ACTSC 372

Investment Science and Corporate Finance

ACTSC 372 Course notes

Edition 1.1

Contents

1

Financial Markets and Net Present Value

1.1 The Financial Markets

Financial markets exists so companies and individuals can adjust consumption across time periods, i.e., can borrow and spend now, and repay with interest later. To begin, we assume financial markets provide risk-free borrowing and lending at a rate r, which at equilibrium is unique.

The purpose of financial markets facilitate borrowing and lending (or investing in general) between participants.

Participants with funds to invest (sometimes called lenders or investors) can either invest directly in firms (called direct finance) or do so via third parties (called Financial Intermediaries) which is called indirect finance.

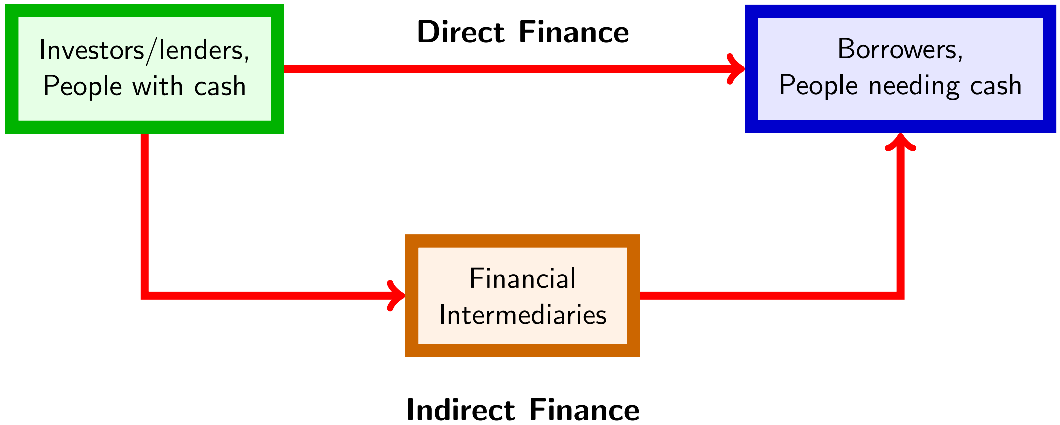
Definition 1.1.1 **Financial intermediaries**

Financial intermediaries are firms that match borrowers and lenders; take deposits and make loans (i.e. banks).

Financial intermediation can take three forms:

- Size intermediation
- Term intermediation
- Risk intermediation

Intermediaries are useful when the desires of borrowers and lenders are not exactly matched. For example, banks can combine the deposits of many small investors into large loans to few borrowers (size intermediation). Individual investors may not be willing to lend for very long term periods, so intermedaries can take funds from short term lenders and re-lend to long term borrowers (term intermediation). Finally, the risk tolerance of the lenders may not match the risk of the borrower, so intermediaries can assume the excess risk and protect the lender (risk intermediation).



1.2 Net Present Value

In order to assess the worthiness of a project, we need to understand its cash flows.

In general, when analyzing cash flows, we need to consider:

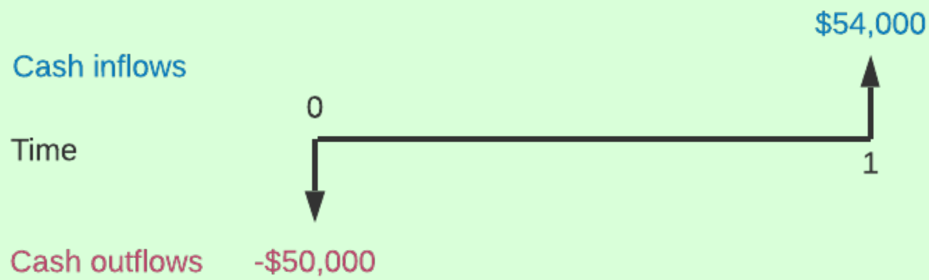
- Amount: how much cash is transacted?
- Timing: when does the transaction happen? Since \$100 received now is better than \$100 received in 1 year, the time value of money becomes important.
- Direction: who is the payer and who is the recipient?
- Likelihood: is the transaction surely to happen or somewhat uncertain?

REMARK

In this class, we refer to a project (or proposal) to mean a combination of cash flows, and a company (or firm) to be a combination of projects.

A project is worth taking if it is at least as desirable as the opportunities available in the financial market. For now, risk-free borrowing and lending are the only market opportunities. So, a project is “good” if it is better than borrowing or lending at the risk free rate.

Example 1 Project A costs $50K now and pays $54K in one year, should you take it?

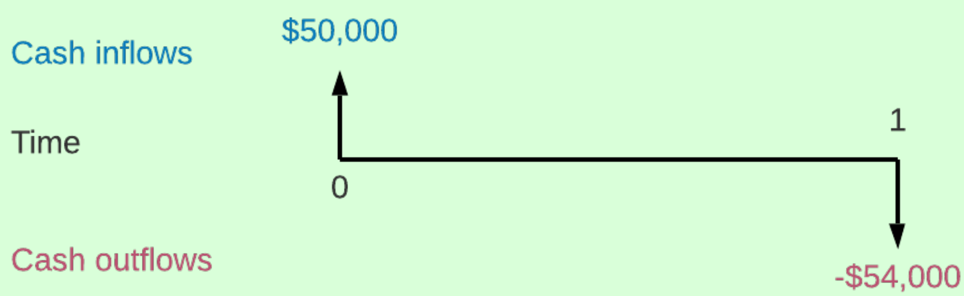


Put another way, Project A offers an 8% return, is it a good deal?

The answer depends on what the deal is in the financial market. If you save $50,000 in a bank with interest rate r, at the end of 1 year, you can withdraw $\$50,000(1 + r)$. So,

- if $r < 8\%$, invest in Project A (i.e., accept project)
- if $r > 8\%$, invest in the market (i.e., reject project and save in the bank)

Example 2 Project B pays $50K now and costs $54K in one year, should you take it? In this case, Project B **charges** 8% interest.



Again, it depends on the interest rate in the financial market. If you borrow $50,000 from a bank at interest rate r, you will need to repay $\$50,000(1 + r)$ in 1 year. So,

- if $r < 8\%$, finance from the market (i.e., reject project and borrow from the bank)
- if $r > 8\%$, invest in the market (i.e., accept project)

In summary, the decision rule can be a bit confusing when comparing Investing Projects (where we pay first and receive later) vs. a Borrowing Project (where we receive first and pay later).

In this course, the vast majority of the projects we will consider will be investing projects.

These complications (and others) can all be addressed by computing the so called Net Present value of the project. We will see that this will become the major tool to determine whether or not to accept a project.

Definition 1.2.1
Net Present Value (NPV)

The Net Present Value (NPV) of a project (i.e., cash flows C_0, C_1, ...) is

$$NPV = C_0 + \sum_{t=1}^{\infty} \frac{C_t}{(1+r)^t} = \sum_{t=0}^{\infty} \frac{C_t}{(1+r)^t}$$

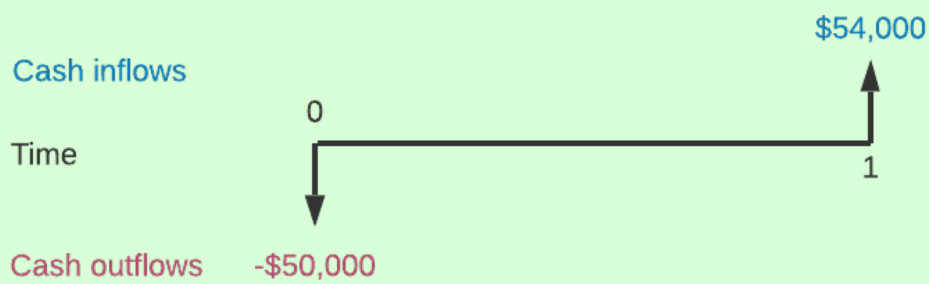
where

- C_t is a cash flow at time t (”+” and ”$-$” for inflows and outflows)
- C_0 is the initial cash flow (often negative for investment projects)
- r is the discount rate (i.e. the rate at which we discount the cash flows)

The choice of discount rate for a given project is a major theme of this course that we will explore further later.

Example 3

Consider Project A again



What is the NPV of this project if the risk-free rate is $r = 8\%$?

$$NPV = -50,000 + \frac{54,000}{1+8\%} = \$0$$

Having an NPV of zero means the project is no better or worse than the market is providing.

EXERCISE

Redo the NPV calculation for $r = 6\%$. Do it again for $r = 10\%$.

This motivates the NPV Rule.

Definition 1.2.2
The NPV Rule

The NPV rule states that a project is "good" if its NPV > 0 and "bad" if its NPV < 0. Given multiple projects, we can rank them in order of the NPVs.

To make calculations easier, it will be helpful to recall the formula for the present value of an annuity. Recall that an **annuity** is an equal stream of cash flows spaced out equally over time. For example, a payment of \$1 each year for 10 years is an annuity.

The present value of an annuity of \$1 received at the end of each period for n periods, discounted at a period rate of r is given by the annuity formula

Formula 1

Annuity Formula

$$a_{\overline{n}|i} = \frac{1 - (1+r)^{-n}}{r}$$

Example 4

Suppose Project A has the following characteristics

- Initial cost = \$1 million
- Cash flows: \$150K at the end of the year for the next 10 years
- Discount rate (also called the **cost of capital**): 8%

Should the project be accepted?

Computing the NPV of the project, we get

$$NPV = -\$1,000,000 + \$150,000 \times a_{\overline{10}|8\%} = \$6,512.21 > 0$$

Since the project's $NPV > 0$, it should be accepted according to the NPV rule.

The NPV rule has some important advantages, including:

- Discounts all cash flows properly
- Takes the time value of money into consideration
- Easy to evaluate NPV for combinations of projects
- Applied to any project, however strange and uneven the pattern of cash flows

However, the NPV rule also has some problems. Recall the formula for NPV:

$$NPV = C_0 + \sum_{t=1}^{\infty} \frac{C_t}{(1+r)^t} = \sum_{t=0}^{\infty} \frac{C_t}{(1+r)^t}$$

In order to actually compute this, what need quite a bit of data.

What are the cash flows? How much and when? (Usually, the initial costs are the easiest to estimate.) Also, at what rate should we discount the cash flows?

We will explore these challenges in future chapters.

2

Investment Rules

2.1 The NPV Rule Revisited

The NPV rule is the main tool corporate analysts uses in order to conduct capital budgeting.

Definition 2.1.1
Capital Budgeting

Capital Budgeting is the decision-making process for accepting or rejecting new projects or investments.

The NPV rule gives us a method to determine which projects to accept and reject. So, the NPV rule becomes our major tool in order to do capital budgeting.

What is not clear is if we have situations where a project may make sense for a firm, but not for the firm's shareholders. In fact, as the next example shows, this is not the case.

Example 1

The Alpha Corporation is considering investing in a risk-free project with

- Initial cost \$100, return of \$107 in one year
- The market interest rate 6%

Should the project be accepted?

$$NPV = -100 + \frac{107}{1.06} = 94 \text{ cents} > 0$$

The NPV rule says to accept the project.

But what about from the shareholder's standpoint? Shareholder's seem to have 2 choices:

1. Reject the project, and instead receive \$100 dividend from the company (basically withdraw their money)
    - Assume this cash flow is reinvested at the discount rate which we assume is available to all investors.
    - At $r = 6\%$ the \$100 now $\Longrightarrow$ \$106 in a year
2. Allow the firm to accept the project, and then receive \$107 dividend from the company in a year

Notice the shareholders should prefer that the firm accept the project too.

To summarize, accepting positive NPV projects benefits the shareholders. As we will see more explicitly later, the value of the firm rises by the NPV of the project. More specifically, the value of a project is a function of the project, and not the investor.

$$NPV > 0 \Leftrightarrow \text{increase in company value} \Leftrightarrow \text{increase in shareholder value}$$

2.2 Payback Period

While NPV is the main tool for capital budgeting, some other investment rules are used as well, including:

1. The Payback Period Rule
2. The Discounted Payback Period Rule
3. The Internal Rate of Return
4. The Profitability index

The payback period rule measures how long does it take a project to "payback" its initial investment.

Example 2 Find the Payback period of the following project:

- Initial cost = \$1 million
- Cash flows: \$150K at the end of the year for each of the next 10 years
- Discount rate: 8%

Solution: The total cash inflow after 6 years is \$0.9 million and is \$1.05 million after 7 years, so $6 <$ payback period < 7. A precise value can be found by linear interpolation.

Some advantages of the Payback Period rule:

- Easy to understand and communicate
- More likely to approve a highly liquid project. (The lower the payback period, the sooner we get all our money back.)

However, there are some disadvantages of the Payback period rule, including:

- Ignores time value of money
- Ignores all the cash flows after the payback period. Two projects with the same payback period look the same even though one may produce large cash flows after the payback period, while the other one does not.
- May result in accepting a project with a negative NPV

- What is the rule? Is a payback period of 4 good or bad? When to accept/reject a project? How to rank multiple projects?

In order to address some of these weaknesses, we can use discounted cash flows instead in order to compute the Discounted Payback Period. So, the discounted payback period is the number of periods in the future when the cumulative discounted cash flows total the initial investment.

EXERCISE

Find the discounted payback period for the project in the last example.

Advantages of the discounted payback period rule include:

- More likely to approve a highly liquid project
- Does not ignore the time value of money

The disadvantages of the discounted payback period rule are similar to the regular payback period rule, including:

- Ignores all the cash flows after the payback period
- What is the rule? When to accept/reject a project? How to rank multiple projects?

2.3 Internal Rate of Return (IRR)

A very popular measure of project performance is the IRR of the project.

Definition 2.3.1
IRR Rule

The ***internal rate of return*** of a project is defined to be the cost of capital such that the NPV of a project is zero.

The typical IRR rule states that given one project, accept it if IRR > required rate of return. This rate of return is usually the discount rate we would have used for NPV or possibly a target rate set by the management (also called the "hurdle rate"). Also, if given multiple projects, we rank them in order of IRR, the higher the better.

Let's do an example:

Example 3

Find the IRR of a project assuming the initial cost is \$25,000 and the cash flows are as follows:

Year	CashFlow
1	\$10,000
2	\$9,000
3	\$8,000

Solution: We need to find i such that

$$NPV = 0$$

Thus

$$0 = -\$25,000 + \frac{\$10,000}{(1+i)} + \frac{\$9,000}{(1+i)^2} + \frac{\$8,000}{(1+i)^3}.$$

We cannot solve this explicitly in general. Instead, we use numerical tools such as Excel in order to approximate the solution.

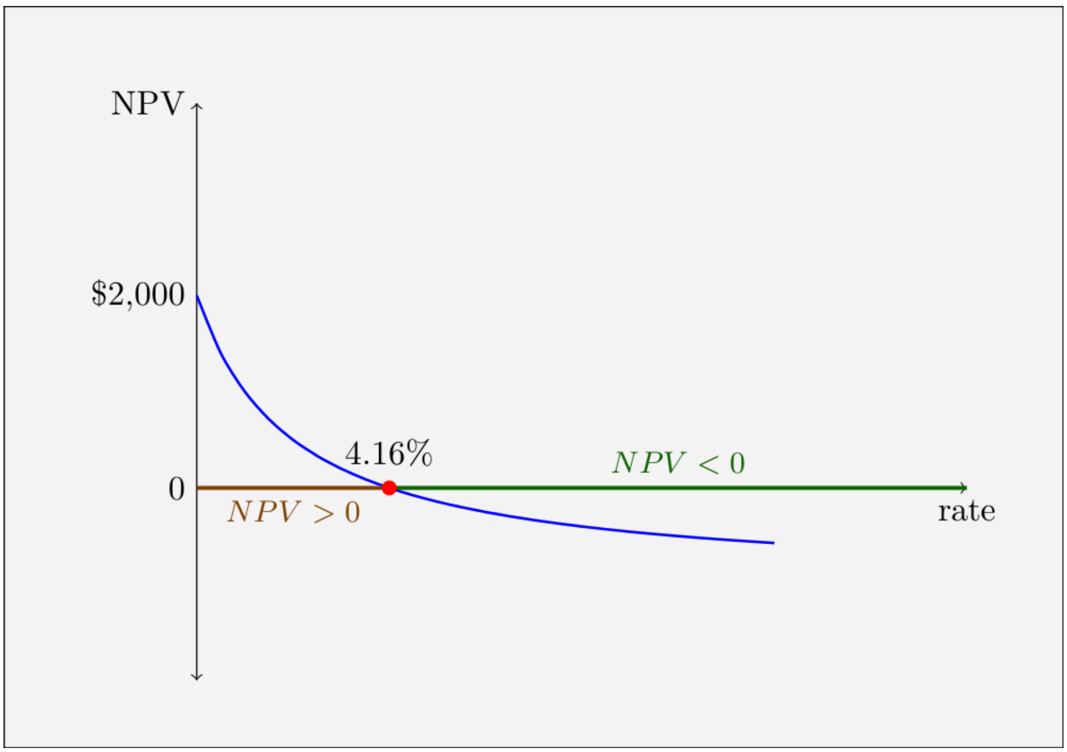
Using Excel, we find

$$i \simeq 4.106\%.$$

Therefore, the IRR for this project is 4.16%

The idea behind this result is that this project returns about 4.16%. If our cost of capital is more than this, we would reject this project, and if the cost of capital is below this, we would accept it.

Here is a graph of the NPV of the project as the cost of capital changes. From the graph, we can see the regions where the NPV is positive and where it is negative.



2.3.1 Problems with IRR

Although popular, the IRR methodology does have some problems. Consider the project having an initial cost of $10,000 with cash flows:

Year	Cash Flow
1	\$5,000
2	\$55,000
3	-\$55,000

Then we can show that the IRR is both 16.3% and 86.8%. What's happening?

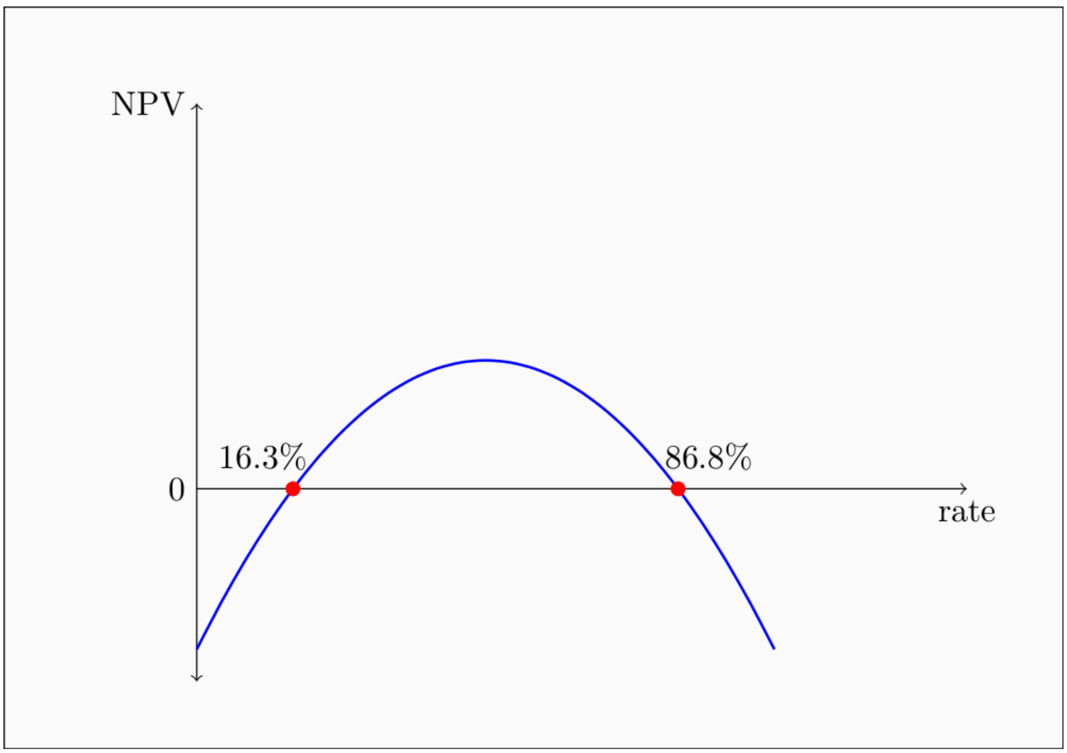
The NPV calculation boils down to solving the polynomial

$$0 = -C + CF_1 x + CF_2 x^2 + \cdots + CF_n x^n,$$

where $x = \frac{1}{1+i}$ and CF_i is the cash flow at time i.

Descartes' Rule of Signs states that the number of positive roots of a polynomial cannot exceed the number of sign changes of the coefficients. So, a typical project has terms like $-++++$ (where the first cash flow – the cost – is negative, and the future cash flows are all positive). Such projects have a unique IRR since the cash flow pattern has only 1 sign change.

However, the project in our last example has cash flows in the pattern $-++-$ (so the are 2 changes of sign). Consequently, the project has 2 IRRs. Essentially, the polynomial has multiple positive roots, as shown in the graph of NPV vs cost of capital below



Here is another example to highlight another problem with the IRR decision rule. This problem relates to the size and timing of the cash flows of different projects.

Example 4 Consider 2 projects, E (early cash flows) and L (late cash flows), with cash flows shown below. Assume both projects cost \$15,000. For each project, calculate (a) the IRR, (b) The NPV if rates are 3%, and (c) the NPV if rates are 8%.

Year	E	L
1	\$10,000	\$3,000
2	5,000	5,000
3	2,000	10,000

Solution: Using Excel (or similar numerical tool), we can find the IRR for each project and find

$$IRR_E = 8.69\% \quad \text{and} \quad IRR_L = 8.01.\%$$

Since Project E has the higher IRR, it looks like the better project.

But, at 3% we have

$$NPV_E = \$1,252 \quad \text{and} \quad NPV_L = \$1,777.$$

In this case, we see that Project L is better.

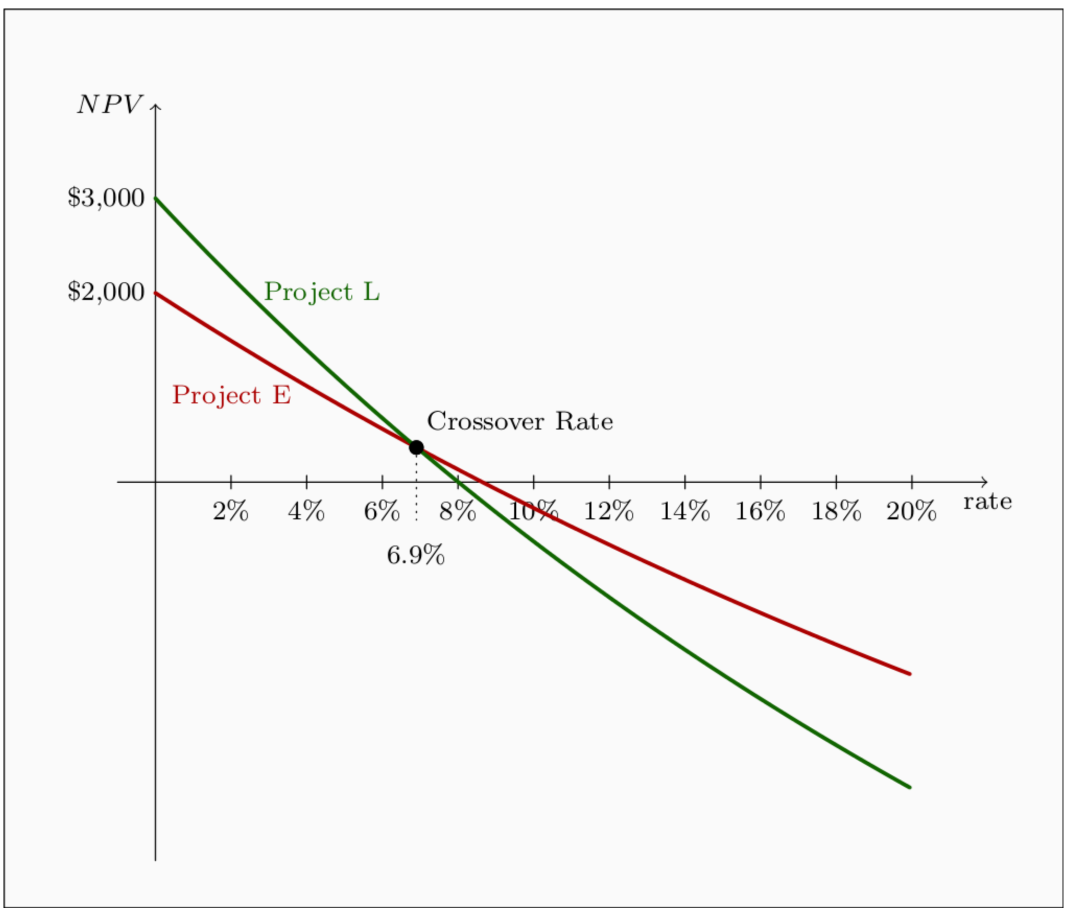
Finally, at 8% we have

$$NPV_E = \$134 \quad \text{and} \quad NPV_L = \$3.$$

In this case, Project E is better.

What's happening? Why does the decision change as the cost of capital change?

Here is a graph of the NPV of the projects as the cost of capital changes.



At a rate of 6.9%, both projects have the same NPV (called the ***crossover rate***). Above that rate, Project E is better, and below that rate Project L is better.

So we can see that the IRR rule does not always provide the best measure of which project we should accept.

A subtle problem with the IRR rule is that it assumes that the cash flows can be reinvested at the IRR, which is typlically not a good assumption. (There is no reason we should expect another project to be available with the same IRR that we can reinvest the cash flows in.) The NPV rule assumes cash flows can be reinvested at the cost of capital, which is generally a better assumption.

There are some other challenges as well. First, a definition.

Definition 2.3.2

Independent and Mutually Exclusive projects

Projects are called independent if selecting one does not affect our ability to accept others. Projects are called mutually exclusive if you can select only one.

Independent projects only need to be "good enough". Mutually exclusive projects must be ranked – we want only the best project.

So, how we we rank projects with the IRR rule? The obvious answer would be to chose the project with the largest IRR. However, consider two mutually exclusive 1-year projects (so you can chose only one of these). Project A has an IRR of 50% and Project B has an IRR of 20%. Project A clearly looks better. But suppose A involves an investment of only $1, while B involves an investment of $100. Project A then returns only 50 cents while Project B returns $20. So, the IRR rules ignores the scale of the project.

In summary, even though there are problems with the IRR, it is a popular rule. Some advantages of the IRR rule include:

- Easy to understand and communicate
- Usually results in the correct decision

Some disadvantages of the IRR rule inclulde:

- Cash flow pattern matters
  - Does not distinguish between investing and borrowing
  - IRR may not exist of there may be multiple IRRs
  - Problems with mutually exclusive investments (scale problem)
  - The timing of the cash flows can matter

In comparing NPV and IRR we see

- NPV assumes the cash flows can be reinvested at the discount rate. This is usually reasonable; it is like saying: you can invest in the market. Essentially, the discount rate is chose to be the rate that is "available" to you.
- The IRR assumes the cash flows can be reinvested at the IRR. This is sometimes reasonable; it is like saying: you can find another project as good as this one. But this is not always the case.
- NPV and IRR usually come to the same decision about a project, except odd cash flow patterns (many sign changes/borrowing vs. lending, and mutually exclusive projects with different scale or timing.)

2.4 Profitability Index

Another decision rule is called the Profitability Index.

