size of the mistake is significant.
    - As long as you stay in the area (-1;1) where the changes in YTM are minor, then the mistakes you make are also smaller. This is because both lines are very close to one another.
    - So, when introducing minor changes it does not matter allot on which line you read of the changes but when introducing large changes it does matter.
- Another interesting point: looking at **under-or overestimation effects** (purple example). Suppose interest rates decline and you use duration instead of the actual price change. Are you conservative or not? (Are you over or underestimate effects?)
    - When interest rates decline than using duration leads you to **underestimate** capital gains because you think that prices will increase to a certain point but if you look at the actual price change it increases to a much higher point.
    - When interest rates increase than using duration means that you will **overestimate** your losses as compared to the actual losses you occur.
    - So, the use of duration leads you to be conservative.

Is duration an accurate proxy of interest rate risk?

Convexity

Modified duration is just a linear approximation to the actual curved relation between prices and yield. This means if we also incorporate n^th^-order effects, then you approximate better the curved relationship. In finance we limit till second-order effects, because after that the effect is small. Second-order effects are also called convexity. You want to include second-order effects when duration is not as accurate. This is whenever the relationship you are analyzing is a **curved one** and when you want to allow for this noncurved function whenever you look at **large changes in interest rates**:

$$dP = \frac{\partial P}{\partial Y} dY + \frac{1}{2}\frac{\partial^2 P}{\partial Y^2}(dY)^2 + h.o.t.$$

$$\frac{dP}{P} = -D^* \times dY + \frac{1}{2}\frac{1}{P}\frac{\partial^2 P}{\partial^2}(dY)^2 + h.o.t.$$

- Only for small changes in yields the proxy is accurate:
  - as yields decrease, modified duration rises so the curve gets steeper on the left
  - as yields increase, modified duration falls so the curve gets flatter on the right
- This property of bonds is known as **convexity**
  - the bond's price after a change in yields is always higher than predicted by duration
  - accounting for convexity thus increases the accuracy of interest rate risk measurement

- Formally, we can quantify convexity as the second derivative of the price yield curve, expressed as a fraction of the bond price
- It is thus the rate of change of the slope of the price-yield curve, expressed as a fraction of the bond price:

$$Convexity = \frac{1}{P}\frac{\partial^2 P}{\partial Y^2}$$

- Consider a T maturity coupon bond, convexity can then be calculated as:

$$Convexity = \frac{1}{P}\left(\left(\sum_{t=1}^{T}\frac{t\times(t+1)\times C}{(1+Y)^{t+2}}\right) + \frac{T\times(T+1)\times M}{(1+Y)^{T+2}}\right)$$

  ⇨ t has an important impact on convexity

Improved measure of interest rate sensitivity

Combining the duration (first order effect) with the convexity (second order effect) allows us to obtain a more accurate measure of interest rate sensitivity:

$$\frac{dP}{P} = -D^* \times dY + \frac{1}{2}\times Convexity \times (dY)^2$$

  ⇨ if ΔY is very small than $(dY)^2$ will be close to zero and then this whole second-order effect will drop out.

Duration-convexity effects

Retake the 4% coupon bond with a face value of 1000 and a maturity of 9 years (coupons paid on an annual basis). The initial YTM on this bond equals 5%. What is the impact on the price of the bond when the YTM decreases to 4%, accounting for convexity?

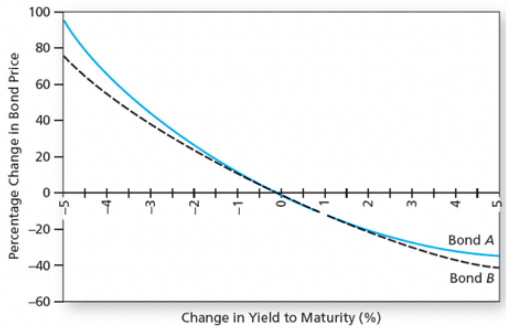
You see that extending to convexity really matters. If I want to calculate the exact impact, I need the part of the table with 4% but if I use duration formula do not need to recalculate the price of my bond with the new YTM, and the second table drops out. So in that case using duration might pay off.

Face value	1000						
Coupon	4%						
Maturity	9 years						
		YTM = 5%				YTM = 4%	
t	CF	PV(CF)	Weight	Convexity	CF	PV(CF)	Weight
1	40	38.095	0.041	69.11	40	38.462	0.038
2	40	36.281	0.039	197.45	40	36.982	0.037
3	40	34.554	0.037	376.09	40	35.560	0.036
4	40	32.908	0.035	596.97	40	34.192	0.034
5	40	31.341	0.034	852.82	40	32.877	0.033
6	40	29.849	0.032	1137.09	40	31.613	0.032
7	40	28.427	0.031	1443.92	40	30.397	0.030
8	40	27.074	0.029	1768.07	40	29.228	0.029
9	1040	670.393	0.722	54725.98	1040	730.690	0.731
Price			928.92			1000	
ΔP/P (exact)			7.65%				
Duration			7.676				
Convexity			65.848				
ΔP/P (D)			7.31%				
ΔP/P (D+C)			7.64%				

Attractiveness of convexity

If you have two bonds with different convexity, which of the two will you prefer the most? Do you like or dislike convexity. There is an asymmetric effect of convexity, this is the curvature. If bond prices/yields decline, duration pushes you to underestimate effects. If yields increase, you overestimate losses when you use duration. This means that you like convexity, because a similar increase in interest rates will lead to a lower decrease in prices whereas a similar decrease in interest rates will lead to a higher increase in prices when it is more convex.

This is also what you see on this curve; if you go to -3 the capital gain you realize for B is slightly limited compared to bond A. On the other side if interest rates increase to 3 than it is bond B that incurs more capital losses and A will realize only a smaller capital loss.



What does it mean if you are not the only one that like convexity? Then convexity will be priced. More convex bonds will be priced higher.

So:

- Convexity is generally considered as attractive due to its asymmetric effect on gains and losses:
  - a % decrease in yields has a greater effect (in absolute terms) on bond prices, than an equal % increase in yields.
  - for a % change in yields we gain more than we loose
- Bonds with higher curvature will therefore be more expensive than bonds with less curvature

Effective duration and effective convexity

- Practitioners will often use the concept of **effective duration** (**effective convexity**) instead of modified duration (and convexity)
- Such effective measures allow to incorporate changes in cash flows due to changes in interest rates
  - this is clearly of interest when intermediate cash flows depend on the general level of interest rates or when options are embedded
    - for example: callable bonds, they depend on the level of interest rate. If rates fall, the bond may be called back (p 508 HB, negative convexity)
  - *In the presence of such options, the future cash flows provided by the bonds are no longer known. If the bond may be called, for example, its cash flow stream may be terminated, and its principal repaid earlier than was initially anticipated.*

Because cash flows are random (they depend on the realization of interest rates), we can hardly take a weighted average of times until each future cash flow, as would be necessary to compute Macaulay's duration.

- The standard formula of duration and convexity assumes coupons to be constant, whenever they are interest rate dependent than these formulas do not longer apply and you should go to effective duration instead of modified duration.

- **Effective duration** is measured as:

$$\frac{P_{-} - P_{+}}{2PdR}$$

  - with $P-$ the price of the bond when interest rates decrease by dR
  - with $P+$ the price of the bond when interest rates increase by dR

- **Effective convexity** is measured as:

$$\frac{P_{+} + P_{-} - 2P}{2P\,(dR)^2}$$

Passive bond management

Passive bond management strategies take price largely as given and focus on **controlling for the risk** in the bond portfolio. This is not the same as the idea of a passive stock strategy. Under passive bond management strategies there are a whole range of strategies that control for risk and there are two ways.

Two main passive bond management strategies can be distinguished:

1. An **indexing strategy** aims at replicating the performance of a given bond index: it targets a risk profile in line with the risk profile of the tracked bond index. This is low effort as you mimic an index.
2. An **immunization strategy** aims at shielding the portfolio from exposure to interest rate fluctuations: it targets a very low, or even zero-risk profile. This is a passive strategy because you want to rule out interest rate risk. This is high effort, so passive may not be understood as low effort.

