

Quantitative Finance

Muchang Bahng

Spring 2024

Contents

1 Ancient History of Finance

Some may argue that this section is not relevant, but to me it is. Later on we will be analyzing data like money supply, interest rates, and stock prices. However, I find myself questioning where all this came from, and why do we actually need this kind of system? To answer this, we need to take a deep dive into history.

Let's develop a bit of a taxonomy of the major periods of **prehistory**.¹

1. (-3000 BCE) The **Stone age** is a prehistoric period during which stone was widely used to make stone tools, lasting for approximately 3.4 million years. It is divided into 3 subcategories.
 - (a) (- 9000 BCE) The **Paleolithic era** is the primitive use of stones.
 - (b) The **Mesolithic era** is transitional period.
 - (c) (9000 - 3000 BCE) The **Neolithic era** is the final stage characterized by agriculture.
2. (3000 - 1200 BCE) The **Bronze age** is characterized by the use of bronze tools, the use of writing in some areas, and the features of urban civilization.
3. (1200 - 550 BCE) The **Iron age** evolved from bronze to iron.

After this, we have **protohistory**. During the Paleolithic era, humans were composed of groups of hunter gatherers (ranging in size from an extended family to no more than 100 people) that traveled around for food and shelter. With humans forming groups, there was some level of specialization of labor, where men might have hunted and women/children gathered. Whether it was within the tribe or between tribes, there was **bartering** already happening between humans.

It is important to distinguish what is and isn't money. For now, let's use the following definition.

Definition 1.1 (Money)

There are four criteria to determine whether something is judged as money.

1. It can be used as a means of payment.^a
2. It can be used as a medium to facilitate indirect exchange, i.e. exchanging object A for money, and money for object B.
3. It can serve as a standard of value.
4. It can be served as a means of storing wealth.

At this point, civilization is not even developed yet, but during the advent of the Neolithic era around 9000 BCE where humans discovered agriculture, various hunter gatherers began to separate from their main groups and begin settlements, starting with those in the Fertile Crescent and others following rapidly.

Definition 1.2 (Cradles of Civilization)

This actually happened independently throughout the world at similar times, giving birth to the following major **cradles of civilization**.

1. Mesopotamia
2. Egypt
3. China
4. India

Mesopotamia and Egypt composed up of a large part of the *Fertile Crescent*, which is the crescent-shaped region of land stretching the two areas.^a

¹Since history is often imprecise, all the dates I mention will be approximate.

^aYou can give a slab of meat for berries.

^aEach cradle went through its own development, but we will focus on one or at most two different civilizations at once. Going through all of them independently is not within the scope of these notes.

1.1 Sumer

With the advent of agriculture, more groups abandoned hunter-gatherer practices to establish permanent settlements that could provide for much larger populations. This had several consequences.

1. The domestication and possession of large animals resulted in a dramatic increase in social inequality since it allowed competition between households and resulted in inherited inequalities of wealth. This led to tribal leaders, chiefs, or small elite groups of people (though this was not pronounced until the Bronze age) which would later form into governments.
2. The surplus of grain and cattle allowed more sophisticated forms of bartering, though a common currency was not established yet and values can still be negotiated.
3. Bookkeeping of agricultural production was established, giving rise to the first accounting practices in humanity that tracked whether surplus had been gained.

The development of Mesopotamia where humanity consisted of multiple independent city-states was known as the *Chalcolithic period* (9000-3500 BCE). Major city-states, whose names will become important, were Uruk, Ur, Kish and Eridu.

Around 3500 BCE, the ancient **Sumerians** (whose origin is not well known but was believed to have migrated) settled into Mesopotamia. While their origin is debated, they had complex societies and are considered the first ever civilization known to man. Let's elaborate on how humanity has evolved.

The society now had rulers that oversaw a decently sophisticated society, hinting at the formation of primitive governments. This was effective for development, since now humans can organize to work together on more sophisticated projects, like walled cities and temples used for religious practices, called *ziggurats*. Therefore, a city would consist of tons of farmland surrounding a walled enclosure. Within the walls would be elevated buildings for citizens to live in, with an even higher elevated ziggurat in the middle. These elevations were necessary because of flooding.

A typical farming season would consist of farmers resting in cities for the summer, and when the weather got cooler, they would go out to repair irrigation channels, plant seeds, and finally harvest their crops. These methods were extremely effective and produced bountiful surplus, which were needed to plant next year's crops, use a livestock feed, or for human consumption.

The problem was storing it, since if you store it in the farmer's houses they can be flooded, resulting in the seeds sprouting too early. Therefore, much of these grains were stored in the elevated ziggurats, and so when it came time to harvest, farmers would go out, get the crops, and carry them back to the ziggurats. But this leads to another problem. From the perspective of a farmer who bothered to transport all that grain, I would want to ensure that I can be the one who eats it or at least in possession of it.

One way to do this was use some form of currency like clay coins. You, as a farmer, can take your grains to a temple and exchange the grain for coins. Later on, you can come back to the temple and buy back the grain. This gave birth to the first banking system in civilization, but there was one problem. It was extremely easy to counterfeit these coins, since they're just little balls of hardened clay.

Definition 1.3 (Bank Deposits in Sumer)

To combat this, the Sumerians had clay "envelopes." Say that one coin is worth 1 pound of barley, and you want to store 10 pounds in the temple. Rather than just receiving 10 coins. You would do the following.

1. You take the 10 pounds to the temple and meet the priest there.
2. The priest mints 10 coins for you and 10 coins for him. He then puts 10 coins in one hardened clay ball envelop and 10 coins in the other.
3. Then, he marks that there were 10 coins the outside of the envelope, hardens the envelope, and gives it back to you.

At this point, you have an envelope to take back home and the priest has one kept in the temple.

Given that you don't break the envelope, you can take it back to the temple at a future time, where both you and the priest can first take a look at the inscriptions to see that they match, then break open the envelopes to see that the coins inside match, and then you receive your 10 pounds of grain. This is essentially a deposit and can't be counterfeited.

From this, we can understand that the origin of banks comes from a need to safely store grains, mainly barley. It turns out that the two most common types of currency was barley and silver.² A **shekel** of silver is worth about 30 liters of wheat/barley, which is almost a year's worth for one person.

There may have been a fixed exchange rate between them, but in general

At this point, the concept of loans took off as well.

Definition 1.4 (Loans and Interests in Sumer)

Mainly between the elite and institutions, those with surplus could lend their resources to others. This can happen between royalty, individuals, or the temple, leading to improved **liquidity** of money. The more efficient flow of money from those who have excess to those in need also increased productivity. These loans would be written in tablets that described

1. the parties involved
2. the quantity of the commodity loaned
3. when it will be paid back and consequences for not paying
4. interest rates, which were explicitly put as a percentage or quantity of the loaned commodity, or if omitted, was presumed to be common knowledge.^a

Definition 1.5 (Taxes)

Sometimes, during special occasions, kings would cancel debts.

The two main

Enshakushanna of Uruk conquered all of Sumer, Akkad, and Hamazi, followed by Eannatum of Lagash who also conquered Sumer. His methods were force and intimidation (see the Stele of the Vultures), and soon after his death, the cities rebelled and the empire again fell apart. Some time later, Lugal-Anne-Mundu of Adab created the first, if short-lived, empire to extend west of Mesopotamia, at least according to historical accounts dated centuries later. The last native Sumerian to rule over most of Sumer before Sargon of Akkad established supremacy was Lugal-Zage-Si.

With the development of power, there were inevitably conflicts between the Sumerian city-states of Kish, Uruk, Ur, and Lagash.

1.2 Akkadian and Assyrian Empire

Then, *Sargon of Akkad* had established himself as ruler in 2350 BCE. Then Assyria and Babylonia within it as an independent city-state with Hammurabi.

1. (770 BCE) Coins are created in China.
2. (700 BCE) System of paper money devised in China.

1.3 Babylonian Empire

Hammurabi.

²It's not quite certain how Sumerians had access to silver since precious metals were hard to find in Mesopotamia. It may have been through trade with neighboring civilizations.

^aFor cattle, which was often traded, the interest would be any offspring that they produced.

2 Modern History of Finance

3 Primer on Economics

The most common principle in finance is that money has a time value. If you were to get \$1 today or \$1 in one year, you should always pick \$1 today since you can take that dollar and deposit it had a bank to receive interest for a year.³ To set up the proper terminology without assumption of knowledge, let's walk through a hypothetical scenario.

Say that party A wants to borrow X dollars from party B. There are some factors to consider here. First, for what period of time will A borrow from B? This is easily agreed upon between the two. Second, what is the interest rate that B will charge A for borrowing? This is dependent on multiple factors, but it ultimately follows the basic principle that riskier loans require higher interest rates.

1. If A has good **creditworthiness**, then B should charge lower interest rates since the risk of default⁴ is low.⁵
2. If the length of time T that A borrows money is longer, then the risk of default may be longer, so interest rates should be higher.
3. What are interest rates for other types of loans? Depending on the context on say the interest rate similar loans, the parties should agree on some reasonable price.

There's a couple questions here (beyond how credit is determined). First, how do we quantify how riskier a longer loan is? If you borrow for 2 years vs 1 year, what should the interest rate be? Second, how are other comparable interest rates determined? More generally, how does the third point even work? It looks like we need more background, both qualitative and quantitative to elaborate on all these things.

3.1 Interest Rates

The amount of money loaned X is called the **principal**, and the interest rate r is represented as a percentage of the principal that needs to be paid in addition *per year*. Therefore, the total amount of money A must pay to B at the end of t years is

$$X(1 + r)^t \quad (1)$$

Simple enough, but this model assumes that the loan compounds annually. This may not hold for certain loans, so really if we assume that across the whole T years it is compounded m times, then the total now becomes

$$X(1 + r/m)^{mt} \quad (2)$$

Using the fact that

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e \quad (3)$$

we can model this with a continuous interest rate.

$$Xe^{rt} \quad (4)$$

It is important to comfortably switch between these two types. The market is always effectively discrete since daily compounding is the highest practical frequency in loans, but continuous compounding (exponential growth) is often mathematically more convenient.

Now note the goals of the two parties A and B.

1. A wants to minimize interest rate payment, so if A finds another lender with lower interest rate (with all else equal), then A will go to the new lender.

³We will assume interest rates to always be nonnegative, though there are some nontrivial cases where it is negative.

⁴This means not being able to pay back a loan.

⁵How one's creditworthiness is determined is a whole other subject, so we will assume that this is known and not delve into the how.

2. B wants to maximize interest rate payment, so if B finds another borrower willing to tolerate a higher interest (with all else equal, like creditworthiness), then B will lend to the new borrower.

There is borrower that everyone has access to: the U.S. government (or could be some other entity). This entity is considered **risk-free**, meaning that they are guaranteed to give you back your money and so is the ideal place to lend to, with their own interest rate r_f (f for risk-free). Therefore, if you are a lender, you should always charge a rate higher than r_f , since if it was lower, you can get higher interest (*and* better creditworthiness) by simply lending to the U.S. government. If you are a borrower, you should expect to pay more than r_f . Therefore we've essentially lower bounded all interest rates by r_f .

$$r > r_f \quad (5)$$

Again, this only holds true given that everything else is equal, such as the time of borrowing. One person might ask why the government doesn't lower its own interest rates even further? By doing this, the government can save money by having to pay less interest. This is true, but this assumes that the government is trying to make a profit, which is not true. The government controls the flow of money, and setting this risk free rate is strictly to do that.