The formula for Profitability Index (PI) is given by

Formula 1

$$PI = \frac{\text{PV of future cash flows}}{\text{Initial cost}} = \frac{\sum_{t=1}^{\infty} \frac{C_t}{(1+r)^t}}{-C_0}$$

Definition 2.4.1
The PI Rule

The PI rule states that given one project, accept it if PI > 1, and given multiple projects, rank the project in order of PI (the higher the better).

You should be able to convince yourself that a project has a $PI > 1$ exactly when $NPV > 0$. So, this rule seems equivalent to the NPV rule.

It in important to get a sense of what the PI is capturing. If a project has a $PI = 1.2$ what does this say? Effectively, it says that the PV of the future cash flows is 20% larger than the cost. So, loosely speaking, the project is generating a 20% excess return over its life.

Some advantages of the PI rule include:

- Somewhat easy to communicate
- Correct decision for independent projects

A disadvantage of the rule is:

- Scale problem. This is largely the same problem as with the IRR, and for largely the same reason. Both of these rules are essentially based on "rates", which necessarily ignore scale.

We will illustrate some of the issues with the PI now. Consider 3 projects A, B and C generating cash flows of C_t at time t. The cash flows, PI and NPV's of each project are summarized in the following table.

Projects	C_0	C_1	C_2	*PI @ 12%*	*NPV @ 12%*
A	-16	60	5	3.59	41.6
B	-20	70	10	3.52	50.5
C	-5	10	10	3.38	11.9

Note that Project A has a higher PI, but lower NPV than project B. This is an example of the scale problem since Project A has a smaller size than Project B.

EXERCISE

Using the above example, answer the following questions:

1. Rank the projects by PIs and by NPVs, are the rankings the same?
2. If you can only pick **one** project, which one would you pick?

Capital rationing: Suppose you only have \$20, which project(s) of the above would you choose? What about \$21?

3 Capital Budgeting

3.1 Net Incremental Cash Flows

A challenge when assessing a project is determining which cash flows should we include in the analysis. The correct cash flows to consider are the net incremental cash flows. That is, the cash flows that occur due to the acceptance of the project. Thinking about this another way, we can ask what is the difference between the cash flows of the entire firm with the project vs. without the project?

Nate that we are interested in **Cash flows**, not accounting income. Therefore, we exclude depreciation, goodwill, unearned capital gains, and other non-cash accounting charges.

Also, we consider **Incremental** cash flows only. Cash flows that change a direct consequence of accepting a project. Note that some costs are tax-deductible and thus have a different after-tax cash flow effect.

Definition 3.1.1 **Net Incremental Cash Flows**

Net Incremental Cash Flows can be thought of as net marginal cash flows that exist due to the acceptance of the project. We exclude **sunk costs** (a cost that occurred in the past), but we include **opportunity costs** (cash flows that would have happened if the project were not accepted) and **side effects** (synergy benefit and erosion costs). Remember that taxes and inflation matter when making decisions.

Estimating cash flows is the most difficult part of capital budgeting. It often starts with balance sheet, income statements, and other such sources. We often need to translate this from earnings to cash flows.

Typically we begin with the Operating Cash Flows (OCF) of the firm

- OCF = EBIT - Taxes + Depreciation
- EBIT = Earnings Before Interests and Taxes = Revenue - Costs - Depreciation
- Taxes = EBIT $\times$ Corporate Tax Rate = EBIT $\times T_c$

After some algebra we get

$$OCF = (\text{Revenue - Costs}) \times (1 - T_c) + \text{Depreciation} \times T_c$$

or

$$OCF = EBIT - \text{Taxes} + \text{Depreciation}$$

Since depreciation is an accounting concept to allocate the initial cost of the project over time, and not a cash cost, it has no effect on OCF. The project cash flows include initial cost when purchased (outflow), and salvage value when fixed assets are sold (inflow) at the end of the project life.

Example 1 Given these estimates based on a company's balance sheet/income statement

Revenues	Costs	Depreciation	Corporate Tax Rate
$1500	$700	$600	34%

Compute the EBIT, and OCF.

Solution: EBIT = $1500-$700-$600 =$200,

Taxes = $200×34% = $68

Thus, OCF = $732 which we can compute in 3 different ways.

Bottom-up approach:

$$\text{OCF} = \text{Net Income} + \text{Depreciation}$$

$$\text{Net Income} = \text{EBIT} - \text{Taxes} = 200 - 68 = \$132$$

$$\text{OCF} = 132 + 600 = \$732$$

Top-down approach:

$$\text{OCF} = \text{Rev - Costs - Taxes}$$

$$\text{OCF} = 1500 - 700 - 68 = \$732$$

Tax shield approach:

$$\text{OCF} = (\text{Rev - Costs}) \times (1 - T_c) + \text{Depre} \times T_c$$

$$\text{OCF} = 1500 - 700 - 68 = \$732$$

Typically (especially in this course) OCF is on an annual basis, and are thus the annual net incremental cash flows.

Definition 3.1.2
Capital Costs

The Capital Cost of a project is the initial investment in plant and equipment to fund the project.

REMARK

In this course, all capital costs occur the day the project is accepted. In real life, these costs may be spread over time.

3.2 Capital Cost Allowance

In Canada, depreciation (in the accounting sense) is not tax deductible. Instead, CCRA permits a deduction called the Capital Cost Allowance (CCA). CCA saves some of the tax costs, hence providing a “tax shield”. CCA is viewed as a “cost” for tax purposes, so it decreases the EBIT, hence reducing taxes.

Since this is not a tax course, we will not cover CCA in detail, but we will cover the basic idea.

Annual CCA amounts are computed on a declining balance basis.This is easiest to see with an example.

Example 2 Build the first 2 rows of the CCA table for an asset purchased for \$1 million that is in a pool with a CCA rate of 20% .

Solution: The table is below

Year	UCC Start	CCA Deduction	UCC End
1	1,000,000	200,000	800,000
2	800,000	160,000	640,000
3	640,000	128,000	512,000
4	512,000	102,400	409,600

Note that the annual deduction is always 20% of the Undepreciated Capital Cost (UCC) at the start of the year, and the UCC at the end of the year is the opening balance less the CCA deduction.

Now that we have the annual CCA deductions, we can determine the annual tax shield these deductions provide.

Example 3 Provide the annual CCA tax shield generated by the asset in the last example, assuming a tax rate of 20%.

Solution: We can create a table of tax savings by year as follows:

Year	CCA deduction	Tax Shield
1	200,000	40,000
2	160,000	32,000
3	128,000	25,600
4	102,400	20,480

Note that the tax shield amount is just 20% of the deduction value. So, these tax savings are essentially positive cash flows we can allocate to the project.

There are a few complications not covered in the last examples. First, once an asset is sold, there is no more CCA tax shield, and as such, the sale price gets deducted from the pool. Secondly, we have the so-called "half-year rule" that states that when you purchase an asset, only 50% of the purchase price can be added to the pool in the first year. The remaining 50% can be added in the second year. Finally, what we are really interested in is the PV of these tax savings, which we can get by discounting each annual amount, then adding all this up.

Using perpetuity formulas, we can derive the PV of the tax shield simplifying our work considerably. It is given by the PVCCATS formula:

Formula 1

$$PVCCATS = \frac{C \cdot d \cdot T_c}{r+d} \times \frac{1+0.5r}{1+r} - \frac{S \cdot d \cdot T_c}{r+d} \times \frac{1}{(1+r)^n}$$

where

- C = original price of the assets
- d = CCA rate that applies to the asset class
- T_c = corporate tax rate
- r = discount rate (k in the text)
- S = Salvage Value
- n = the period when assets are sold

Example 4 Consider the purchase of a delivery truck as specified as follows:

- Initial cost \$30,000, lasts for 5 years, then sold for \$1,500
- CCA class with 20% rate, 10% discount rate, 40% corporate tax

Compute the PVCCATS.

Solution: We have

$$PVCCATS = \frac{\$30,000 \times 20\% \times 40\%}{10\% + 20\%} \times \frac{1+0.5 \times 10\%}{1+10\%} - \frac{\$1,500 \times 20\% \times 40\%}{10\% + 20\%} \times \frac{1}{(1+10\%)^5}$$

Solving gives $PVCCATS = \$7,388$

3.3 Equivalent Annual Costs

There are cases where a direct NPV analysis misses the point a bit. Consider a case where a firm must have some type of safety equipment. Option A costs \$10,000 up front, plus \$1000 per year for 5-year, after which the equipment needs to be replaced. Option B costs \$8,000 up front plus \$1,400 per year but lasts 7 years.

A direct NPV analysis would show that both have negative NPV's and thus both are rejected. But that is not an option since we must by law have this equipment. Given that, we can chose the one with the highest NPV, even if it happens to be negative. But if Option A happens to have a higher NPV, we may be fooled into believing it is cheaper. But that may be wrong since it only last 5 years, while Option B lasts 7. We need a better approach.

Instead of NPV, we look at the **equivalent annual costs (EACs)**, which spreads the NPV to each year in a "financially equivalent" way. This can be found with the *EAC* formula.

Definition 3.3.1

Equivalent Annual Costs

Given an uneven stream of cash flows with a given NPV, the EAC of those cash flows is given by

$$NPV = EAC \times a_{\overline{n}|r} \Rightarrow EAC = \frac{NPV}{a_{\overline{n}|r}}$$

where n is the life of the project and r is our discount rate.

REMARK

- NPV is the PV of uneven cash flows
- EAC is the financially equivalent equal annual amount

Example 5

Compute the *EAC* of the two options A and B stated above assuming a discount rate of 10%.

Solution: Computing the NPV of the two options gives:

$$NPV_A = -\$10,000 - \$1,000 \times a_{\overline{5}|10\%} = \$13,790.79$$

$$NPV_B = -\$8,000 - \$1,400 \times a_{\overline{7}|10\%} = \$14,815.79$$

At the moment, A looks cheaper. Now we compute the equivalent annual costs of both.

$$EAC_A = \frac{\$13,790.79}{a_{\overline{5}|10\%}} = \$3,637.98$$

$$EAC_B = \frac{\$14,815.79}{a_{\overline{7}|10\%}} = \$3,043.24$$

So, option B is in fact cheaper.

REMARK

Note that if the NPV value is negative, the negative sign indicates that, mathematically, EACs are cash outflows. In practice, however, EACs are usually quoted as positive numbers.

3.4 Inflation and Capital Budgeting

The purchasing power of $1 changes through inflation. In **nominal terms**, cash flows are expressed in dollar values, while **real terms**, cash flows are expressed in purchasing power.

Effects of inflation must be included in capital budgeting, especially in long term projects. Real interest rates, nominal interest rates and inlfation rates are connect via the following formula.

Definition 3.4.1
Real Rate

The real interest rate is given by the formula

$$(1 + \text{Nominal Rate}) = (1 + \text{Real Rate}) \times (1 + \text{Inflation Rate})$$

which can be approximated by

$$\text{Real Rate} \cong \text{Nominal Rate} - \text{Inflation Rate}$$

when the rates are small.

Example 6

Inflation rate is 6%, nominal interest rate is 10%. Cash flows are $100 now in real terms at the end of each of the next 2 years. Calculate PV in nominal terms and real terms.

Solution:

In nominal terms:

$$C_1 = 100(1.06), C_2 = 100(1.06)^2$$

$$PV = \frac{C_1}{1.10} + \frac{C_2}{(1.10)^2} = 189.22$$

In real terms:

$$C_1 = C_2 = 100, i_{real} = \frac{1.10}{1.06} - 1 = 3.78\%$$

$$PV = \frac{C_1}{1.0378} + \frac{C_2}{(1.0378)^2} = 189.22$$

When dealing with inflation, **consistency** is key. Discount real cash flows at real rates. Discount nominal cash flows at nominal rates. You can convert cash flows to match discount rates or convert discount rates to match cash flows. You must pick either real or nominal by convention, but not both. DO NOT mix and match.

Note that CCA tax shields must be calculated in nominal terms, since once the asset is purchased, the entire CCA table is determined regardless of what future inflation might be.

3.5 Risk Analysis

Contrary to our previous examples, in real life, future cash flows are uncertain. There are several methods to evaluate projects with uncertain cash flows

1. Decision Trees
2. Scenario Analysis
3. Sensitivity Analysis
4. Breakeven Analysis
5. Monte Carlo Simulation
6. Valuation of Real Options

Uncertainty assessment is sometimes known as **risk analysis**.

Let's start with a simple example.

Example 7 MathSoc is considering a project to install solar panels on the MC roof.

- Initial cost = \$1 million
- Annual operating costs = \$10K
- Panels last 15 years, after which will be sold for scrap at \$100K
- Each sunny hour can generate \$100 revenue
- Based on weather forecasts, 1,500 sunny hours/yr is expected

Should the solar panel project be installed? Assume $r = 10\%$ and ignore taxes.

Solution: Let's compute the NPV.

$$\text{Inital cost} = -\$1\text{million}$$
$$PV(Revenues) = 1500hrs/year \times \$100/hr \times a_{\overline{15}|10\%} = \$1,140,912$$
$$PV(Costs) = -10K/year \times a_{\overline{15}|10\%} = -\$76,061$$
$$PV(Salvage) = \frac{\$100K}{(1+10\%)^{15}} = \$23,939$$
$$\text{Project NPV} = -\$1 \text{ million} + \$1,140,912 - \$76,061 + \$23,939 = \$88,790$$

So the project is a **go** according to the NPV rule.

What can go wrong with this project? Perhaps a failed panel installation? Not as sunny as we expected? Risk analysis allows us to dig deeper into the project analysis.

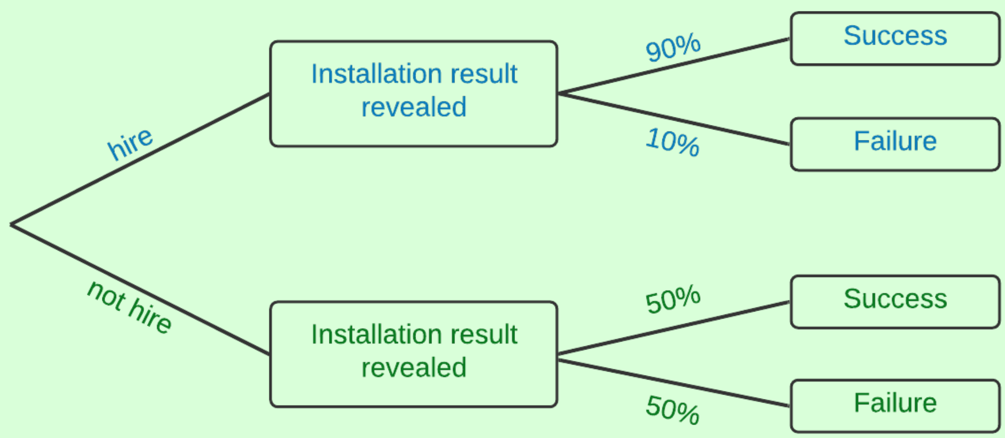
We can use decision trees, which are graphical representations of decision and information, in order to help identify the best course of action.

Example 8

Revisiting the solar panel example, suppose for simplicity that project is abandon (with an NPV of zero) if the installation fails. Consider two options

- Hire experienced technician for \$10K, who has a 90% chance of installing successfully (blue)
- Student intern costs \$0, but only installs successfully 50% of the time (green)

The flow or action/information can be represented by a decision tree



Compare the **expected NPVs** of the two options.

Hire technician:

If success: $NPV = \$88,790 - \$10,000 = \$78,790$

If failure: $NPV = -\$10,000$

$$\text{Expected NPV} = 0.9 \times \$78,790 + 0.1 \times (-\$10,000) = \$69,911$$

Don't hire technician:

If success: $NPV = \$88,790$

If failure: $NPV = \$0$

$$\text{Expected NPV} = 0.5 \times \$88,790 + 0.5 \times 0 = \$44,395$$

$\$69,911 > \$44,395$, the experienced technician should be hired.

Example 9

Suppose the panels are successfully installed. What if the weather is not as sunny as we expected?
Consider 3 scenarios:

- Cloudy: 1100 hours of sun/year

- Average: 1500 hours of sun/year
- Sunny: 1700 hours of sun/year

What are the NPVs of the project in these scenarios?

Initial Cost	Annual Cost	Project Life	Salvage	PV(Costs)
$1 mil	$10K/year	15Years	$100K	-$1,052,122

	Cloudy	Average	Sunny
Sunny hours	1100/year	1500/year	1700/year
Annual Revenue	$110,000/year	$150,000/year	$170,000/year
PV (Revenue)	$836,669	$1,140,912	$1,293,034
NPV	$215,453	$88,790	$240,912

Another popular item in the risk analysis tool kit is Break-Even analysis. There are different types of B/E analyses, but we focus on NPV break-even. The idea is we set up NPV equation ($NPV = 0$), identify the variable of interest, and solve for unknown. This determines the necessary input level to reach a break-even point.

A useful quantity is the Contribution Margin, which revels how much 1 unit of product sales contributes to net income. More formally

Definition 3.5.1
Contribution Margin

The contribution margin is increase in net income per unit increase in sales (usually after-tax). Defined by

$$(\text{Unit Revenue} - \text{Unit Variable Cost}) \times (1 - T_c)$$

Here is a simple example of how it can be applied.

Example 10

A donut sells for $1.20 at Tim Hortons. Assume each donut costs $0.70 to make. Compute the contribution margin. If each Tim Horton location incurs monthly fixed costs of $100,000, how many donuts do they need to sell to break even? (Ignore taxes.)

Solution: The contribution margin is given by

$$\$1.20 - \$0.70 = \$0.50$$

So, each unit contributes $0.50 to the bottom line.

If fixed costs are $100,000, we need to sell enough product to cover these costs. Thus the break-even equals

$$B/E = \frac{\$100,000}{\$0.50} = 200,000$$

So, they need to sell 200,000 donuts per month.

REMARK

In the last example, the fixed costs were all equal in each period, making the break-even analysis easy. In the more realistic cases that the costs are not equal in each period, we need to do a more thorough NPV analysis.

In the examples above, there were a small number of possible outcomes we wanted to consider so it was possible to analyze them individually. In some cases, there are too many future scenarios to work this way. A method to deal with this is called Monte Carlo simulation where we simulate a sample of these possible scenarios. This technique is explored in greater detail in courses such as STAT 340 or STAT 341.

Finally, projects may have options to expand or abandon that have additional value. We will explore this near the end of the course.

4 Return and Risk

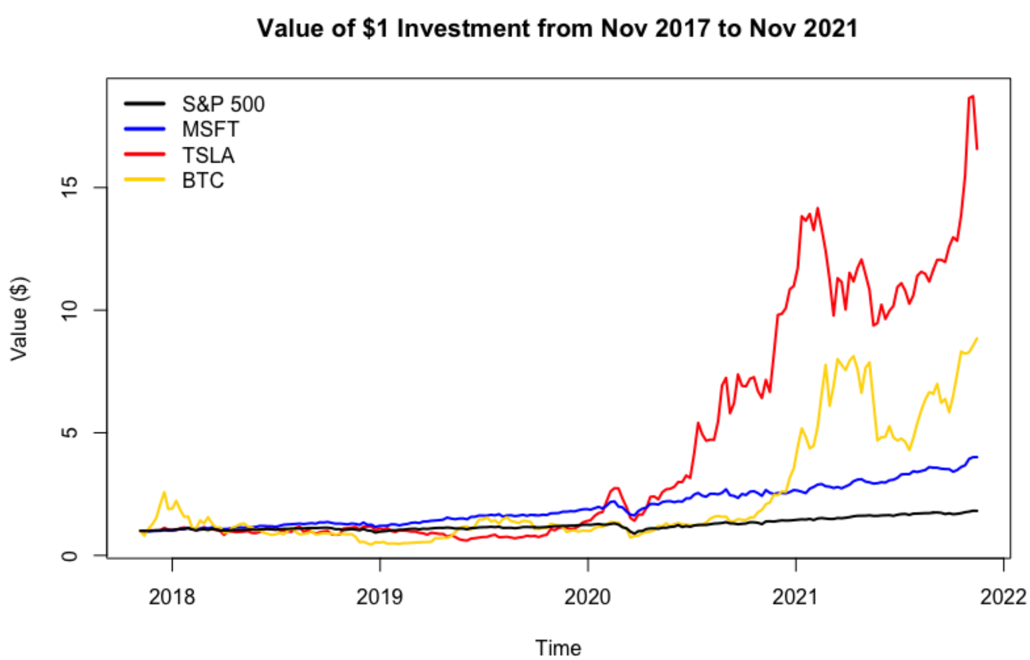
4.1 Return and Risk Measures

The financial market is more than just risk-free borrowing and lending. We can invest in the financial market in instruments such as: cash, stocks, exchange traded funds (ETFs), bonds, options, commodities, real estate and so forth.

Each type of investment has a different risk-return trade off, but in general

$$\text{High risk} \implies \text{high return}$$

$$\text{Low risk} \implies \text{low return}$$



We need to be a bit more precise about what we mean by "returns" on an investment. We may be interested in dollar returns or percentage returns.

Definition 4.1.1
Returns

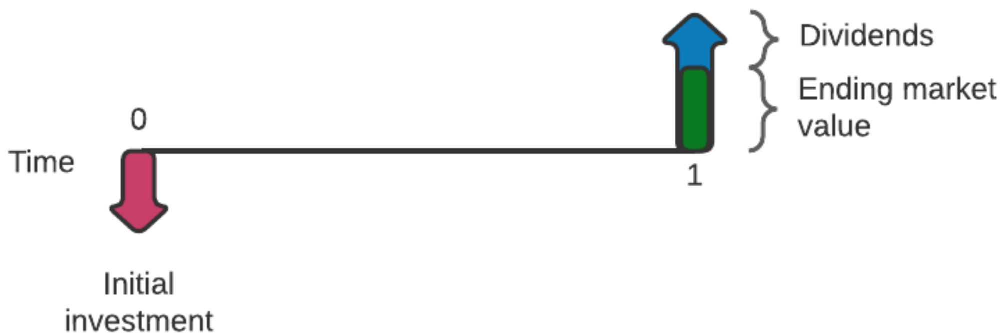
The dollar returns is given by

$$\begin{aligned}\$\text{Return} &= \text{Div}_{t+1} + (P_{t+1} - P_t) \\ &= \text{Dividend} + \text{Change in market value}\end{aligned}$$

Where Div_{t+1} are the dividends received over the holding period, and P_t is the price, or value, of the asset at time t. Note that the change in market value is also called the capital gain (or loss) over the holding period.

The percentage return is given by

$$\begin{aligned}\%\text{Return} &= \frac{\$\text{Return}}{\text{Initial price}} \\ &= \frac{\text{Div}_{t+1}}{\text{Initial price}} + \frac{(P_{t+1} - P_t)}{\text{Initial price}} \\ &= \text{Dividend yield} + \text{Capital gains yeild}\end{aligned}$$



We generally wish to analyze returns of assets over many periods and compare them with other investments. It turns out there are several ways to do this. First we define the average (or mean) return.

Formula 1

Let R_i be the %-return over the i-th period, for i =1, 2, ..., n. The mean return (or average return) is given by the formula

$$\mu = \frac{1}{n}\sum_{i=1}^{n} R_i$$

Alternatively, we may be interested in the holding period return on on assets over several periods. While the average return is an arithmetic mean, the holding period return is a geometric mean of the period returns. This is easiest to see with an example.

Example 1 \$1,000 is invested for 3 years. The annual returns are 10%, 5%, and 8%.

(a) What is the total holding period return over the three years?

Solution:

$$R_{13} = (1 + 10\%)(1 + 5\%)(1 + 8\%) - 1 = 24.7\%$$

(b) What is the annualized return over the three years?

Solution:

$$\bar{R}_{13} = \sqrt[3]{(1 + 10\%)(1 + 5\%)(1 + 8\%)} - 1 = 7.65\%$$

Note that the mean return equals

$$\frac{10\% + 5\% + 8\%}{3} = 7.67\%$$

which is slightly higher. In fact, the arithmetic mean is always higher.

REMARK

It will be important to know which mean to use when. In general, if we want to know what the equivalent average return a given investor earned in an asset over an extended period, we use the holding period return (i.e. the geometric average)

If however, we are asking what is the most likely return a random investor would have realized if they invested in the asset for 1 random period, that would be the average return. So, the average return is more of a statistical measure.

Risk measures are much harder to define. We will primarily use the variance of returns, or essentially equivalently, we use the standard deviation (more commonly called the "volatility").

Formula 2 Let R_i be the %-return of an asset A over the i-th period, for i =1, 2, ..., n and let μ be the mean return. The variance of the returns over that period is given by

$$\sigma_A^2 = Var(R_A) = \frac{1}{n}\sum_{i=1}^{n}(R_i - \mu_A)^2$$

and the volatility is given by

$$\sigma_A = \sqrt{Var(R_A)}.$$

Note that these are just the usual formulas for the variance and standard deviation for the random variable R_A.

While variance is probably the most common risk measure, it has its significant drawbacks such as symmetric penalty for unexpected losses and profits, and risk measurement around the centre of the distribution. Although we will not cover other risk metrics in this course, we include a few others for interests sake.

Formula 3

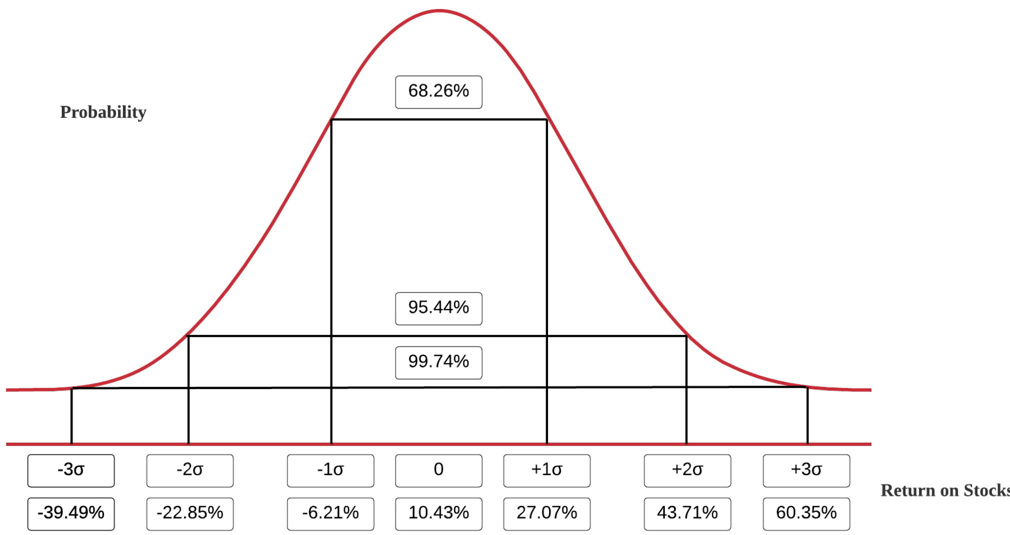
Let R be a return random variable (assumed continuous). We define

Downside Semi-Variance: $\sigma_-^2 = E[min(0, R - E[R])^2]$
Value-at-Risk (quantile of the return distribution): $VaR_\alpha = \sup\{x|\Pr(R \leq x) \leq 1 - \alpha\}$
Conditional Value-at-Risk (conditional tail expectation): $CVaR_\alpha = E[R|R \leq VaR_\alpha]$

Note: Some sign adjustments may be needed for a loss random variable. You will need to note the difference between "$\leq$" versus "$<$" if the random variable is discrete.

4.2 Normally Distributed Returns

Normally distributed asset returns should mostly be around its means.



If we assume assets returns are normally distributed, then the asset returns should centered mostly around its means.