Bond index funds

- Bond index funds track the composition of a broad market bond index
- You cannot just extend a stock index strategy to the bond market. It is a similar strategy as used in equity index funds, but with specific problems
  - bond indices often include thousands of issues, many of which are **infrequently** traded and then it will be hard for you to buy and sell them in the market and you will only be able to sell them at a discount
  - bond indices turn over a lot due to bonds maturing; this requires a lot of rebalancing
  - bonds generate considerable intermediate income which requires reinvestment; this complicates the job of the index fund manager
- because of these problems it is not easy to replicate this strategy (so most traders mimic an index)
- These problems require modifications to traditional tracking
  - a bond index fund will **restrict its portfolio to a smaller set of representative bonds**, matching the characteristics (e.g. maturity, coupon rates, credit risk) of the bond index
  - this is achieved via **stratified sampling** or a **cellular approach**

Example of bond stratification

- we can stratify here along maturity and along the sector
- this means that in this first sell you will put all bond issues e.g., of treasury in the column of strategy with a term to maturity of less than 1 year
    - than you know that 12.1% of your strategy should be a treasury of <1 year

Sector / Term to Maturity	Treasury	Agency	Mortgage-backed	Industrial	Finance	Utility	Yankee
<1 year	12.1%						
1–3 years	5.4%						
3–5 years			4.1%				
5–7 years							
7–10 years		0.1%					
10–15 years							
15–30 years			9.2%			3.4%	
30+ years							

- By stratifying the broad universe of bonds into a limited set of key characteristics, bonds having identical key characteristics are considered homogeneous/identical
- The bond manager then composes a portfolio according to the portfolio weights in the different representative cells, by investing in an instrument that matches the cell's characteristics

Immunization

If you have a bond you are exposed to credit risk and interest rate risk. You can hedge against credit risk through credit derivatives (see earlier).

- Immunization strategies **insulate the portfolio from interest rate risk**
- It is a typical strategy used by banks and LT investors (e.g. pension funds) to protect their market value from interest rate volatility
- The basic principle of immunization is the matching of interest rate risk of assets and liabilities
    - by matching the duration of the assets and liabilities in a portfolio, price risk and reinvestment risk exactly cancel out
    - the value of the assets will then track the value of liabilities when interest rates change

Need for immunization
Assume an insurance company markets a guaranteed investment contract for $10,000, with a 5 year to maturity and guaranteed interest rate of 8%. The insurance company funds this liability with a $10,000 investment in 8% coupon bonds with a maturity of 7 years that sells at par. What is the final net worth of the insurance company when the interest rate stays unchanged at 8%. What happens when interest rates decrease to 7%, or increase to 9%?

LIABILITY (at 8%)	
current value	10000
coupon rate	0
time to maturity	5
duration	5
Total accumulated value at T	$=10000*(1+0.08)^5 = 14693.28$

ASSETS (at 8%)			
current value		10000	
coupon rate		0.08	
time to maturity		7	
duration		5.62	
Payment number	Years remaining until obligation	Payment at t	Accumulated value of invested payment after 5 years
1	4	800	$800 * 1.08^4 = 1088.39$
2	3	800	$800 * 1.08^3 = 1007.77$
3	2	800	$800 * 1.08^2 = 933.12$
4	1	800	$800 * 1.08^1 = 864$
5	0	800	$800 * 1.08^0 = 800$
sale of bond	0	$\frac{800}{1.08}+\frac{10800}{1.08^2}$	$10000 * 1.08^0 = 10000$
Total accumulated value at T			$= 14693.28$

ASSETS (at 7%)			
face value		10000	
coupon rate		0.08	
time to maturity		7	
Payment number	Years remaining until obligation	Payment at t	Accumulated value of invested payment after 5 years
1	4	800	$800 * 1.07^4 = 1048.64$
2	3	800	$800 * 1.07^3 = 980.03$
3	2	800	$800 * 1.07^2 = 915.92$
4	1	800	$800 * 1.07^1 = 856$
5	0	800	$800 * 1.07^0 = 800$
sale of bond	0	$\frac{800}{1.07}+\frac{10800}{1.07^2}$	$10180.8 * 1.07^0 = 10180.8$
Total accumulated value after 5 years			$= 14781.39$

ASSETS (at 9%)			
face value		10000	
coupon rate		0.08	
time to maturity		7	
Payment number	Years remaining until obligation	Payment at t	Accumulated value of invested payment after 5 years
1	4	800	$800 * 1.09^4 = 1129.27$
2	3	800	$800 * 1.09^3 = 1036.02$
3	2	800	$800 * 1.09^2 = 950.48$
4	1	800	$800 * 1.09^1 = 872.00$
5	0	800	$800 * 1.09^0 = 800.00$
sale of bond	0	$\frac{800}{1.09}+\frac{10800}{1.09^2}$	$9824.09 * 1.09^0 = 9824.09$
Total accumulated value after 5 years			$= 14611.86$

As long as interest rates remain at 8%, the coupon bond will allow you to fulfill your obligation. **(Left panel)**

(**Right panel**) Here the interest rates do not stay the same.

- Interest rate decrease: The reinvestment value of your coupons will be lower as compared to the other scenario (*reinvestment rate risk)*. But at the end you will be able to sell the bond above par (*price risk)*
- Interest rate increase: The reinvestment value of your coupons will be higher as compared to the other scenario. But at the end you will be able to sell the bond below par. Here in total, you will have slightly less money as compared to what you need. So, we have an interest rate risk.

To summarize: when interest rates change the fund no longer grows to the targeted value

- when interest rates decrease the coupon bond gains value, but reinvestment income decreases.
- when interest rates increase the coupon bond loses value, but reinvestment income increases.
    - ⇨ fixed-income investors face two offsetting types of interest rate risk: *price risk* and *reinvestment rate risk*.
    - ⇨ can we find a portfolio that allows us to immunize against interest rate changes?

Only when duration is appropriately chosen, these two effects offset one another:
In particular: For an equal duration of assets and liabilities, price risk and investment risk compensate

Instead of funding the liability with a $10,000 investment in 8% coupon bonds with a maturity of 7 years that sells at par, the insurance company funds itself with a $10,000 investment in 8% coupon bonds with a maturity of 6 years that sells at par. What is the final net worth of the insurance company when the interest rate stays unchanged at 8%. What happens when interest rates decrease to 7%, or increase to 9%?
Why 6 years? If you invest in this bond, you effectively invest in a bond with a duration equal to 5.

ASSETS (at 8%)			
face value		10000	
coupon rate		0.08	
time to maturity		6	
duration		5	
Payment number	Years remaining until obligation	Payment at t	Accumulated value of invested payment after 5 years
1	4	800	$800 * 1.08^4 = 1088.391168$
2	3	800	$800 * 1.08^3 = 1007.7696$
3	2	800	$800 * 1.08^2 = 933.12$
4	1	800	$800 * 1.08^1 = 864$
5	0	800	$800 * 1.08^0 = 800$
sale of bond	0	$\frac{10800}{1.08}$	$10000 * 1.08^0 = 10000$
Total accumulated value at T			= 14693.28077

ASSETS (at 7%)			
face value		10000	
coupon rate		0.08	
time to maturity		6	
Payment number	Years remaining until obligation	Payment at t	Accumulated value of invested payment after 5 years
1	4	800	$800 * 1.07^4 = 1048.636808$
2	3	800	$800 * 1.07^4 = 980.0344$
3	2	800	$800 * 1.07^4 = 915.92$
4	1	800	$800 * 1.07^4 = 856$
5	0	800	$800 * 1.07^4 = 800$
sale of bond	0	$\frac{10800}{1.07}$	$10000 * 1.07^4 = 10093.45794$
Total accumulated value after 5 years			= 14694.04915

ASSETS (at 9%)			
face value		10000	
coupon rate		0.08	
time to maturity		6	
Payment number	Years remaining until obligation	Payment at t	Accumulated value of invested payment after 5 years
1	4	800	$800 * 1.09^4 = 1129.265288$
2	3	800	$800 * 1.09^4 = 1036.0232$
3	2	800	$800 * 1.09^4 = 950.48$
4	1	800	$800 * 1.09^4 = 872$
5	0	800	$800 * 1.09^4 = 800$
sale of bond	0	$\frac{10800}{1.09}$	$10000 * 1.09^0 = 9908.256881$
Total accumulated value after 5 years			= 14696.02537

(**Left panel**) When interest rates remain the same, the portfolio grows to a value equal to a liability value of our portfolio. But the problem is when interest rates changes:

(**Right panel**) When we have a decline of interest rate to 7%, we earn less coupons. On the other hand, we have a positive effect because at the end we can sell the bond above par. The total value is almost identical as the total value of our liabilities. So, the investment effect and the price effect neutralize one another. A similar effect is observed when interest rate increases to 9%. The coupons were higher, and these were not high enough to fulfill our obligations, but in this case it is enough. We earned more on our coupons, and it is just enough to make sure that the loss we make at the end by selling below par, is offset such that the total accumulated value is sufficient and equals more or less the value of our liabilities.