Example 3.1 (Home Loans)

Mortgages are loans to individuals to buy a house, which last either 15 years and 30 years. Historically, the home loan rates are documented below^a, where we can see that the 30 year rates are consistently higher. In order to understand why rates are the way they are, we will need to read further.

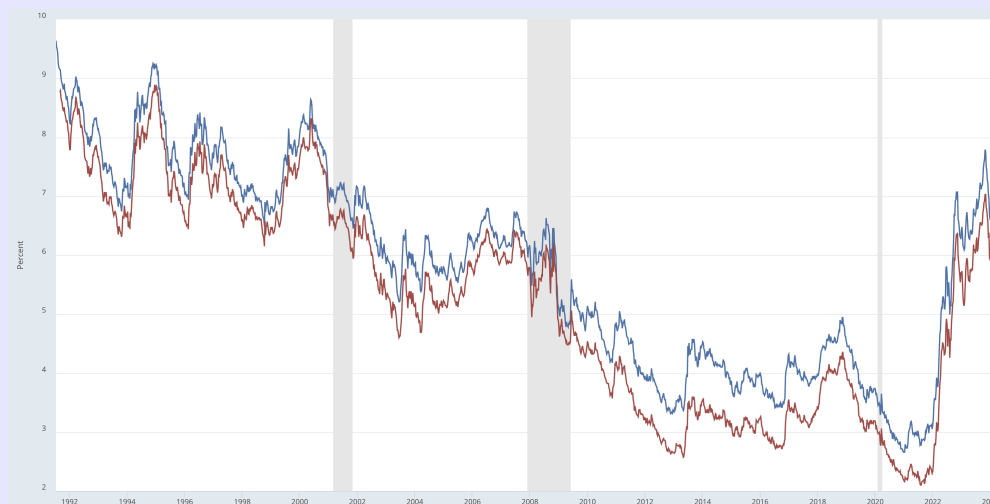


Figure 1: 15-Year and 30-Year Fixed Rate Mortgage Average in the United States [FRED]

There are countless loans, such as auto loans, home loans, student loans, bonds, etc. So how are these interest rates actually determined? To answer this, we must delve a little into macroeconomics. We will first talk about how banks work, the monetary policy, and the federal funds rate which then determines all other interest rates.

3.2 Banks and Monetary Supply

Most financial institutions such as banks make money from interest they earn by lending out money to other clients. It is clear that the interest rate paid by the bank on the money they borrow is less than the rate charged on the money they lend. Therefore, the majority of the deposits are not held by the bank.

^aThe shaded areas represent economic recessions.

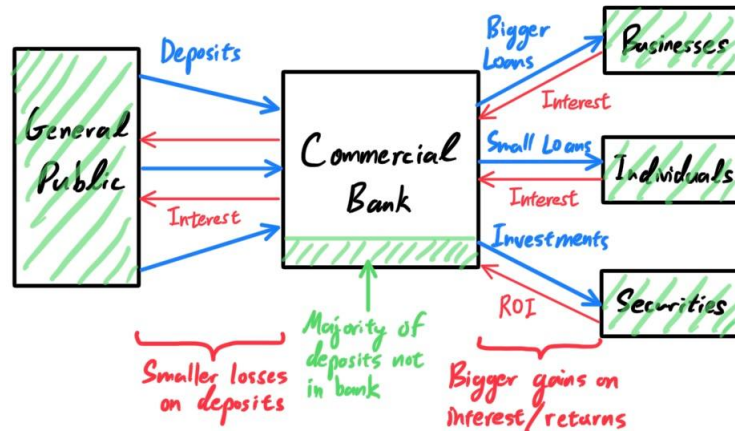


Figure 2: Banks charge higher interest on loans than they pay on deposits to make a profit.

How one keeps this money is important. We start off with the most liquid vs the least liquid.

Definition 3.1 (Physical Currency)

The **physical currency** represents banknotes and coins. This is clearly extremely liquid since if you physically have it, you can use it whenever you want.

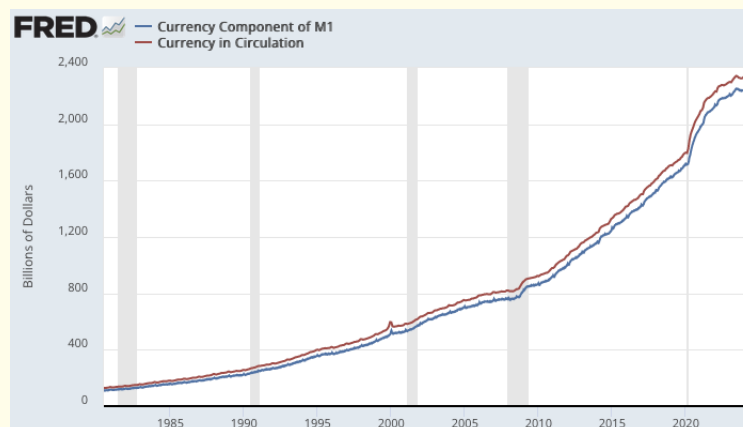


Figure 3: Currency in circulation

Definition 3.2 (Demand Deposits)

Demand deposits refer to funds that can be withdrawn immediately at the holder's request without prior notice given to the bank. This is pretty much a checkings account.

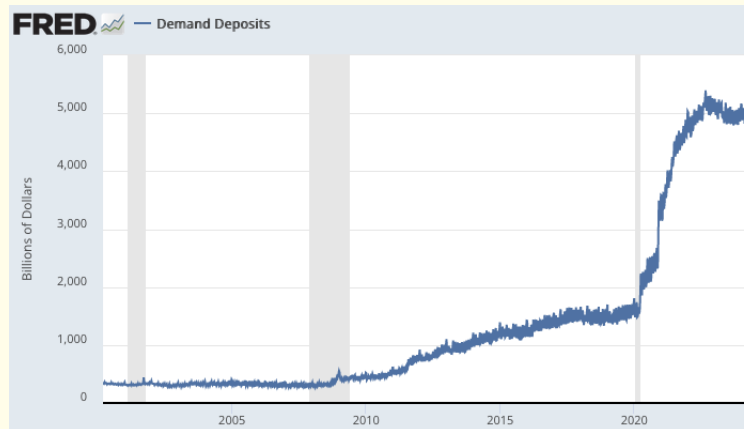


Figure 4: Demand deposits.

Definition 3.3 (Savings Deposits)

A **savings deposit** is less liquid than a checkings account since there are usually restrictions on how often you can withdraw from it. However, where this resides is a bit fuzzy, since savings accounts are also very liquid.^a

Since demand and savings deposits can be converted into real money in the near future, they are referred to as *near money*.

Definition 3.4 (Time/Term Deposits)

Time/Term deposits, on the other hand, require the customer to lock in their money at a bank. This is further categorized into how long the term is (7 days to 2 years are considered *short-term* and beyond are *long-term*) and how big the deposit is (\$100k or more are considered large deposits).

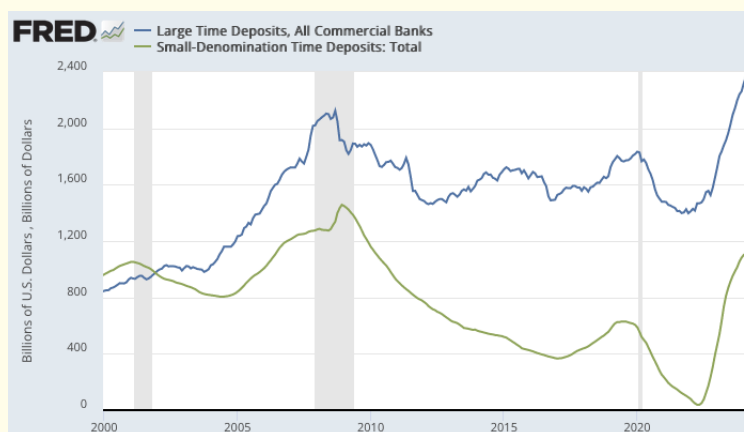


Figure 5: Time deposits

Term deposits are referred to as *near, near money*.

^aFor example, at Bank of America you can withdraw up to 3 times without prior notice.

Definition 3.5 (Levels of Monetary Supply)

In the U.S. (or any other nation), there is a certain amount of money flowing through the economy, known as the **monetary supply**. There are many subcategories of this, which we will elaborate now.

1. **M0** is the total of all physical currency in public circulation and held by the public along with bank reserves^a.
2. **M1** is M0 plus demand deposits and savings deposits. This is the most widely used metric of money supply.
3. **M2** is M1 plus short-term time deposits.
4. **M3** is M2 + large time deposits and *money market funds*, which are funds that invest in short-term securities.^b

You can see the monetary supply in the U.S., which are updated weekly. The M3 is updated quarterly. Since it consists of other assets that are less volatile, it is generally less volatile than M1. Note that prior to May 2020, savings accounts were not included in M1, but due to increased liquidity it is now included.

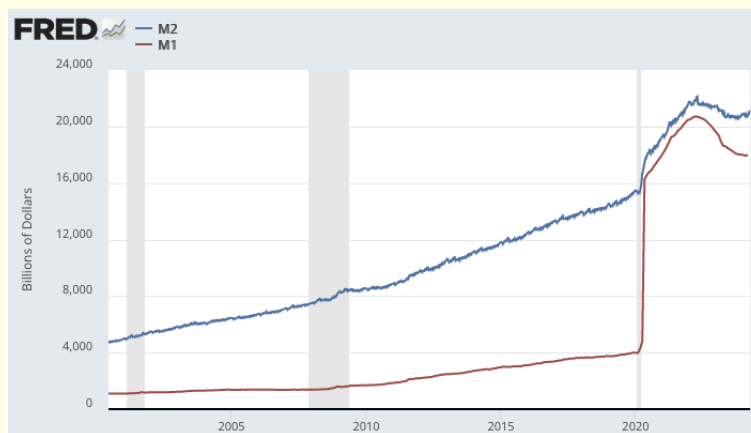


Figure 6: M1 and M2 monetary supply.

3.3 Federal Funds Rate

This can be risky since if a lot of people want to withdraw their money from the bank at once, the banks may not have enough liquidity to pay them back, leading to panic. In order to minimize the possibility of these bank runs, the Fed legally requires banks to have a required **reserve ratio** (the amount of funds that a bank holds in reserve to ensure that it is able to meet liabilities in case of sudden withdrawals), which have historically been between 0% and 10% of total deposits. These **federal funds** (the minimum amount required to have) are stored in the regional Federal Reserve banks (so the central bank acts as the banks' bank). The end-of-the-day balances in each bank's account, averaged over two-week reserve maintenance periods, are used to determine whether the bank meets its reserve requirements.

^aThese are accounts at the central bank that can be exchanged for physical currency, which will be elaborated later

^bThis is discontinued since 2006 by the fed since it provided no useful indicators.