Indeed **IF** returns (R_i's) are normally distributed, then μ and σ fully characterize the return distribution. The normality assumption makes mathematical analysis much easier.

Example 2

Assume a stock has a mean return of 10% and a volatility of 15%. What is the probability that if we buy this stock, we will lose money over the next year.

Solution: Let R be our return on the stock. We are asking to find $P(R < 0)$. If we assume R is normally distributed, we can use the usual techniques:

$$\begin{aligned} P(R < 0) &= P\left(\frac{R-\mu}{\sigma} < \frac{0-\mu}{\sigma}\right) \\ &= P\left(Z < \frac{0 - 10\%}{15\%}\right) \\ &= P(Z < -0.6667) \\ &\sim 25\% \end{aligned}$$

where Z is a standard unit normal variable.

In practice, “risk” can depend on higher moments such as skewness, kurtosis, etc. If R_i is normal, then all moments can be calculated using μ and σ, and VaR and CVaR can be calculated using μ and σ alone.

HOWEVER stock returns are NOT normally distributed. A heavier tail indicates that extreme returns are likely, while asymmetric suggest that stock prices are skewed (can be right or left). This is a serious problem for risk management.

5 Capital Asset Pricing Model

5.1 Portfolio Optimization with Two Risky Assets

In general, investors do not buy only one asset, instead they buy multiple assets and combine them into a portfolio. We will be interested in computing the portfolio statistics.

We begin with 2 assets A and B, which we form into a portfolio

$$P = tA + (1-t)B.$$

Here t represents the weight invested in asset A with the residual $(1-t)$ invested in B. Let R_A and R_B be the returns of the assets, which we think of as random variables. Also we let μ_A and μ_B be the mean returns (also denoted $E(R_A)$ and $E(R_B)$ which makes the connection to expected values more clear.)

Since the mean (or expected value) of a linear combination of random variables equals the linear combination of the means, we get

$$\mu_P = t\mu_A + (1-t)\mu_B$$

or equivalently

$$E(R_P) = tE(R_A) + (1-t)E(R_B).$$

The volatility is more complicated. It is not true that the variance of a linear combination equals the linear combination of the variances, instead we get

$$Var(R_P) = t^2 Var(R_A) + (1-t)^2 Var(R_B) + 2t(1-t)Cov(R_A, R_B)$$

where $Cov(R_A, R_B)$ is the covariance of the returns. Using the formula for the correlation between random variables

$$\rho_{AB} = \frac{Cov(R_A, R_B)}{\sigma_A \sigma_B}$$

and substituting this into the above equation, we can re-write things a bit more compactly as

Formula 1 The variance of a 2-asset portfolio is given by

$$\sigma_P^2 = t^2\sigma_A^2 + (1-t)^2\sigma_B^2 + 2t(1-t)\rho_{AB}\sigma_A\sigma_B$$

This is an extremely useful formula we will make use of repeatedly.

Example 1 Consider 2 assets, A and B. Assume the mean returns are given by $E(R_A) = 5\%$ and $E(R_B) = 12\%$. Also assume the asset volatilities are given by $\sigma_A = 10\%$ and $\sigma_B = 20\%$ and that the returns are independent. Find the portfolio statistics for the portfolio consisting of 75% invested in asset A and then remainder in B.

Solution: The portfolio return is given by

$$E(R_P) = 75\% \times 5\% + 25\% \times 12\% = 6.75\%.$$

Since the returns are independent, we have the correlation of the returns, $\rho_{AB} = 0$. Thus, the portfolio variance is given by

$$\begin{aligned} Var(R_P) = \sigma_P^2 &= 75\%^2 \times 10\%^2 + 25\%^2 \times 20\%^2 \\ &= 0.0813 \end{aligned}$$

Thus $\sigma_P = 9.014\%$.

REMARK

This example motivates everything that follows. Note that the portfolio return is higher than the asset A returns, but has a lower risk (i.e. volatility) than either asset. In this sense, the portfolio is clearly superior to asset A.

We wish to explore what happens to the volatility of the portfolio as we change the weights invested in the assets.

Definition 5.1.1 **Feasible Set** The feasibility set, also referred to as an opportunity set, is the set of portfolios that can be constructed from a given set of assets.

Definition 5.1.2 **Efficient Set** The efficient set, also referred to as the efficient frontier, is the set of optimal portfolios. A portfolio is called optimal if no other feasible portfolio can have both high return and lower risk.

In the example above, asset A is not optimal since the portfolio has both lower risk and higher return. It is not yet clear if the portfolio itself is optimal.

With only 2 risky assets, the feasibility set and the efficient set coincide. This is not true in general and definitely not true in practice!

5.2 Efficient Frontier

Let's now turn our attention to multi-asset portfolios. We define the following

- Random return vector: $R = [R_1, ..., R_N]^T$
- Mean return vector: $\mu = [\mu_1, ..., \mu_N]^T$
- Variance-Covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1N}\sigma_1\sigma_N \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2N}\sigma_2\sigma_N \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N1}\sigma_N\sigma_1 & \rho_{N2}\sigma_N\sigma_2 & \cdots & \sigma_N^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{bmatrix}$$

  where

  - $\sigma_{ij} = Cov[R_i, R_j] = \rho_{ij}\sigma_i\sigma_j = \sigma_{ji}$
  - $\sigma_{ii} = Cov[R_i, R_i] = Var[R_i] = \sigma_i^2$
  - Σ is $N \times N$, symmetric and positive definite
- Portfolio weight vector: $w = [w_1, ..., w_N]^T$
- Vector of ones for convenience: $e = [1, ..., 1]^T$
- Budget constraint (sum of weights equal to 1): $\sum_{i=1}^N w_i = e^T w = 1$

For an asset pool μ, Σ and fixed portfolio weight w, what is the portfolio mean and variance?

Formula 2 The portfolio return is given by

$$\mu_P = \sum_{i=1}^N w_i\mu_i = w^T\mu$$

and the portfolio variance is given by

$$\sigma_P^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} = w^T \Sigma w$$

A key fact is that diversification reduces portfolio risk.

Formula 3 The portfolio volatility satisfies

$$\sigma_P = \sqrt{w^T\Sigma w} \leq \sum_{i=1}^N |w_i| \cdot \sigma_i$$

Thus, the portfolio volatility is less than or equal to the weighted average of the asset volatilities.

The Markovitz's mean-variance question is then: How do we construct optimal portfolio(s) w^* for a given set of N risky assets?

The following important assumptions are made:

- Single period model
- Assets are perfectly divisible
- No transaction costs and no taxes
- We wish to maximize return, minimize variance, or both
- Mean returns μ and covariance Σ are known
- The variance-covariance matrix Σ is positive definite
- Not all expected returns are equal (μ is not a multiple of e)

Note that maximizing return μ_p and/or minimizing risk σ_P^2 is essentially the same as trying to find efficient (i.e. optimal) portfolios.

Maximizing portfolio return for given risk tolerance, σ_{max}^2, is equivalent to the following optimization problem:

Formula 4

$$\max_w w^T\mu \text{ s.t. } w^T\Sigma w = \sigma_{max}^2 \text{ and } e^T w = 1$$

Minimizing portfolio risk for given level of return, μ_{min} is equivalent to this optimization problem:

Formula 5

$$\min_w w^T\Sigma w \text{ s.t. } w^T\mu = \mu_{min} \text{ and } e^T w = 1$$

To maximize risk-adjusted return for given trade-off parameter, τ, we have this optimization problem

Formula 6

$$\max_w \tau \cdot w^T\mu - \frac{1}{2} \cdot w^T\Sigma w \text{ s.t. } e^T w = 1$$

We will use the method of Lagrange multiplier to solve for this last optimization problem. More specifically, we want to solve for the optimal portfolio w_{opt} in

$$\max_w \tau \cdot w^T\mu - \frac{1}{2} w^T\Sigma w \quad \text{s.t.} \quad e^T w = 1.$$

The Lagrangian function is defined as

$$\mathcal{L}(w,\lambda) = \tau \cdot w^T\mu - \frac{1}{2}w^T\Sigma w - \lambda(e^T w - 1)$$

Taking partial derivatives, the first optimality condition is

$$\nabla_w \mathcal{L}(w,\lambda) = \tau\mu - \Sigma w - \lambda e = 0$$

Solving for w gives

$$w = \tau\Sigma^{-1}\mu - \lambda\Sigma^{-1}e.$$

The second optimality condition is

$$\nabla_\lambda \mathcal{L}(w,\lambda) = -(e^T w - 1) = 0 \implies e^T w = 1.$$

Substituting in w from above gives

$$\tau e^T\Sigma^{-1}\mu - \lambda e^T\Sigma^{-1}e = 1$$

Solving for λ gives

$$\lambda = \tau\frac{e^T\Sigma^{-1}\mu}{e^T\Sigma^{-1}e} - \frac{1}{e^T\Sigma^{-1}e}.$$

Substituting λ into the formula for w gives

$$\begin{aligned}
w_{opt} &= \tau\Sigma^{-1}\mu - \left(\tau\frac{e^T\Sigma^{-1}\mu}{e^T\Sigma^{-1}e} - \frac{1}{e^T\Sigma^{-1}e}\right)\Sigma^{-1}e \\
&= \frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e} + \tau\left(\Sigma^{-1}\mu - \frac{e^T\Sigma^{-1}\mu}{e^T\Sigma^{-1}e}\Sigma^{-1}e\right) \\
&= w_m + \tau w_z
\end{aligned}$$

where

$$w_m = \frac{\Sigma^{-1}e}{e^T\Sigma^{-1}e} \quad \text{and} \quad w_z = \Sigma^{-1}\mu - \frac{e^T\Sigma^{-1}\mu}{e^T\Sigma^{-1}e}\Sigma^{-1}e$$

By setting $\tau = 0$, the optimization problem reduces to finding the minimum variance portfolio, and thus w_m is in fact the minimum variance portfolio.

A quick calculation reveals that $e^T w_z = 0$. This is sometimes stated that the portfolio w_z is "self financing". This takes us to

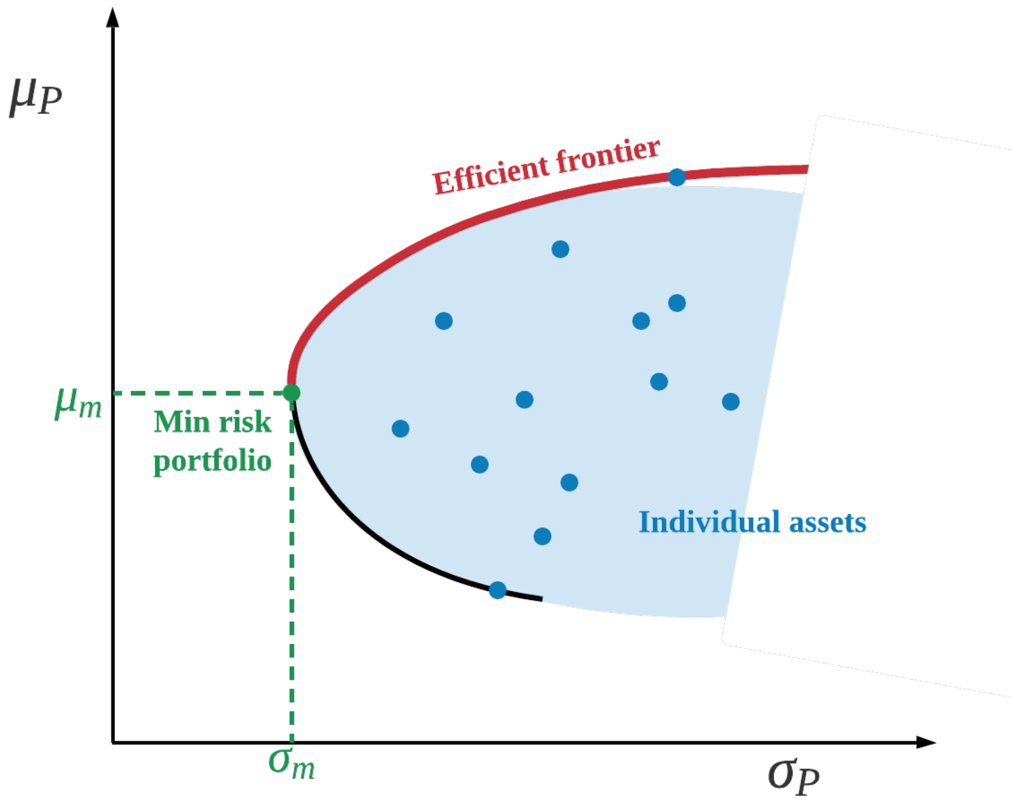
Theorem 7 **(Two-Fund Theorem)**

The formula

$$w_{opt} = w_m + \tau w_z$$

states that any efficient portfolio can be replicated by two portfolios, w_m and w_z where w_m is the minimum variance portfolio, and w_z is self financing. These two portfolios can "generate" the whole efficient frontier.

Furthermore, if we let μ_{opt} be the return on an optimal portfolio, then it can be shown that the set of efficient portfolios is a hyperbola in (σ, μ)-space.



In summary, an investor seeking efficient portfolios needs only invest in combinations of two portfolios that are known to be efficient. This remains true under all Markovitz's assumptions.

From the individual investor's point of view, if he/she believes in Markovitz's solution, they need just two efficient portfolios to invest optimally. From the investment banks perspective, if all investors believe in the same efficient frontier, then two efficient products can provide a complete investment service.

5.3 Inclusion of a Risk-Free Asset

The assumption that Σ is positive definite means no security is risk-free. If there was a risk-free asset, Σ would have one row/column of zeros.

Now we will consider allowing the investment in risk-free asset along with N risky securities in a portfolio w, and a fraction w_0 of wealth in a risk-free asset with return r_f.

If we let w_0 be the fraction of wealth is invested in the risk-free asset r_f, we are lead to the following optimization problem

$$\max_{w} \tau(\mu^T w + w_0 r_f) - \frac{1}{2} \cdot w^T \Sigma w \quad \text{s.t. } e^T w + w_0 = 1$$

This can be solved using Lagrange multipliers similar to as before.

However, there is an alternative more graphical way to attack this problem.

Say we invest w_0 in risk-free asset (r_f) and $1 - w_0$ in risky asset Q. We have

$$\mu_P = (1 - w_0)\mu_Q + w_0 r_f \text{ and } \sigma_P = (1 - w_0)\sigma_Q$$

Substitute $w_0 = 1 - \frac{\sigma_P}{\sigma_Q}$ into the first equation, we have

$$\mu_P = r_f + \frac{\mu_Q - r_f}{\sigma_Q}\sigma_P$$

which is a straight line with intercept r_f and slope $\frac{\mu_Q - r_f}{\sigma_Q}$, known as the **Capital Allocation Line (CAL)**. Different risky assets Q result in in CALs with different slopes.

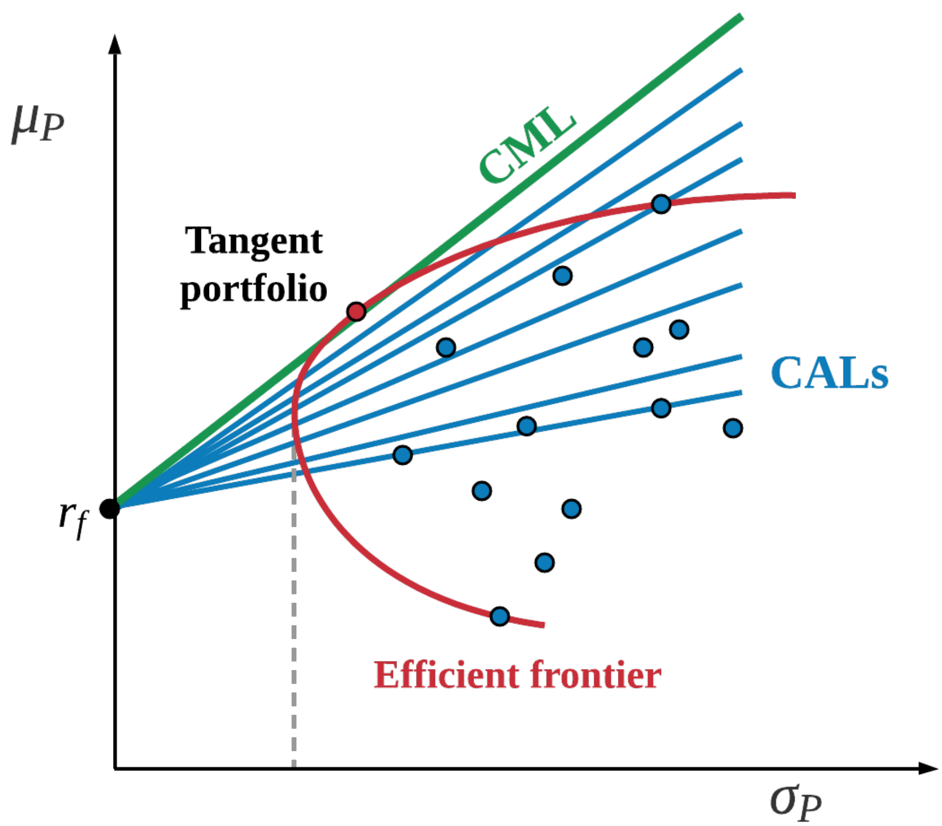
Definition 5.3.1
Sharpe Ratio

The ratio

$$\frac{\mu_Q - r_f}{\sigma_Q}$$

is called the Sharpe Ratio of Q. It capture the excess return per unit risk of the portfolio

A higher slope (or equivalently, a higher Sharpe Ratio) is better as it indicates a greater excess return per unit risk. The steepest CAL is referred to as the **Capital Market Line (CML)**.

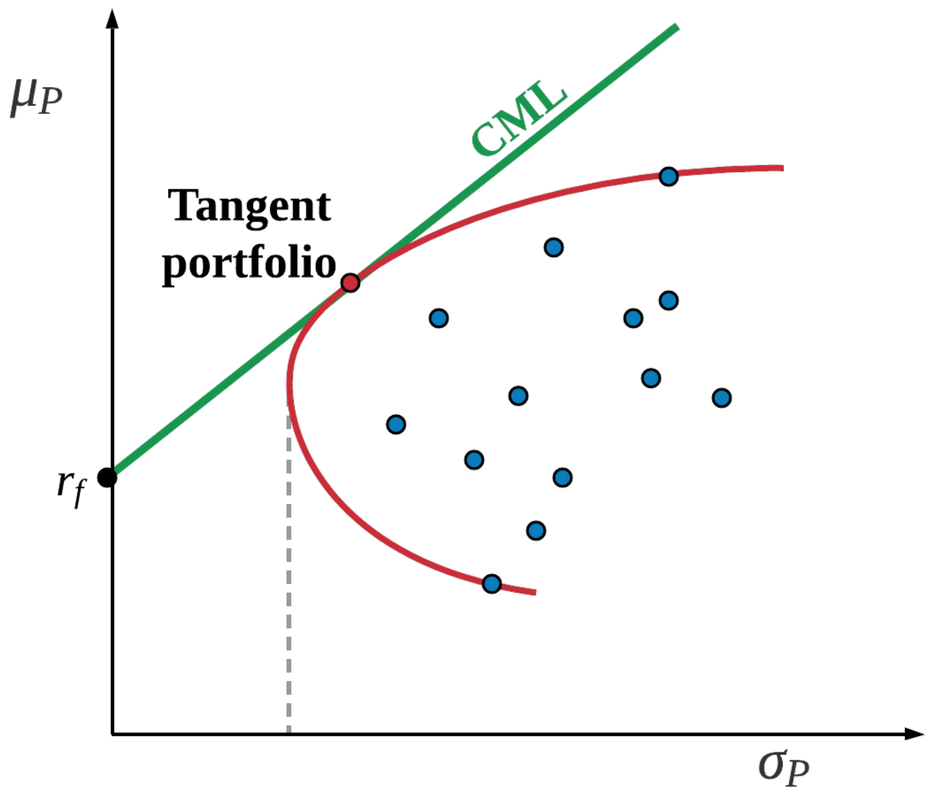


From the graph, we can see that the tangent portfolio (called the Market Portfolio) is the unique optimal portfolio of risky assets. Taking combinations of this Market Portfolio with the risk free asset gives a new efficient frontier. This is summarized as

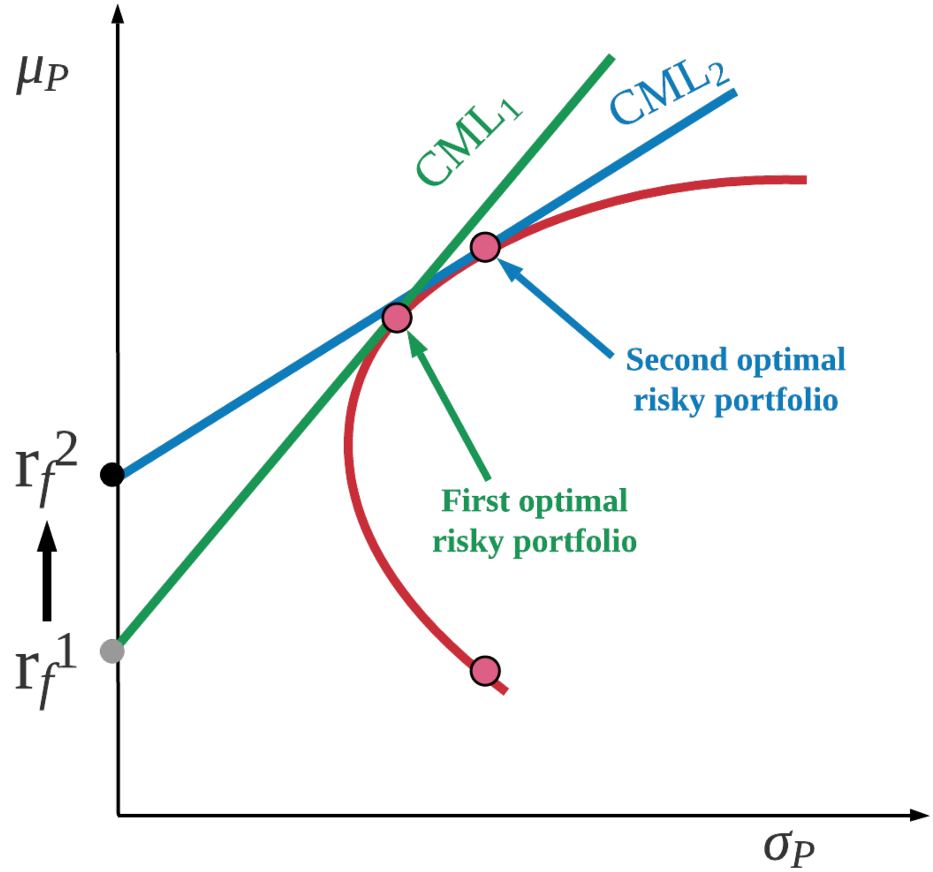
Theorem 8 **(One-Fund Theorem)**

There is a unique optimal risky portfolio M such that any efficient portfolio can be constructed as a combination of M and the risk-free asset r_f. The new efficient frontier is a line, the called the Capital Market Line (CML).

A "smart" investor should hold a portfolio on the CML. The investors risk tolerance is then reflected in the weight w_0.



Note that the tangent portfolio changes with r_f. If r_f increases, then σ_M and μ_M increases and the slope of CML decreases.



In summary portfolio selection can be separated into two stages:

1. First, determine the tangent portfolio M
2. Second. decide the optimal mix of tangent portfolio and the risk-free asset based on the investors risk-preferences

This is sometimes called the "separation principle" since the determination of the optimal risky portfolio, M, is separated from the preferences of the individual.

REMARK

Regarding the Separation Principle:

- Homogeneous expectations is a key assumption, but this is almost always false
- It also assumes all investors care about expected return and variance, which is also almost always false. Some investors may care about skewness, kurtosis, variance in losses, etc.
- The CML depends on the risk-free rate

EXERCISE

Determine if the portfolio in Example 1 of this Chapter is efficient.

5.4 Capital Asset Pricing Model

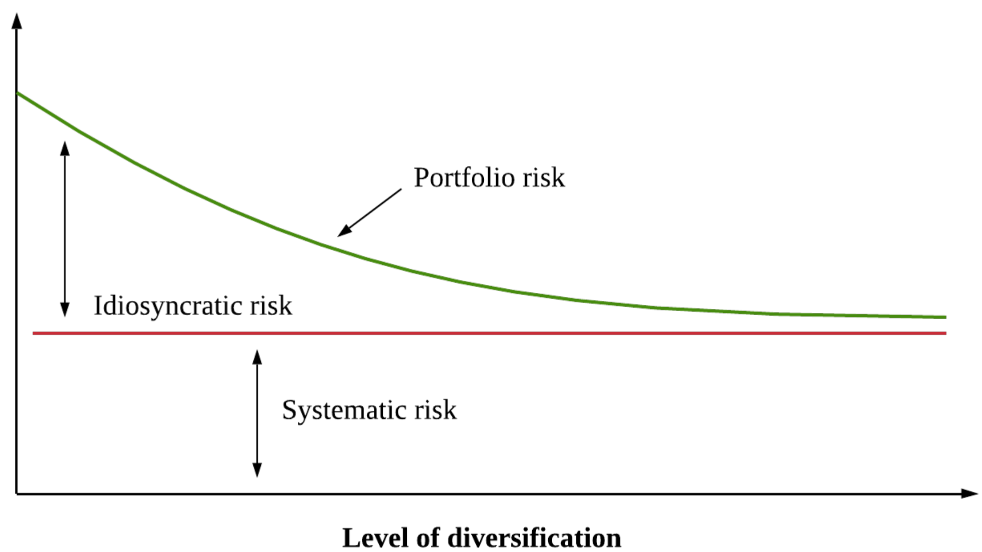
By adding securities, we are able to reduce the risk of the minimum variance portfolio. This is essentially the diversification effect. However, in general, the risk cannot be reduced to zero – some risk remains. This remaining risk, called "systematic risk", cannot be diversified away by adding more securities.

In fact, the risk of a security comes from two sources:

- Systematic risk (a.k.a. market risk, non-diversifiable risk)
- Non-systematic risk (a.k.a. specific risk, diversifiable risk, idiosyncratic risk)

Since all risky assets contain systematic risk, adding more of them to our portfolio will not diversify that risk away, we can only diversity away the non-systematic risks.

Therefore, in a well-diversified portfolio, only systematic risk matters. A key consequence of this is that investors should only expect to be compensated for systematic risk. Investors should not expect to receive excess return for assuming non-systematic risk since that can be diversified away.



Given a security A, how to we quantify the systematic risk that it contains? The answer lies in its correlation to the market portfolio.

Definition 5.4.1
Beta

Given a security A and the market portfolio M, we define the security's beta, β_A by

$$\beta_A = \frac{Cov[R_A, R_M]}{Var[R_M]} = \frac{\sigma_{AM}}{\sigma_M^2}$$

β_A is then our measure of the systematic risk of the security A. This then leads to the celebrated CAPM formula.

Theorem 9

CAPM The expected returns of an asset A and the market portfolio M are connected via the formula

$$\mu_A = r_f + \beta_A(\mu_M - r_f)$$

where r_f is the risk-free rate.

Definition 5.4.2
Risk Premium

The value $\mu_M - r_f$ is called the market risk premium and the quantity $\beta_A(\mu_M - r_f)$ is called the risk premium of asset A.

CAPM is based on the following assumptions:

- All investors are rational mean–variance optimizers
- Investors' planning horizon is a single period
- Investors use identical input lists (i.e. have the same expectations)
- All assets are publicly traded; short positions are allowed (i.e. can short lower expected return securities), investors can lend/borrow at risk–free rate
- All information is publicly available
- No taxes or transaction costs

REMARK

Note that CAPM says that the total risk of an asset A, σ_A, is in fact irrelevant in a diversified portfolio. Since we are able to diversify away the non-systematic risk, only the systematic risk, captured by β_A, matters.