The graph here under presents the present values of the bond and the single-payment obligation as a function of the interest rate. At the current rate of 8%, the values are equal, and the obligation is fully funded by the bond. As interest rates change, the change in value of both the asset and the obligation is equal, so the obligation remains fully funded. For *greater changes* in the interest rate, however, the present value curves *diverge*. This reflects the fact that the fund actually shows a small *surplus* at market interest rates other than 8%.

*So, this matching is not perfect, and this means that immunization, which is based upon duration is just a proxy. We still have **surpluses** at the beginning and at the end.*

Should we be worried with the small deviations in accumulated values?

This is driven by the fact that the liability is a ZCB, and this is a CB. These two are different instruments so the **convexity** is different. We have been focused on duration and thus only first-order price effects.

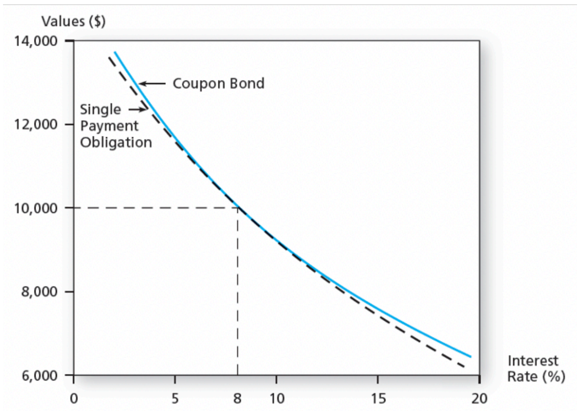


Figure 16.10 Immunization. The coupon bond fully funds the obligation at an interest of 8%. Moreover, the present value curves are tangent at 8%, so the obligation will remain fully funded even if rates change by a small amount.

Even if the obligation was immunized, there are surpluses in the fund because of the **convexity**. The graph shows that the coupon bond has greater convexity than the obligation it funds. Hence, when rates move substantially, the bond value exceeds the present value of the obligation by a noticeable amount.

Can the fund manager now rest until maturity? (Once he/she has immunized the portfolio)

No (this goes counter the idea that immunization is a passive bond strategy). Whenever time passes by (even when interest rates stay the same) you will have to **rebalance** your portfolio because it will no longer be immune. In that sense the effort that immunization requires is high and passive does not mean low effort!

- An immunization strategy requires rebalancing for two reasons
    - the portfolio is immune at a given moment in time, for small changes in the interest rate
    - even if interest rates do not change, portfolio duration changes due to passage over time
- So even though immunization is a passive strategy, it does require close monitoring

Need to rebalance
An insurance company has an obligation of $19,487 in 7 years. The market interest rate is 10%. They fund this using a 3-year ZCB and a perpetuity paying annual coupons. How can the obligation be immunized? Assume now that after 1 year interest rates are still at 10%. Is the obligation still immunized? If not, what actions are required?

Step 3: you want a duration of 7 so 7 = $w_1D_1+(1-w_1)D_2$ and you know the durations so solving this you come with a weight of 0.5.
After 1 year the duration of the portfolio does not equal 6. How can you make sure that it equals 6? 6 = $w_1D_1+(1-w_1)D_2$ but then the value of ZCB and of the perpetuity changes which means that you need to rebalance your portfolio. How? The coupon of 500 can be reinvested then I get 6000. Still not enough I need 111 so I sell part of my perpetuity to get 111.

This allows me to immunize again my portfolio. **Here you see that just as time passes by you need to rebalance your portfolio.**

TODAY		AFTER 1 YEAR	
time to maturity obligation	7	time to maturity obligation	6
time to maturity ZCB	3	time to maturity ZCB	2
time to maturity perpetuity (C=0.1)	∞	time to maturity perpetuity (C = 0.1)	∞
interest rate	0.1	interest rate	0.1

1. Calculate the duration of the obligation		1. Calculate the duration of the obligation	
value after 7 years	19487	value at maturity	19487
value today	10000	value after 1 year	11000
duration obligation	7	duration obligation	6

2. Calculate the duration of the asset portfolio		2. Calculate the duration of the asset portfolio	
duration ZCB	3	duration ZCB	2
duration perpetuity	11	duration perpetuity	11

3. Find the corresponding asset mix		3. Find the corresponding asset mix	
asset duration	7	asset duration	6.00
weight ZCB	0.50	weight ZCB	0.56
weight perpetuity	0.50	weight perpetuity	0.44

3. Fully fund the obligation		3. Fully fund the obligation	
value ZCB today	5000	value ZCB after 1 year	6111
value ZCB after 3 year	6655	value perpetuity after 1 year	4889
value perpetuity today	5000		

	value		coupon		extra investment
ZCB	5500	+	500	+	111
perpetuity	5000			+	-111

Cash flow matching

- Clearly, in general, immunization strategies are quite time-expensive due to the continuous rebalancing that is needed to keep the portfolio immunized over time
- This problem can be solved by focusing on a specific immunization strategy called **cash-flow matching**
  - by matching the cash flows the portfolio is automatically immunized from interest rate risk
  - you match for example by buying a ZCB with face value equal to the projected cash *outlay (the obligation)*
  - then there is automatically immunization of the interest rate risk because the CF from the bond and the obligation exactly offset each other
- When applied on a multi period basis such cash flow matching is called **dedication strategy;** match all future cashflows of all moments in time = '*sit back and relax*' scenario however the effort to construct such a dedication strategy is tremendous and the cost might also be expensive
- Such dedication strategies have the advantage that it is a 'once-and-for-all approach' (this means that once the CFs are matched there is no need for rebalancing), but this comes at a cost of highly reduced bond choice:
  - if you invest in very particular bonds because you need to match perfectly your liability you will look for very specific instruments and the chance that these will be expensive is realistic.
  - Or you will even not be able to find instruments to match your CFs.
- This is why in practice we don't often see these dedication strategies even though they are *'sit back and relax'* strategies

Active bond management

- Just like active equity portfolio management, bond portfolios can also be managed actively: looking for over- or underprices assets in the market
- The success of such strategy (again) crucially depends upon how valuable the information is, on which one trades

Different active bond management strategies can distinguished

1. **Substitution swap**: exchange a bond for a nearly identical substitute that is under priced
2. **Intermarket swap**: exchange bonds of one market for bonds of another that is under priced
3. **Rate anticipation swap**: exchange varying maturity bonds when interest rate changes are anticipated. If you believe interest rates are going to increase (bond prices go down), then you want to minimize the cost and you will be short and thus shorten the duration of your bond portfolio.

    ⇨ **These three are called alpha strategies**

4. **Pure yield pickup swap**: exchange low-yield bonds for high-yield bonds (together with the higher risk!)

Chapter 10: Applied portfolio management

Readings
BKM – Chapter 24: Portfolio performance evaluation

Up till now we have typically been talking about ex-ante decisions. You would like to evaluate to what extend your forecasts were accurate and if your portfolio delivered what you thought it will deliver. Not only in terms of returns but also in terms of risk profile.

Performance evaluation: traditional approach

After investing, we need to evaluate the investment decisions that were taken: is it acceptable? Is it in line with expectations?

This poses two challenges:
1. How to measure (average) returns when investing for longer periods?
2. How to account for the riskiness of the investment?

Average returns

Multiple return concepts have been introduced...
- To obtain an unbiased forecast of expected returns = **arithmetic average** of HPRs
- To calculate actual performance = **geometric average** of HPRs (=EAR if expressed on an annual basis)

Which one is most appropriate when you want to summarize the actual returns that you earned? You should have a preference for the geometric average as it is indeed the measure to calculate actual returns earned. An arithmetic can be used ex ante when for example you have a time series of returns, and you want to come up based on that with a forecast.

Both concepts can be labelled **time-weighted average returns** as they are based on year-by-year rates of return where each return has equal weight. Since this is based on the yearly rates of returns, this means that each year gets an equal weight in that average. You do not account for the amount that was invested at the moment that a particular return was earned.

There is no impact of 'amounts invested' and thus cash in- and outflows.

Suppose I earn past period a return of -9% per year, but I had very few money invested in the market VS a situation where I earned a return of -9% per year but I had allot of money invested. In terms of actual performance that might matter. You might want to account for the amount you had invested. Then you might go to a dollar-weighted rate of return.

To account for cash in- and outflows we need to calculate the **dollar-weighted rate of return** = internal rate of return (IRR) calculated using the DCF approach. (The way we value bonds, this accounts for the amount of money that was invested, is the same)

Such return is dollar-weighted since amounts invested, and thus cash in- and outflows at each period, impact the average rate of return.