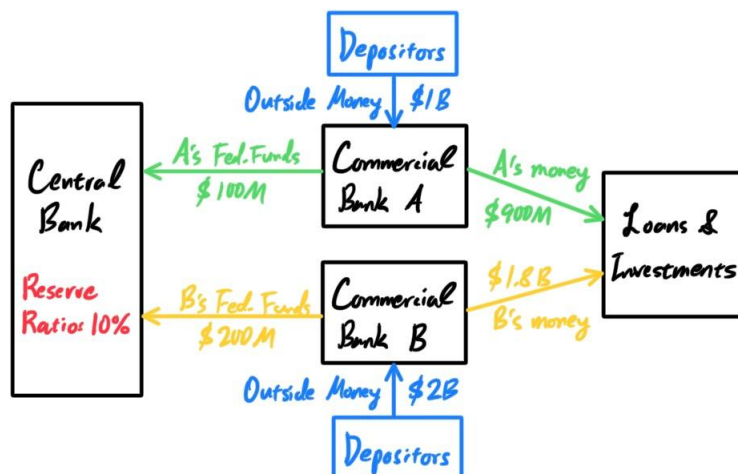


Figure 7: If the reserve ratio is 10%, commercial bank A must deposit \$100M in the Fed and can lend out \$900M. Bank B must deposit \$200M in the Fed and can lend out \$800M.

Definition 3.6 (IOER and Discount Rate)

There is a dynamic relationship between the federal bank and the many commercial banks. They can lend money to each other, and depending on who lends who, different rates are charged. The two rates below are directly controlled by the Fed.

1. The **interest rate on reserve balances (IORB)** is a combination of the **interest rate on required balances (IORR)** and **interest rate on excess reserves (IOER)**, which are the rates of interest that the central bank pays commercial banks on required reserves and excess reserves. These are directly controlled by the Fed.
2. The **discount rate** is the rate of interest that commercial banks pay the central bank. Directly controlled by the Fed.

Recognize that each bank wants to hold the optimal amount of federal funds. If it holds too few federal funds, it might not be able to honor its obligations to other banks (e.g., when they show up with checks written by its customers) or might fall below the legal amount of reserves it is required to hold. If it holds too many federal funds, it will earn less interest on its assets than it otherwise would on other investments. Hence, each bank is trying to hold the optimal amount of reserves, given its unique position and the opportunity cost of holding reserves. So, there will inevitably be banks that hold too many federal funds that can lend money to banks that hold too few federal funds. Therefore, this money can be lent between banks, with the promise that those funds are paid back overnight.

Definition 3.7 (Federal Funds Rate)

This financial market in which interbank lending occurs in the U.S. is called the federal funds market, and the interest rate the lending bank can charge is referred to as the federal funds rate (FFR). There are two variations of this rate:

1. The actual interest rate for a loan between two banks is determined through negotiations between them. The weighted average of this rate across all such transactions on a certain night is the federal funds effective rate (FFER).
2. The federal funds target rate (FFTR) is a (usually quarter percent) range set by the FOMC (usually 8 times a year) as a guidepost, which they enforce by open market operations and adjustments in IOR/IOERs. The target rate is almost what is meant by the media referring to the Fed "changing interest rates." The actual FFR generally lies within the range of that target rate.

Note that the FFR is the benchmark interest rate of the economy, since all other interest rates on loans from commercial banks are dependent on it. That is, a higher FFR will mean that commercial banks will have to raise interest rates on all of their loans (to profit accordingly), and a lower FFR will mean that banks will lower their interest rates on loans. Therefore consumers should care about the FFR since it influences how much they pay to borrow and how much they're paid to save.

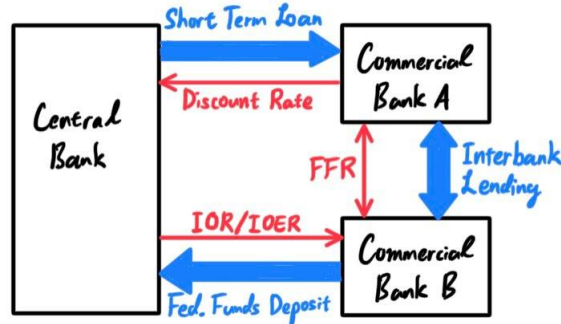


Figure 8: A visual diagram of the three interest rates: IOER, Discount, and FFR.

Theorem 3.1 (FFR)

There are simple relationships between these rates. If I am a commercial bank that needs money to satisfy my overnight reserve requirements, then I can either borrow from the Fed (with the discount rate) or other commercial banks (with FFR).

1. Since borrowing from the Fed with the fixed discount rate is always an option, other commercial banks should naturally have an FFR lower than the discount ($FFR < \text{Discount}$).
2. If I am a commercial bank that has extra money, then I can either loan it to the Fed at the IOER rate or other commercial banks at the FFR. The FFR should naturally be higher than the IOER ($IOER < FFR$).

Combining these two gets

$$IOER < FFR < \text{Discount} \quad (6)$$

Since the Fed has direct control over IOER and Discount, it has control over the FFR, which lies in between the two.

Question 3.1 ()

IOER has been actually greater than FFR consistently in the past 10 years. Why? Read this.

Since the FFR is really the interest rate at the bottom of the chain, all other interest rates are essentially functions of the FFR.

3.4 Monetary Policy

Now, we describe how the Fed uses these policies to control the monetary supply. We can interpret some sum of money "being" or "flowing" in the economy if it is available to use by commercial banks to invest in other securities, i.e., not in the federal funds. Also note that increasing the money in the economy allows commercial banks to have more cash reserves on hand, which induces them to lend them out at lower rates, making loans easier to get in the overnight market and thus decreasing the FFR. Keep in mind that

$$\begin{aligned} \text{More money in economy} &\iff \text{FFR decreases,} \\ \text{Less money in economy} &\iff \text{FFR increases.} \end{aligned}$$

Definition 3.8 (Open Market Operations)

The Fed can finally conduct **open market operations (OMO)**, which means to buy and sell government securities in the open market.

1. Buying government bonds (with possibly newly printed money) from a group of banks (or other entities in the market) → banks get more money by selling the bonds → more money supply in commercial banks & economy ↔ FFR decreases,
2. Selling government bonds to banks (or other entities in the market) → banks lose money by buying the bonds → less money supply in commercial banks & economy ↔ FFR increases.

Definition 3.9 (Modifying Reserve Requirements)

The simplest thing the Fed can do is change the reserve requirements that banks must keep in the central bank.

1. Lowering reserve requirements → Lower percentage of bank deposits must be in federal reserve → More money supply in commercial banks & economy ↔ FFR decreases,
2. Raising reserve requirements → Greater percentage of bank deposits must be in federal reserve → Less money supply in commercial banks & economy ↔ FFR increases.

Definition 3.10 (Changing Discount Rate)

The Fed can also change the discount rate that banks pay on short-term loans from the Federal Reserve bank.

1. Lowering discount rate → Easier for commercial banks to borrow from Fed → more money supply in commercial banks & economy ↔ FFR decreases,
2. Raising discount rate → Harder for commercial banks to borrow from Fed → less money supply in commercial banks & economy ↔ FFR increases.

Definition 3.11 (Changing IOR/IOER Rates)

It can change the IOR/IOER. It's really just the IOER since the IOR rate applies only to mandatory reserves, while the IOER is on excess reserves.

1. Lowering IOER rate → Banks prefer to lend funds out rather than keep it in Fed → more money supply in commercial banks & economy ↔ FFR decreases,
2. Raising IOER rate → Banks prefer to keep funds in Fed than lend to potentially risky borrower → less money supply in commercial banks & economy ↔ FFR increases.

3.4.1 Money Multiplier Effect

Recall that the $M0$ monetary supply, or the monetary base, is the total amount of money that the Fed has injected into the U.S. economy. For this reason, the monetary base is also called **high-powered money**. However, the $M1$ and $M2$ supplies are usually much larger than the $M0$. As of 2021, the $M0$ supply is USD 6.4T, while the $M1$ is USD 20.6T and $M2$ is 21.6T. Where did this extra money come from? The **multiplier effect**, which exists in the **fractional reserve banking system**, tells us that infusions of capital will have a broad snowball effect on various aspects of economic activity. That is, if the government injects a certain amount of money into the economy, it turns out that the money supply increases by multiple factors of what was put in.

Suppose that the government prints USD 10B and credits an additional USD 90B (i.e., credits loans by selling bonds) to banks and other depository institutions. At first, it might seem like the economy just received a monetary influx of USD 100B, but that is actually a very small percentage of the actual money creation. Let us assume that the reserve requirement is 10% (that is, all depository institutions must have 10% of their deposits kept in the federal reserve).

1. The banks get USD 100B. With the 10% requirement, they must keep 10B in reserves and can lend out 90B in investments or loans, which goes into the hand of institutions and individuals. There is a total of 100B in money supply: 90B lent out and 10B in reserves.
2. These institutions/individuals will deposit the 90B receivings in banks. The banks get USD 90B. With the 10% requirement, they must keep 9B in reserves and can lend out 81B in investments or loans, which goes into the hand of institutions and individuals. There is a total of 190B in money supply: $90+81=171$ B lent out and $10+9=19$ B in reserves.
3. These institutions/individuals will deposit the 81B receivings in banks. The banks get USD 81B. With the 10% requirement, they must keep 8.1B in reserves and can lend out 72.9B in investments or loans, which goes into the hand of institutions and individuals. There is a total of 271B in money supply: $90+81+72.9=243.9$ B lent out and $10+9+8.1=27.1$ B in reserves.
4. These institutions/individuals will deposit the 72.9B receivings in banks. The banks get 72.9B USD...

Calculations for the first few deposit/lending cycle is shown. Note that money supply is all of the amount lent out until now plus all the reserves accumulated until now.

Bank	Deposit Amount	Amount Lent Out	Money Supply	Reserves	Total Reserves
1	\$100.00	\$90.00	\$100.00	\$10.00	\$10.00
2	\$90.00	\$81.00	\$190.00	\$9.00	\$19.00
3	\$81.00	\$72.90	\$271.00	\$8.10	\$27.10
4	\$72.90	\$65.61	\$343.90	\$7.29	\$34.39
5	\$65.61	\$59.05	\$409.51	\$6.56	\$40.95
6	\$59.05	\$53.14	\$468.56	\$5.90	\$46.86
7	\$53.14	\$47.83	\$521.70	\$5.31	\$52.17
...

This continues on in the form of a geometric series with a common ratio of 0.9 (1 minus the reserve ratio of 10%). After an infinite time, the total money supply converges 10 times USD 100B to become USD 1 trillion.

One subtle distinction that we must mention is that in the ideal case where banks loan out precisely 90% of their deposits and the economy deposits 100% of their savings, this multiplier is actually called the **deposit multiplier**. It is indeed the case that the reserve requirement is about 10%, but the $M1/M2$ supplies are not even close to 10 times the $M0$ supply. This is because of imperfections due to excess reserves, savings, and conversions to cash by consumers, the actual multiplier, called the **money multiplier**, is ultimately less. Borrowers do not spend all of the money received from bank loans. If they did, and if banks loaned out every possible dollar beyond the minimum reserve requirements, then the deposit multiplier and the money multiplier would be close to exactly equivalent.

4 Markets

If we have two parties that want to exchange goods, they can meet up, decide on a price, and exchange the good. However, when it comes to many players, we must make some sort of exchange. We go over the construction of one.

4.1 Counterparties and Clearing Houses

Think about when one party buys stocks or bonds. Their downside is limited to the amount they paid on the stock, but there are many other cases where they may not be protected.

Definition 4.1 (Default)

If a borrower fails to pay back a debt or cover a loss, then they **default**.