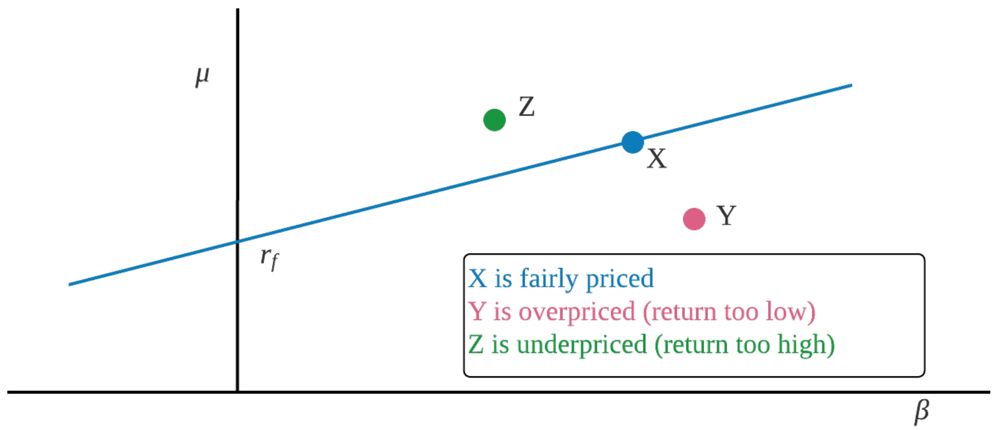
5.5 The Security Market Line (SML)

The CAPM formula

$$\mu_A = r_f + \beta_A(\mu_M - r_f)$$

generates a line in the (β, μ)-plane called the Security Market Line (SML). The SML is an upward straight line with intercept at r_f and the slope is given by $\mu_M - r_f$.

In theory, all securities lie on the SML. By plotting actual securities and comparing them with the SML, we can determine if a given security is correctly-priced.



This observation leads to a performance metric for securities.

Definition 5.5.1
Traynor Ratio

The quantity

$$\frac{\mu_A - r_f}{\beta_A}$$

is called the Traynor ratio of security A.

The Traynor ratio measures the excess return of a security, for each unit of systematic risk, as given by beta. All securities on the SML have the same Traynor ratio. Securities with a higher Traynor ratio have excess risk-adjusted returns, and can be identified as being underpriced.

Here is a summary of some betas of large companies.

Company	Beta 5Y Monthly
Microsoft (MSFT)	0.80
Alphabet Inc. (GOOG)	1.03
Tesla (TSLA)	1.89
3M Company (MMM)	0.95
Royal Bank of Canada (RY)	0.79
Exxon Mobil Corporation (XOM)	1.35

Finding beta in practice

The CAPM formula indicates a relationship between the mean returns μ_M and μ_A

$$\mu_A = r_f + \beta_A(\mu_M - r_f)$$

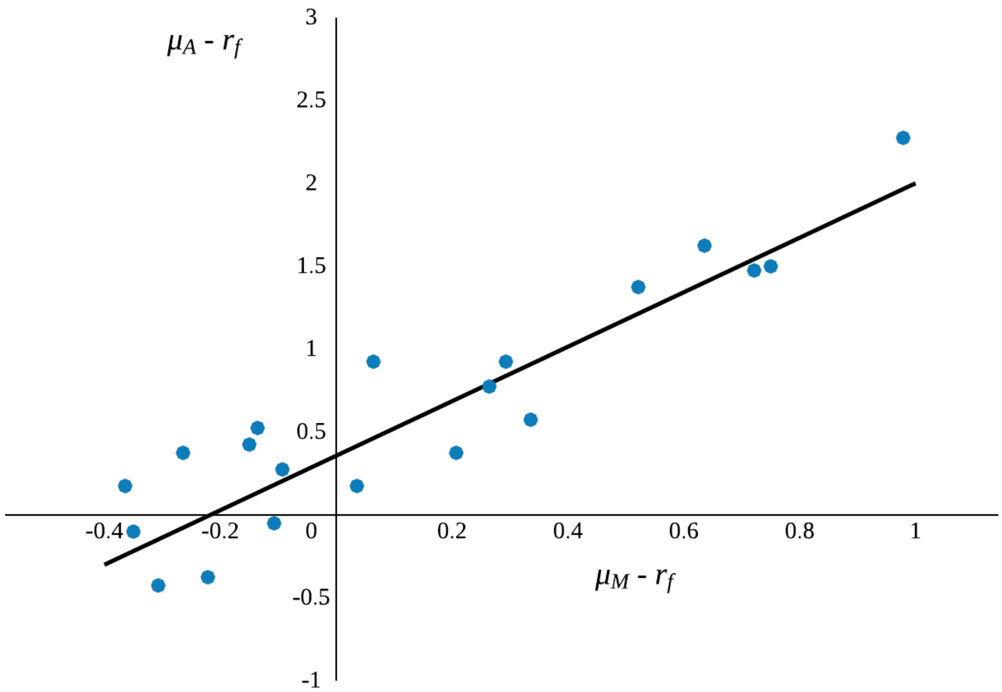
or equivalently

$$\mu_A - r_f = \beta_A(\mu_M - r_f)$$

In order to actual find beta in practice, we can look at the actual historical returns and consider the regression equation

$$R_{A_i} - r_f = \alpha_A + \beta_A(R_{M_i} - r_f) + \epsilon_A$$

The slope of the regression line is then β_A.



Although this is a popular technique, there are a few problems with implementing this regression in practice:

- True market portfolio is not observable. Typically, practitioners use a proxy such as the S&P500 as the market portfolio.
- The risk-free rate may change over time, and it is not clear which risk-free rate we should use – the T-bill rate, the 10-year government bond yields?
- Changing the time period over which the analysis is performed may result in radically different betas. Essentially betas can change over time. We will explore possible causes for this in later chapters

In theory, the value α_A should be zero, as all securities should sit on the SML. In practice it is not. This value is called **Jensen's alpha**. If alpha is positive, then this suggests that the security has excess risk adjusted returns above that predicted by CAPM. This is why investors are always "seeking alpha".

6

Arbitrage Pricing Theory (APT)

6.1 Risk Decomposition

The CAPM formula states that the expected return on a security is a function of the return on the market portfolio. Put another way, the returns on securities are a function of the performance of the market. The Arbitrage Pricing Theory extends this idea to include additional factors.

We start with the observation that the actual return on a security can be decomposed into two parts

$$\text{Actual Return} = \text{Expected Return} + \text{Unexpected Return (due to risk)}$$

The unexpected return (or risk) can also be decomposed into two parts

$$\text{Risk} = \text{Systematic risk} + \text{Idiosyncratic risk}$$

The idea here is that there are unexpected events that affect all securities to some degree. These are systematic risks. Other events tend to only affect a given specific security. These are the idiosyncratic risks.

The APT model of asset returns is can be summarized by:

Formula 1 The actual returns on an asset A, R_A can be modeled by

$$R_A = \mu_A + m_A + \epsilon_A$$

where m is the excess systematic return due to systematic risks, and ϵ_A is the idiosyncratic return due to idiosyncratic risks.

Example 1 Consider a crop farm:

- μ: expected revenue given expected weather conditions

- m: unexpected drought due to global warming. This drought is likely to affect everyone to some extent and thus it forms a systematic risk.
- ϵ: accidental death of farm owner. This adverse event is likely to affect only the farm in question, and is thus an idiosyncratic risk

Example 2 Consider an oil company:

- μ: expected revenue given expected production level and oil price
- m: unexpected spike of oil price due to war. Again, this is likely a systematic event.
- ϵ: sudden change of the comapny's CEO. A likely idiosyncratic risk

6.2 Factor Model

APT uses a factor model for systematic risk.

We begin by assuming there are k different factors, $F_1, ..., F_k$, that affect all assets to some extent (these are the systematic factors). However, not all assets are affected equally given a systematic event. We let $\beta_{A,i}$ be the measure of the sensitivity of asset A to factor i – called "factor betas".

This leads to the equation
$$m_A = \beta_{A,1}F_1 + ... + \beta_{A,k}F_k$$

Some simple examples highlighting the impact of factor betas:

- Car manufacturer's returns are likely more affected by steel prices than wheat prices; while a farmer's returns would likely be the opposite.
- Oil prices affects Exxon and Air Canada in opposite directions

Some common factors include:

- Macro economic factors: GDP, inflation, oil price, gold price, etc.
- Natural disasters: earthquakes, hurricanes, global warming, etc.
- Social-political factors: fiscal policies, warfare, etc.
- Pandemic: COVID-19

Note that the Factor Model assumes linear sensitivity to risk factors. Putting all, this together, we have the formula

Formula 2

$$R_A = \mu_A + \beta_{A,1}F_1 + ... + \beta_{A,k}F_k + \epsilon_A$$

where F_j = j-th systematic risk factor = $\Big($actual value - expected value$\Big)$

It is common to assume the following

- $E(F_i) = 0$ and $E(\epsilon_A) = 0$. Hence, all the expected return is captured in the μ_A term.
- We assume the factors are independent, and thus $Cov(F_i, F_j) = 0$ if $i \neq j$.
- We assume the idiosyncratic factor is independent from the systematic factors and the idiosyncratic factors of other assets. Hence $Cov(F_i, \epsilon_A) = 0$ for all i, and $Cov(\epsilon_A, \epsilon_B) = 0$ for all other assets B.

Definition 6.2.1 **Market Model**

The Market Model is one-factor model where the market is the only risk factor. Here we have

$$R_A = \mu_A + \beta_{A,M}(R_M - \mu_M) + \epsilon_A$$

This looks very similar to the CAPM model

$$\mu_A = r_f + \beta_A(\mu_m - r_f)$$

The main difference is that the APT model deals with actual returns, and CAPM models expected returns. As we will see, we can derive CAPM from the APT model, and the beta values coincide.

This connection is established in the following APT formula statement

Theorem 3

(APT) We assume a universe of N assets. Then the expected return on asset i satisfies

$$\mu_i = r_f + \beta_{i1}\gamma_1 + ...\beta_{ik}\gamma_k$$

where r_f is the risk-free rate, and the values $\gamma_1, \dots \gamma_k$ are the same for all assets. These values are called "factor risk premiums".

Let us return to the Market Model. In this case, the market portfolio is the only factor. However, M is also an asset, which has $\beta_M = 1$. Thus

$$\mu_M = r_f + \gamma_1 \implies \gamma_1 = \mu_M - r_f$$

Therefore, for all assets we have

$$\mu_i = r_f + \beta_i(\mu_M - r_f)$$

which brings us back to CAPM

6.3 APT and Portfolios

Consider a portfolio assembled by investing in 2 assets with weights w_1 and w_2. Then we have

$$\begin{aligned}
R_P &= w_1 R_1 + w_2 R_2 \\
&= w_1\Big(\mu_1 + \beta_{11}F_1 + \cdots + \beta_{1k}F_k + \epsilon_1\Big) + w_2\Big(\mu_2 + \beta_{21}F_1 + \cdots + \beta_{2k}F_k + \epsilon_2\Big) \\
&= \Big(w_1\mu_1 + w_2\mu_2\Big) + \Big(w_1\beta_{11} + w_2\beta_{21}\Big)F_1 \cdots + \Big(w_1\beta_{1k} + w_2\beta_{2k}\Big)F_k + \Big(w_1\epsilon_1 + w_2\epsilon_2\Big) \\
&= \mu_P + \beta_{P,1}F_1 + \cdots \beta_{P,k}F_k + \epsilon_P
\end{aligned}$$

From this we conclude:

- The expected return on the portfolio is a weighted average of the expected returns of the constituents.
- The factor betas of the portfolio is a weighted average of the factor betas of the constituents.
- The idiosyncratic risk of the portfolio is a weighted average of the idiosyncratic risks of the constituents.

Also, if we assume the factors and idiosyncratic risks are all independent, we have for all assets

$$\begin{aligned}
Var(R_A) &= Var(\mu_A + \beta_1 F_1 + \cdots \beta_k F_k + \epsilon_A) \\
&= \beta_1^2 Var(F_1) + \cdots \beta_k^2 Var(F_k) + Var(\epsilon_A)
\end{aligned}$$

which can make it easier to compute portfolio statistics.

Finally, suppose we have an equally weighted portfolio in a large number of assets, then

$$\epsilon_P = \frac{1}{N}\sum_{j=1}^{N} \epsilon_j$$

Since the expected value of each ϵ_j is zero, if we assume reasonable conditions, a law of large numbers argument suggest that as $N \to \infty$ we have $\epsilon_P \to 0$. Thus, in large portfolios, the idiosyncratic risk is all diversified away, leaving only the systematic risk behind.

	CAPM	**APT**
Systematic Risk Factor	• Unique market portfolio • Derived mathematically	• Multiple risk factors • Selected by users
Return Model	• Known μ & Σ • Derived equilibrium return	• Model return statistically • Derived expected return μ
Asset Prices	• All fairly priced assets should lie on the SML	• Assets can deviate from the SML due to idiosyncratic risk
Main Challenges	• Estimation of μ & Σ	• Selection of systematic risk factors

Let's put all this together with an example:

Suppose the following data is known for 3 stocks which follow a 2 factor APT model.

i	β_{i1}	β_{i2}	μ_i	$\sigma^2_{\epsilon i}$	ϵ_i
Stock A	1.0	-1.5	10%	0.01	5%
Stock B	0.5	1.0	7%	0.0144	1%
Stock C	0.75	1.25	9%	0.0225	-2%

	Actual	**Expected**	**Var[F]**
F_1	7%	5%	0.02
F_2	1%	2%	0.05

Let's compute the actual returns for each stock: For Stock A, we have

$$\begin{aligned}R_A &= \mu_a + \beta_{A1}F_1 + \beta_{A2}F_2 + \epsilon_A \\ &= 10\% + (1)(7\% - 5\%) + (1.5)(1\% - 2\%) + 5\% \\ &= 15.5\%\end{aligned}$$

Similar calculations show $R_B = 8\%$ and $R_C = 4.25\%$

Now let's compute the variance of the returns for each stock. For Asset A we have

$$\begin{aligned}Var(R_A) &= Var(\mu_a + \beta_{A1}F_1 + \beta_{A2}F_2 + \epsilon_A) \\ &= \beta_{A1}^2 Var(F_1) + \beta_{A2}^2 Var(F_2) + Var(\epsilon_A) \\ &= (1^2)(0.02) + (-1.5)^2(0.05) + 0.01 \\ &= 0.1425\end{aligned}$$

Similar calculations show $Var(R_B) = 0.0694$ and $Var(R_C) = 0.11875$.

Let's construct a portfolio P using securities A and B only, such that $\beta_{P1} = 0$. Let w_A and w_B be the weights in stocks A and B. Then we have the system of equations

$$\begin{aligned} w_A + w_B &= 1 \\ \beta_{P1} = w_A(1) + w_B(0.5) &= 0 \end{aligned}$$

Solving gives $w_A = -1$ and $w_B = 2$.

Finally, we'll construct a portfolio that has $\beta_{P1} = \beta_{P2} = 0$. Let w_A, w_B and w_C be the weights in stocks A, B and C respectively. Then we have the three equations:

$$\begin{aligned} w_A + w_B + w_C &= 1 \\ \beta_{P1} = w_A(1) + w_B(0.5) + w_C(0.75) &= 0 \\ \beta_{P2} = w_A(-1.5) + w_B(1) + w_C(1.25) &= 0 \end{aligned}$$

Solving gives $w_A = 16.7\%$, $w_B = 316.7\%$ and $w_C = -233.3\%$. Since this portfolio has no sensitivity to the systematic risk factors, its expected return should be the risk-free rate (assuming we are in equilibrium).

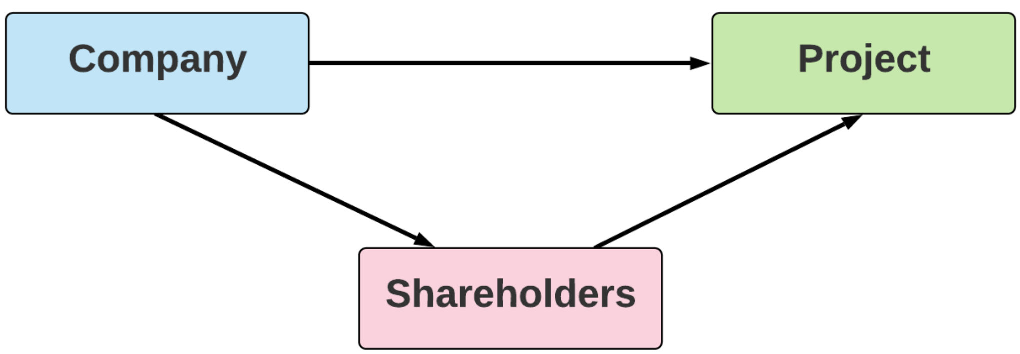
Thus $r_f = \mu_P = 2.83\%$

7 Risk, Return, and Capital Budgeting

7.1 Project Beta

Company Z is considering a project. The company has two options:

1. Invest in the project directly
2. Pay dividends to the shareholders and let the shareholders invest



Whether or not the company sits between the shareholders and the project cannot change project value. The project's value depends on its cash flow pattern, not investor identity. Therefore, company and shareholders should use the same discount rate for valuation.

In particular, the discount rate for a project is driven primarily by the **project beta** – not the company's beta.

How do we compute the beta of a project? It depends on the project.

If the project "looks like" an extension of the company, use the **company's beta**. If the project looks very different from the company, typically an **industry beta** is used. If Project P is a **combination** of n projects with known $\beta_1, ..., \beta_n$, then $\beta_P = \sum_{i=1}^{n} w_i \beta_i$, where the weights sum to 1.

Example 1

Tim Horton's is opening a new location. Suppose CAPM holds and

- Company's beta = 2
- Market premium = 7%
- Risk free rate = 2%

What discount rate should we us to value the new location?

Solution:

$$r_T = r_f + \beta_T(\mu_M - r_f) = 2\% + 2 \times 7\% = 16\%$$

7.2 Determinants of Beta

Different companies and projects have different betas due to a couple of different factors such as

- Business risk: cyclicality of revenues and operating leverage
- Financial risk: financial leverage

Cyclicality of Revenues

Highly cyclical stocks have high betas. For example, high-tech firms and retailers fluctuate with the business cycle. Other companies that deal with things such as utilities and funeral homes are much less dependent upon the business cycle.

Note that cyclicality $\neq$ variability. Stocks with high volatilities do not necessarily need to have high betas. For example, movie studios have revenues that are highly variable but low beta due to the fact that it really depends on the movie quality and not the business cycle.

This is a more qualitative idea than quantitative calculation.

Operating Leverage

Definition 7.2.1 **Operating Leverage**

Operating leverage is defined by

$$OL = \frac{\text{\% change in EBIT}}{\text{\% change in sales}}$$

It can be shown that a high proportion of fixed cost leads to a high operating leverage, and thus a high risk. A high operating leverage indicates that profits are more sensitive to sales, and magnifies the effect of cyclicality.

Here is a stylized example to illustrate the effects of operating leverage.

Example 2

Project A generate \$250 revenue and costs \$50 annually, in perpetuity. $\beta_a = 0.6$, $r_f = 5\%$, and $\mu_M = 10\%$. Assume CAPM holds.

If the costs are 100% variable costs, what is the NPV of this project?

Solution:

$$r_A = r_f + \beta_A(\mu_M - r_f) = 8\%$$

$$NPV = \frac{\$250 - \$50}{8\%} = \$2,500$$

If the costs are 100% fixed costs, what is the NPV of this project?

Solution: Fixed costs are "risk-free" (you have to pay no matter what), so

$$NPV = \frac{\$250}{8\%} - \frac{\$50}{5\%} = \$2,125$$

If the costs are 100% fixed costs, what is the new project beta?

Solution:

- Revenue's beta remains $\beta_{rev} = 0.6$
- Fixed costs' beta is $\beta_{FC} = 0$

Project A is a combination of revenue and costs, weighted by NPVs

$$NPV_{rev} = \frac{\$250}{8\%} = \$3,125, \quad NPV_{FC} = \frac{-\$50}{5\%} = -\$1,000$$

$$w_{rev} = \frac{\$3,125}{\$3,125 - \$1,000} = 1.47, \quad w_{FC} = \frac{-\$1,000}{\$3,125 - \$1,000} = -0.47$$

The project's new beta has become:

$$\beta_A = w_{rev}\beta_{rev} + w_{FC}\beta_{FC} = 1.47 \cdot 0.6 = 0.88$$

EXERCISE

Recompute the project NPV in the 100% fixed cost scenario given the new project beta.

Financial Leverage

Definition 7.2.2
Debt-to-Equity Ratio

Defined by the formula

$$\text{DE ratio} = \frac{\text{Debt}}{\text{Equity}}$$

Asset is a portfolio of Debt and Equity, based on the balance sheet equation. Usually $\beta_{Equity} \gg \beta_D \approx 0$, so

Formula 1

$$\beta_E \approx \beta_{Assets}\left(1+\frac{D}{E}\right)$$

As the DE ratio increases, β_E increases, and thus stock has a higher risk.

We will explore the "optimal" DE ratio later in the course.

Selected Betas of Major US Industries

Industries	No. of Firms	Beta	D/E Ratio
Trucking	388	1.41	33.76%
Advertising	2	0.24	77.50%
Utility (Water)	598	1.54	40.46%
Metals & Mining	52	0.98	23.86%
Steel	169	1.26	50.24%
Shipbuilding & Marine	129	1.18	62.15%
Telecom. Services	265	1.28	83.14%
Retail (Online)	101	1.14	7.15%

7.3 Weight Average Cost of Capital

To compute beta in practice, you will need to run a regression of stock returns against the market returns. Since stock = equity, β_E is obtained from regression.

Note that CAPM and APT are computed by the equity beta and not the asset beta.

However, most companies and projects are financed by a mix of debt and equity. Therefore, we need a "blended" discount rate called the weighted adjusted cost of capital.

Definition 7.3.1

Weighted Average Cost of Capital (WACC)

The WACC is defined by the formula

$$r_{WACC} = \frac{E}{A}r_E + \frac{D}{A}r_D(1-T_C)$$

where

- r_E: cost of equity, calculated by CAPM or APT
- r_D: cost of debt, typically equal to the interest rate on the company's debt
- T_C: corporate tax rate

The reason the cost of debt is multiplied by $1-T_C$ is because interest costs are tax deductible. Consequently, the factor $r_D(1-T_C)$ represents the *after-tax* cost of debt.

Investment rules, capital budgeting, CAPM/APT can all be connected through the WACC formula. WACC can be viewed as the company's **overall return on assets**. However, it is important to note that subdivisions of a company may have different WACCs of their own; different divisions may have different risks, betas, and r_E's.

Example 3 Recall Tim Horton's beta = 2, market risk premium = 7%, r_f = 2%. In addition, Tim Horton's has a DE ratio of 2, it pays 25% corporate tax, its cost of debt is 3%. Compute WACC for the expansion project.

Solution:

$$r_{WACC} = \frac{E}{A} r_E + \frac{D}{A} r_D (1 - T_C)$$

$$r_{WACC} = \frac{1}{3}(16\%) + \frac{2}{3}(3\%)(1 - 25\%) = 6.83\%$$

We would use this cost of capital to discount cash flows of projects that have the same risk as the overall firm and are financed at the same mixed of debt and equity as the overall firm.

Suppose the EBIT remains constant in perpetuity, then

$$\text{Asset Value} = OCF \times a_{\overline{\infty}|r_{WACC}} = \frac{OCF}{r_{WACC}}$$

A lower WACC indicates a lower cost of capital and thus, higher asset values. Therefore, managers wish to have a low WACC. This is an important task in risk management.

There are two additional factors that affect the cost of capital: liquidity and adverse selection.

Liquidity means different things to different people in different contexts, but in this section we define liquidity as the ease and cost with which investors can trade a security.

Since investors think of their final returns as the net of all costs, including transaction costs, management fees, etc., all else equal, stocks with high liquidity usually have a lower WACC.

In real life, we do not have access to all relevant information, we refer to this as information asymmetry. If I believe you know more than I do, I am reluctant to trade with you. I believe that by knowing more than me, you are more likely to take advantage of me. If you are willing to trade with me, I might suspect you know something I don't know. In compensation, I will sell to you at a higher price than I would buy from you. This leads to wider bid-ask spreads and less trades, thus causing lower liquidity.

Companies can do a couple of things to reduce the cost of capital. They can disclose more information, thus reducing information asymmetry. They can also try to make their stock easier to trade by doing a stock split, list on more exchanges, and online trading.

8 Efficient Market Hypothesis

8.1 Efficient Market Hypothesis (EMH)

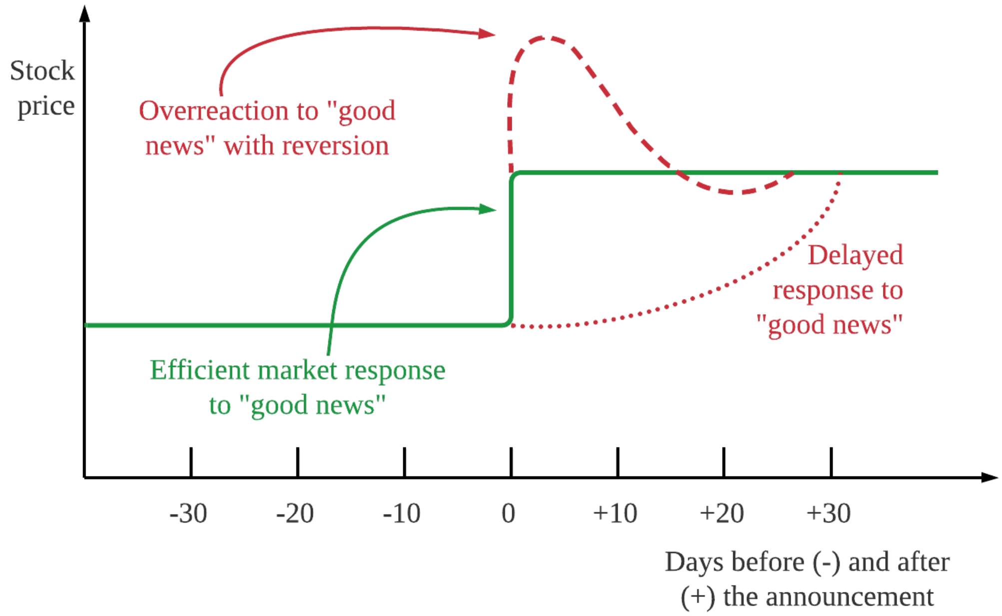
Definition 8.1.1 **Efficient Market Hypothesis**

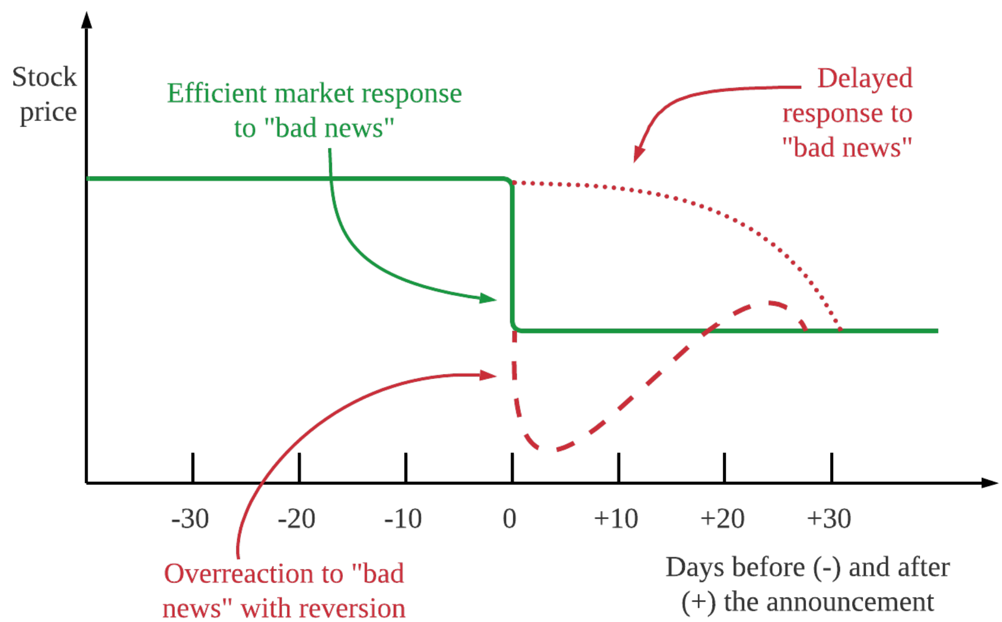
The EMH states that an efficient capital market is one in which stock prices fully reflect available information. It focuses on "information efficiency" rather than "transactional efficiency".