Can you rank these three concepts? NO
→ the geometric average will be smaller or equal to the arithmetic average; the larger the variability, the lower the geometric average will be.
→ the dollar-weighted rate of return cannot be ranked relative to the geometric average or arithmetic average; but you can make a prediction whether it will be larger or smaller → it will be smaller when lower returns are earned when more money is invested, and will be larger when higher returns are earned when more money is invested

Example: average returns
You bought 5 shares of Apple in January 2019 at $67 each. In January 2020 you bought 2 additional shares at $77.4. You also received your 2019 dividend of $3 per share. In January 2021 you sold all of your stocks at $132 each, after having received a dividend of $1.8 per share.
Calculate the arithmetic average, the geometric average and the dollar-weighted rate of return.
You start by listing the general in and outflows per year:

	P	n
Jan 2019	67	5
Jan 2020	74.4	2 + 5.3
Jan 2021	132	7 + 7.18

From 67 → 74.4 = 0.2 increase
From 74.4 → 132 = 0.73 increase

Both the arithmetic and geometric average, are based on year-by-year rates:

$$R_A = \frac{0.2 + 0.73}{2} = 46.43\%$$

$$R_G = (1.2 \times 1.73)^{\frac{1}{2}} - 1 = 44.03\%$$

The dollar-weighted rate of return is based on cashflows:

$$0 = -335 + \frac{-139.8}{1 + R_{IRR}} + \frac{936.6}{(1 + R_{IRR})^2}$$

$$R_{IRR} = 47.64\%$$

-335 = 5 x 67
-139.8 = you buy 2 shares and receive dividends

Which is more appropriate: geometric average or dollar-weighted average?
In summary: if you would like to summarize the actual performance which one of the three should you use? Arithmetic drops out as already mentioned. For the rest it depends. In practice we see that most of the time the geometric average is used. That is appropriate when you assume that the fund manager has no control over in and outflows (which is typically the case in a mutual fund). If however you have a control over in and outflows than you could use this IRR concept (suppose you actively engage in timing).
So:
- When comparing performance across fund managers, the geometric average is more appropriate since cash in- and outflows are not under their control.
- On a more individual level, when cash in- and outflows are controlled and can be 'timed' an IRR concept can be useful.

Risk correction
Of course, return is just one element of performance, also risk involved should be measured

Two approaches are used:
- comparison universe: compare rates of return with those of other investment funds with similar risk characteristics
    - you make a subset of investments in terms of riskiness of these investments and then you look at the returns earned by each of the investments
    - is often used in the industry.
    - While it gives a first idea of relative performance, it might hide specific and important risk exposures. For example, within a particular universe, some managers may concentrate on particular subgroups, so that portfolio characteristics are not truly comparable.
    - So, a amore precise measure is needed:

- risk-adjusted performance measure

To unravel them, we need **risk-adjusted measures**
A wide range of risk-adjusted performance measures exist:

1. Absolute performance measures: the individual investment risk-return characteristics are summarized and we do not compare with some kind of benchmark. Here you can rank different investment in terms of increasing or decreasing attractiveness.
2. Relative performance measures: investment risk-return characteristics are benchmarked
3. Industry fund rating systems

Risk correction: absolute performance measures

The most traditional ('grandfather') performance measure is the **Sharpe ratio**:

$$SR = \frac{R - R_F}{\sigma}$$

measure where, in practice, returns are annualized (EAR) and σ is the volatility of returns (ignoring any variation in risk-free rates)

This is an ex-post measure now in this case, how does that influence the calculation of your SR? It influences both the numerator as the denominator. The numerator was a risk premium but now it is the **effectively realized return**. Also, the calculation of volatility will differ; you should calculate the volatility of the **population** and not of the sample (as in the ex-ante setting).

Interpretation: investors with mean-variance utility prefer investments with higher rewards per unit of risk taken (SR = slope of the CAL)→ rank investments according to SR
The SR is then an absolute measure of performance, a method that allows you to rank different methods.

While the SR in itself does not account for any benchmark comparison, one can define 'acceptable' performance of an investment by comparing it to the SR of a passive market portfolio. You could incorporate a relative measure of performance by using a benchmark by calculating the SR of a passive portfolio and then ranking the SRs vis a vis this benchmark.

Numerous variations to the above Sharpe exist:

1. **Roy ratio**: $\frac{R-R_T}{\sigma}$ with R_T a minimum target return
2. **Revised Sharpe ratio**: $S = \frac{R-R_F}{\sigma_{R-R_F}}$
3. **Adjusted Sharpe ratio**: $SR \times \left[1 + \frac{\zeta}{6} \times SR - \left(\frac{\kappa-3}{24}\right) \times SR^2\right]$ with ζ the skewness of returns and κ its kurtosis
4. **MAD ratio**: $\frac{R-R_F}{MAD}$ with MAD the mean absolute deviation

- Roy ratio: reward-to-risk measure with the difference that the reward is computed in a different way; instead of computing excess returns you will calculate the return vis a vis some target return
- Revised SR: here we divide by **risk of the excess returns**. This accounts for time variation in R_f.
- Adjusted SR: the SR fits in normal returns, it takes into account expected returns (higher moments) and volatilities (lower moments). Kurtosis are ignored which is fine as long as your returns are normally distributed. But we do know that the returns are not. The adjusted SR accounts for this observed skewness and kurtosis.
- MAD ratio: is also a reward-to-risk where risk is MAD → difference with volatility is that outliers are dampened

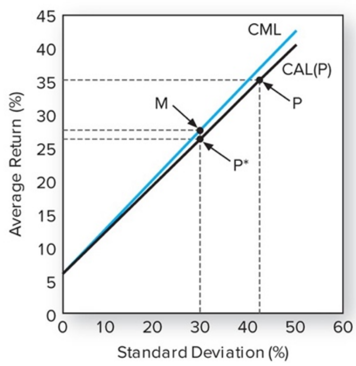
Drawback: all of these measures allow to rank portfolios, but do not really allow to quantify the size of the performance
→ compute **risk-adjusted returns:** we can do this by making sure that you compare the returns of the investments conditional upon equal levels of risk. This is what risk-adjusted returns do;

Most commonly used and easy-to-interpret measure of risk-adjusted return is the **M^2**-measure (named after Leah and Franco Modigliani). It compares the return of your portfolio and those of the market but after making sure that the riskiness of your portfolio matches this of the market:

$$M^2 = R^* - R_m$$

with the (active) portfolio to evaluate with return R and risk σ, R^* a portfolio combining the active portfolio and the risk-free rate with risk $\sigma_p{}^* = \sigma_m$ the risk of the passive market portfolio.

Figure 24.2 The M^2 of portfolio P is negative even though its average return was greater than that of the market index.



Suppose we have some kind of portfolio P. We calculate the CAL(P) and the CML of the market portfolio M. Based on the SR the M is more attractive. But we want another measure, an expression in terms of returns earned. How can I do this? Instead of evaluating P I evaluate P* which is a combination of portfolio P with R_f where the riskiness (volatility) is the same as M. Moving on the CAL does not impact your SR. Now I can just look at realized returns. This is a risk-adjusted return measure.

Risk correction: relative performance measures

While an absolute measure of performance allows to rank overall risky portfolios, a relative measure is more appropriate when ranking portfolios that will be mixed

1. **Information ratio**: to evaluate the contribution of an active portfolio to a benchmark portfolio, passive portfolio (cf pragmatic approach to portfolio selection). This about which assets I want to add to my existing portfolio whereby you focus on the level of idiosyncratic risk.
2. **Treynor ratio**: to evaluate the components of a well-diversified complete risky portfolio. This is used in a context where you want to add a number of well-diversified portfolios together in one overall portfolio (see example HB 819); whereby you focus on systematic risk. And a well-diversified portfolio is already well diversified and does not have idiosyncratic risk. What is the risk focus of a well-diversified portfolio then? The systematic risk.

The **Information ratio** is also a reward to risk ratio:

$$IR = \frac{\alpha}{\sigma_\epsilon}$$

with reward measured as a free lunch return, and risk as tracking error (=idiosyncratic risk)

Intuition: IR trades of the extra return from active investing versus the additional risk it brings

Extra explanation if you don't understand: The information ratio is yet another version of a reward-to-risk ratio. In this context, the reward is the alpha of the active position. It is the expected return on that incremental port- folio over and above the risk premium that would normally correspond to its systematic risk. On the other hand, the incremental position tilts the total risky portfolio away from the passive index, and therefore exposes it to risk that could, in principle, be diversified. The information ratio quantifies the trade-off between alpha and diversifiable risk.