Example 4.1 (Risks Associated with Trades)

1. The simplest case is when they are in debt. They want to borrow X for a certain time, and repay it back in a year at some interest rate $X(1 + r)$. However, they may not have the money by then.
2. When they short a stock, they can borrow money from another party that owns the stock, sell it, and buy it back at a lower price. However, if the price of the stock rises, then they must pay even more money to buy it back, and they may not have enough money to.
3. When they buy a stock on margin, they borrow money from a broker to buy the stock, and they must pay back the broker with interest. However, the price of the stock may fall and the borrower, even after liquidating the stock, may not have enough money to pay back the broker.
4. Someone who longs a forward contract may not have enough money to pay the difference between the forward price and the spot price.

Definition 4.2 (Counterparty)

A **counterparty** is the other party in a financial transaction.

Note that in a vanilla exchange, the two parties trading are counterparties to each other. This is called a **bilateral exchange**. As we have seen, this is risky since if one party defaults, the other party is left with nothing. Therefore, the markets introduce a middleman that takes care of this risk.

Definition 4.3 (Clearing House)

The **clearing house** is an entity that is counterparty to all parties. This means that if party A and party B want to trade, they must go through the clearing house. The clearing house buys from B and sells to A. Therefore, the counterparty risk is now within the clearing house.

It seems that we have simply moved the risk from the traders to the clearinghouse, and this is correct. The traders are glad that the risk is gone from them since they will always be paid by the clearinghouse, and the clearinghouse can add additional regulations on trading to reduce the risk of default. The first regulation they should impose are margin requirements.

Definition 4.4 (Margin)

Margin is simply cash or some other asset (stocks, bonds, etc.) that can be used as collateral during a trade. Note that this does not mean that the collateral must be invested into anything. It is just restricted to sit there in your bank account.

1. Cash is the most liquid form of margin, and is the most common form of margin.
2. Bonds are also a common form of margin, and Treasury bonds are usually valued at 90% of their market value.
3. Stocks are also a common form of margin, and are usually valued at 50% of their market value.

The margin requirements are generally very inflexible unless you have a good relationship with your broker and are a big trader. Now that we've seen the important roles that a clearing house plays, we look at the development of more complex markets, namely, the exchange and OTC markets.

4.2 Exchange Markets

In a trading floor or pit, there are groups of traders huddled at certain physical locations that trade a certain asset, like corn futures. Back then, there would be individual traders that have their own small capital that they can trade with, called **locals**, but they have been dominated by traders that work for large financial institutions, causing locals to be either pushed out or joined with the firms. Market making firms like Akuna

have pit traders, who have headsets and tablets, that are in real-time communication with the screen traders upstairs.

Standing in the perimeter of the trader circle are the **brokers**, who relay orders between the traders in the pit and outside of the pit. These brokers also have devices that allow them to communicate with **off floor** parties, who are individuals that are interested in these products.

Example 4.2 (Off Floor Parties)

Some examples of off floor parties are:

1. Hedge funds, who have devised trading strategies that they want to implement, e.g. using drones to look at the harvest quantity. Now they want to gain exposure to the corn market.
2. Banks might also want to use client money to develop a trading strategy.
3. Consumers such as Kellogs (which needs a lot of corn to make cereal) might be interested in buying contracts to hedge their underlying exposure.
4. Producers like corn farmers might want to buy some contract to lock in the price in which they can sell it at.
5. Large funds like pension funds or family money.
6. Retail investors like small, non-professional traders that might dabble in options.
7. Other market makers might be connected to these brokers

There is specific terminology that we should introduce first.

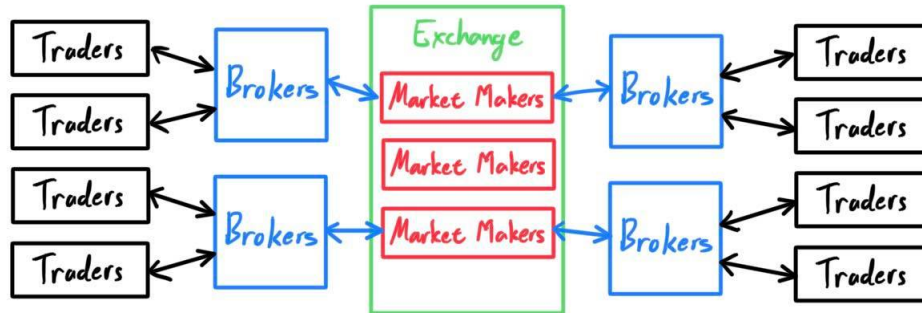
Definition 4.5 (Bid, Offer)

Now if some off floor party decides to sell some calls, they will call their broker and ask to quote some prices. This means that they are asking for bid and offer prices. These traders might quote the following prices:

When more people want to buy and sell things, they must go to an **exchange**, which is literally a physical space that is leased to various parties (e.g. large banks, hedge funds) so that they can exchange goods. These exchanges also collect money through fees from companies to be listed on NYSE. There are two big exchanges, the NYSE and the NASDAQ, in the U.S., both located in New York.

Us retail traders are not direct participants of the exchange; the parties that participate there are called the **market makers**, or **liquidity providers**. Market makers, usually large banks or financial institutions (like hedge funds), make sure that there is enough trading volume to ensure liquidity in the market. A buyer and a seller must meet together to complete a deal, and to ensure that this happens smoothly (i.e. provides liquidity), a market maker buys stocks (through the individual's brokerage) and sells them to the corresponding recipient. Essentially, they provide a pool of shares and act as intermediaries between them. They profit from the bid-ask spread, though sudden volatility is always a risk. For example, if a crash happens (i.e. a market maker buys a stock and it tanks before they can sell it), then the market makers get screwed because they are left with an undervalued stock.

The trader cannot directly trade through the market makers. They must contact their **broker**, which is another company that acts as an intermediary. If I want to buy a stock, my buy order gets sent to my broker, which gets sent to the market makers, which pairs me up with a seller through their own broker, and the transaction is completed. These brokers make money through commissions from the traders and from trading in dark pools, which we'll talk about later. They also give access to traders the forecasts of analyst reports for companies and other research.



Clearly, there can be some shady stuff going on here, but luckily, the SEC and the FINRA consistently regulate the markets to ensure fairness for the little players (retail investors). One of the most important regulations is the 2005 Regulation NMS (National Market System), which required exchanges to publish the best bid and offer price for each stock, required them to route orders to the trading venue with the best price, and had set the minimal price quotation increment to \$0.01. This ensured transparency and protection of the investor with the best price execution (but this can be taken advantage of by high frequency traders).

Definition 4.6 (Exchange Markets)

Exchange markets are markets that exchange *standardized* contracts and is *highly regulated* (to protect investors) by a central authority. A buyer and seller come together to trade a contract, and they enter the contract through a **clearing house**, which is the *counterparty* to all trades. That is, both the buyer and seller enter into a contract with the clearing house. Doing this reduces the counterparty risk by ensuring that both sides have adequate *margin*. For example, if the buyer wins the bet and the seller can't pay, the clearing house will pay the buyer.

Question 4.1 ()

Talk about counterparties, clearinghouses.

4.2.1 Market Making and High Frequency Trading

A significant mover of markets are **high frequency traders**, or HFTs. HFT is a type of algorithmic trading that are characterized by high speeds and high turnover rates. It is the primary type of algorithmic trading and consists of about 50% of all equity trades in the U.S. There is an average of about 1 trader per 10 milliseconds. Whether HFT is beneficial is very controversial:

1. Some argue that it is good since it increases liquidity and lowers transaction costs for retail investors. Even for potentially risky stocks (e.g. very overvalued), HFTs will always make a market out of them.
2. It can be considered unfair because it gives a huge advantage to HFT firms through front-running and other short-term strategies. This also increases the probability of **flash-crashes** (e.g. 2010 flash crash).

Unsurprisingly, many HFT firms are market makers due to their ability to provide liquidity.

The primary form of HFT is **scalping**, which profits off of small price changes. The multiple benefits and drawbacks are easy to spot:

1. It requires a strict exit strategy since one large loss could eliminate the many small gains the trader worked to obtain.
2. It also requires a much higher ratio of winning trades vs losing ones, while keeping profits roughly equal or slightly bigger than losses.
3. A brief exposure to the market diminishes the probability of running into an adverse event like a crash.

4. Smaller moves are easier to obtain, so there are plenty of opportunities to exploit.

This strategy is extremely sensitive that even the physical location of the firm relative to the exchange matters. Heavy money in infrastructure is invested.

4.3 Over the Counter Markets

Definition 4.7 (OTC Markets)

In **over the counter markets** (OTC), there is also an interaction between the buyer and the seller, but it is done in two ways.

1. In *bilateral clearing*, the buyer and seller are the counterparties to each other, and so there is indeed the risk of default on one side.
2. Post-2008, there has been a push to clear through a *central counterparty*. Therefore, there has been a slow push to clear OTC derivative trading to reduce this counterparty risk.

Finally, the OTC market is much larger (about 5 to 10 times larger in terms of the principal dollars underlying the assets) than the exchange and is less regulated since most of the players are institutional investors (they don't need as much protection but need more flexibility).

4.4 Dark Pools

The inefficient aspect about markets that they are sensitive to large **block trades**, which are defined to be a trader involving at least 10,000 shares or at least \$200,000 (though they can get much larger). They are usually made by institutional investors and are often privately negotiated in order to prevent market price changes and fluctuations. When an institutional investor would like to make a sell block trade, they can either

1. sell it on the exchange, where it will cause downwards pressure on the price, causing **slippage** (think of the limit order book: this sell order would clear out all buy orders past the sell price) and perhaps affecting the wider market. Even worse, once an order to sell a huge block has been filled, an investor can submit a buy order at a lower price in hopes that the block sell will hit his lowered price (this is an example of front running).
2. sell it privately off an exchange, but finding other parties to buy such a large amount is difficult.

Either way is quite unfavorable. The market impact of a sale of one million shares in Company XYZ could still be sizable regardless of which option the investor chose since it was not possible to keep the identity or intention of the investor secret in a stock exchange transaction.

These investors can trade in **dark pools**, which are privately organized financial exchanges for trading securities, also an **alternative trading system** (ATS). They were created originally to facilitate block trading by institutional investors who did not wish to impact the markets with their large orders. Most importantly, they keep trades anonymous. They allow traders to make block trades without having to publicize who they are, the buy/sell price, or the number of shares traded.

4.5 Market Manipulation

Market manipulation is a type of market abuse where there is a deliberate attempt to interfere with the free and fair operation of the market.

1. **Stock Bashing**: Creating fake news about a company to drive prices down and get shares for a cheaper price. The perpetrators sometimes work directly for Investor Relations firms who have convertible notes that convert for more shares the lower the bid or ask price is. Thus, the bashers can drive the stock price down by convincing shareholders that they have bought a worthless security.
2. **Pump and Dump**: Misleadingly promoting a company to drive share prices up (the "pump", e.g. with bogus emails to investors). When the stock price reaches a target level, the promoter sells their

shares (the "dump").