If we accept the EMH, then there are significant implications for investors and firms, including.

- Information is reflected in security prices immediately once it is known.
- Knowing information after it is publicly released is useless.
- Firms should expect to receive the fair value for securities, they cannot profit from fooling investors about their stock prices.

The following reaction graphs show alternatives to how the market might react to good news and bad news, respectively.





The foundations of market efficiency are built upon investor rationality, independence deviation from rationality, and absence of arbitrage. Investor rationality assumes investors adjust their estimates from new information in a rational manner. Note that investors could react to the same information differently, which is where independent deviation from rationality comes in. With lots of investors, different reactions will “offset” each other. Stupidities cancel each other out, smart decisions remain profitable. EMH does not require rational individuals, but a rational population overall. Any existing arbitrage opportunity will be taken in large quantity and vanish quickly.

8.2 Types of Market Efficiency

There are three types of market efficiency that we will discuss:

- Weak Form
- Semi-strong Form
- Strong Form

Weak Form Efficiency

If we believe in weak form efficiency, then security prices reflect all information contained in past price and volume data. Often, weak form efficiency is represented as

Formula 1

$$P_t = P_{t-1} + \text{Expected return} + \text{random error}$$

The random error comes from the arrival of **unanticipated** news. Anticipated news forms part of the expected return. By definition, this information arrives randomly. Thus, if this model holds, stock prices are said to follow a random walk, without a discernable pattern (i.e. not predictable).

"Technical traders" try to look for the patterns in past price and volume data by examining quantities such as 10-day, 30-day, and 90-day moving averages, levels of "support" and "resistance", or "following the big guy" by looking at volume. We can summarize this as

$$\text{Weakly efficient market} \Rightarrow \text{technical analysis is useless}$$

$$\text{If technical analysis works} \Rightarrow \text{market " < " weakly efficient}$$

REMARK

In the case of the second implication, the market is called *inefficient*.

Semi-Strong Form Efficiency

If we believe in semi-strong form efficiency, then security prices reflect all publicly available information. Publicly available information includes historical price and volume information, company financial statements and annual reports, public information about competitors, interest rates, GDP data, and weather forecasts.

"Fundamental traders" use public information to select investments. Sometimes, such traders use "styles" such as value investing and growth investing. They use macro-economic data to make tactical asset allocation decisions. We can summarize this as

$$\text{Semi-strong efficient market} \Rightarrow \text{fundamental analysis is useless}$$

$$\text{If fundamental analysis works} \Rightarrow \text{market " < " semi-strong efficient}$$

REMARK

In the case of the second implication, the market may be weakly efficient or inefficient.

Strong Form Efficiency

If we believe strong form efficiency, then security prices reflect all information—public and private. If news becomes known to someone – anyone – the market instantly reacts to it rationally. The stock price will always be equal to its true underlying value and hence, insider trading cannot profit. This is the “best” market.

But if markets are strong form efficient, why trade stocks at all? In reality, almost no one believes strong form EMH as it is difficult to believe that insider trading cannot profit.

8.3 Misconceptions of the EMH

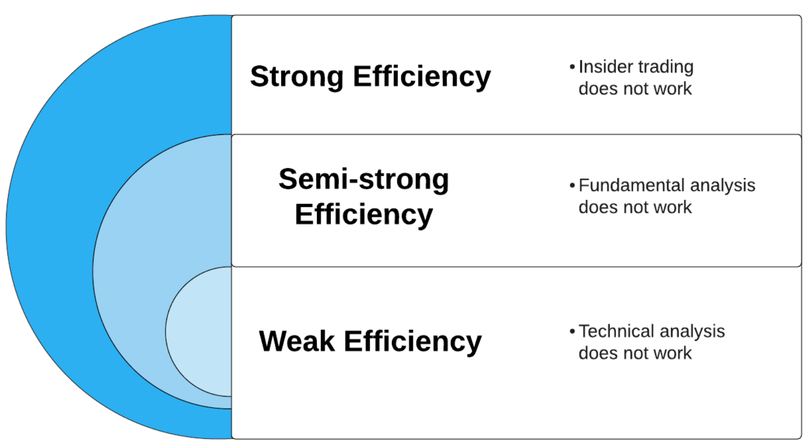
There are a number of misconceptions regarding the EMH.

The first is the idea that it is not possible to make money in the market. EMH says prices reflect information quickly and accurately, so there is no consistent way of “outsmarting” the market. However, stocks do have an expect return that investors earn on average by bearing a certain amount of risk.

Another misconception is that a monkey can make just as good of an investment decision as any investor. Oddly, this is basically what EMH says, but investors need to select portfolios that match their risk preferences.

One other common misconception is that prices are random for no reason. We know that prices reflect information and are driven by new (and random) information – so there are reasons, they just arrive randomly.

A summary of EMH: the level of information reflected on stock prices



Traders may not want the EMH to be true, especially not in its strong form, as they want to be able to beat the market! However, no one can outsmart and beat a strongly efficient market.

8.4 EMH Anomalies and Issues

Is the EMH true? This is actually a complicated and contentious issue.

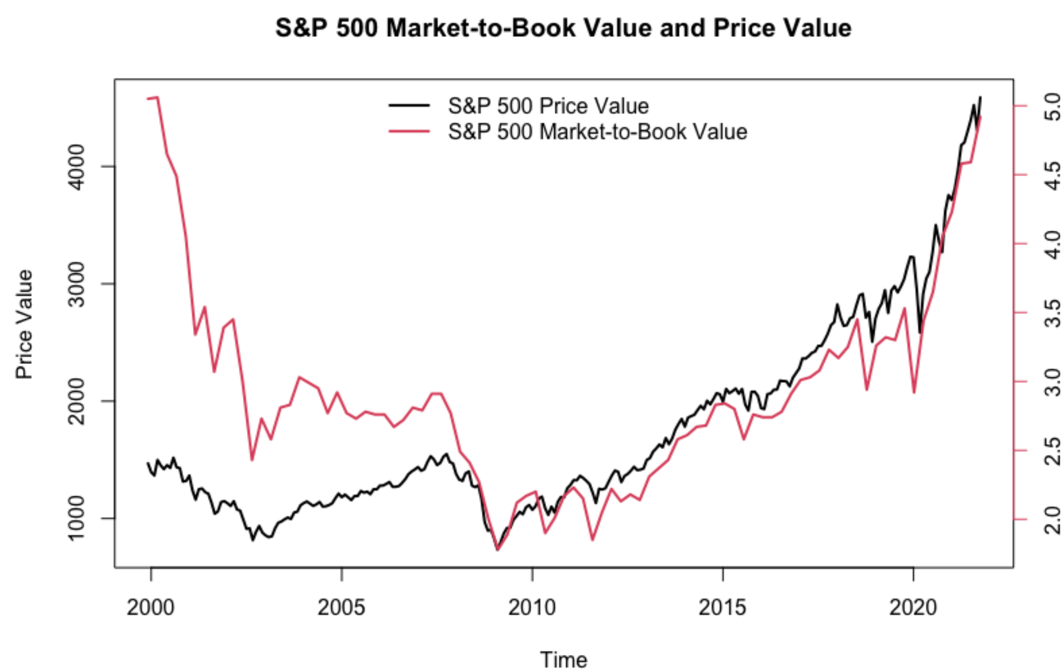
There is some evidence that small firms outperform large firms and that stocks with below-average price-to-book ratios tend to outperform the market. This suggests markets are not semi-strong efficient.

A serious issue with any studies regarding the EMH is the so-called "joint hypothesis problem". Suppose you believe a given strategy "works" which would disprove the EMH. By "works", it is not enough to prove that the strategy makes any money (T-bills make money), but that it makes *abnormal* profits. But what is abnormal? For that we need a model of normal returns, say CAPM. But then are we testing if the EMH is true or if CAPM is true? It is hard to separate the EMH from the underlying risk-adjusted return model you are using to test it.

REMARK

For the next several chapters, for the most part we will assume EMH. We will also make a number of other assumptions. Securities will be traded at their "fair value" or the "model value". Companies cannot "time" the issuing stocks, payment of dividends, etc., in order to create value. Companies cannot affect the value of their stock by altering accounting policies. Firms can sell an arbitrary number of stocks and bonds without depressing prices.

There are an array of studies that try to either prove or disprove the EMH, and many investors believe they have trading "rules" that can beat the market. We will certainly not provide an exhaustive list here, but the graph below gives an example of such an analysis. Some people ("value traders") see the market-to-book value ratio as an indicator of market "cheapness". When that ratio is low, stocks are cheap and it is a good time to buy, while the opposite holds if the ratio is high.



REMARK

Suppose this analysis is correct. Which form(s) (if any) of the EMH are violated?

9 Long-Term Financing

If there is not enough cash on hand a company can pay for a project by issuing new equity (stocks) or issuing new debt (bonds). There are different types of equity and debt a firm can issue, each with distinct features.

9.1 Equity

Generally, firms are able to issue new shares to raise capital. Public companies can issue shares via an **Initial Public Offering (IPO)** or by a *Subsequent Offering*, also confusingly called a *Secondary Offering* on the primary market. Shares of common stock are the fundamental ownership units. The *Articles of Incorporation* often indicates the maximum number of shares a company can issue, which limits management's ability to "dilute" existing shareholders.

Classes of Shares

It is important to note that not all shares are created equal. The two major share classes are common shares and preferred shares (or preferreds). Each of these classes may also have sub-classes; each sub-class has distinct features that serve different purposes.

Within common shares, there are regular, superior, and limited voting which different primarily in voting rights. Each regular common share only has 1 vote, each superior class may have, say, 100 votes, and each limited voting class may have, say, 0.5 votes. The point of this all has to do with voting control; the more votes a share carries, the more valuable is likely is. Superior shares are often held by the family that initially owned the company prior to the IPO.

Preferred shares often look more like debt than equity, and are sometimes thought of as a "hybrid security". Preferred shares have a "stated dividend" that is usually expected to be paid (like interest on a loan) at regular intervals – usually quarterly. However, skipping preferred dividend by itself will not lead to bankruptcy. So, they are similar to debt in that they carry regular distributions to investors, but they are like equity in that these payments are not legal obligations of the firm.

Cumulative vs. non-cumulative

What happens if a company skips a preferred dividend? With cumulative dividends, the dividend is "remembered". With non-cumulative dividends, the skipped dividend is "forgotten". Note that in any case, no dividends can be paid to common shareholders if there are outstanding preferred share dividends in arrears. This is one of the reasons why it is called preferred - they get preference in the payment of dividends.

Fixed rate, floating rate and reset preferreds

Fixed rate preferreds pay a fixed dividend (e.g., quarterly). Floating rate preferreds pay according to a benchmark – often something like T-bill rates + 2%, where the spread is a function of the issuer's creditworthiness.

Reset preferreds are a mixture between fixed and floating. You receive a fixed dividend initially, then after a certain period of time, it "resets" against a benchmark. For example, the dividend might be set at the 5-year government bond yield plus 1% for the next 5-years. Then, 5-years from now, the dividend resets to the then current 5-year government bond yield plus 1% for the subsequent 5 year period.

Convertible, redeemable, and retractable shares

Preferred shares may come with some other features as well. A convertible perferred share can be converted to common shares (or possibly another class of preferred share) at the option of the holder. Essentially, they consist of a preferred share plus a call option (or a warrant) on the other shares. A redeemable perferred share (also called callable) can be repurchased at the option of the issuer at a pre-set price after a certain date. Somewhat the opposite is a retractable preferred share which gives the holder the option to redeem the preferred at par, on or after some date. (So, redeemable is at the option of the issuer, and retractable is at the option of the investor.) Retractable preferreds are called "hard retractable" if the issuer must give cash, whereas "soft retractable" means the issuer can give cash or equivalent common shares.

9.2 Shareholders' Rights

Shareholders vote to appoint members to the Board of Directors. The Board then hires and supervises management. Thus, voting is the only way shareholders can affect the management. Shareholders' rights is currently a hot topic as the common voting mechanisms are not considered to be very effective; effective meaning that management maximizes shareholders value.

This brings about the topic of straight versus cumulative voting. Straight voting is by far the most common mechanism. In straight voting, each share typically carries one vote. Shareholders vote on directors essentially one at a time. With straight voting, it is possible for a majority shareholder (51%) to select the entire board. In comparison, cumulative voting works in the sense that if there are N seats available on the board, each share carries N votes. Shareholders can concentrate votes on as many or as few candidates as they wish. This way, minority shareholders are able to secure some seats.

Definition 9.2.1
Proxy Vote

A proxy is the granting of your voting rights to someone else.

To gain voting control without owning many shares, a shareholder can solicit proxies from others, essentially creating a "proxy battle". Most actual voting forms have a default option to select management as your proxy. Doing so, management can become entrenched. This is a serious issue for corporate governance.

Profits of a company are distributed to shareholders via dividends. Although stocks often have a “stated dividend”, these are not legal obligations. The mechanism is that the board meets regularly and decides if and how much of a dividend to pay. Should the Board chose not to pay a dividend, the shareholders cannot sue for payment, since dividends are not legal obligations of the firm, but simply a distribution of profits.

9.3 Debt

Debtholders are not owners of the firm. They have no voting right and no control over the company. To protect their interest, debtholders rely on the loan contract (or indenture). The terms of these agreements include the size of the loan, interest rate, repayment terms, and protective covenants such as limiting dividend payments or other protections for lenders.

Unpaid principal and/or interest can lead to bankruptcy. Companies may have incentive to issue more debt since interest payments are tax deductible while payments of dividends are not. From the company’s perspective, tax status favours debt (can get tax deduction) and legal status favours equity (avoid bankruptcy due to default). We explore the optimal mix of debt and equity in subsequent chapters.

There are many different terms to distinguish all the different types of debts with different features.

- Bond = debt that is secured by actual property
- Mortgage = loan secured by real estate
- Debenture = debt that is not secured by property (relies on issuer credit)
- Note = debt that has a short term to maturity
- Trade payable = very short term loan, usually no interest charge, granted by a supplier as part of day-to-day operations

Convertible debentures are bonds that can be converted to common at the option of the holder. These are often traded on stock exchanges. These are essentially bonds + call options (or more accurately, warrants). Contingent convertibles, or CoCos, are complicated bonds that are automatically converted to equity on under certain conditions. For example, for banks, if the capital ratio is too low, then CoCos must convert to common shares to prevent bankruptcy.

In the event of a bankruptcy, investors get paid **in order** as follows:

1. Senior secured debtholder
2. Junior debt debtholder (subordinated, unsecured)
3. Trade payables debtholders
4. Preferred shareholders
5. Common shareholders

This preference in liquidation value gives some additional protection to debtholders.

10 Capital Structure

10.1 M & M Proposition: No Taxes

We can think of the value of a firm as the value of its assets, or equivalently, the value of the debt plus equity. That is

$$V = D + E$$

The main question for this chapter is to determine if changing how the company is financed (i.e. what the mix of debt and equity is), changes firm value, and if so, what is the optimal mix?

This brings us to Modigliani and Miller's capital structure theories, which is comprised of MM Propositions I & II (no taxes) and MM Propositions I & II (with taxes).

M & M Proposition I (No Taxes)

For the rest of this chapter, we assume the EMH and a frictionless markets. We also assume that the interest rate is the same for both borrowing and lending, and it is the same for individuals and corporations.

Under these assumptions, the value of the firm is unaffected by financial leverage, i.e.,

Formula 1

$$V_L = V_U$$

To prove this proposition, we need to look at the cash flows that these alternative capital structures generate.

In the unlevered firm, the cash flow to the investors (in this case, just the equity holders) is given by

$$CF_E = EBIT$$

In the case of leverage, the cash flows to the investors are

$$CF_E = EBIT - r_D D$$
$$CF_D = r_D D$$

These last equations add up to $EBIT$ exactly as in the unlevered case. Therefore, the cash flow to investors is exactly the same, and the risk is the same (the assets that generated the cash flows is the same), so the present value of the cash flows must be the same.

Thus the value of the levered and unlevered firms agree.

Recall the formula for $WACC$

Formula 2

$$r_{WACC} = \frac{E}{A} r_L + \frac{D}{A} r_D$$

This is equivalent to

$$A \cdot r_{WACC} = E \cdot r_E + D \cdot r_D$$

The terms on the right are the cash flows to the equity holders and debt holders respectively. The MM proposition says the present values of these cash flows are independent of the capital structure. Thus, the present value of the left hand side is independent of capital structure.

Thus, if for simplicity we assume the cash flows are a perpetuity, we have

Formula 3

$$A = V_L = V_U = \frac{EBIT}{WACC} \Rightarrow WACC = \frac{EBIT}{V_L} = \frac{EBIT}{V_U}$$

Put another way, if the company value is independent of capital structure, then so is WACC (EBIT is assumed to be fixed).

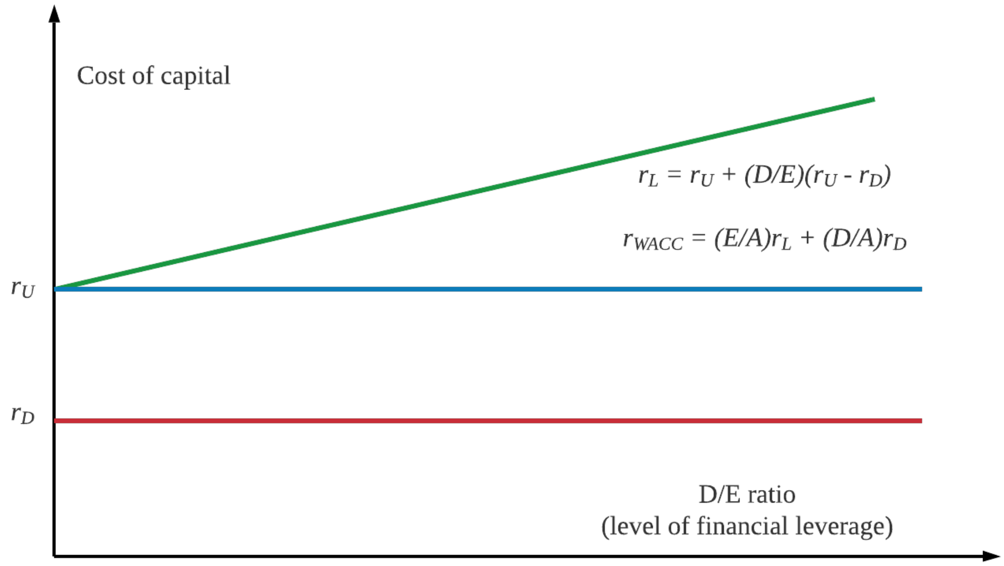
Reorganizing the WACC formula, we get

Formula 4

$$r_L = r_U + \frac{D}{E}(r_U - r_D)$$

where

- r_L = cost of equity for levered firm
- r_U = cost of equity for all-equity firm (also its WACC)
- r_D = cost of debt



Looking a the WACC formula, it seems that intuitively, by adding debt, which usually has a low cost (i.e. r_D is usually small) we should reduce the WACC since we are increasing the percentage of assets financed by cheap debt. However, as debt increases, so does risk to equity holders, and thus r_E should rise – and a rising r_E should increase $WACC$. M&M basically says that these effects exactly cancel each other out yielding a constant $WACC$.

10.2 M & M: With Taxes

Surprisingly, the conclusions change when we include the effects of corporate taxes since interest on debt is tax deductible. Increased debt has then has two effects: higher proportion of low-cost debt (offset exactly by increased equity risk as before), and higher tax deductions. The additional tax reduction has value.

M & M Proposition I (With Taxes)

Firm value **increases** with financial leverage

Formula 5

$$V_L = V_U + D \cdot T_C$$

The added term $D \cdot T_C$ represents the present value of the tax savings generated by the tax deduction on interest payments. Those tax deductions have value, and they form the source of the increase in firm value as debt levels rise.

M & M Proposition II (With Taxes)

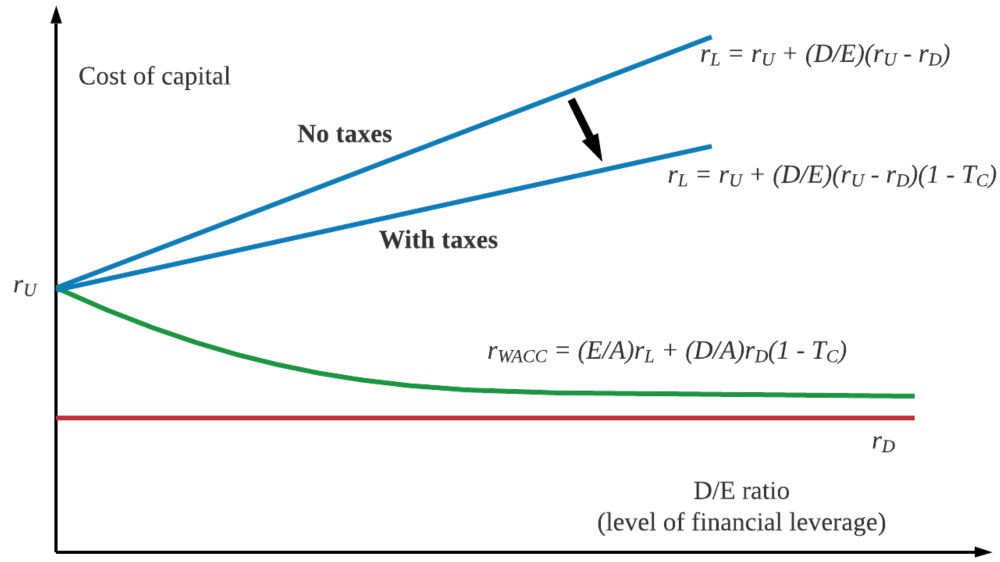
The cost of equity increases as a result of the increase of debt according to the formula

Formula 6

$$r_L = r_U + \frac{D}{E}(r_U - r_D)(1 - T_C)$$

The formula is very similar to the M&M II without taxes, the only difference is the additional $(1 - T_C)$ at the end.

Here, the cost of equity increases, but this increase is offset by the increased proportion of low-cost debt, then further offset by the interest tax shield.



11 Limits to the Use of Debt

11.1 Agency Costs

With taxes, M&M propositions suggest that debt lowers WACC and increases firm value. Does that mean the more debt the better!? There is nothing wrong with the math and logic, but something is missing – 100% debt = bankruptcy – so that can't be right. The key is that **the possibility of bankruptcy has negative effect on firm value.**

REMARK

This is not in regards to the risk of bankruptcy itself (shareholders and bondholders are fairly compensated by higher cost of equity/debt capital), but the actual cost of bankruptcy (a.k.a., the costs of financial distress).

There are direct and indirect costs of financial distress. Direct costs include things like legal, accounting, and administrative costs. These could be large in magnitude but tend to be a small percentage of firm value. Indirect costs take form through things like lost sales (e.g., lost brand value, lost customer loyalty, etc.) and agency costs due to conflicts of interests.

What are agency costs? The shareholder is the "agent" of bondholders in owning the company. Management is the "agent" of shareholders in managing the company. Conflicts of interest cause agency costs which can decrease firm value. Shareholders' selfish investment strategies (near bankruptcy) include the incentive to take large risks, the incentive towards underinvestment, and "milking the company".

Incentive to Take Large Risks

When a firm is close to bankruptcy, the shareholder's viewpoint is along the lines of *"there is nothing to lose anyway"*, so why not go for a big gamble.

Example 1 A firm owes its bondholders \$100, consider 2 projects.

	Project A		**Project B**	
Scenarios	**Recession**	**Boom**	**Recession**	**Boom**
Firm Value	\$100	\$200	\$50	\$250
Bond	\$100	\$100	\$50	\$100
Equity	\$0	\$100	\$0	\$150

Project B is riskier based on the range of firm values. Suppose recession and boom are equally likely. Let's compute the expected values of the bonds and the equity.

Solution:

Project A: bond = 100, equity = 50

Project B: bond = 75, equity = 75

To the shareholders, the choice is clear. The choice to accept project B advantages equity holders at the expense of bondholders. This sub-optimal decision is an agency cost borne by debtholders.

Milking the Company

The idea behind this is that if you know the company is about to go bankrupt, why not go out with a bang! Pay a big, final, killer dividend, and have a huge party.

11.2 Managing/Reducing Costs of Debt

Bondholders are aware of these issues (i.e., agency cost by shareholders). In fair compensation, bondholders will demand a high interest rate on the loans. This increase in the cost of debt is bad for shareholders. Managing the cost of debt is actually better for everyone. The real question is: How?

Definition 11.2.1
Covenants

Covenants are terms or conditions that the borrower agrees to. They are written into the loan documents or loan indentures.

There are two main categories of covenants: positive covenants and negative covenants. Positive covenants are things the borrower (company) must do. Some examples of positive covenants include: insure assets, maintain good conditions of assets, provide audited financial statements, maintain a minimum level of working capital, and allowing redemption in the event of a merger, spinoff or other major corporate change. Negative covenants are things the borrower (company) cannot do. Some examples of negative covenants include: not to pay dividend (above a limit), not to issue more debt (above a limit), not to issue more senior debt (above a limit), and not to sell assets without paying bondholders with proceeds.

The whole idea is that the covenants are usually in the interests of the bondholders. If these covenants make the bondholders believe the agency costs have been reduced, they may agree to a lower interest rate, which benefits shareholders.

The second category of agency costs are the costs borne by shareholders due to management's behaviour.