Active portfolios with highest IR are most attractive as they maximize the overall SR:

$$SR^2 = SR_m^2 + \left[\frac{\alpha}{\sigma_\epsilon}\right]^2$$

with SR_m the Sharpe ratio of the passive market portfolio

Also the **Treynor ratio** is a reward to risk ratio:

$$T = \frac{R - R_F}{\beta}$$

with reward measured as excess return, and risk as systematic risk (ignoring specific risk)

Intuition: Treynor trades off the excess return from an investment versus the systematic risk it entails

Portfolios with highest Treynor (or steepest T-line) are most attractive

Figure 24.3 Treynor measures of two portfolios and the market index.

On the Y-axis you have rate of returns and on the X-axis, you have Beta's.
Here you draw Traynor lines. If you have here M, the line that combines the assets with the market portfolio are all combinations of the Rf and M but with varying Beta exposures = SML

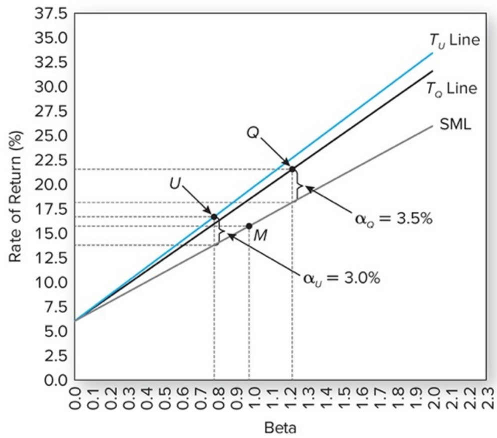
How to evaluate it? The higher the slope of the Traynor line, the more attractive the portfolio will be. If you look roughly at the numbers you will see that the return earned on portfolio q is higher than the return earned on portfolio u.
In addition, the alpha earned by q is larger than the alpha earned by u.

Portfolio U
β : 0.8
R : 17 %
α : 3 %

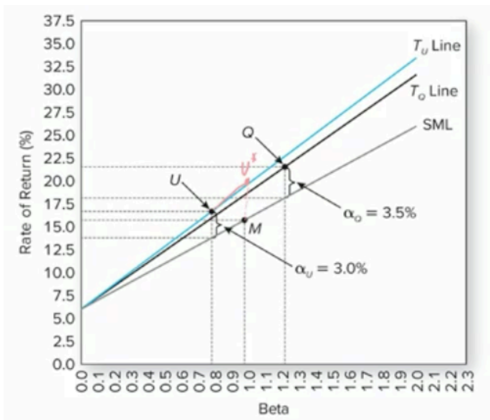
Portfolio Q
β : 1.3
R : 22 %
α : 3.5 %

Market
R : 16 %

Tbill
R_f : 6 %



In line with the M^2 you could also come up with a T^2 adjusted measure to measure the advantage of each portfolio compared to the market index. By making sure that portfolio u has the same risk profile as the market. You move on the Traynor line such that you end up with the same Beta. Then you can compare the returns earned by the market M and the returns earned by U*. You make sure that the riskiness is identical (ß). Here you see that portfolio U*, given an amount of systematic risk, outperforms the market = it gives us alpha.



Main take-away: do not be overwhelmed by initial information in terms of attractiveness of an investment. It is not because returns are higher and because alpha is higher, that the portfolio is necessarily attractive. It depends on the measure you are using.

Performance Measure	Definition	Application
Sharpe	$\frac{\text{Excess return}}{\text{Standard deviation}}$	When choosing among portfolios competing for the overall risky portfolio
Treynor	$\frac{\text{Excess return}}{\text{Beta}}$	When ranking many portfolios that will be mixed to form the overall risky portfolio
Information ratio	$\frac{\text{Alpha}}{\text{Residual standard deviation}}$	When evaluating a portfolio to be mixed with the benchmark portfolio

What about pure alpha?

Much emphasis is put on Alpha. Alpha is necessary to outperform, superior performance requires alpha, but largest alpha is not necessarily most attractive.

The above risk-adjusted performance measures all reflect alpha presence, but also show how its attractiveness depends on the risk involved:

	Sharpe	IR	Treynor
Relation to α	$\frac{E(R_P)-R_F}{\sigma_P}$ $= \frac{\alpha_P}{\sigma_P} + \rho SR_m$	$\frac{\alpha}{\sigma_\epsilon}$	$\frac{R_P-R_F}{\beta}$ $= \frac{\alpha_P}{\beta_P} + T_m$
Improvement vs market	$SR_P - SR_m$ $= \frac{\alpha_P}{\sigma_P} - (1-\rho)SR_m$	$\frac{\alpha}{\sigma_\epsilon}$	$T_P - T_m$ $= \frac{\alpha_P}{\beta_P}$

This table summarizes indicates how these performance measures are linked to Alpha and how you can calculate the improvement of a particular investment vis a vis the market. You will also be able to see immediately what the main drivers are of this improved performance. For example, in terms of the SR you see that it is all about the alpha that a particular investment realizes vis a vis total variance. **But** also, the correlation with the market matters. For the IR it is straightforward, and the improvement is driven by this ratio. Finally, the Treynor ratio you can rewrite this in function of Alpha. When you look at the improvement vis a vis the market you see that this is driven by Alpha and **also** by the systematic risk. You **should not** necessarily go for that investment with the largest alpha, it depends on the purpose of your investment.

So, while positive alpha is necessary, it is not sufficient to guarantee that a portfolio will outperform the index: Taking advantage of mispricing means departing from full diversification, which entails a cost in terms of nonsystematic risk. A mutual fund can achieve a positive alpha, yet, at the same time, its volatility may increase to a level at which its Sharpe ratio will actually fall.

Risk correction: industry fund rating systems
(Not go in detail; know the concepts and the two common systems.)
A number of private companies provide mutual fund performance data:

1. Morningstar
2. Lipper leaders

Since these ratings are influential and drive sales of retail investors, good understanding of the rating systems is necessary. These are important as retail investors really take into account the ratings, they get via these two. CFs flows in to funds with allot of stars and flows out of funds with few funds. They are really important in the market.

Morningstar has created a proper risk-adjusted performance measure MRAR:

$$MRAR(\gamma) = \left[\frac{1}{N} \times \sum_{t=1}^{T} \left(\frac{1+R_t}{1+R_{F,t}}\right)^{-\gamma}\right]^{-\frac{12}{\gamma}} - 1$$

with γ = risk aversion (=2 for Morningstar)

Interpretation: MRAR gives the certainty equivalent excess return of the portfolio (funds with volatile returns are penalized). There is a log utility analysis behind this. Funds are separated in subsets.

By ranking the risk-adjusted performance among peers, funds are assigned a 'star rating'. This means that the peer group you are assigned to is very important. It could be that in a particular peer group you are the best in class whilst in another peer group you are not best in class. Belonging to a good peer group is very important. There is some arbitrage there where fund managers discuss with Morningstar…

Lipper leaders provide estimates of performance via five key metrics:

1. total returns
2. consistent returns
3. preservation
4. expense
5. tax efficiency (for US funds)

By ranking the performance metrics among peers, funds are assigned a rating (scale 5 to 1)

Challenges

Risk-adjusted measures are intuitive and attractive to evaluate portfolio performance, but important challenges remain:

1. Which measure is most appropriate?
2. What is the impact of time-varying risk?
    - All these measures assumes that the strategy is constant during the investment horizon, but this is not necessarily the case.

Which measure is appropriate?

There is a multitude of (risk-adjusted) performance measures:

- some easy to calculate; others less
- some easy to understand; others less
- sometimes contradicting one another: if you rank portfolios for example according to Treynor and you rank according IR than you can end up with different rankings

Important: define the (risk-adjusted) performance metrics

1. if you define your risk evaluation measures you have to do this ex ante (no datamining); do not decide this ex post when you want to do this revaluation (to avoid that you pick the one that will make your results look best)
2. consistent with the investment objectives (eg absolute vs relative)
3. aligned with investment strategy peculiarities (eg exotic strategies call for more complex metrics)
4. well understood and informative (eg reward to risk ratios to obtain ranking vs risk-adjusted returns to quantify the relative performance)
5. with a focus on one or at max a few (to avoid conflicting objectives): do not calculate a multitude of measures. Some people do thiss but then calculate an average of all the results, your results will then be okay, but it does not say allot.