3. **Spoofing / Layering:** When a trader places a bid or offer on a stock with the intent to cancel before execution, these fake order trick other market participants by creating the false impression of heavy buying or selling pressure. Layering is an advanced form of spoofing with multiple orders that are "layered." For example, let's say that I want to buy stock XYZ, which is priced at \$100. If I wanted to use layering, I would put a sell order of 100 stocks at \$101, 100 at \$102, and 100 at \$103. An algorithm or trader might see this and believe that there is selling pressure and jump in front of these orders before they go down. They might sell at the market price of \$100, which may bring the price of XYZ a bit lower, to say \$99.95. I can buy at this price now.
4. **Front Running / Tailgating:** This is the practice of entering into a transaction that takes advantage of nonpublic knowledge of a large "block" pending transaction that will influence the price of an underlying security. Say a broker gets an order from a major client to buy 500,000 shares of XYZ Co. Such a huge purchase is bound to drive up the price of the stock immediately, at least in the short term. The broker sets aside the request for a minute and first buys some XYZ stock for their own personal portfolio or through accounts of relatives. Then the client's order is put through. The broker immediately sells the XYZ shares and pockets a profit. This can be used by HFT firms, since with the knowledge of large block buy orders from institutions, they can "buy up" all the orders in the market and sell it at a slightly higher price to the institutions.
5. **Quote Stuffing:** This is the practice of quickly entering and withdrawing a large number of orders in an attempt to flood the market, used by HFT firms. This can create confusion in the market (by filling up bandwidth and increasing latency of the data feed lines) and trading opportunities for high-speed algorithmic traders.

Another illegal act, which isn't market manipulation, is *insider trading*, which refers to a company insider who trades on advanced knowledge of corporate activities.

5 Bond Markets

Introduce to IOUs and why people need to borrow.

Definition 5.1 (Bond)

A **bond** is a financial instrument that represents a loan made by an investor to a borrower (typically corporate or governmental). The borrower agrees to pay interest on the loan and to repay the principal at a later date.

1. The **face value** of a bond is the amount that the bond issuer agrees to repay the bondholder at the maturity date.
2. The **coupon rate** is the interest rate that the issuer will pay on the face value of the bond.
3. The **maturity date** is the date on which the issuer will repay the bondholder the face value of the bond.
4. The **yield** is the rate of return earned on a bond.

5.1 Primary Market and Types of Bonds

Definition 5.2 (Treasury Bonds)

Treasury bonds are issued by the U.S. government and are considered to be the safest bonds available.

1. The **Treasury bill** is a short-term debt obligation backed by the U.S. government with a maturity of less than one year.
2. The **Treasury note** is a marketable U.S. government debt security with a fixed interest rate

and a maturity between one and 10 years.

3. The **10-year Treasury bond** is a bond issued by the U.S. Treasury that matures in 10 years.
4. The **30-year Treasury bond** is a bond issued by the U.S. Treasury that matures in 30 years.
5. The **Treasury Inflation-Protected Securities (TIPS)** is a bond issued by the U.S. Treasury that is indexed to inflation to protect investors from the negative effects of inflation.

Treasury bills, notes, and bonds are issued through an electronic bill Dutch auction.⁶ A **Dutch auction** is an auction in which the price of something offered is determined after taking in all bids to arrive at the highest price at which the total offering can be sold. The bond auction is open to the public, both institutional and individual investors. 24 **primary dealers** (financial institutions and brokerages) are required to participate. It is divided into competitive and noncompetitive bids.

- **Competitive bids** are bids where the buyer specifies the rate that they will accept. It is mainly submitted by institutional investors, and competitive bidding is limited to 35% of the offering amount for each bidder.
- **Noncompetitive bidding** are bids where the buyer agrees to accept the rate determined at the auction (same rate as that determined by competitive bids). It is submitted often by individuals, and noncompetitive bidding is limited to \$5 million per auction.

Example 5.1 (Dutch Auction for Treasury Bonds)

It is best to explain through an example: Suppose the Treasury seeks to raise \$9 million in two-year notes with a 5% coupon. Entities can bid up to 30 days in advance of the auction, and let us assume that the submitted bids are as follows:

- \$1 million at 4.79%
- \$2.5 million at 4.85%
- \$2 million at 4.96%
- \$1.5 million at 5%
- \$3 million at 5.07%
- \$1 million at 5.1%
- \$5 million at 5.5%

The bids with the lowest yield will be accepted first since the issuer will prefer to pay lower yields to its bond investors. In this case, since the Treasury is looking to raise \$9 million, it will accept the bids with the lowest yield up to 5.07%. At this mark, only \$2 million of the \$3 million bid will be approved. All bids above the 5.07% yield will be rejected, and bids below will be accepted. This auction is cleared at 5.07%, and all successful bidders receive the 5.07% yield.

Definition 5.3 (Zero Coupon Bonds)

Zero-coupon bonds do not pay coupon payments and are instead issued at a discount to their par value that will generate a return once the bondholder is paid the full face value when the bond matures (e.g., U.S. Treasury bills).

Definition 5.4 (Zero Coupon Bond)

A **zero coupon bond (ZCB)** with maturity T is an asset that pays 1 at time T (and nothing else).

Let $V(t, T)$ be value at current time t of a ZCB with maturity T , for $t \leq T$. Then, it must be the case that $V(T, T) = 1$. Accounting for interest rates and the discounted value of money, we get the following result.

⁶The website for upcoming auctions can be found here.

Definition 5.5 (Types of Bonds by Issuer)

Bonds can be categorized by the entities that issue them.

1. **Corporate bonds** are issued by companies. Companies issue bonds rather than seek bank loans for debt financing in many cases because bond markets offer more favorable terms and lower interest rates.
2. **Municipal bonds** are issued by states and municipalities (local governments).
3. **Government bonds**, or **treasury bonds**, are issued by the U.S. Treasury.
4. **Agency bonds** are those issued by government-affiliated organizations (e.g., Fannie Mae).

Definition 5.6 (Types of Bonds by Grade)

Bonds can also be categorized by their credit ratings.

1. **Investment-grade bonds** are those rated BBB or higher by Standard & Poor's or Baa3 or higher by Moody's.
2. **High-yield bonds** are those rated below investment grade (BB+ or lower by Standard & Poor's or Ba1 or lower by Moody's).
3. **Junk bonds** are those rated below investment grade.

Definition 5.7 (Other Categories of Bonds)

Bonds can have other characteristics that distinguish them from each other.

1. **Convertible bonds** are bonds that gives the bondholder the option to convert their debt into stock at some point, depending on conditions.
2. **Reverse convertible bonds** are bonds that now give the *issuer* the option to convert.
3. **Callable bonds** is one that can be called back by the company before it matures, for instance, when interest rates become too disadvantageous for the issuers.
4. **Puttable bonds** allow the bondholders to put (or sell) the bond back to the company before maturity.

5.1.1 Daycount Conventions

Another important concept to mention is the daycount conventions in the financial markets. Note that since a year consists of 365 days, when doing discrete compounding we cannot easily calculate r/m . For example, if we wanted to do quarterly payments, we would have $m = 4$ but we cannot split 365 into 4 parts. Note the following days in each quarter.

Table 1: Number of days between quarterly dates

Period	Days
16 March 2011 to 16 June 2011	92
16 June 2011 to 16 September 2011	92
16 September 2011 to 16 December 2011	91
16 December 2011 to 16 March 2012	91

There are lots of imperfections and therefore there are many different market conventions. We list some of them.

Definition 5.8 (Act/365)

The **actual 365** (act/365) daycount convention calculates the quarterly interest rate by dividing by the number of days by 365.

Definition 5.9 (Act/360)

The **actual 360** (act/360) daycount convention calculates the quarterly interest by dividing the number of days by 360.

Definition 5.10 (Act/Act)

The **actual actual** (act/act) convention calculates the quarterly interest by dividing by the actual number of days in each year. This accounts for leap years.

Definition 5.11 (30/360)

The **30 360** (30/360) convention assumes that there are 30 days in each month and was designed to produce equal amounts of interest payment each month or quarter. It calculates the quarterly interest by simply $120/360 = 1/4$.

Example 5.2 (Accrual Factors for Each Quarter)

The **accrual factor** is the convention to divide a year into equal time periods, denoted by α . For each quarter above, we can get

5.2 Valuation**Theorem 5.1 (Value of Zero Coupon Bond)**

The value of a zero coupon bond at time t with maturity T with continuously compounded interest rate is

$$V(t, T) = e^{-r(T-t)} \quad (7)$$

and with discretely compounded interest rate is

$$V(t, T) = \frac{1}{(1 + r_A)^{T-t}} \quad (8)$$

where r_A is the annually compounded rate.

Proof.

We prove this result by *discounting cashflows back to today*, which means to take whatever cashflow in the future, and discounting it at the interest rate to current time. If you had $e^{-r(T-t)}$ at time t , then you can invest it at the interest rate r , compounded continuously, to get

$$e^{-r(T-t)} \cdot e^{r(T-t)} = 1 \quad (9)$$

after time $T - t$, which is what the ZCB pays.

If we assume that the price P is a good approximate of the value V , then we can substitute it into the

equation to get an estimate of the interest rate.

$$P(t, T)(1 + r_A)^{(T-t)} = 1 \text{ or } P(t, T)e^{r(T-t)} = 1 \quad (10)$$

Note that r is not constant across time, so it should really be written as

$$P(t, T)(1 + r_A(t, T))^{(T-t)} = 1 \text{ or } P(t, T)e^{r(t, T)(T-t)} = 1 \quad (11)$$

5.3 Annuities

Definition 5.12 (Annuity)

An **annuity** is a series of fixed cashflows C at specified times T_i for $i = 1, 2, \dots, n$.

Theorem 5.2 (Value of Annuity)

The value V at current time $t \leq T_1$ of an annuity can be calculated by summing the value of the zero coupon bonds at each expiry date T_i .

$$V(t, \mathbf{T}) = C \sum_{i=1}^n Z(t, T_i) \quad (12)$$

6 Stock Markets

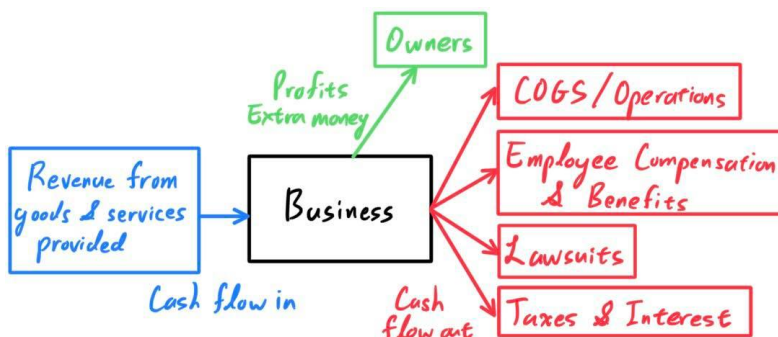
Definition 6.1 (Stocks)

Stocks are a type of security that represents a share in the ownership of a corporation. That is, owning a stock is equivalent to owning its relative position in the company and its profits, and possibly voting rights.

But what exactly does it mean to “own” a company? To explain this, we can interpret a business to be some sort of entity with input and output cash flows.

1. The majority of the cash flows flowing into the business is the revenue earned from the goods and services that the business provides.
2. The majority of the cash flows flowing out of the business are the costs, such as COGS, employment compensation and benefits, operations, lawsuits, taxes, interest on debts, etc.

Ideally, the company will have a net profit, or a net positive cash flow, which means that the cash flows in is greater than the cash flows out. All this profit, i.e. this extra money, now goes to the owners of this business through **dividends** (or may be reinvested into the company). Note that it is not the case that the profits go to the CEO. The CEO is a glorified version of an employee, but still in the end an employee, with an income just like every other employee, albeit a very high one.