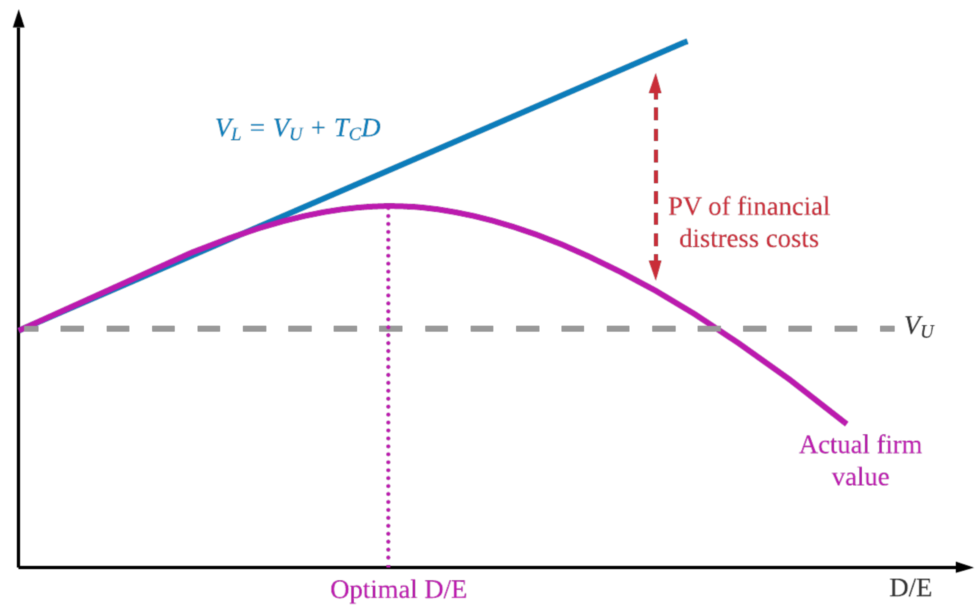
Often management has little ownership in the firm and has different incentives than shareholders. This misalignment of incentives is the source of agency costs of equity.

How can shareholders manage this? One way may be to reduce the free cash flow available to management.

Definition 11.2.2
Free Cash Flow Hypothesis

The free cash flow hypothesis states that management tends to waste more money when there is lots of "free cash" around. If the firm issues more debt, then the increase in interest charges reduces the free cash flow, thus reduces managements opportunity to waste.

We are left with some contradictions: Increase in debt brings tax shield advantage (which is good for firm value) but also increase in costs of financial distress (which is bad). Lower debt levels reduces costs of financial distress (which is good), but according to the free-cash flow hypothesis, this increases agency costs to shareholders (which is bad). Considering this tradeoff, what is the optimal D/E ratio? The idea is clear, but there is no widely accepted formula yet. This is known as the "trade-off theory" in capital structure.



Given this complicated trade-off, what do firms actually do?

Some firms have a target debt-to-equity ratio. In some way, this follows the trade-off theory. Other firms, however, do not. They seem to follow the so called "pecking-order theory".

Definition 11.2.3
Pecking-Order Theory

The pecking order theory states that when cash is needed for investment:

1. Use internal financing first
2. Then issue debt
3. Finally issue equity

This suggest profitable firms, which have lots of internally generated cash available, will use less debt.

There is one final point to make. In real life, investors know less about a firm than its management does – there is an "information asymetry" between investors and management.

Investors try to close that gap by listening to and looking at management's actions to infer what they really know about the firm. It seem reasonable that if management is pessimistic about the future, they would not want high debt levels they cannot support which could lead to bankruptcy. Optimistic management may do the opposite and have high debt levels which increase firm values.

Therefore, if management increases debt levels, the market may take this as a signal that they are optimistic about the future. Thus, firm value can increase by more than predicted by theory. Management can unfortunately use this knowledge to try to fool investors.

The act of the market trying to read the mind's of management by studying their behaviour of this type is often collected under the term "signalling theory".

12

Capital Budgeting for Levered Firms

12.1 Capital Budgeting for Levered Firms

When it comes to capital budgeting for levered firms, we have two main methods.

- The Adjusted present Value (APV) method: Here we recognize the project value as the unlevered value plus the present value effects of the financing.
- The Weighted Average Cost of Capital (WACC) method: This method discounts all cash flows at the WACC to reflect the combination of business risks and financing effects.

Example 1 ABC Inc is considering a project with

- Initial cost \$1 million
- Generates EBIT of \$200,000 per year forever
- ABC finances the projects using \$300,000 in debt
- ABC pays tax at a 25% rate
- ABC's unlevered return on equity is 10% ad borrowing rate is 5%

In notation: $C_0 = -1mil$, $EBIT = 200K$, $D = 300K$, $T_C = 25\%$, $r_U = 10\%$, $r_D = 5\%$

Should ABC Inv accept the project?

Solution:

The APV method:

Given a debt level, the APV of a project is

$$NPV_{APV} = NPV_{unlevered} + PV_{financing}$$

Compare this to MM Proposition I (with taxes)

$$V_L = V_U + T_C D$$

Find the NPV_{APV} of the previous example

$$NPV_{unlevered} = C_0 + \frac{EBIT(1-T_C)}{r_U} = 500K$$

$$PV_{financing} = T_C D = 75K$$

$$NPV_{APV} = NPV_{unlevered} + PV_{financing} = 575K$$

The WACC method:

Use r_{WACC} to discount unlevered future cash flows

$$NPV_{WACC} = C_0 + \frac{EBIT(1-T_C)}{r_{WACC}}$$

In order to calculate r_{WACC}, D/E ratio is needed. Total firm value after accpetance of the project is \$1,575K, \$300K of which is debt.

$$\frac{D}{E} = \frac{300}{1575-300} = 23.53\%$$

$$\frac{E}{D+E} = 80.95\%$$

$$\frac{D}{D+E} = 19.05\%$$

Find the NPV_{WACC} in the previous example

$$r_L = r_U + \frac{D}{E}(r_U - r_D)(1-T_C) \approx 10.88\%$$

$$r_{WACC} = \frac{E}{D+E} r_L + \frac{D}{D+E} r_D(1-T_C) = 9.52\%$$

$$NPV_{WACC} = C_0 + \frac{EBIT(1-T_C)}{r_{WACC}} \approx \$575K$$

So when should we use which method?

- APV method is used when the debt level is known in dollars.
- The WACC method is used when the debt-to-equity ratio is known. The WACC method is by far the most commonly used method.

12.2 Low Interest Loans

There is one special case where we need to use the APV method, and that is when the firm is given a low interest loan to pursue the project. Such low interest loans are often given by governments in order to promote some sort of economic activity.

Let r be the market interest rate for the firm and let r^* be the low interest rate. It would seem that the annual value of this low interest loan should be $D(r-r^*)$ where D is the loan amount (i.e. the value is the reduction in interest paid). However, this is not correct. Since we pay less interest, we also get a smaller tax deduction on that interest. So we win on low interest payments but lose on low tax deductions. We need to combine these effects.

Formula 1 The annual net value of a low interest loan is given by

$$D\big(r - r^*(1-T)\big)$$

The total value of this financing effect over the life of the project is the sum of the present values of these annual amounts, discounted at the market rate r. In the case of a perpetuity, we have

$$PV_{financing} = \frac{D\big(r - r^*(1-T)\big)}{r}$$

The way to think about the formula is that the first term $D \times r$ would be the net cost of the loan in an environment of no subsidies – either low interest loan subsidy or tax deduction subsidy. The second term $D \times (r^*(1-T)$ is the net cost of the loan in the presence of all subsidies. The difference must then equal the value of all subsidies combined.

For illustration, there consider the special case of a perpetuity where $r = r^*$ (i.e. so there is no low interest subsidy). We then get

$$PV_{financing} = \frac{D\big(r - r(1-T)\big)}{r} = D \times T$$

which is back to our original formula.

REMARK

Consider the case where the low interest loan is a perpetuity at a rate $r^* = 0$. Interpret that result.

13

Dividends

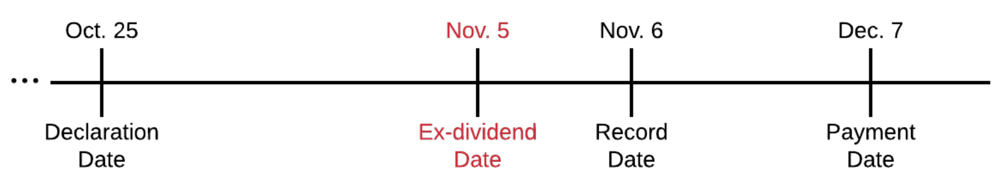
13.1 Types of Dividends

A dividend is a distribution of earnings to shareholders. There are a number of types of dividends: regular cash dividends, stock dividends, and stock repurchases. With regular cash dividends, public companies often pay quarterly. Sometimes, an extra or special cash dividend is declared. In extreme cases, a liquidating dividend is paid. With stock dividends, original shareholders are issued new stocks instead of cash and consequently, the number of shares outstanding increases. With stock repurchases, the company will buy back some of its own stock, thus the number of shares outstanding decreases.

There are a number of important dividend dates:

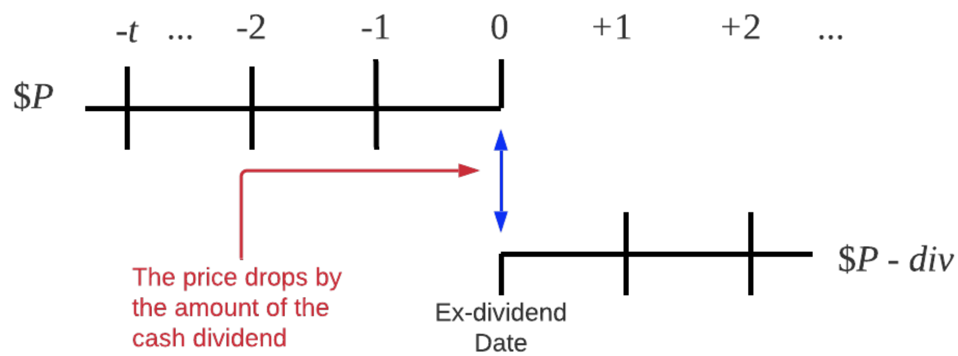
- **Declaration date:** Board of Directors declares the dividend
- **Record date:** Who owns the stock on this date gets the dividend
- **Ex-Dividend date:** The first day that the seller of a share is entitled to the dividend (one business day prior to record date)
- **Payment date:** When the dividend is actually paid

The difference between the ex-dividend date and the record date is due to the delayed settlement period. When a purchaser buy a stock on an exchange, the transaction actually does not finalize (or "settle") until 2 days later. This is sometimes called T+2 settlement. Different assets on different exchanges in different countries may operate on a different delayed settlement cycle. Take a look a the cash dividend timeline below.



The ex-dividend date is the one that matters the most. If you sell on or after the ex-dividend, you get the dividend. If you buy on or after the ex-dividend, you don't get the dividend.

In a perfect world (i.e., EMH), the stock price will fall by the amount of the dividend on the ex-dividend date. Tax effects are complicated and to be discussed later.



13.2 Modigliani and Miller Indifference Proposition

The management makes the decision on dividend policy, which consists of deciding whether to pay (start paying) dividends and how much the dividends should be. They may also consider if there is an "optimal dividend policy" to maximize firm value.

In a perfect world where EMH holds, there are no transaction costs or taxes, and the investment policy of the firm is fixed, then the dividend policy is irrelevant and a firm cannot create value by changing its dividend policy.

We will try to illustrate this as follows: Assume we have a firm where

- Cash flows are \$100K today and \$110K in 1 year with certainty
- After 1 year, the firm will dissolve
- 10K shares outstanding
- Assume $r_U = 10\%$

The current dividend policy is to pay out all earnings as dividends (a so-called "100% payout ratio"). That is, dividends are \$100K today, 110K in 1 year.

Suppose management thinks firm value will increase if they increase today's dividend to \$110K, then to whatever is available next year.

What are firm values under these two policies?

Current policy:

Firm value is total discounted CFs

$$V_1 = \$100K + \frac{\$110K}{1 + 10\%} = \$200,000$$

Each shareholder receives \$10 today and \$11 in 1 year, price per share is

$$PPS_1 = \$10 + \frac{\$11}{1 + 10\%} = \$20.00$$

On ex-dividend date, stock price will fall by \$10 to \$10.00. It will be \$11 just before the final liquidating dividend of \$11 at time 1.

Alternative policy:

In order to fund the extra \$10,000 in dividends, the firm issues \$10,000 shares at the new ex-dividend price. At the moment, we don't know what this price is, since this change in policy may have changed firm value today. (To assume the new ex-dividend price is \$9 is to assume no change in value, which is what we are trying to prove.)

Let S be the new ex-dividend price. Then the number of shares issued equals

$$\frac{\$10,000}{S}.$$

But, the ex-dividend price must be the present value of the dividends at time 1, which equals the present value of the \$110k divided by the total number of shares outstanding. Thus

$$S = \frac{\$110,000/(1+10\%)}{10,000 + \frac{\$10{,}000}{S}}.$$

Solving gives $S = \$9.00$. Thus, the price just before the stock went ex-dividend is $\$9+\$11 = \$20$ just as before. So, the new policy made no difference.

Note that the new dividend at time 1 then equals \$9.90. Thus each original shareholder receives \$11 today and \$9.90 in 1 year, yielding a price per share of

$$PPS_2 = \$11 + \frac{\$9.9}{1+10\%} = \$20$$

again just as before.

Another way to see this conclusion is that the company undertook a zero NPV investment, and as such, there was no change in value.

There is yet an alternative, perhaps more practical way to see what is going on. Suppose you initially own 100 shares. The company pays you \$1,000 today, and \$1,100 next year. But you wish to receive dividends according to the new policy which would be \$1,100 today and \$990 in 1 year.

In order to generate the extra cash today to meet your preferred cash flow pattern, sell \$100 worth of stock ex-dividend (after receiving the dividend). Now you hold $100 - \frac{\$100}{\$10} = 90$ shares. Your next dividend payment, in one year, will be $90 \times \$11 = \$990$. This is exactly the cash flow that you preferred. This strategy is called "homemade dividend", which is selling stocks to replicate a cash dividend. By buying and/or selling shares, a shareholder can essentially replicate any dividend pattern they wish. Since the investor can do this for themselves, there can be no value creation by the firm in doing this for them.

Therefore, investors are indifferent to dividend policy. They can re-do or un-do whatever dividend policy they prefer. Some implications (in theory) are as follows:

- Dividend policy cannot create value for the firm
- Changing dividend policy has no effect on stock price
- Management should not spend time thinking about changing it

However, companies do alter dividend policy, so there must be some reason.

Clientele Effect

In practice, it may not be cost efficient to sell small number of shares in order to generate a homemade dividend. Investors (such as retirees) who prefer a stream of current income may prefer high-dividend paying stocks for this reason. If there is a mismatch between the different types of investors in the market (demand), and the different types of dividend policies in the market (supply), a firm might be able to create value by switching to a new policy. However, at equilibrium, supply = demand, and hence, different investors (e.g., the clienteles) are satisfied. In this case, a firm cannot create value by changing policy.

13.3 Stock Splits and Stock Repurchases

Stock Splits

Stock dividends and stock splits are essentially the same thing: shareholders receive additional shares instead of cash. They are called differently by practical conventions. Stock dividends are commonly expressed as a percentage, while stock splits are usually expressed as a ratio.

- 20% stock dividend: 50 shares ⇒ 60 shares
- 2:1 stock split: 50 shares ⇒ 100 shares
- 1:4 reverse split: 100 shares ⇒ 25 shares

Assuming the EMH, the price per share will adjust to reflect unchanged total firm value, so why bother? There are a number of practical reasons. It is believed that shares have a "trading range" and by staying in that range, liquidity and marketability may be improved. Reverse splits increase the share price to a more "respectable level" and above some "minimum price".

Stock Repurchases

Stock repurchases allow companies to buy back some of its own stock, reducing total number of outstanding shares. Shareholders who do not sell own a higher percentage of the company – this is the opposite of dilution. The share price appreciates and shareholders get appreciated shares instead of dividends. Recently, share repurchase has become an important way of distributing earnings to shareholders.

While dividends are set essentially arbitrarily by the Board, dividends are viewed as "commitments" whereas stock repurchases are "one-time" decisions. Thus, companies may prefer repurchases over cash dividends to maintain flexibility.

Signalling Theory

If a company increases its dividend, the market may see this as a signal that management believes it can sustain the higher dividend for a long time. This belief in the optimism of management is good news. Consequently, an announcement of a dividend increase may result in a higher stock price, in contradiction to theory.

Tax Effects

Not all investors pay tax on the same basis. In Canada, corporations do not pay taxes on cash dividends received, and as such, they may prefer cash dividends. Individual investors who hold stocks for a long period must pay tax on dividends annually, as opposed to paying tax on perhaps a larger capital gain some time in the future. Thus, they may prefer repurchases. This is another example of the clientele effect.

Management Bonuses

In many cases, management bonuses are a function of stock price. Since the stock price falls on the ex-dividend date by the dividend amount, management has an incentive to use stock repurchases instead to maintain a high stock price.

Free Cash-Flow Hypothesis

Finally, if we accept the free-cash flow hypothesis that management has more opportunity to waste money if there is lots of free-cash flow available, then an increase in dividends can decrease free-cash flow and thus reduce agency costs.

In summary, MM proposition says dividend policy does not matter, it is neither for or against paying cash dividend. However, practical considerations may impact a firms dividend policy

14 Options

14.1 Introduction to Options

So far, when thinking about project or firm values, we have assumed the project consisted of physical assets of some kind that generate positive cash flows. We now turn to something different.

Definition 14.1.1 (Financial) Derivatives

A Financial Derivative is an asset whose values depends on some other asset (or collection of assets).

In this chapter, we will focus on options.

Options give the holder the **right**, but *not the obligation*, to buy or sell an asset at a pre-specified price on or before a pre-specified date. The pre-specified price and date are agreed upon today. The transaction takes place on (or perhaps before) a given future date.

Definition 14.1.2 Call and Put Options

A call option gives the holder the right, but not the obligation, to buy an asset. A put option gives the holder the right to sell an asset.

The asset itself is called the "underlying". So, an option to buy Royal Bank stock has Royal Bank shares as the underlying.

The "optionality" always has value to the holder. In the worst case, they simply do not exercise (i.e., walk away from the option). This right to the holder is the obligation of the seller. Consequently, the seller charges premiums to sell options.

We define the following variables:

- S_t: price of underlying asset at time $0 \leq t \leq T$
- K: Strike of the option. Also called the exercise price.
- T: Expiration time (or date) of the option

Since we would exercise a call option only when the asset price is above the exercise price, and we would only exercise a put if the asset price is below the exercise price, the payoff (i.e. the value of the option if exercised) of these options is given by the formualas

Formula 1

$$\text{Payoff}_{Call} = \max(0, S_T - K)$$

and

$$\text{Payoff}_{Put} = \max(0, K - S_T)$$

Definition 14.1.3

Moneyness and Intrinsic value

An option is called at-the-money if the strike price equals the current asset price.

An option is called in-the-money if there is a positive value if it is exercised immediately

An option is called out-of-the-money if the holder suffers a loss from immediate exercise. (The option may still have time-value because it may become in the money later.)

The value of an in-the-money option if exercised today is called the intrinsic value. The difference between the intrinsic value and the market value of the option is called its time value.

Example 1

A stock trading at \$49 has put and call options with a strike of \$45 trading at \$1.10 and \$5.25 respectively. Compute the intrinsic and time values for each option.

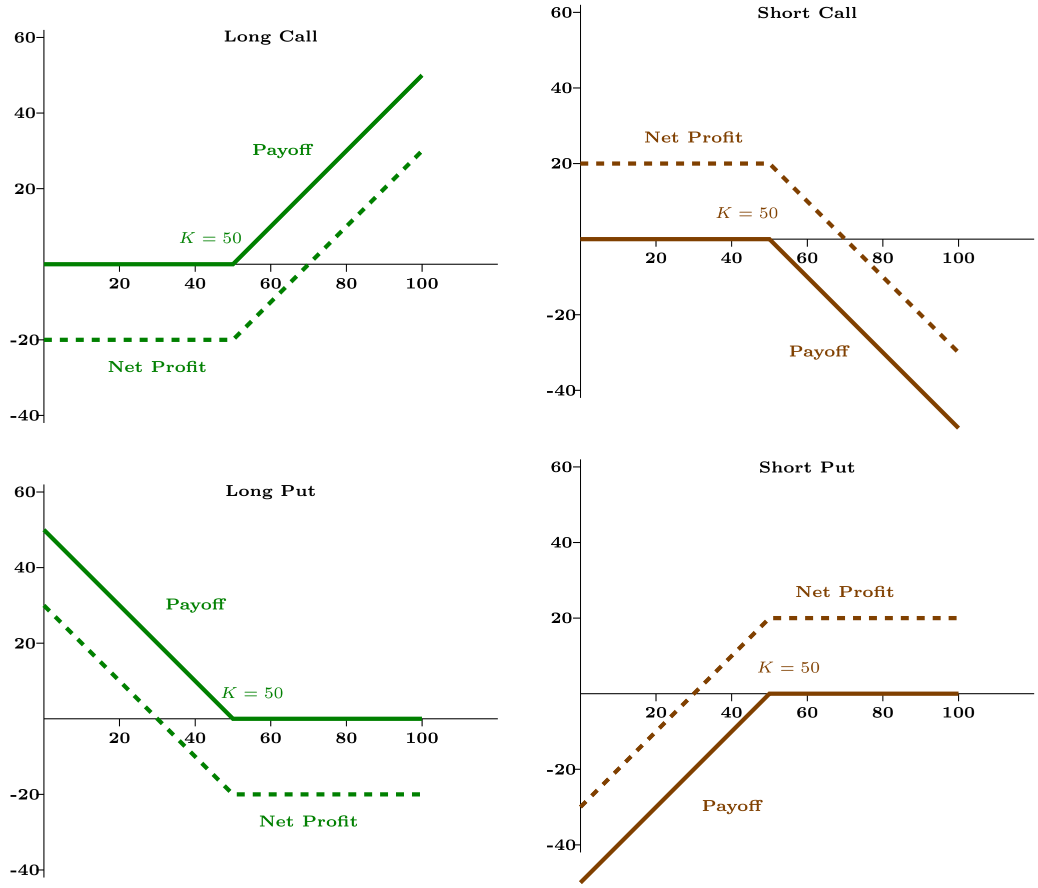
Solution: The put option value if exercised immediately equals $\max(0, \$45 - \$49) = 0$. Thus it is out-of-the-money and has intrinsic value of zero. Hence its time value is \$1.10.

The call option intrinsic value is then $\max(0, \$49 - \$45) = \$4$. Its time value is then $\$5.25 - \$4 = \$1.25$.

REMARK

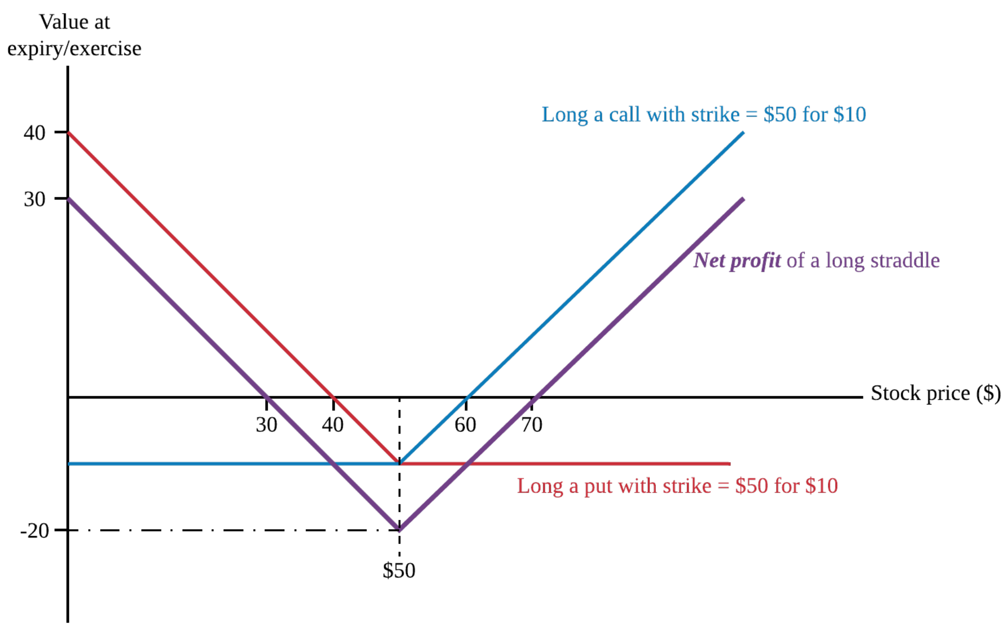
Note that there is a slight difference between so-called European and American options. With European options, they can be exercised only at expiry. Whereas American options can be exercised at any time up to expiry. There are also Bermudian options which can be exercise only on specified dates between now and expiry.

We can graph the payoff of an option (or basket of options) against the stock price to illustrate regions of profitability.



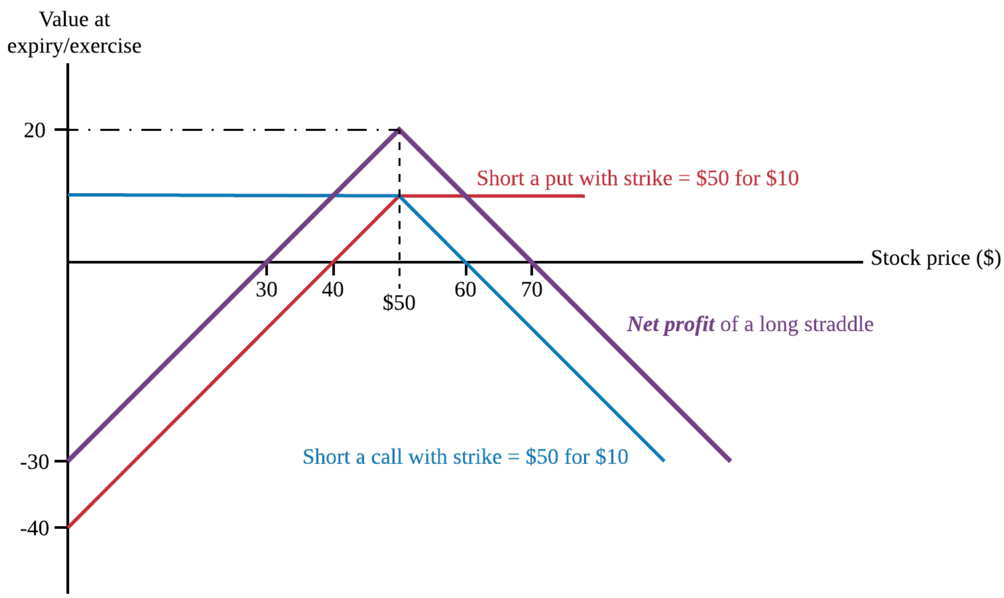
Options can be combined into portfolios to give more complex payoffs. Consider the portfolio consisting of buying both a put and a call with the same strike price and expiry date. Such a portfolio is called a **straddle**.

The payoff diagram of a straddle is shown below:



This long straddle makes a profit if the stock price moves $20 away from the strike. It is more likely to make profit if the underlying has high volatility.

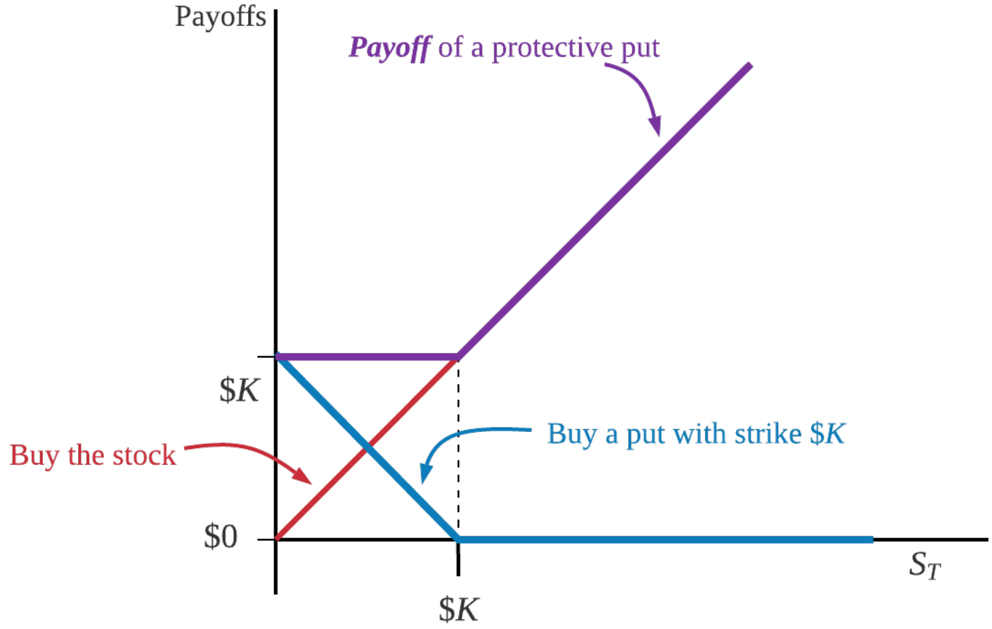
The opposite of a long straddle would be a short straddle where the traders sells both the put and call options.