To ensure consistency, standardization and investor confidence, Global Investment Standards (GIPS®) exist. These are ethical standards based upon **fair representation** and **full disclosure** to promote **self-regulation**. As an asset/fund manager you can adhere to the GIPS. That is, you agree to compute portfolio performances according to the principles that have been given by GIPS.

The standards define e.g.:

- fundamentals of compliance
- input data used
- calculation methodology
- reporting requirements

Time-varying risk

The different risk-adjusted measures assume that portfolio risk is constant over the evaluation period.
However, in practice, portfolio risk changes for two reasons:

1. **Changing portfolio composition**
2. **Market timing**

Whatever the reason is these results in a different risk profile in different periods.

Changing portfolio composition

When the portfolio composition/investment strategy is significantly changed during the evaluation period, this can **bias estimates of risk-adjusted performance** measures.

Even worse: an agency problem can arise when such biases are ’gamed’ by the portfolio manager

When a change in strategy is misinterpreted as risk...

Example: Changing portfolio risk

Assume a manager follows a low-risk strategy in year 1, and switches to a high-risk strategy in year 2, with following realized returns:

Quarter	Realized excess return
Q1	-1%
Q2	3%
Q3	-1%
Q4	3%
Q5	-9%
Q6	27%
Q7	-9%
Q8	27%

The Sharpe ratio of the market index is 0.4. Is the manager performing well? (note: use the geometric average in the calculation of SR!)

The Sharpe ratio (p.a.) over the 2 years equals:

$$SR = \frac{0.042}{0.134} = 0.312$$

which is less than the SR of the market index.

However, we observe the following year-by-year Sharpe ratio's:

$$SR_1 = \frac{0.010}{0.02} = 0.49 \quad \text{and} \quad SR_2 = \frac{0.075}{0.18} = 0.42$$

which is higher than the market index Sharpe.

The difference in return dynamics over the 2 years is interpreted as risk (volatility), while it reflects a change in investment strategy.

(Use population volatility!)

- *If you calculate the SR (which is lower than the SR_m) over the total period, you see that it is not performing well*
- *If you calculate this over separate periods:*
    - *Year 1: almost 50%*
    - *Year 2: 42%*
    - *In both cases the SR is above the SR of the market: so you will conclude that over one year it performs well but over the two year period it doesn't*
- *Volatility is misunderstood as risk, while the change in results is largely due to a change in investment strategy.*

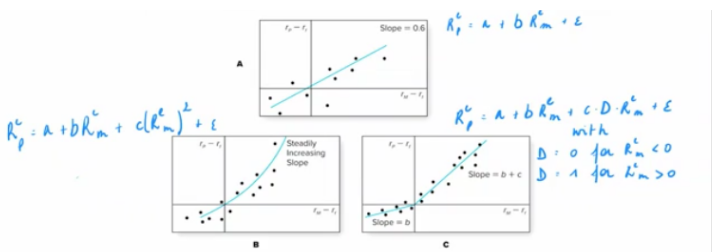
Such misunderstanding of investment strategies pursued, can be 'abused' by managers to manipulate their performance by changing investment strategies within the evaluation period. Managers who observe portfolio performance on a going forward basis, might decide to adjust the capital allocation decision (increasing or decreasing leverage) to change the relative importance of already observed performance numbers (amplify or dampen)

However: such arbitrary variation in leverage and risk is utility-reducing for investors and only benefits the manager who aims at maximizing the Sharpe (no matter how).

Note: the previously introduced MRAR does not allow for such manipulation

Market timing

More generally, all changes in the capital asset allocation (arbitrary or not) lead to changes in portfolio risk...thus also **market timing** where managers **go in and out of the market depending on their expectations**. *In its pure form, market timing involves shifting funds between a market-index portfolio and a safe asset, depending on whether the index is expected to outperform the safe asset.*



Panel A: No market timing, because beta is constant.
Panel B: Market timing, beta increases with expected market excess returns. He is stepping into the market when results are good, and he is stepping out of the market when results are bad. You can easily calculate this by $c(R^e_m)^2$
Panel C: Market timing with only two values of beta. Here we include dummy variables.

What is the impact of market timing?
Whenever a portfolio manager tries to time the market, it is important to account for that in your model. Otherwise, your estimates will not be accurate.

Example: Market timing

To understand market timing, run:

$$R_i^e = \alpha_i + \beta_i R_m^e + \tau_i (R_m^e)^2 + \epsilon_i$$

Estimate	Portfolio P		Portfolio Q	
Alpha	1.63%	1.77%	5.26%	-2.29%
Beta	0.70	0.70	1.4	1.1
Timing		0.00		0.10
R-square	91%	91%	64%	98%

- Portfolio P did not attempt market timing
- Portfolio Q was successful in market timing (which is offset by unsuccessful selection)

⇨ Here first results are with no market timing and second results with (in each portfolio)
⇨ Here we see that P does not doe market timing, but Q does and in that cases the estimates changes dramatically; market timing 0.1 means that the beta, in the case of market is doing well, increases. Alpha is negative; we see here that initially alpha was positive, but the model did not fit as there was market timing involved and using the correct model shows that the alpha is actually negative.
⇨ Trying to game the market is for a large part offset by unsuccessful selection.
⇨ This shows you that if someone is pursuing such a strategy than you should account for this otherwise your standard estimated will be biased.

This example illustrates the inadequacy of conventional performance evaluation techniques that assume constant mean returns and constant risk. The market timer constantly shifts beta and mean return, moving into and out of the market. So, market timing presents another instance in which portfolio composition and risk change over time, complicating the effort to evaluate performance. The simple SCL does not capture this so this shows that it is important to find a way to capture this market timing.

Is the potential of market timing big?
Empirical evidence on successful market timing is sparse; however, its potential is huge

Strategy	Bills	Equities	Perfect Timer
Terminal value	$20	$5,271	$755,809
Arithmetic average	3.39%	11.49%	16.41%
Standard deviation	3.14%	20.04%	13.44%
Geometric average	3.34%	9.76%	15.85%
Maximum	14.71%	57.35%	57.35%
Minimum*	−0.02%	−44.04%	0.00%
Skew	1.05	−0.39	0.75
Kurtosis	1.01	0.07	−0.07
LPSD	0.00%	13.10%	0.00%

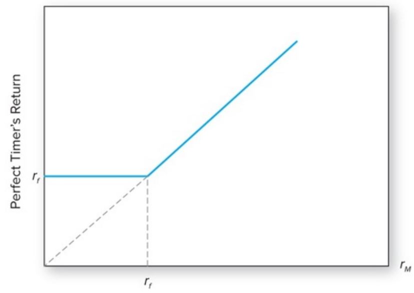
*A negative rate on "bills" was observed in 1940. The Treasury security used in the data series in these early years was actually not a T-bill but a T-bond with 30 days to maturity.

Table 24.5
Performance of bills, equities, and perfect (annual) market timers. Initial investment = $1

Here you see a strategy where you are either always invested in T-bills or always invested in equities (when the equity market is doing better than T-bill market) and then perfect timer (invested in T-bill when stock market is doing worse than the T bills and otherwise if stock is outperforming T-bills). Look at the terminal value. The potential of perfectly timer is huge. Also, other key characteristics; one number that we highlight is the SD of the perfect timer; of an equity is pretty high, of bills low, but for the perfect timer it is at 13.44%. Is this the riskiness of the perfect timer? This is a high volatility, does this mean that this strategy is risky? Can you have a negative return? Look at the minimum return, it is zero. With equity you could go negative. = risk and that is captured in the volatility. But with the perfect timer strategy their volatility is not necessarily risk. Because you always get the best of the two so you are bounded at the lower end, you will always at least have the T-bill or better. So, this volatility is the deviation on the upside and not the downside. So is volatility an appropriate measure of risk? It accounts for lower but also higher than expected (which is not risk but an opportunity). So, for the perfect timer case it is positive deviations. This is an example in which you see the weakness of the volatility as a measure of risk.

How can we explain the potential of market timing?
The value of perfect market timing can be appreciated as a call option on the risky investment (at zero cost)

Figure 24.9 Rate of return of a perfect market timer as a function of the rate of return on the market index



If you are able to time the market it is as if you are invested in T-bills, but you have a call option on the stock market. This means whatever happens you have the T-bill return but when the stock market is doing well, you exercise the call option and get active on the market. You can't go below the steady line, only up and this upward sloping line is the SD.

Skill to perfectly time the market = being long a call option on the market + investing in a T-bill.