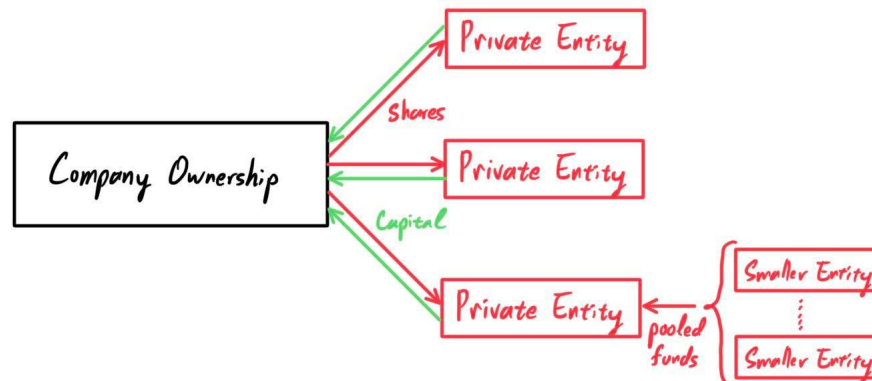


Obviously, the owner has a bigger role to play rather than collect profits. They must also decide on things. But since there are usually many owners, these decisions are made through votes in meetings like the annual shareholder meeting. The ownership of a stock really possible rights to earn dividends and possible rights to vote. We will go through the specifics of these powers later.

6.1 Primary Market and Types of Stocks

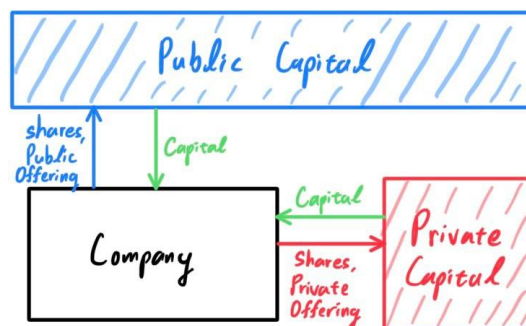
A company gets its funding from debt (e.g. bonds) and equity (stocks). If the prospect of paying interest on debts are not suitable, a company may issue stocks to investors. In other words, the company is selling parts of itself in exchange for capital. The number of shares, type of shares, and the prices for these shares is entirely determined by the seller and buyer. For example, a business may need to raise \$100 million in capital and can choose to issue 10 million shares at \$10 each.

For early-stage businesses, these are mainly funded by **venture capital** firms and **angel investors** (but is not limited to them) through different rounds roughly ordered: seed round, Series A, Series B, Series C, etc. After due diligence, these venture firms, as early investors, may get extra privileges, such as anti-dilution protection, guaranteed board seats, liquidation preferences, priority dividends. Note that upon multiple rounds of issuance, more and more stocks will circulate through more hands, causing **stock dilution**. No investor would want this and therefore stocks must be issued carefully. At this point, the ownership of a company is held within the hands of private individuals and firms.

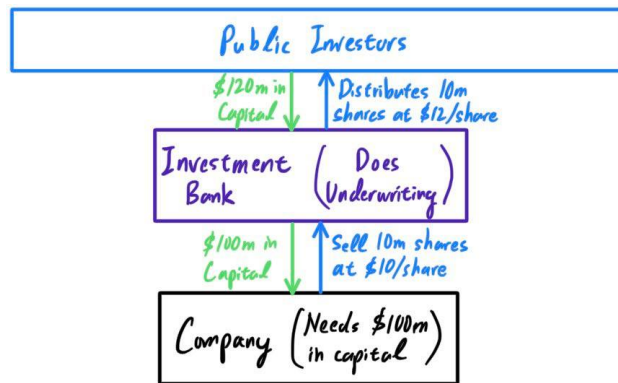


One final characteristic to note is the **nominal value**, or **par value**, of a company's stock. Like the face value of a bond when issued, the nominal value of the stock is its stated value. It is an arbitrary value assigned for balance sheet purposes when the company is issuing share capital, and is typically \$1 or less. It has little to no bearing on the stock's market price, so no need to worry about this number.

Perhaps after a few years, the firm has grown to the point where funding on an even larger scale to support even more expansion is needed. Private sources may be too restrictive or small, and so companies may need to tap into the capital of the general public. Thus, they can do a **public offering**, an issuance of the stock to the general public rather than private entities. This process of an **initial public offering (IPO)** and the company being listed on a stock exchange is what is referred to as a company "going public."



Investment banks, through underwriting services, facilitate public offerings.



Definition 6.2 (Additional Actions on Stocks)

Throughout the years, companies may do the following:

1. Additional **stock issuance** to both public and private investors.
2. **Stock splits** to increase liquidity if the stock price is too high or for other reasons. **Reverse stock splits** may also be done.
3. **Stock buyback** where the company uses its institutional funds to buy back shares from the market to increase its **treasury shares**. This may be used for anti-dilution purposes or to retain voting control. We can think of these as the opposite as a stock issuance, since treasury shares do not have dividend rights nor voting rights.
4. Issue **restricted stock units (RSUs)** to employees. The grant is restricted because it is subject to a vesting schedule, which can be based on length of employment or performance goals.

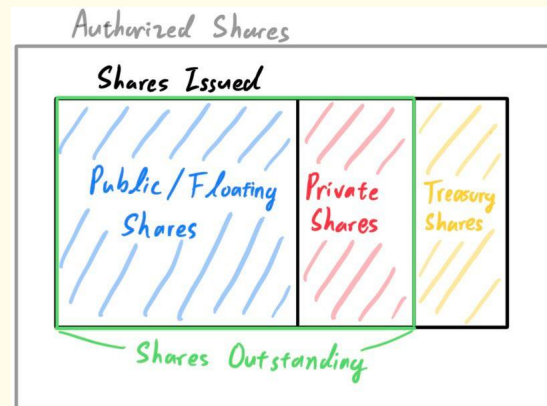
Note that if an entity (individual or organization) has the majority ownership or a **controlling interest** (at least 50% in voting rights) of a company, then they have the power to do pretty much whatever they want with the company.

Definition 6.3 (Types of Circulating Stocks)

With these, we can organize the number of stocks circulating in different levels, from biggest to smallest:

1. Authorized shares refers to the maximum number of shares that a corporation is legally permitted to issue. Companies don't usually get close to this number due to market conditions, which will be explained later.
2. Shares issued refers to the total number of shares issued, including all public shares, private shares, and treasury shares.
3. Outstanding shares refer to a company's stock currently held by all shareholders, both public and private shares (including RSUs). It does not include treasury stocks however.
4. Floating shares are the shares considered available for the general public. Moreover, the floating percentage represents the portion of outstanding shares that are floating:

$$\text{Floating Percentage} = \frac{\text{Floating Shares}}{\text{Shares Outstanding}}$$



Definition 6.4 (Properties and Powers of Stocks)

Stocks may also have the following characteristics below. Generally, we work with common and preferred stocks, but other ones may be labeled using letters (e.g. Class A, Class B stocks) that have certain qualities mentioned below.

1. **Voting Power:** Nonvoting shares have no voting power while executive shares can be worth 10 votes.
2. **Payment Priority:** Deferred shares are set as a lower priority for dividends and corporate assets.
3. **Cumulative shares** can accumulate dividend payments that have been deferred due to low profits in the past.
4. **Convertible shares** can be converted into different forms of financial assets (e.g. preferred shares may be convertible into common ones, common shares may be convertible into corporate bonds).

Definition 6.5 (Stock Classes and Types)

Companies may issue different classes of stocks, each with different rights and privileges.

1. **Class A shares** may have more voting rights than Class B shares.
2. **Class B shares** may have more dividend rights than Class A shares.
3. **Class C shares** may have more liquidation rights than Class A shares.
4. **Common stocks** are the most common type of stock. Common stock shareholders receive dividends in proportion to the amount of profit generated each year, and they have voting rights. They are usually the lowest priority over a company's income, after the creditors and preferred stock shareholders.
5. **Preferred stock** are much less common and liquid. Preferred stock shareholders receive fixed dividends and no voting rights. They receive payments after bondholders but before common shareholders.

6.2 Valuation

6.3 Secondary Market

6.3.1 Stock Orders

Definition 6.6 (Stock Orders)

The two most common orders one can do is the market order and the limit order.

1. A **market order** just buys or sells securities at the market price. This ensures that the transaction will be completed, but the price at which you buy or sell may not be what you want.
2. A **limit order** buys or sells at least at a certain price. It ensures that you get the price you want for a transaction, but it may not be carried out always.
 - (a) A buy limit order at \$X tells the broker to buy a stock at \$X or lower.
 - (b) A sell limit order at \$x tells the broker to sell a stock at \$ or higher.
3. A **stop loss order** is used to limit an individual's loss or lock in a profit on a stock position.
 - (a) If an investor buys a stock at \$100, they can make a stop loss at \$95. This means that when the stock price reaches \$95, then it will automatically place a market order to sell the stock (the price may not be fulfilled at exactly \$95 due to volatility).
 - (b) If an investor has a short position, on the stock, then they can put a stop loss order at \$105, essentially telling the broker to buy the stock when the price reaches \$105.

In addition to holding a long position, we can **short sell**, or short, a stock. Given that 100 shares of AAPL are each priced at \$100, we can borrow shares from some investor (probably an institutional investor), sell it on the market for \$100, and hope for the price to drop so we can buy it back. This is pretty symmetric to longing a stock, and many hedge funds use a **long-short equity strategy** involving a combination of longs and shorts, but there are additional risks.

1. You must pay additional interest for borrowing stocks.
2. Your potential losses is not bounded.
3. If you have too much unrealized losses as the stock price goes up, then you may not have enough **margin** (cash) in your account to even buy back to stock. To prevent this, your broker may issue a **margin call**, which forces you to buy the stocks back, locking in your loss, unless you add more capital to your account. This margin call may not happen right at the point when your free cash is not enough to buy back all shorted shares. Rather, your brokerage may run some complicated statistical simulations, calculate some confidence interval, and then decide on a threshold that determines whether you will get a margin call.

The third risk is the deadliest, and at worst, you can get **short-squeezed**. Let's explain how this works. Say you and a bunch of other investors shorted GME. GME prices start going up, and it reaches a point where some investors have to buy back the GME shares due to a margin call. The investors buy back the GME at market price, which adds more orders to the limit order book, causing the price to go even higher. This higher price causes more short-sellers to buy back GME due to additional margin calls, causing GME to go even higher, and so on. This positive feedback loop is extremely deadly, killing off many short sellers.

Sometimes, the ethics of short selling are called into question. People who want to ban shorting state that short selling can drive down the stock price, which can be bad because it doesn't spur the economy and reduces optimism in stock markets. Short selling means that stock is being borrowed and sold on the market, increasing supply, and therefore (all else being equal) decreasing price. However, these short sellers keep the stock price more in line with reality, and bad companies are punished by short sellers.