This short straddle makes a profit if the stock price stays within $20 of the strike. It is more likely to make profit if the underlying has low volatility.

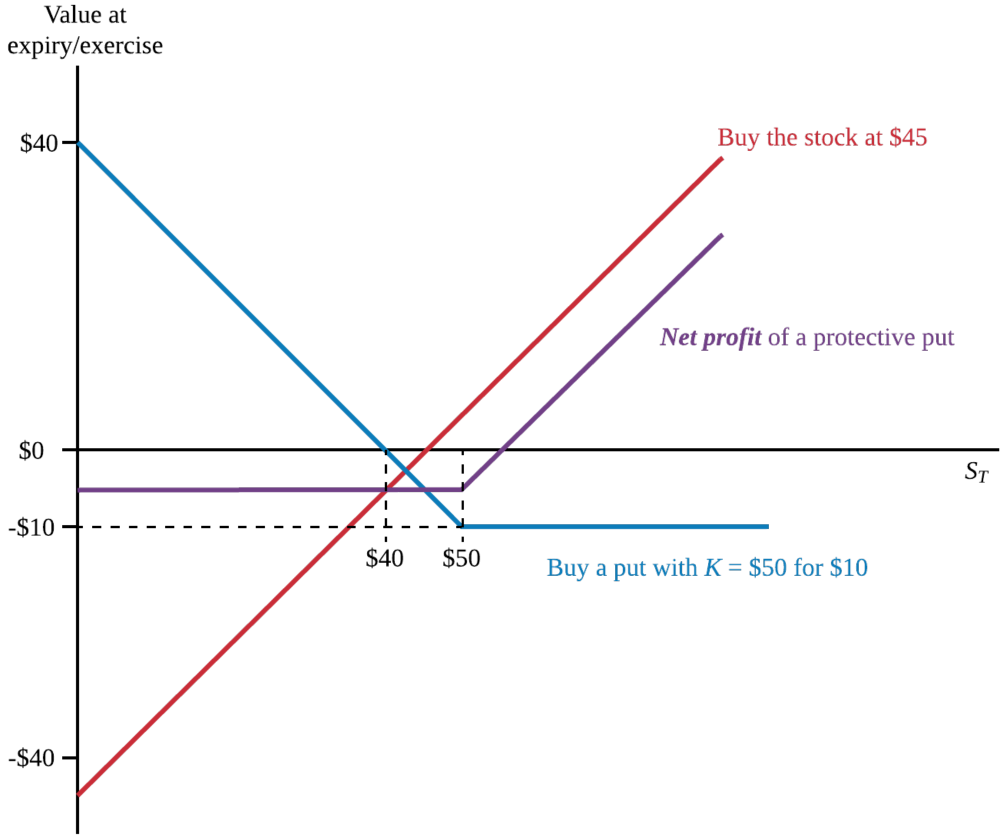
Call and puts are building blocks for complex option strategies, which are created to meet different risk appetites. Option strategies can *re-create* other options or trades, and are often called “synthetic options” or “synthetic positions”, or just “synthetics”. These synthetic options might be cheaper (leading to possible arbitrage trades), or an have other tax, regulatory or liquidity advantages.

A common option strategy is called a **protective put**, where the underlying stock is combined with a long put. This limits the downside risk, while still allowing for upside gains. Essentially, this is like a form of “stock insurance”.

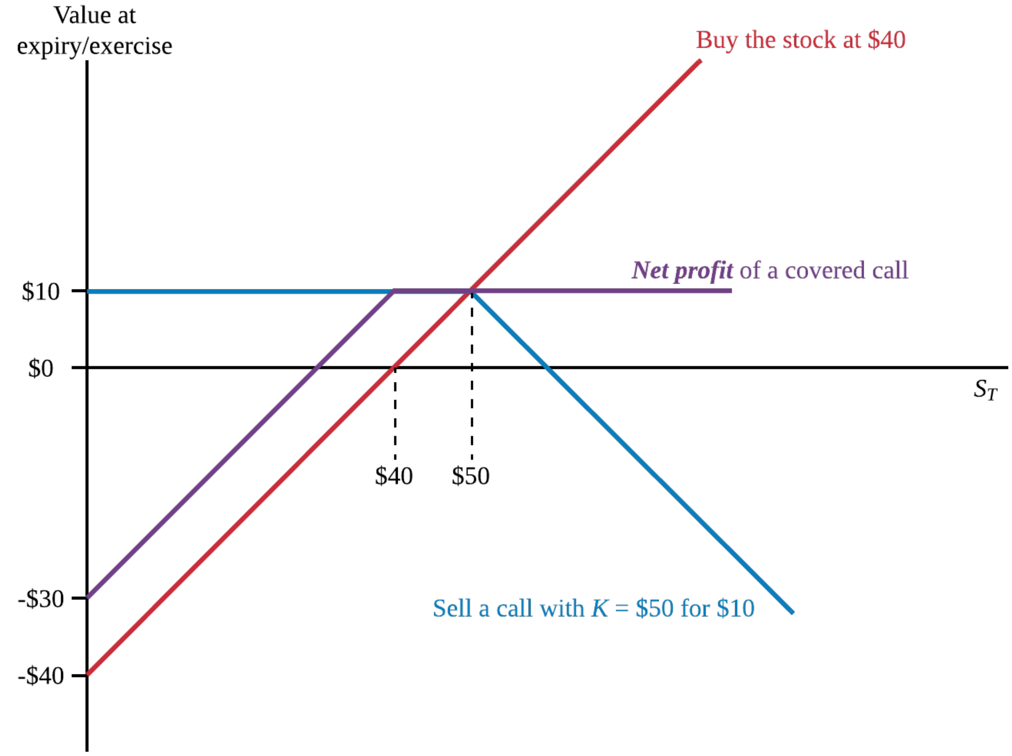


Note that this has the same basic payoff shape as a call option (so this trade can be thought of as a “synthetic” call), and in fact has the same payoff as long call combined with a zero-coupon bond that pays $\$K$ at time T.

Since investors are buying puts in the protective put trade, they must pay option premium, leading to the following example profit diagram.



Another common strategy is called a **covered call** where a short call option is combined with a long position in the underlying. The sale of the option generates premium for the seller. The consequence, however, is that the seller loses out on large upside moves in the stock.



There are many other strategies including so-called a bull spread, bear spread, strangle, butterfly, iron condor, and so on.

As we saw above, a protective put strategy at a strike of K has the same payoff as a portfolio consisting of a call option with a strike of K and a zero coupon bond that pays K at expiry. Since they have the same payoffs, they must have the same value, otherwise arbitrage exists.

This leads to Put-Call parity.

Definition 14.1.4
Put-Call Parity

Put-Call Parity is defined by the formula

$$P_t + S_t = C_t + Ke^{-r(T-t)}$$

where P_t, C_t are the prices of put and call with the same underlying, strike, and maturity. S_t is the price of the underlying, and r is the risk-free interest rate on zero-coupon bonds maturing at time T.

Rearrangements of this formula lead to replicating portfolios and synthetic options, which we will do below. This is the *most fundamental option relationship*.

The formula above actually applies to stocks that do not pay dividends. There are other forms of put-call parity in more general cases

Formula 2

In general, if we let $V_t(X)$ be the time-t price of the time-T payoff X, then

$$P_t + V_t(S_T) = C_t + V_t(K) \text{ or } C_t - P_t = V_t(S_T) - V_t(K)$$

So, if the stock pays continuous dividend at a rate of δ and risk-free asset pays continuous interest at a rate r, we have

$$C_t - P_t = S_t e^{-\delta(T-t)} - Ke^{-r(T-t)}$$

If the stock pays discrete dividends and the risk-free asset pays discrete coupons, we have

$$C_t - P_t = [S_t - PV_t(Div_{t,T})] - [K - PV_t(Coupon_{t,T})]$$

In this course, we will largely focus on the case where the stock pays no dividends, and the risk-free asset pays no coupons. In other words, the formula given in Definition 14.1.4 will apply to our cases.

Put-call parity can be used to create an array of synthetic assets.

Definition 14.1.5
Synthetic Portfolios

In each of these examples, we can create synthetic version of the asset on the left, by trading the portfolio on the right.

Synthetic Stock

Owning a share at time-T without buying any share at time 0.

$$S_0 = C(K,T) - P(K,T) + Ke^{-rT}$$

Synthetic T-bill

Borrow (or lend) at the risk-free rate from the stock market.

$$Ke^{-rT} = S_0 + P(K,T) - C(K,T)$$

Synthetic Options

Having the effect of a long call/put by trading a put/call.

$$C(K,T) = S_0 + P(K,T) - Ke^{-rT}$$

$$P(K,T) = Ke^{-rT} + C(K,T) - S_0$$

14.2 Option Pricing

Like put-call parity, some option price properties hold for any pricing model. If a pricing model produces option prices that violate these properties, the model is wrong. For example, American options cannot be cheaper than the corresponding European options since American-style options can be exercised at maturity, with more possible exercising times. A call option price C must satisfy $\max(S_t - K, 0) \leq C \leq S_t$ since exercising a call gives you a share at best or the option payoff at worst.

There are several other relations connecting the strike prices and the option value that must hold.

For strike prices $K_1 < K_2 < K_3$, the price of the corresponding call options must satisfy:

- $C(K_2) \leq C(K_1)$ (Monotonic)
- $C(K_1) - C(K_2) \leq K_2 - K_1$ (Steepness)
- $\frac{C(K_2)-C(K_1)}{K_2-K_1} \leq \frac{C(K_3)-C(K_2)}{K_3-K_2}$ (Convexity)

Similar results hold for put prices. For $K_1 < K_2 < K_3$, the put price P satisfies the following properties:

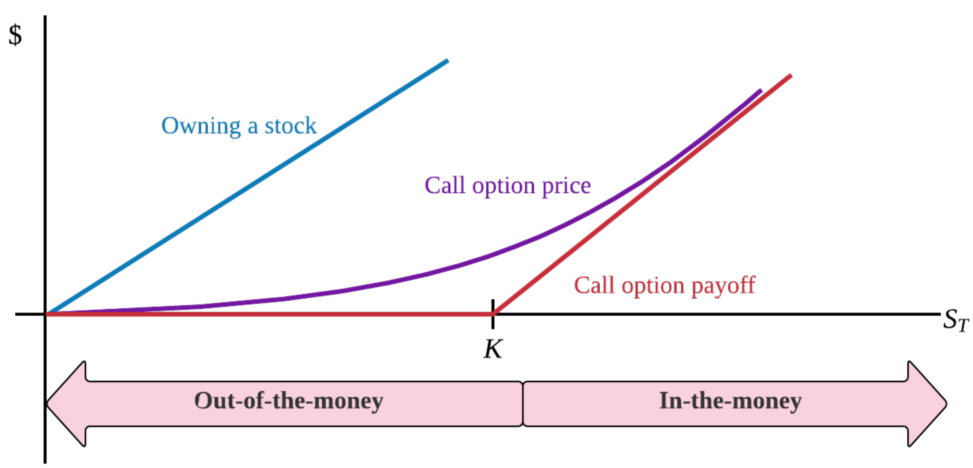
- $P(K_1) \leq P(K_2)$
- $P(K_2) - P(K_1) \leq K_2 - K_1$
- $\frac{P_t(K_2)-P_t(K_1)}{K_2-K_1} \leq \frac{P_t(K_3)-P_t(K_2)}{K_3-K_2}$

EXERCISE

Try proving the inequalities by constructing arbitrage portfolios.

While the above relations compare the value of an option as a function of different strike prices, it is often more interesting to have a sense of the value of an option as the value of the underlying changes.

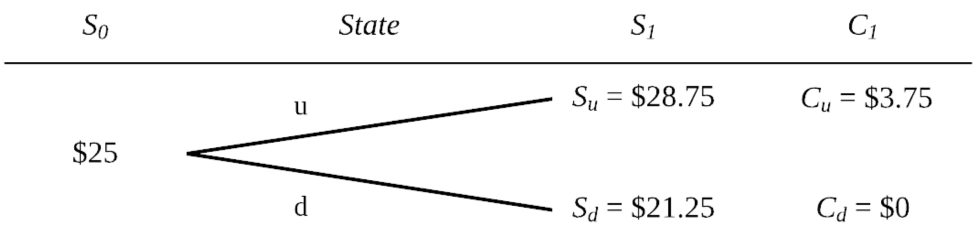
Generally, the value of a call option prior to expiry as a function of the value of the underlying is similar to what is shown below.



Actually valuing options prior to maturity is a challenging problem. We begin with a simplified model that turns out to be surprisingly useful.

Binomial Option Pricing Model

Suppose a stock is worth \$25 today and in one period will either be worth 15% more or 15% less. The continuously compounded risk-free rate is 5%. What is the value of an at-the-money call option?



We will try to find C_0 by constructing a replicating portfolio. This is a portfolio that replicates the payoffs C_1. Note that C_0 equals the initial cost of the replicating portfolio.

At $t = 0$ we hold Δ units of stock and $\$B$ amount of zero-coupon-bond. Depending on the option, Δ and/or $\$B$ may be positive or negative. To replicate the payoff at $t = 1$, we solve

$$\begin{cases} (28.75)\Delta + e^r B = 3.75 \\ (21.25)\Delta + e^r B = 0 \end{cases} \Rightarrow \begin{cases} \Delta = 0.5 \\ B = -\$10.11 \end{cases}$$

The price of the option is

$$C_0 = S_0\Delta + B = \$2.38$$

What are the probabilities of the states? It's not needed to price the option!

Option pricing by valuing the replicating portfolios usually consist of "primitive securities" that are easier to value. The replicating portfolio is indeed a *hedging or replicating portfolio*. The last example illustrates *Delta-hedging*

Formula 3

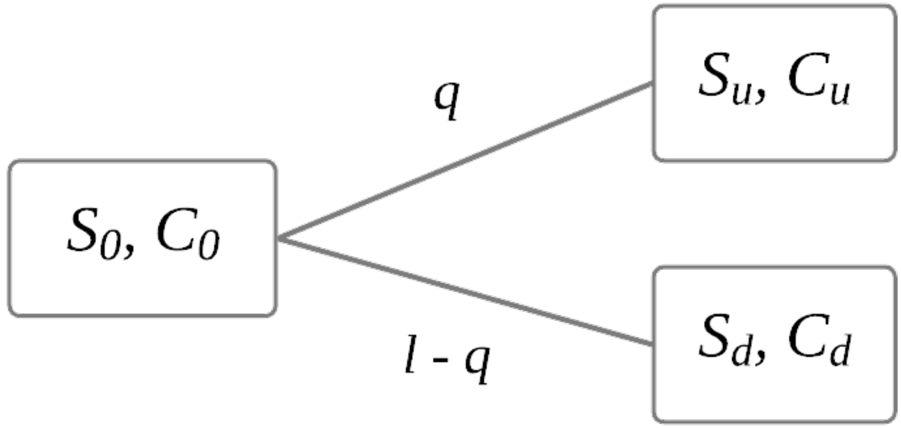
$$\Delta = \frac{C_u - C_d}{S_u - S_d}$$

Holding Δ units of S_0 hedges against changes of C_1 due to changes in S_1. The hedging portfolio replicates the payoff of the option in all economic states (in this binomial model).

We will now show a completely different way to value options that is sometimes more computationally convenient.

Risk-Neutral Valuation

We start with the binomial tree of values. We are looking for a formula for C_0.



Naively, we might expect the value of the option to simply be the expected present value of future cash flows, which would be given by

Formula 4

$$C_0 = \frac{1}{1+r}[q \times C_u + (1-q)C_d]$$

or, using continuous compounding of interest rates:

$$C_0 = e^{-rt}[q \times C_u + (1-q)C_d]$$

It turns out that this works if we choose the transition probability q such that it satisfies

Formula 5

$$S_0 = \frac{1}{1+r}[q \times S_u + (1-q)S_d]$$

or

$$S_0 = e^{-rt}[q \times S_u + (1-q)S_d]$$

if using continuous compounding of rates.

The value q chosen this way is called a **risk-neutral** probability.

Risk-neutral valuation is a convenient mathematical tool. In this risk neutral world, all assets, risky or not, are expected to earn the risk-free rate.

EXERCISE

Compute q and then C_0 using the data from the previous example. Did you get the same option value?

The Black-Scholes Model

The binomial model seems extremely artificial since in the real world, there are far more than 2 possible outcomes in the future. However, we can extend the model by adding more time steps, which then increases the number of possible future outcomes. As we increase the number of time steps indefinitely, the model converges towards a continuous model. In this continuous world (under suitable assumptions) there are closed-form formulas for certain option values. Below is the so-called Black-Scholes formula for the value of a call option (sometimes also called the "Black-Scholes model price")

Formula 6 The price of a call option is given by

$$C = S_0 \cdot N(d_1) - Ke^{-rT} \cdot N(d_2)$$

where

$$d_1 = \frac{\ln(\frac{S_0}{K}) + \left(r + \frac{\sigma^2}{2}T\right)}{\sigma\sqrt{T}}, \text{ and } d_2 = d_1 - \sigma\sqrt{T}$$

where

- S_0: current stock price
- K: strike price
- r: annualized continuously compounded risk-free rate
- σ: annualized volatility of stock return
- T: time to maturity (in years)
- $N(\cdot)$: standard normal CDF

REMARK

In the limit as we increase the number of time steps in the binomial model we presented above, the distribution of asset prices converges to a log-normal distribution. The Black-Scholes formula applies to assets that follow such a distribution.

This log-normal distribution observation can be useful in pricing some exotic options. We can take sample prices from a lognormal distribution and price the option via essentially a simulation method. Using simulation is a topic for courses such as STAT 340 or STAT 341.

14.3 Options and Corporate Finance

Recall our basic balance sheet equation $V = D + E$. Option theory gives us a whole new way to think about this relationship.

For a limited liability company, when the debt matures

- If $V > D$, the shareholders receive $V - D$
- If $V \leq D$, the company is bankrupt so the shareholders receive 0 and the debtholders receive the residual value.

Therefore, the shareholders receives $\max(0, V - D)$, which is the payoff of a call option on the firm with a strike price of D! So, one view of the firm is that the shareholders have a call option on the company, which is viewed as belonging to the bondholders. Shareholders exercise the call if the company performs well and has higher value than D. Since call option values increase as volatility increases, shareholders may see an advantage to high volatility – i.e. taking big risks. Debt holders, as sellers of this option have the opposite view. This reflects the agency problem introduced earlier.

This new viewpoint gives a completely different perspective on valuing equity. Although we will not cover it here, note that the firm is bankrupt when this option is out of the money. So, we can determine the probability of bankruptcy by computing the probability that the option is out of the money, which we can do with the log-normal model.

Executive Stock Options (ESOs)

Executive stock options are technically warrants on the employer's shares – although will not pursue that distinction here. They typically are inalienable, meaning they can only be exercised but not sold. The claimed purpose of these options is to align shareholders and CEO's interests. There are a number of criticisms associated with them, such as "option backdating" scandals which you can read about with a quick Google search.

As the next example shows, these options can have considerable value to recipients.

Example 2

The CEO of XYZ Inc. is granted 2 million at-the-money stock options. These options all have a strike price of \$50 and 4-year expiration. Assume that the risk-free rate is 5% and the annual volatility of XYZ's equity is 40%.

Calculate the option value using the Black-Scholes formula.

Solution: The Black-Sholes formula is

$$C_0 = S_0 \times N(d_1) - Ke^{-rT} \times N(d_2)$$

where

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}, \; d_2 = d_1 - \sigma\sqrt{T}$$

Here, we have the values

- At-the-money: $K = S = \$50$, $T = 4$, $r = 5\%$, $\sigma = 40\%$
- $\ln\left(\frac{S}{K}\right) = \ln(1) = 0$
- $d_1 = 0.65$, $d_2 = -0.15$
- $N(d_1) = 0.7422$, $N(d_2) = 0.4404$

$$C = (50)(0.7422) - (50)(e^{-0.05\times4})(0.4404) = \$19.08$$

Total option value = 2 mil × \$19.08 = \$38.16 mil.

14.4 Other Option Applications

An insurance contract can be viewed in terms of option theory. Consider a car with car insurance. The car value insured equals $\$X$, say \$90K, and the monthly premium equal \$P, say \$300/month, and the deductible equals \$D, say \$2K. If the car is damaged and its value falls below \$88K, you make an insurance claim which will repair or replace the car, bring its value back to \$88k. This is like a protective put strategy. You pay premiums for the right to get paid when asset value falls below a threshold. Essentially, you bought an asset of value $\$X$ and buy a put with strike $K = \$(X - D)$. Although similar to an put, there are some important differences. For one, you have monthly premiums for insurance versus the single option premium paid up-front. Individual car insurance is also not tradable in the financial market. Finally, making insurance claims can be more complicated than exercising options.

Options may also appear as part of capital budgeting. An important option when valuing a start-up is *the option to expand*. Such option could turn a negative NPV to positive.

Example 3 Albert wants to open a new restaurant and he has done his research:

- Initial cost is \$70,000
- PV of future cash flows is \$50,000

The NPV is -\$20,000, so it seems Albert should give up on this idea, but...

- If the first one is profitable, Albert plans to open 30 more restaurants after 4 years
- Assume the annual discount rate is 10% and the volatility of future cash flows is 50%

Albert should value the option to expand too. For example, using the Black-Scholes model. Consider the expansion:

- The total cost should we chose to expand is $30 \times \$70,000 = \$2,100,000$
- The expected PV of future cash flows as of year 4 is $30 \times \$50,000 = \$1,500,000$
- The expected PV of future cash flows as of today is $\frac{\$1,500,000}{(1+10\%)^4} = \$1,024,520$

But those are just the *expected* future cash flows. The actual cash flows may be higher if the first restaurant is successful. The expansion is essentially a call option on the future cash flows.

- Current value of the underlying asset is $S_0 = \$1,024,520$
- Strike $K = \$2,100,000$
- Maturity $T = 4$ years
- Annual volatility $\sigma = 50\%$
- Annual risk-free rate $r = 10\%$

We thus have sufficient information to apply the Black-Scholes formula.

Recall the pilot restaurant's $NPV = -\$20,000$ and the option to expand is a call option with $S_0 = \$1,024,520$, $K = \$2,100,000$, $T = 4$, $\sigma = 50\%$, $r = 10\%$.

$d_1 = \frac{\ln\left(\frac{S_0}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} = 0.18$, $d_2 = d_1 - \sigma\sqrt{T} = -0.82$

$N(d_1) = 0.5714$, $N(d_2) = 0.2061$

$$C = (1024520)(0.5714) - (2100000)(e^{-0.1\times 4})(0.2061) = \$295,289$$

So NPV + call option ≥ 0, should proceed with the pilot restaurant.

In the example above, some assumptions for the Black-Scholes model may not hold, for example

- The underlying asset is not traded
- The underlying asset price does may not follow a continuous log-normal process
- The underlying asset price's volatility may not be constant over time, and our estimate of 50% may have been wrong to start with

The pilot restaurant may fail, but the option takes advantage of the upside potential. They view the future cash flows as the underlying asset with some volatility,

REMARK

An interesting paradox (when attracting venture capital): opening a single restaurant is unlikely to attract any investor, potentially opening 30 more such restaurants can be a very attractive investment.

14.5 Exotic Options

Some options have complex (or exotic) payoffs, like *path-dependent payoffs*, where the payoff depends not only on the final price S_T, but on its path to S_T too. Here is a list of a variety of exotic options

Definition 14.5.1
Some Exotics

Bermudan Option

The option has specific exercise times, (possibly monthly or annually), before maturity.

Binary or Digital Option

These options have a fixed payoff (say \$1) if the underlying value is greater or lower than a threshold. For example, a digital call with a strike of \$25 will pay \$1 as long as the underlying is valued above \$25 at expiry.

Asian Options

The payoff depends on the average price $\overline{S_T}$ prior to maturity. For a floating-strike Asian call $\max(0, S_T - \overline{S_T})$, and for a fixed-strike Asian call $\max(0, \overline{S_T} - K)$.

Lookback Option

The payoff depends on the max price S_{max} or min price S_{min} prior to maturity. With a floating lookback call $\max(0, S_T - S_{min})$, and with a fixed lookback call $\max(0, S_{max} - K)$.

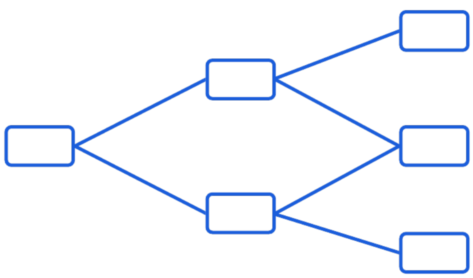
Barrier Option

Payoff is activated/deactivated when its underlying reaches some barriers.

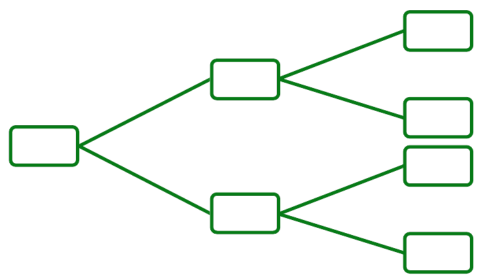
The Black-Scholes model cannot be used to price many such exotic options. Instead, the binomial option pricing model can be used as an alternative. It is flexible, easy to visualize, and can be extended to multiple periods.

When constructing a binomial tree, the tree can either rejoin (or recombine) or not, depending on the model and what is needed to price the given security.