Assume a call option on the market (currently valued at S_0) with exercise price $X = S_0(1 + R_F)$ *(this is the terminal value of the T-bill).* This means that you won't exercise your call if the market is doing poor, only when it is doing well such that your total payoff is either the lower bound determined by the T-bill or the upward potential of the market. The T-Bill earns a certain rate of return R_F

	$\mathbf{S_T < X}$	$\mathbf{S_T \geq X}$
TBill	$S_0(1 + R_F)$	$S_0(1 + R_F)$
Call	0	$S_T - X$
Total	$S_0(1 + R_F)$	S_T

Can I value this?
→ Use **option pricing** to value perfect market timing skill

The value of a call option can be derived using **Black Scholes** as:

$$C = S_0 N(d_1) - X \exp^{-R_F T} N(d_2)$$

With

$$d_1 = \frac{\ln(S_0/X) + (R_F + \sigma^2/2)\,T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

Assume now a call option on the market with an initial value $S_0 = 1$ and an exercise price $X = \$1 \times exp^{RFT}$

What is the market value per $ of assets?

The **market value of a perfect timer** call option on a $1 market portfolio reduces to:
You should use that N(-X) = 1 – N(X) to calculate this

$$C = 2N\left(0.5\sigma\sqrt{T}\right) - 1$$

Example: Market value of perfect timer call option

Assume a single year, and a market volatility of 25%. What is then the value of perfect timing?

Solution: Market value of perfect timer call option

For T $= 1$, $\sigma_M = 25\%$:

$$C = 2N\,(0.125) - 1 = 2 \times 0.55 - 1 = 10\%$$

→ 10% per dollar invested is the value of being able to time the market.

What about the market value of <u>imperfect</u> market timing?
To account for imperfect timing as opposed to perfect timing, we need to correct for **timing ability** = proportion of correct forecasts of bull markets (step into the market) and of bear markets (step out of the market and into T-Bills).
Define:

P_1 = proportion of correct bull forecasts
P_2 = proportion of correct bear forecasts

Then:

Timing ability = $P_1 + P_2 - 1$

The **market value of an imperfect timer** call option on a $1 market portfolio is then the market value of the perfect times *times* timing ability:

$$(P_1 + P_2 - 1) \times C = (P_1 + P_2 - 1) \times (2N\left(0.5\sigma\sqrt{T}\right) - 1)$$

Example: Market value of imperfect timer call option

Assume a single year, and a market volatility of 25%. A fund manager flips a coin to predict bull and bear markets. What is then the value of imperfect timing of this fund manager?

Solution: Market value of imperfect timer call option

For T $= 1$, $\sigma = 25\%$:

$$C = 10\%$$

A fund manager that flips a coin to predict market timing, predicts bull and bear markets completely randomly. As a consequence, half of bull and bear markets will be predicted correctly:

$$P_1 + P_2 - 1 = 0.5 + 0.5 - 1 = 0$$

As there is no timing ability, the market value of the imperfect timer is also zero.

Drivers of performance

→ more specific information about what led to over or underperformance or the returns that you got
While traditional evaluation methods measure the risk-return trade-off of an investment, it is not an evaluation of the investment decisions made, nor how they contribute to the performance.
Two procedures exist to explain the **performance of portfolio allocation decisions taken**:

1. Style analysis
2. Performance attribution

Style analysis

It tries to identify the style of your investment strategy. For example a small cap investor his style is small cap.
Style analysis tries to determine the exposures of a portfolio to a set of indices I_k representing a range of asset classes:

$$R_i = \alpha_i + \sum_{k=1}^{K} \beta_k I_k + \epsilon_i \qquad \text{with} \quad \forall k : \beta_k \geq 0 \quad \text{and} \quad \sum_{k=1}^{K} \beta_k = 1$$

This way it gives the implicit allocation (exposure β_k) to each *style*, with α_i the average success of security selection.
The R-square of such regression measures the percentage return variability that is explained by the set of styles accounted for and thus that can be attributed to style (the portion not explained is attributed to security selection within styles and timing). Then you will have a part that is not explained by the styles. This part can be attributed either to selection skill within style or to market timing. If you want to put a number on the success of security selection you should look at the unexplained part of the regression as well on the intercept here. Very often people will focus on this unexplained part because when you run such a regression mostly these R squared are quite high and then they would conclude that most of the part is explained by the style and people are not good in selecting. But we should not forget that there is also an alpha, and you should account for this one.

Example:

Table 24.4

Style analysis for Fidelity's Magellan Fund

Style Portfolio	Regression Coefficient
T-bill	0
Small cap	0
Medium cap	35
Large cap	61
High P/E (growth)	5
Medium P/E	0
Low P/E (value)	0
Total	100
R-square	97.5

Source: Authors' calculations. Return data for Magellan obtained from **finance.yahoo.com/funds** and return data for style portfolios obtained from the Web page of Professor Kenneth French: **mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.**

Performance attribution

Where style analysis gives broad insights into the drivers of realized returns, **performance attribution** goes further and aims at detecting **the sources of superior (inferior) out- (under) performance**. This detects whether it was asset allocation decisions rather than security asset selection decisions that led to out or underperformance.

By comparing the portfolio's weights and return to the ones of a 'bogey' one can determine the fund's performance attributable to:

- asset allocation decisions
- sector/security selection decisions

The bogey is designed to measure the returns the portfolio manager would earn if he or she were to follow a completely passive strategy. “Passive” in this context has two attributes:

- First, it means that the allocation of funds across broad asset classes is set in accord with a notion of “usual,” or neutral, allocation across sectors.
- Second, it means that *within* each asset class, the portfolio manager holds an indexed portfolio, such as the S&P 500 index for the equity sector.

In such a manner, the passive strategy used as a performance benchmark rules out asset allocation as well as security selection decisions. Any departure of the manager’s return from the passive benchmark must be due to either *asset allocation bets* (departures from the neutral allocation across markets) or *security selection bets* (departures from the passive index within asset classes).

Assume a portfolio *P* and a bogey *B*, with w_k the weight allocated to asset class *k*, and R_k its corresponding return.

Return attributable to **asset allocation decisions**:

$$R_A = \sum_{k=1}^{K} (w_k^P - w_k^B) R_k^B$$

These are the weights you allocate to your portfolio compared to the bogey. This allows you to look if you deviate from the bogey, that you allocate more wealth to the asset class and the asset allocation decision.

Return attributable to **sector/security selection decisions**:

$$R_S = \sum_{k=1}^{K} w_k^P (R_k^P - R_k^B)$$

For a given allocation to a particular asset class k you look at the return you earn compared on the returns of the same asset class of the bogey of the benchmark portfolio. If you earned higher return than you made smart selection decisions.

You see both differences in allocation and selection decisions because the allocation differs, and the returns earned on the different asset classes differ. The question is than did they pay off? And which of the decisions pays off more. What contributed to the excess return of 1.37%?

Example: Performance attribution (BKM p. 837)

A managed portfolio, with return 5.34%, has following characteristics:

Asset class	Weight	Return
Stocks	70%	7.28%
Bonds	7%	1.89%
Cash	23%	0.48%

The bogey, with return 3.97%, looks as follows:

Asset class	Weight	Return
Stocks	60%	5.81%
Bonds	30%	1.45%
Cash	10%	0.48%

The excess return of the managed portfolio equals 5.34%-3.97% = 1.37%

First, we calculate the contribution of opting for a different **allocation decision**:

A. First column (1) compared to second column (2); shows that we deviate from the benchmark. Was this smart? In the last column we calculate the contribution to performance. We see that the contribution was 0.31 (of the 1.37%). This means that the rest must be contributed by selection decisions.

B. Here we look at the contribution again. Comparing the first two columns shows that you earned more than the Bogey on the equity and more on the fixed income. You see that most is contributed by the choices you made by respect to the equity stake (in the last column)

Table 24.7 Performance attribution

A. Contribution of asset allocation to performance

Market	(1) Actual Weight in Market	(2) Benchmark Weight in Market	(3) Active or Excess Weight	(4) Index Return (%)	(5) = (3) × (4) Contribution to Performance (%)
Equity	0.70	0.60	0.10	5.81	0.5810
Fixed-income	0.07	0.30	−0.23	1.45	−0.3335
Cash	0.23	0.10	0.13	0.48	0.0624
Contribution of asset allocation					0.3099

B. Contribution of selection to total performance

Market	(1) Portfolio Performance (%)	(2) Index Performance (%)	(3) Excess Performance (%)	(4) Portfolio Weight	(5) = (3) × (4) Contribution (%)
Equity	7.28	5.81	1.47	0.70	1.03
Fixed-income	1.89	1.45	0.44	0.07	0.03
Contribution of selection within markets					1.06

Next: take it one step further by analyzing the **sector and security choices**. Zoom into the equity stake what drives this result?