Finally, the entities that drive this short selling are **prime brokers**, which are large financial institutions (usually investment banks) that provide financial management to mainly hedge funds. They act on behalf of the short seller, locating the assets to be sold short for them and providing them with a margin account (which must hold capital to the sum of at least 150% of the value of the initial transaction). They enable

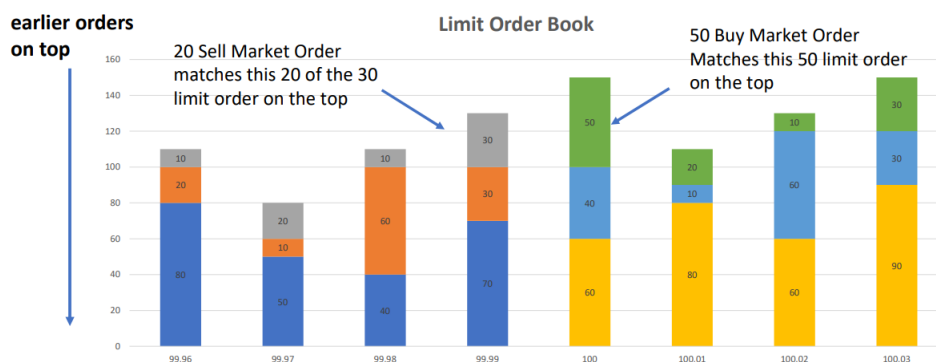
hedge funds to borrow large amounts of stocks from institutional investors to short-sell them and allow them to access large amounts of margin from commercial banks. The prime brokerage makes money through commissions. Prime brokers can also loan capital to investors to increase their leverage, and they also provide financial research and analytics to their clients.

6.3.2 Stock Prices and Limit Order Books

Now how is the share price determined? There are two large paradigms for this. The first is the **fundamental approach**, which attempts to calculate the intrinsic value for the company by calculating its discounted future cash flows.

$$\text{Present Value} = \sum_{n=1}^{\infty} \frac{\text{Cash Flow}_1}{(1+R)^n}$$

However, we will concern ourselves with the second approach: the **market approach**, which claims that the stock price is determined by supply and demand. This can be seen by looking at the **limit order book** (LoB), which shows all pending limit orders at the current time. This allows us to see where the prices are concentrated at and how strong the demand is versus the supply. In the diagram below, the right side represents the sell limit orders while the left side represents the buy limit orders. If we submit 50 buy market orders, then the best price is executed.



The market makers profit off of the **bid-ask spread**, which is the difference in the buy and sell price. If a seller wants to sell at \$99.99 and a buyer wants to buy at \$100.00, then the market maker can buy the stock from the seller for \$99.99 and sell it to the buyer at \$100.00, making a \$0.01 profit.

6.4 Fundamental Stock Valuation

We will denote company metrics in bold letters and stock metrics in regular font.

Definition 6.7 (P/E Ratio)

Given a company XYZ, say that its earnings (profits) are **E** and its market cap **P**. Then, its price to earnings ratio is

$$\frac{P}{E}$$

which is a multiple of how much the company is worth given its earnings. Theoretically, it should be much greater than 1, since if it had a PE ratio of 1, then it can earn its entire worth annually, allowing the investor to break even on their investment within a year.

Definition 6.8 (Earnings per Share)

Given a company XYZ, say that its earnings are \mathbf{E} and its number of shares outstanding \mathbf{S} . Then its earnings per share is

$$\frac{E}{S}$$

which is basically telling us how much a company is earning per share. Note that given the stock price P and the EPS, note that the P/E ratio is simply

$$\frac{P}{E} = \frac{P/S}{E/S} = \frac{P}{EPS}$$

Definition 6.9 (PEG Ratio)

The PEG ratio is the P/E ratio, but now it accounts for the growth rate of the company earnings. Say that company XYZ has earnings \mathbf{E} this year and \mathbf{E}_0 last year. Then, its earnings growth can be calculated as $\mathbf{G} = \mathbf{E}/\mathbf{E}_0 - 1$. Then, its PEG ratio is

$$\frac{P/E}{G}$$

Definition 6.10 (P/S Ratio)

The P/S ratio is like the P/E ratio, but now we use the company's sales \mathbf{S} , or revenue. This is particularly useful for when the company does not have net profits every year.

$$\frac{P}{S}$$

Definition 6.11 (ROIC)**Definition 6.12 (ROA)**

7 Stock Trading Strategies

Let us introduce some valuation metrics that are used for analyzing a portfolio. We will usually work with some sort of time series of the form $\{X_i\}_{i=0}^n$ (i.e. a discrete-time stochastic process), but if we are only worried about the initial and final price X_0 and X_n , then we will just mention them.

Definition 7.1 (Stock Price)

Given a stock P , we can construct a discrete-time stochastic process $\{P_i\}$, where P_i are random variables.

Definition 7.2 (Dividend)

Given time i, j with $i < j$, the dividend that a stock P pays within time interval $[i, j]$ is the random variable

$$D_{[i,j]}$$

Definition 7.3 (Total Return)

Given stock price $\{P_i\}$, the amount of money you would have made from time i to J is the random variable

$$P_j - P_i + D_{[i,j]}$$

The **total return** during that period is defined

$$R_{[i,j]} = \frac{P_j - P_i + D}{P_i} = \frac{P_j - P_i}{P_i} + \frac{D_{[i,j]}}{P_i}$$

which is the price return plus the dividend rate. If this is a single-step return, i.e. $j = i + 1$, then we can write this as shorthand:

$$R_i = R_{[i-1,i]} = \frac{P_i - P_{i-1}}{P_{i-1}} \approx \ln \left(\frac{P_i}{P_{i-1}} \right) \text{ for } i = 1, \dots, n$$

Sometimes, we refer to **return** as the total return without the dividend payment.

Most of the time, we work in **log returns** as an appropriate approximation of the return. This is mathematically justified since if we fix P_i and assume that P_j is close to P_i , we can take the Taylor expansion of $\ln(P_j)$ to be

$$\ln(P_j) \approx \ln(P_i) + \frac{1}{P_i}(P_j - P_i) \implies \frac{P_j - P_i}{P_i} \approx \ln(P_j) - \ln(P_i) = \ln \left(\frac{P_j}{P_i} \right)$$

Almost always, we will work with log returns and $R_{[i,j]}$ will denote the log return from time i to j . It has its advantages:

1. It is time-additive: If we have prices P_i, P_j, P_k for $i < j < k$, we can see that

$$R_{[i,j]} + R_{[j,k]} = \log \left(\frac{P_j}{P_i} \right) + \log \left(\frac{P_k}{P_j} \right) = \log \left(\frac{P_k}{P_i} \right) = R_{[i,k]}$$

2. Symmetricity: It is well known that a stock going down 50% and then going up 50% does not return the stock to its original price, despite it "looking" like it did. This effect is killed when looking at log returns. If $P_j = P_k(1 - k)$ for $0 < k < 1$, then we know that it must scale up by a factor of $\frac{1}{1-k} \neq 1 + k$ to get the original price. Indeed, it is clearly the case that

$$\log(1 - k) + \log(1 + k) = \log(1 - k^2) < 0$$

and in fact is always less than 0 (so we always lose money). This is due to the concavity of this function.

3. It focuses on relative change, which is typically invariant on the underlying price of the stock.

These properties allow us to focus on the returns to calculate our metrics.

Definition 7.4 (Expected Return)

The **expected return** of a stock from time i to j is simply the expected value

$$\mathbb{E}[R_{[i,j]}] = \mathbb{E}\left[\frac{P_j - P_i}{P_i}\right] \approx \mathbb{E}[\log(P_j) - \log(P_i)]$$

Definition 7.5 (Stock Volatility)

Given a stock P , its return $R_{[i,j]}$ is a random variable. Its standard deviation is called the **volatility**.

$$\sigma = \sqrt{\text{Var}(R_{[i,j]})}$$

If we do have data on the stock $\{P_i\}_{i=0}^n$ and its log returns $\{R_i\}_{i=1}^n$, the volatility can be estimated simply by taking the sample standard deviation of this data

$$\sigma = \sigma(\{R_i\}_{i=1}^n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_i - \mu)^2} \text{ where } \mu = \frac{1}{n} \sum_{i=1}^n R_i$$

Definition 7.6 (Sharpe Ratio)

The **Sharpe ratio** is a measurement of risk-adjusted return. Given that you have some stock price with total return R and volatility σ , the Sharpe ratio is defined as the random variable

$$S = \frac{R - R_f}{\sigma}$$

where R_f is risk-free return (e.g. the return you earn on U.S. Treasury bonds), which can be considered as a constant random variable. Sometimes, R_f may not be constant and may be some time-series data itself.

The Sharpe ratio is one of the most widely used metrics, and it is usually based on an annual horizon. While it is straightforward to compute the annual returns, it is more difficult to compute the annual volatility of a stock because it is sensitive to the time scale (i.e. computing it using daily data will result in much higher volatility than computing in minutely data). In practice, people follow the Square-root-of-Time rule, which takes the daily volatility by a factor of square root of 260, which is approximately the number of trading days in a year. Therefore, the invariant annual Sharpe ratio is

$$S = \frac{R - R_f}{\sigma \sqrt{260}} \quad (13)$$

7.1 Momentum Strategies

Momentum strategies can be divided into momentum trending, which bets that the stock price will continue following the trend, and momentum reversing, which bets that the stock price will revert back to some mean.

7.1.1 Moving Averages

The simple moving average and the exponential moving averages represent where the stock's price is at average.

Definition 7.7 (Simple Moving Average)

Let us have some stock price $\{P_i\}_{i=0}^N$ and fix some lookback parameter L . Then, the L -period **simple moving average (SMA)** of the stock is the average of prices of the past L periods, $\{S_i\}_{i=L-1}^N$, where

$$S_i = \frac{1}{L} \sum_{j=i-L+1}^i P_j$$

Definition 7.8 (Exponential Moving Average)

We can build momentum strategies by comparing two different MAs of lookback periods $L_1 < L_2$. We can compare the current price to the moving averages, but the current price can be thought of as the 1-period moving average anyways. The L_2 -MA is thought of as the long term trend of the stock, while the L_1 -MA is the short-term trend. The follow are essentially MA strategies.

1. Momentum Trending:

- (a) If the L_1 -MA is above the L_2 -MA, then the stock has momentum upwards. \implies Long.
- (b) If L_1 -MA is below the L_2 -MA, then the stock has momentum downwards. \implies Short.

2. Momentum Reversing:

- (a) If the L_1 -MA is above the L_2 -MA, then the stock has momentum upwards. \implies Short.
- (b) If L_1 -MA is below the L_2 -MA, then the stock has momentum downwards. \implies Long.

7.1.2 Bollinger Bands**Definition 7.9 (Bollinger Bands)**

Let us have some stock price $\{P_i\}_{i=0}^N$ and fix some lookback parameter L , along with some z -score Z . We compute the standard deviation of the prices P_i in the past L periods to get $\{\sigma_i\}_{i=L-1}^N$.

1. The **middle band** is defined to be the L -period SMA, which we will call

$$\{\mathcal{M}_i\}_{i=L-1}^N$$

2. The **upper band** is defined to be the band that is Z standard deviations above the middle band.

$$\{\mathcal{U}_i\}_{i=L-1}^N = \{\mathcal{M}_i + Z\sigma_i\}_{i=L-1}^N$$

3. The **lower band** is defined to be the band that is Z standard deviations below the middle band.

$$\{\mathcal{L}_i\}_{i=L-1}^N = \{\mathcal{M}_i - Z\sigma_i\}_{i=L-1}^N$$

Now the algorithm for Bollinger bands is very simple.

1. In a momentum following case, if P_i crosses below \mathcal{L}_i , we short, and if P_i crosses above \mathcal{U}_i , we long.
2. In a momentum reversing case, if P_i crosses below \mathcal{L}_i , we long, and if P_i crosses above \mathcal{U}_i , we short. The upper band acts as a resistance level, and the lower band acts as a support level.