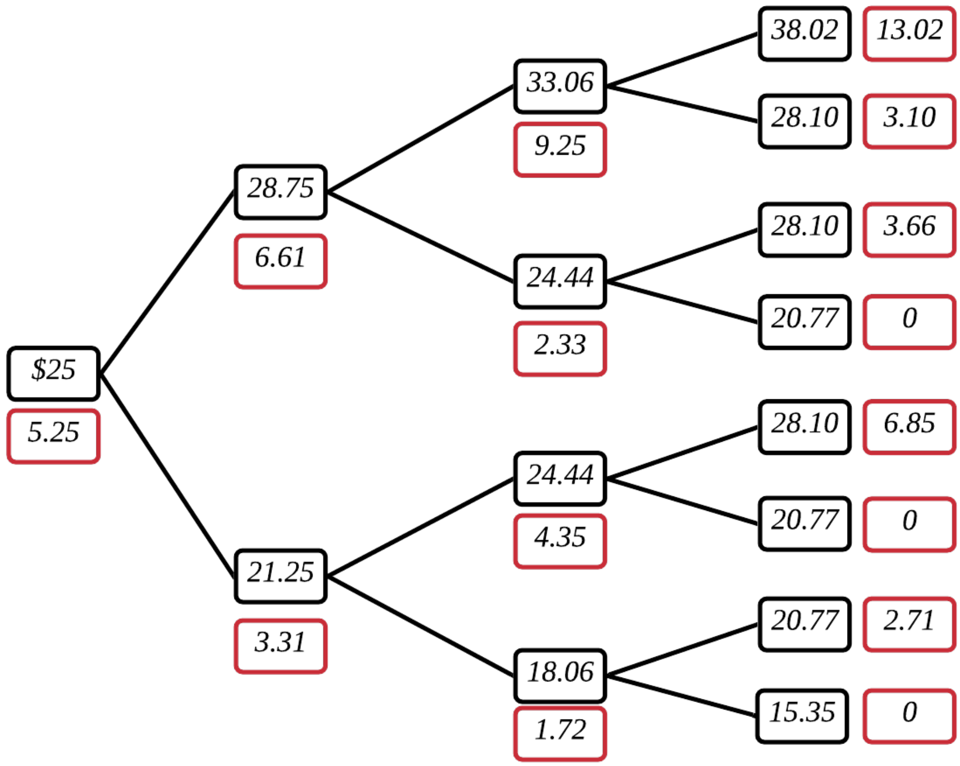
Rejoining tree (lattice)



Non-rejoining tree



Consider a three-period binomial model of lookback call option prices. The strike of this option is $\min(S_t)$ over the life of the option. Here, we assume asset prices either rise or fall by 15% each period. Assume $r = 5\%$ compounded continuously. The terminal values are the computed payoffs. The earlier values are computed with risk neutral valuation.



REMARK

Non-recombining trees have the unfortunate property that there are 2^n nodes after n time steps. This can quickly get computationally unwieldy. Simulation provides an alternative approach in such cases.

15 Forwards

15.1 Forward Contracts

Loosely speaking, there are two types of transactions (or contracts): spot transactions and forward transactions.

Definition 15.1.1
Spot and Forward Transactions

A spot transaction is one where the terms of the deal, the payment of money and the delivery of the commodity occur essentially simultaneously.

A forward transaction is one where the terms of the deal are determined today, but the payment and the delivery of the commodity occur at some point in the future.

Note that a forward contract is not an option. Once the terms are finalized, the buyer must buy and the seller must sell and deliver the goods.

Definition 15.1.2
Spot and Forward Prices

The spot price is the price of a commodity under a spot transaction (i.e. for delivery today) while the forward price is the price we agree today for delivery later. We often specify this more precisely by stating a "1-year forward price" to represent the forward price for a commodity for delivery in 1 year, for example.

At first glance, there does not seem to be a reason for the forward price to be different from the spot price, but this is not the case. Suppose as a source of example, the spot and 1-year forward price of gold are both \$1,000 per ounce. Suppose further, the 1-year risk free interest rate is 5% (compounded annually). The smart gold investor would then sell their gold today for \$1000 and simultaneously enter into a 1-year forward contract to buy gold for \$1000 in 1 year. They then invest the \$1000 proceeds from the spot sale which will grow to \$1050 in 1 year. Then, once the year has passed, they purchase the gold under the forward contract for \$1000, leaving them with 1 oz of gold plus \$50. If they simply held the gold, in 1 years time, they would have the gold and no cash.

Many forward price relationships can be established via arbitrage trades. Suppose as above the 1-year forward price for gold is \$1100. The arbitrage trade would be to borrow \$1000 now to buy gold in the spot market and then sell a forward contract at \$1100. In 1 year, we sell the gold for \$1100 and repay the loan now worth \$1050 (including interest) yielding an arbitrage profit of \$50. Thus, the forward price cannot be higher than \$1050. If it is lower, then the opposite trade generates arbitrage profits.

This motivates the forward price relation given by

Formula 1

Forward Price 1:

The forward price for an asset for delivery in $t-$ years is given by

$$F = S \times (1 + r)^t$$

where S is the current spot price, and r the risk free interest rate.

If r is expressed in continuous compounding, we get the formula

$$F = Se^{rt}.$$

15.2 Dividend Paying Stocks

Suppose the asset in question is a stock currently priced at \$100 that pays a semi-annual dividend of \$2.00. Note that the holder of the actual stock would receive the dividends, while the holder of a 1-year forward contract would not. In this case, the forward contract seems less valuable then owing the actual stock. Therefore, we need to adjust the pricing formula accordingly.

Formula 2

Forward Price 2:

The forward price for an asset for delivery in $t-$ years is given by

$$F = (S - D) \times (1 + r)^t$$

where S is the current spot price, r the risk free interest rate and D is the present value of the dividend payments.

If r is expressed in continuous compounding, we get the formula

$$F = (S - D)e^{rt}.$$

If in addition we assume a continuous dividend paid at a rate of d, we get the formula

$$F = Se^{(r-d)t}.$$

Example 1

Find the 1-year forward price of a stock priced at \$100 today that pays a semi-annual dividend of \$2.00. Assume a risk free rate of 6% compounded continuously. (Assume delivery of the stock occurs on the ex-dividend date of the second dividend.)

Solution: First we compute the PV of the 2 dividends that the forward purchaser will not receive.

$$PV(Div_1) = 2e^{-6\% \times 0.5} = \$1.94 \quad \text{and} \quad PV(Div_2) = 2e^{-6\%} = 1.88$$

Thus $D = 1.94 + 1.88 = 3.82$. Therefore, the forward price equals

$$F = (100 - 3.82) \times e^{6\%} = \$102.12.$$

REMARK

This pricing model applies to any asset that generates cash flows, including stocks that pay dividends and bonds that pay coupons.

15.3 Storage costs

A further complication arises when the asset requires storage that may cost money. Consider oil as an example. A holder of oil will need to store it in a depot somewhere, and may also have to pay insurance to provide protection should some disaster occur. This costs money that the forward purchaser would not incur, making the forward contract seem relatively more valuable. We need to adjust for this.

We call all these extra costs associated with actually holding an asset the carrying costs.

Formula 3

Forward Price 3:

The forward price for an asset for delivery in $t-$ years is given by

$$F = Se^{(r+q)\times t}.$$

where S is the current spot price, r the risk free continuously compounded interest rate and q is the continuously compounded additional carrying costs.

REMARK

The key observation in all of the formulas is the idea that in order to prevent arbitrage, the future value of all the cash flows associated with owing the asset (both inflows and outflows) should equal the forward price. Applying this principle yields forward price formulas in an array of scenarios.

15.4 Additional Complications

The formulas in the earlier sections ignore some complicated realities that make real-world forward pricing more challenging. These include:

1. Counterparty risk. This is the risk that one of the parties defaults on their obligations under the forward contract. For example, a buyer who enters into a 1-year forward contract to buy gold at \$1000 per oz *thinks* they will be able to buy gold for \$1000, but if the seller defaults, they will not be able to do so.

2. Uncertain dividends. In the case of a dividend paying stock, we may have a forecast for the dividend payments, but dividends are not guaranteed – they can change. Thus, we may enter into a forward agreement with a given forecast of cash flows that turns out not to occur.

3. Uncertain carrying costs. Like with dividends, we may have a forecast for what the carrying costs will be, but the actual costs may vary.

4. Uncertain interest rates. Interest rates can change, so the rate at which we priced the forward contact may change over its life. This risk grows as the length of the contract grows.

5. Storability. All of the formulas above are derived from arbitrage trades that require the asset to be both storable and shortable. Commodities like electricity have neither of these properties making this logic invalid. Similar problems occur with assets that are perishable, such as milk.

6. Convenience yield. These trades above assume the investor has no particular use for the commodity over the life of the contract. In some cases, there can be value in actually holding the commodity, as you can take advantage of temporary shortages and spot price spikes. The advantage to holding the actual commodity is sometimes called the convenience yield.

Some of these risks can be mitigated somewhat. In the case of counterparty risks, traders often need to post collateral in order to enter into a contract. The futures markets are designed to address this risk, but these contracts are beyond the scope of this course.

16 Utility Theory

16.1 Introduction to Utility Theory

So far, all our analysis of valuing projects and securities can be thought of as sitting under the broader umbrella of making decisions under uncertainty. We will explore this from a different perspective here.

Definition 16.1.1
Expected Value Principle

The expected value principle states that you should choose the investment with the highest expected value.

Note that expected value is sometimes called the actuarial value or fair value.

If the expected value principle is true, then

- One is willing to pay up to $E[X]$ in a gamble with random payoff X
- One is indifferent between receiving $E[X]$ or random payoff X

Note that this is largely what we have been doing all a long. We chose projects only if they have a positive expected NPV, and the higher the expected NPV the better.

However, there are important examples where this thinking seems to fall apart. Consider insurance. Insurance companies will only sell insurance if the expected value from doing so is positive. Therefore, all the customers must be buying a negative expected value instrument. So, no one should buy insurance. But people do. Why?

Consider the St. Petersburg Paradox, which was proposed by Nicolaus Bernoulli in 1713 and solved by Daniel Bernoulli (Nicolaus' brother) in 1738.

Here, let X be the "game" or "gamble" that works as follows: We flip a fair coin. If it turns up heads (H) you win \$2. If it turns up tails (T), we flip again. We continue to flip until the first H appears. At that point, you win \$$2^n$ where n is the number of flips. (So, TTH wins \$8). How much should you pay to play this game? We compute the expected payoff. We let p_k be the probability that the first H appears on flip k, then

$$E(X) = \sum_{k=1}^{\infty} 2^k p_k = \sum_{k=1}^{\infty} \frac{2^k}{2^k} = \infty.$$

So, this game seems to have an arbitrarily large payoff. Buy who would pay a massive sum of money to play a game that has a 50% change of only generating \$2? Something's wrong.

We will look at these problems from the economic principle of expected utility, which was formalized by John Von Neumann and Oskar Morgenstern around 1944, leading to the Von Neumann-Morgenstern expected utility theory. This is an important topic in economics, actuarial science, and philosophy (so-called utilitarianism).

Definition 16.1.2
Utility Theory

Utility theory suggest that instead of measuring the monetary value of a gamble, measure the "level of satisfaction" (i.e., happiness, usefulness, enjoyment, pleasure, etc.). What matters to a decision maker is the "utility" of consuming money, not the monetary amount itself.

The expected utility theory should be able to explain why one additional dollar means less to a rich person than to a poor person (note that a similar economic concept is diminishing marginal return), and why sometimes it might make sense to lose a certain amount to avoid uncertainty.

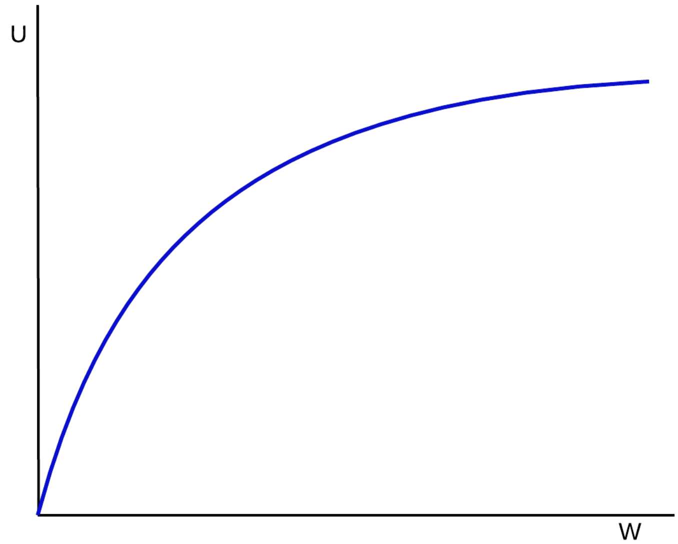
Definition 16.1.3
Utility Function

Mathematically, we define a utility function $u(\cdot)$, where $u(x)$ is the "utility" of consuming wealth $\$x$. A unit of utility is called a **utile**. 1 utile has no real meaning; the relative magnitude is important.

In this class, we follow some general properties of utility functions:

- The more the better, so $u(\cdot)$ is increasing. For two amounts $\$x > \$y$, $u(x) > u(y)$. If $u(\cdot)$ is differentiable, then $u'(w) > 0$.
- Decreasing marginal utility of wealth, so $u(\cdot)$ is concave. If an investor is wealthier, each additional dollar has less utility. If $u(\cdot)$ is twice differentiable, then $u''(w) < 0$.
- Instead of maximizing expected value $E[X]$, one may maximize expected utility $E[u(X)]$.

For convenience we often assume $u(0) = 0$ (so no money has zero utility), yielding a typical utility function of the form



For ease of discussion, will often use the term lottery to describe any uncertain payoff.

Definition 16.1.4
Lottery

A **lottery** or **gamble** will represent a set of cash flows along with the associated probability distribution.

Simple Lottery

A simple lottery is denoted by $L(x, y; p)$, meaning payoff x with probability p and pay off y with probability $1 - p$.

Compound Lottery

A compound lottery can be expressed as $L(X, Y; p)$ where X and Y are themselves lotteries (simple or compound).

16.2 Axioms of Cardinal Utility

Our actual goal here is to try to understand why individuals make the decisions they do, and how we can better make decisions now. In other words, we are seeking a "decision rule" that determines when to select one option over another. That rule can be captured in a preference relation

Definition 16.2.1
Preference Relation

A preference relation is an ordering of opportunities by individuals based on their preferences.

Given options X and Y (typically, these will be lotteries) we write

- $X \succ Y$: X is strictly preferred to Y
- $X \sim Y$: X is indifferent to Y
- $X \succeq Y$: X is weakly preferred to Y

Preference relations are used to rank investment opportunities.

It turns out, to have a sensible theory, we need an individual's preference relation to follow some axioms:

Axiom 1: Completeness/Comparability

Given any two lotteries X and Y, exactly one of the following holds:

- $X \succ Y$
- $X \prec Y$
- $X \sim Y$

The decision maker can rank all investment opportunities. One can always make a decision on two alternatives.

Axiom 2: Transitivity

Given any three lotteries, X, Y, and Z, the following relationships must hold:

- $X \succ Y \ \& \ Y \succ Z \Rightarrow X \succ Z$
- $X \sim Y \ \& \ Y \sim Z \Rightarrow X \sim Z$

Axiom 3: Continuity

If $X \succ Y \succeq Z$ or $X \succeq Y \succ Z$, then there exists a unique probability p between 0 and 1 such that $Y \sim L(X, Z; p)$.

Axiom 4: Independence/Substitution

Given two lotteries X and Y. For any lottery Z and any probability $0 < p < 1$

- $X \succ Y \Leftrightarrow L(X, Z; p) \succ L(Y, Z; p)$
- $X \sim Y \Leftrightarrow L(X, Z; p) \sim L(Y, Z; p)$

Mix each of the two lotteries (X and Y) with a common lottery (Z). The preference ordering of the resulting mixtures is independent of the third lottery Z and the mixing probability p. This is very useful in establishing preferences between more complex alternatives based on preferences between simpler alternatives.

Axiom 5: Monotonicity

Given four lotteries X, Y, Z, and W, and probabilities p_1, p_2. Suppose $X \succeq Y \succeq Z$ and $X \succeq W \succeq Z$ with $Y \sim L(X, Z; p_1)$ and $W \sim L(X, Z; p_2)$, then

- $p_1 > p_2$ if and only if $Y \succ W$
- $p_1 = p_2$ if and only if $Y \sim W$
- $p_1 < p_2$ if and only if $Y \prec W$

The connection between utility functions and preference relations is established in the following theorem

Theorem 1 Preference relation $\succeq$ satisfies Axioms 1-5 (and other technical assumptions) if and only if there exists a utility function $u(x)$ such that

$$E[u(X)] \geq E[u(Y)] \text{ if and only if } X \succeq Y$$

Furthermore, $u(x)$ is unique up to an affine transformation. That is, the utility function

$$u^*(x) = a \cdot u(x) + b,\ a > 0$$

expresses exactly the same preferences as $u(x)$ does.

The power of this theorem is that once we have an individual's utility function, we can determine which outcomes they prefer simply by comparing expected utilities. And we can find the best alternative by maximizing expected utility.

REMARK

1. Expected value is a special case of expected utility. For utility function $u(x) = x \Rightarrow E[u(X)] = E[X]$
2. Individuals can make different decisions because their utility functions may be different

16.3 Utility Functions

Let's look at some decision making given some explicit utility functions.

Example 1 Consider the following lotteries:

- A: Payoffs of \$25, \$64, and \$100 with probabilities 30%, 10%, and 60%
- B: Payoffs of \$16, \$36, and \$121 with probabilities 30%, 10%, and 60%

Which lottery is preferred if $u(x) = \sqrt{x}$?

Solution: Given a utility function $u(\cdot)$ and lottery X, a rational investor maximizes $E[u(X)]$.

$$E[u(A)] = (0.3)(\sqrt{25}) + (0.1)(\sqrt{64}) + (0.6)(\sqrt{100}) = 8.3$$
$$E[u(B)] = (0.3)(\sqrt{16}) + (0.1)(\sqrt{36}) + (0.6)(\sqrt{121}) = 8.4$$

So $A \prec B$, and lottery B is preferred.

EXERCISE

Which lottery is preferred if $u(x) = \ln(x)$?

Different utility functions allow for different preferences!

The philosophy of utility theory now allows us to resolve the St. Petersburg Paradox

Suppose your utility function is $u(x) = \ln(x)$, how much are you willing to pay for the St. Petersburg Paradox game? We compute the expected utility

$$\begin{aligned} E[u(X)] &= \frac{u(\$2)}{2} + \frac{u(\$4)}{4} + \frac{u(\$8)}{8} + \cdots \\ &= \sum_{n=1}^{\infty} \frac{\ln(2^n)}{2^n} \\ &= \ln(2) \sum_{n=1}^{\infty} \frac{n}{2^n} \\ &= \ln(2) \times 2 = \ln(4) \end{aligned}$$

So, the game provides $\ln(4)$ utiles. We need to pay $\$x$, where

$$u(x) = \ln(x) = \ln(4) \Rightarrow x = \$4$$

If $u(x) = \ln(x)$ is truly your utility function, then you will pay at most \$4. Indeed, the game does not have infinite "value".

Definition 16.3.1
Certain Equivalent

Given lottery X, the fixed monetary amount with the same utility as X is called the **certainty equivalent** of X, denoted by $CE(X)$. For the same X, **different** $u(\cdot)$ **can result in different** $CE(X)$.

In the St. Petersburg paradox's payoff above, the certainty equivalent is \$4.

The connection between certain equivalents and gambles can be summarized with the formula

Formula 2

$$u(CE(X)) = E[u(X)]$$

or, typically equivalently

$$CE(X) = u^{-1}(E[u(X)])$$

So,

- Maximizing $E[u(X)] \Leftrightarrow$ maximize $u(CE(X))$
- If $u(x)$ is increasing, then maximizing $u(CE(X)) \Leftrightarrow$ maximize $CE[X]$, so maximizing $E[u(X)] \Leftrightarrow$ maximizing $CE[X]$

Let's revisit the previous example, in terms of certain equivalents.

Example 2

Consider the following lotteries:

- A: Payoffs of \$25, \$64, and \$100 with probabilities 30%, 10%, and 60%
- B: Payoffs of \$16, \$36, and \$121 with probabilities 30%, 10%, and 60%

Calculate $CE(A)$ and $CE(B)$ if $u(x) = \sqrt{x}$.

Solution: We have calculated that $E[u(A)] = 8.3$ and $E[u(B)] = 8.4$. Since $u^{-1}(x) = x^2$ we get

$$CE(A) = u^{-1}(E[u(A)]) = 8.3^2 = \$68.89$$

$$CE(B) = u^{-1}(E[u(B)]) = 8.4^2 = \$70.56$$

16.4 Risk Aversion

Once we know a person's utility function, we can determine their preferences, and hence their responses to risk.

Definition 16.4.1
Risk Averse

Let X be a random payoff and $E[X]$ be its actuarial value. An investor is called **risk averse** if and only if

$$u(E[X]) > E[u(X)], \text{ or equivalently, } E[X] > CE(X)$$

They are called **risk seeking** if

$$u(E[X]) < E[u(X)], \text{ or equivalently, } E[X] < CE(X)$$

and are called **risk neutral** if

$$u(E[X]) = E[u(X)], \text{ or equivalently, } E[X] = CE(X)$$

So, risk averse people prefer the certain equivalent to the gamble (they will give up value to get certainty – i.e. remove risk), while risk seeking people prefer the gamble to the certain equivalent.

For everything that follows, we assume $u'(x) > 0$, i.e. an increasing utility function, as the alternative seems totally irrational.

It can be shown with Jensen's inequality that an individual is risk averse if and only if $u(x)$ is concave (i.e. $u''(x) < 0$), and they are risk neutral if and only if $u''(x) = 0$.

We can illustrate this with the following example

Example 3 Calculate the expected utility of the lottery $X \sim L(\$2, \$0; 50\%)$ under the following 3 utility functions $u_1(x) = \sqrt{x}$, $u_2(x) = x$, $u_3(x) = x^2$

Aside: $u(E[X]) = u(\$1) = 1$ utile, for all 3 utility functions.

Solution:

$$E[u_1(X)] = (0.5)(\sqrt{0}) + (0.5)(\sqrt{2}) = 0.71 < 1$$

Risk averse utility, $X \prec E[X]$.

$$E[u_1(X)] = (0.5)(0) + (0.5)(2) = 1$$

Risk neutral utility, $X \sim E[X]$.

$$E[u_1(X)] = (0.5)(0^2) + (0.5)(2^2) = 2 > 1$$

Risk seeking utility, $X \succ E[X]$.

We can measure the level of risk aversion shown by a decision-maker via the the (Arrow-Pratt measure of) Absolute Risk Aversion (ARA) which is defined by

Formula 3

$$ARA(x) = \frac{u''(x)}{u'(x)}$$

The (Arrow-Pratt measure of) Relative Risk Aversion (RRA) is defined by

Formula 4

$$RRA(x) = x \cdot ARA(x)$$

These measures are named after Kenneth J. Arrow (1972's Nobel Memorial Prize in Economics) and John W. Pratt (Fellow of the American Statistical Association).

The idea is that the higher the value of the risk aversion measures, the more risk averse a decision maker is. Loosely speaking, the absolute measure of risk aversion measures how risk averse a decision maker is when the gambles are presented in absolute (i.e. dollar value) terms, while the relative measure of risk aversion measures how risk averse a decision maker is when the gambles are presented in relative (i.e. percentage change) terms

Example 4 Find the ARA/RRA for

- $u_1(x) = \sqrt{x}$

$$u_1'(x) = \frac{1}{2x^{1/2}},\ u_1''(x) = -\frac{1}{4x^{3/2}}$$

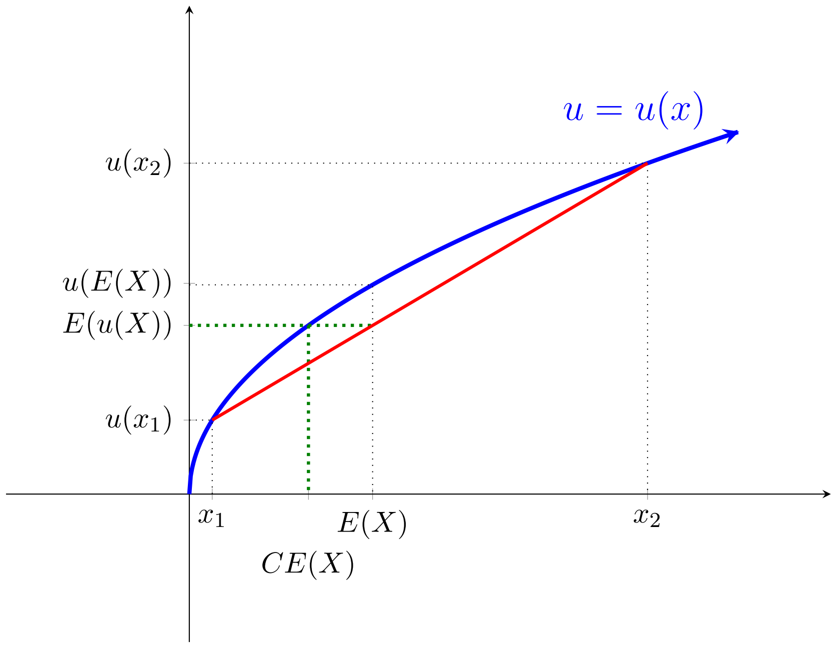
  Thus $ARA = \frac{1}{2x}$ and $RRA = \frac{1}{2}$.

- $u_2(x) = -e^{-0.0001x}$

$$u_2'(x) = -10^{-3}u_2(x),\ \ u_2''(w) = 10^{-6}u_2(x)$$

  Thus $ARA = 10^{-3}$ and $RRA = 10^{-3}x$.

This idea can be captured graphically as follows.



Here X is a gamble with outcomes x_1 and x_2. Notice that since $u'' < 0$, we have $CE(X) < E(X)$. The risk measure $ARA(x)$ captures the "curvature" of the utility function, and hence the risk aversion of the decision maker.

EXERCISE

What does the corresponding graph reveal in the case $u(x) = x$? What about when $u'' > 0$?

Quadratic Utility

A utility function of the form

$$u(x) = x - \alpha x^2, \ x < \frac{1}{2\alpha}, \ \alpha > 0$$

is called quadratic utility. Note that

$$u'(x) = 1 - 2\alpha x, \ u''(x) = -2\alpha$$

Since we want $u'(x) > 0$ this function is bounded above for $x < \frac{1}{2\alpha}$. Essentially, this only works for "small" gambles.

Note the following connection with mean-variance portfolio optimization (recall $\sigma_X^2 = E[X^2] - \mu_X^2$) To maximize utility, we consider expressions of the form

$$E[u(X) = E[X - \alpha X^2] = (\mu_X - \alpha\mu_X^2) - \alpha\sigma_X^2 \sim \mu_X - \alpha\sigma_X^2$$

as long as μ_X is small. But maximizing the expression on the right is exactly mean-variance optimization!!

Thus, maximizing quadratic utility $\Leftrightarrow$ mean-variance optimization. There is no assumption made on the distribution of X. It turns out that by using the second-order Taylor expansion, for "small gambles" any utility function is (approximately) quadratic. Therefore, for all investors, Mean-variance optimization makes sense, at least mathematically, when stock price changes are small (e.g., frequent rebalancing based on daily/monthly data).

Exponential Utility

A utility function of the form

$$u(x) = e^{-\alpha x},\ \alpha > 0$$

is called **exponential utility.** In order to compute expected utilities, we are lead to expressions of the form

$$E[u(X)] = E[-e^{-\alpha X}] = -M_X(-\alpha)$$

where $M_X(t)$ is the moment generating function (M.G.F.) of X. The MGF of any random variable X is defined as $M_X(t) = E[e^{tX}]$

Recall $CE(X) = u^{-1}(E[u(X)])$. Consider an exponential utility $u(x) = e^{-\alpha x}$ for some $\alpha > 0$, then

$$u^{-1}(x) = -\frac{\ln(-x)}{\alpha}$$

Now, suppose X is normally distributed, i.e., $X \sim N(\mu, \sigma^2)$, then the MGF of X is given by

$$M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

Therefore, if both of the above conditions hold, then

$$\begin{aligned} CE(X) &= u^{-1}(-M_X(-\alpha)) \\ &= -\frac{\ln\left(\exp\left(-\mu\alpha + \frac{1}{2}\sigma^2\alpha^2\right)\right)}{\alpha} \\ &= \mu - \frac{\alpha}{2}\sigma^2 \end{aligned}$$

So, maximizing utility is once again back to a mean-variance optimization problem.

In summary, if an investor has exponential utility and is facing normally distributed gambles, or for any investor facing small gambles, then mean-variance optimization is the correct decision rule.