The excess return of the managed equity portfolio equals 1.47%
Your equity stake consists of different subclasses. Maybe you allocated more to particular sectors and less to other sectors. So, the selection within the equity class can then be explained by asset allocation decisions within the equity class or is it selection decisions within the equity class? We redo the same analysis but within the equity stake.
We compare with a bogey, here the S&P500.
Then: the remaining excess return 1.47%-1.29% = 0.18% is attributed to security selection within sectors

Table 24.8 Sector selection within the equity market

Sector	(1) Beginning of Month Weights (%): Portfolio	(2) Beginning of Month Weights (%): S&P 500	(3) Active Weight (%)	(4) Sector Return (%)	(5) = (3) × (4) Sector Allocation Contribution
Basic materials	1.96	8.3	−6.34	6.9	−0.4375
Business services	7.84	4.1	3.74	7.0	0.2618
Capital goods	1.87	7.8	−5.93	4.1	−0.2431
Consumer cyclical	8.47	12.5	−4.03	8.8	0.3546
Consumer noncyclical	40.37	20.4	19.97	10.0	1.9970
Credit sensitive	24.01	21.8	2.21	5.0	0.1105
Energy	13.53	14.2	−0.67	2.6	−0.0174
Technology	1.95	10.9	−8.95	0.3	−0.0269
Total					1.2898

Summing up all attributions...

Table 24.9 Portfolio attribution: summary

		Contribution (basis points)
1. Asset allocation		31
2. Selection		
a. Equity excess return (basis points)		
i. Sector allocation	129	
ii. Security selection	18	
	147 × 0.70 (portfolio weight) =	102.9
b. Fixed-income excess return	44 × 0.07 (portfolio weight) =	3.1
Total excess return of portfolio		137.0

Guest lecture: Belfius

What is asset management?

Why is it interesting to have an asset management in a bank? Because of the low interest rates → made I very hard for banks to make money

It's also sales and marketing

We have different business models

Business model

High net worth investors: have allot of net worth is a separate catg because mostly interested in other products

Channels: In Belgium asset management goes mainly through the banks

Distribution wrappers:
- open ended funds
- mandates: portfolio managed specifically for you

…
→ different legal wrappers to put your investment in it

Asset managers: not only picking the best stocks but also diversify
Making also the right investment
Reducing transaction cost: not optimal if individual investors do it on their own, bank can be more efficient

Overview of funds

Funds categories
Traditional:
- products actively managed with the ambition of generating alpha against a benchmark index
  - index
  - fixed income
  - balanced
  - money market

Alternative
- real estate: products investing mainly in real estate
- private equity: products that directly invest in pricate companies, or tat negage in buyouts of public companies
- hedge funds: products invested in alpha generating strategies, allowed to use leverage and sell short
- structured & other: products structured to provide a specifi risk/ return profile or invested in assets

Passive:
- products managed with no objective to generate alpja
- includes index funds, passive etfs, passive mandates

value chain

asset management value chain
sales & marketing
- product dev
  - competitive intelligence and customer research
  - product development/ sourcing and protyping
  - feedback from RMs
- marketing
- sales/sales support

investment management
- rsearch
- portfolio managm
- trading

operations
- investment operations
- fund amdin
- reporting

support functions
management & admin
technology

How to build a risk framework?

Efficient portfolio theory:
- why do asset managers have multiple funds?
  - Because markets are not fully efficient and different risk types
  - Asset management is also about telling a story, telling a dream
- Is Markowitz still used?
  - For what type of funds?
    - Stock picking funds/equity funds? No.
    - Real estate funds? No.
    - It is not used for everything
  - For what type of activity?
    - For a very different risk profile, ?

How a fund is born
When will you use this technique?
First fund conceptualization: do some research, look at client needs, global offering and gabs, funds definition, reference allocations with example portfolios (might use MW to optimize) and the prospectus
Then fund creation: creation of the investment universe, creation/adaption of risk framework, selection of underlying, building of the refeence portfolio (might use MW to optimize), commercial material
Then fund operations: legal structure, contracting, set up accounts and information channels

What are the various risks?
- Equity risk
- Interest rate risk
- Commodity risk

- Credit risk
- Settlement risk
- ….

If someone says you need to build a risk framework you will talk about this
Which one are really captured by the MW?

- Equity risk
- Credit risk? No, tailor risk so not captured by MW
- Interest rate risk
- Commodity risk
- Echange rate risk
- Performance risk
- Capital risk
- …

Example: Volatility based fund → based on what risk you take

Markowitz implementation

Objectives & notations

Objective: finding the asset allocation that gives the highest expected return for a given level of risk or for each level of risk

Definition: the efficient frontier is the set of optimal portfolio that gives the highest expected returns for each level of risk

The efficient frontier can also be represented as the differnt asset allocations by level of risk

We need:

- Weights
- Expected returns of each asset class
- Covariance matrix

Calculate ER of the portfolio and the volatility of the portfolio
Then we maximize the ER of the portfolio
Then maybe with some constraints:

- Linear constraints
- Non linear constraints
- But if you have too many constraints there will maybe be not enough margin

Inputs

- ER of each asset class
    - What methodology would you suggest to get ER?
        - Historical returns
    - What time horizon would you target?
        - Not always clear
        - It really depends on the type of investment
        - 1-10 years
- Covariance matrix between each asset classes
    - What methodology would you suggest?
        -
- Risk limits: risk exposures must be correctly identified

ER methodology

- Observed historical returns
- Expert based (forward looking expectations)
- Black-litterman (reverse engineering based on observed market cap)
    - Issues with black-litterman:
        - How to define market cap of each asset classes?
        - Does it make sense if actors have different definition of risk?
        - Regulatory impact on asset allocation: is it optimal or forced by regulation?

Time horizon

- No clear answer, market practice range mostly betwn 1-5y but the time hprizpn should male sense with regards to the rebalancing policy

Covariance matrix methodology:

- Observed historical volatilities/correlations
- Forward to market implied volatility from the options market
- Mathematical forecasting method

Time windows & time horizon

- Distinction should be made between
    - Frequency of the observed data used to make estimates
    - Time windows used to collect data
    - Horizon of the prediction
    - Scaling of the calculated value

Technical limits

Several limitations exiss when one is trying to apply the markowits model
Example: concentration on a few asset classes (example with target SD of 10%)
→ how do solve this issue?

- Include lower limits on some or each asset classes
    - Will result in investing in poor asset classes and remove the added value of the optimizer
- Add a non linear constraint
    - Better but could still result in investing in poor asset classes
- "resampled" efficient frontier
    - More complex and can solve both the concentration issue and the sensibility to exp returns. Will be the focus of the next slides

If you run an optimization with all the asset class with the same SD, how to solve it?
Resample.
Monte carlo on expected returns: using assumption that the distribution of the exp returns is?

- Randomly draw ERs of each asset class
- Run the optimization using the generated ERs
- Compute the average weights of the results of each simulation

Discretization issue

Convergence issue
In this example there was no issue

Market limits

When the condition in the market does not allow you to get what you wanted

i. A solution under all constraints might simply not exist!
ii. A solution might exist but gives a higher level of risk for no added value

iii. A solution might exist but gives lower return

→ Main message: you really need to take into account the market when you do your investment and your MW and also try not to look at just one point

Business limits

- Rebalancing costs
- Tax impacts

Rebalancing: distinction should be mate between;

- Rebalancing **to** the optimal portfolio
- Rebalancing **of** the optimal portfolio
    - How would you take into account rebalancing costs to compute the best portfolio starting from the current portfolio? Reduce the return;
        - two ideas
            - Taking into account cost in the function to optimize
            - Fixing a limit of the maximum rebalancing percentage allowed

Summary

Importance of the inputs: ERs, vol, corr, risk limits

Technical limits

- Concentration on a few asset classes
- Sensitivity to ERs
- Convergace issues
- discretization

Market limits

- impossible target
- added risk for no added value
- added risk for decrease in ER

→ you should check the overall frontier

Business limits

- rebalancing costs
- tax impacts

Implementation monitoring

Monitoring of rules

- regulatory rules
- prospectus rules
- internal rules

→ daily review through a rule engine that notifies expectations

follow up indivators/stress tests

- risk indicators: vol, vaR
- asset allocations
- currency allocations
- ratings counterparties
- ratings underlying ficed income products

→ follow up of indicators and their historic movements