7.1.3 Relative Strength Index

The relative strength index indicates the momentum or lack of it.

Definition 7.10 (Relative Strength Index)

Let us have some stock price $\{P_i\}_{i=0}^N$. Let us define the period changes as $\{D_i\}_{i=1}^N$ with $D_i = P_i - P_{i-1}$. Then, the total gain and total loss can be defined as

$$D_{\text{gain}} = \sum_{D_i > 0} |P_i - P_{i-1}|, \quad D_{\text{loss}} = \sum_{D_i < 0} |P_i - P_{i-1}|$$

Then, the **relative strength index (RSI)** is defined to be

$$RSI = 100 \frac{D_{\text{gain}}}{D_{\text{gain}} + D_{\text{loss}}}$$

which ranges in $[0, 100]$, where 0 is extremely bearish and 100 is extremely bullish. Roughly, the RSI being 30 means that for every 1 increase in a period, there are 2 decreases in other periods. The RSI being 70 means that for every 1 decrease, there are 2 increases.

The algorithm is quite simple:

1. In a momentum following case, if the RSI is below 30, we short, and if the RSI is above 70, we long.
2. In a momentum reversing case, if the RSI is below 30, we long, and if the RSI is above 70, we short.

7.1.4 Determining Momentum Trending or Reversing

Now how can we quantitatively decide whether we should use a momentum trending or reverting strategy? We can measure this behavior of a stock by looking at its autocorrelation.

Definition 7.11 (Autocorrelation)

Let us have some stock price $\{P_i\}_{i=0}^N$ and fix some lookback parameter L . We construct a lagged version $\{P_i\}_{i=0}^{N-L}$ and compute the correlation of this lagged time series with the original.

$$\text{Corr}(\{P_i\}_{i=0}^{N-L}, \{P_i\}_{i=L}^N)$$

is called the L -period **autocorrelation** of the stock price.

We can determine which strategy to use as follows. Let us assume that L is small.

1. If this autocorrelation is high (near 1), then this indicates that if the stock moves in a certain direction, it is likely to move in that same direction L periods later. So we should use a momentum following strategy.
2. If it is negative (near -1), it indicates that if the stock moves in a certain direction, it is likely to move in the opposite direction L periods later. So we should use a momentum reverting strategy.

Trend following stocks tend to be growth stocks, while trend reversing ones are the traditional conservative companies.

7.2 Pairs Trading**7.2.1 Naive Approach**

Intuitively, pairs trading takes two stocks of very similar companies and bets that they will rise and fall together. If one rises and the other falls, then we can long the one that falls and short the one that rises, ultimately betting on the way that they will converge. To do this, take two sets of stock prices

$$\{P_i\}_{i=0}^N \text{ and } \{Q_i\}_{i=0}^N$$

We first talk about the naive approach, which assumes that if P rises by 5%, then Q will also rise by 5%. Beginners just look at their common ratios $\{P_i/Q_i\}_{i=0}^N$ and calculate whether this time series diverge or not. Remember that this is not symmetric, so we must log both of them and look at the log of the quotient, i.e. the difference of the logs.

$$\left\{ \log \left(\frac{P_i}{Q_i} \right) \right\}_{i=0}^N = \{\log(P_i) - \log(Q_i)\}_{i=0}^N$$

If this time series diverges too much from its average, then we can long and short accordingly. This is another way of saying that the time series of returns are

However, this model is too simplistic, since we are limited by the assumption that if P rises by 5%, then Q will also rise by the same 5%. Companies P and Q may use similar supply chains, but perhaps P is more dependent on it than Q . Therefore, if there are supply chain problems, then the effect on P may be say, twice as much as that on Q . So, if P goes down by 5%, then Q may go down by 2.5%.

7.2.2 Sophisticated Pairs Trading

Ultimately, our basis assumption is that the returns are linearly correlated in the following relationship.

$$\frac{\Delta P}{P} = \beta \frac{\Delta Q}{Q} \iff \log(P_j) - \log(P_i) = \beta (\log(Q_j) - \log(Q_i))$$

This results in the model

$$\log(P) = \beta \log(Q) + \alpha + \epsilon$$

which can be seen to be equivalent because taking the change over time $[i, j]$ on both sides gives

$$\Delta \log(P) = \beta \Delta \log(Q) + \Delta \epsilon \iff \frac{\Delta P}{P} = \beta \frac{\Delta Q}{Q} + \Delta \epsilon$$

So, if we look at

$$\{\epsilon_i\}_{i=0}^n = \{\log(P_i) - \beta \log(Q_i) - \alpha\}_{i=0}^n$$

we expect this to be a 0-mean time series. Let the standard deviation be $\sigma = \sigma(\{\epsilon_i\}_{i=0}^n)$ and let us fix some Z-score threshold. Then

1. If $\epsilon_i > Z\sigma$, then short P and long Q .
2. If $\epsilon_i < -Z\sigma$, then long P and short Q .

7.2.3 Long/Short Market Weights

But how much should we long or short? Let's look at a couple scenarios:

1. If $\beta = 1$, and we shorted \$99 of P and long \$1 of Q , then a 5% rise in P would result in 5% rise in Q , but since we shorted much more of P , we would have a net loss. To mitigate this risk, we should have a weight of $P : Q = 1 : 1$.
2. If $\beta = 10$, and we shorted equally \$50 of P and long \$50 of Q , then a 10% rise in P would result in a 1% rise in Q . But since we had equal weights in P and Q , this scenario would cause 10 times more losses in P than gains in Q , resulting in a net loss. To mitigate this risk, we should have a weight $P : Q = 1 : 10$.

So, we need to be careful of setting the ratio of our market weights of the stocks P and Q :

$$\lambda = \frac{MV_Q}{MV_P}$$

Intuitively, we can see that our ratio should be $1 : \lambda = 1 : \beta$, i.e. $\lambda = \beta$, but let's formalize this with some mathematical derivation. Our portfolio value is

$$V = MV_P + MV_Q$$

where $MV_P = n_P P$ and $MV_Q = n_Q Q$, where n_P, n_Q are the number of shares of P, Q . If we are longing P and shorting Q , then $n_P > 0$ and $n_Q < 0$. Then, our change in portfolio value V is

$$\begin{aligned}\Delta V &= \Delta MV_P + \Delta MV_Q \\ &= MV_P \left[\frac{\Delta MV_P}{MV_P} + \frac{MV_Q}{MV_P} \frac{\Delta MV_Q}{MV_Q} \right] \\ &= MV_P \left[\frac{\Delta P}{P} + \lambda \frac{\Delta Q}{Q} \right] \\ &= MV_P \left[\beta \frac{\Delta Q}{Q} + \Delta \epsilon + \lambda \frac{\Delta Q}{Q} \right] \\ &= MV_P \left[(\beta + \lambda) \frac{\Delta Q}{Q} + \Delta \epsilon \right]\end{aligned}$$

Therefore, our change in portfolio value depends on the terms in the last equation. We don't want the change ΔQ to have any effect on the performance, so we set $\lambda = -\beta$, ultimately resulting in

$$\Delta V = MV_P \Delta \epsilon$$

and now our performance is purely dependent on $\Delta \epsilon$. We would like ΔV to be positive, so if we are longing P , i.e. $MV_P > 0$, then we want $\Delta \epsilon$ to also be positive. Likewise, if we are shorting P , then $MV_P < 0$ and so we want $\Delta \epsilon < 0$.

7.2.4 Choosing Correct Stocks

So, how do we ensure that $\Delta \epsilon$ behaves this way? Remember that $\{\epsilon_i\}$ is 0-mean time series. However, we want to impose the additional condition that it is *mean-reverting* as well. That is, we don't want it to diverge or oscillate too frequently ϵ_i , since if it did then there is the risk of $\{\log(P_i) - \beta \log(Q_i) - \alpha\}_{i=0}^n$ swinging too widely, resulting in losses. In other words, if $\epsilon_i > 0$, then we want $\Delta \epsilon_{i+1} = \epsilon_{i+1} - \epsilon_i < 0$, and if $\epsilon_i < 0$, then $\Delta \epsilon_{i+1} > 0$, so that the ϵ_i 's tend to "go back towards 0." More formally, we can plot all ϵ_{i-1} 's with the $\Delta \epsilon_i$'s, and look at a potentially linear relationship

$$\Delta \epsilon_i = \alpha \epsilon_{i-1} + \xi_i$$

To be mean reverting, we want to test that $\alpha < 0$ with adequate statistical significance, i.e. with p-value 5%. By looking at not just the previous ϵ_{i-1} but also the last p ϵ_i 's we can develop the general linear relationship

$$\Delta \epsilon_i = \xi_i + \sum_{j=1}^p \alpha_j \epsilon_{i-j}$$

This is called the **Augmented Dickey Fuller (ADF)** test, which tells us whether a time series is mean-reverting or not.

7.2.5 Entry and Exit Points

8 Stock Portfolio Analysis

We will define a portfolio as a collection of K stocks with prices labeled as a time series data

$$\{P_{1,i}\}_{i=1}^N, \{P_{2,i}\}_{i=1}^N, \dots, \{P_{N,i}\}_{i=1}^N \quad (14)$$

with one-step returns $\{R_{k,i}\}_{i=1}^N$ for each $k = 1, \dots, K$ and total return over the period $[0, N]$ as R^k . For simplicity, let us assume that we buy all stocks at time 0 and hold until time N , and so we can consider a one-period market with $\{P_{k,0}, P_{k,1}\}$ for $k = 1, \dots, K$ and corresponding returns

$$\{R_k\} = \{\log(P_{k,1}) - \log(P_{k,0})\} \quad (15)$$

all random variables.

Definition 8.1 (Market Value of Portfolio)

The **market value of a portfolio** $\{P_1, \dots, P_K\}$ is the random variable

$$V = \sum_k MV_k = \sum_k n_k P_k \text{ which is } \sum_k n_k P_{k,0} \text{ at time } t = 0 \quad (16)$$

The market weights for each stock is determined as

$$w_k = \frac{MV_k}{V} \quad (17)$$

where $\sum w_k = 1$. We will denote $\mathbf{w} = (w_1, \dots, w_K)$. If we allow unlimited shorting, then the possible values of \mathbf{w} are $\{\mathbf{w} \in \mathbb{R}^K \mid \mathbf{w} \cdot \mathbf{1} = 1\}$, but if we only allow all-long positions, then this domain reduces to $\{\mathbf{w} \in \mathbb{R}^K \mid \mathbf{w} \cdot \mathbf{1} = 1, 0 \leq w_k \leq 1\}$.

Definition 8.2 (Portfolio Return)

The **portfolio return vector** is the random K vector of returns $\mathbf{R} = (R_1, \dots, R_K)^T$, which may be correlated. We can take its expectation by taking the component expectations over the sample space w , or by integrating the component mappings $e_k : \mathbb{R}^K \rightarrow \mathbb{R}$ over the measure λ induced by \mathbf{R}

$$\mathbb{E}[\mathbf{R}] := \begin{pmatrix} \mathbb{E}[R_1] \\ \vdots \\ \mathbb{E}[R_K] \end{pmatrix} = \begin{pmatrix} \mathbb{E}_\lambda(e_1) \\ \vdots \\ \mathbb{E}_\lambda(e_K) \end{pmatrix} \quad (18)$$

The **portfolio return** is defined as the random variable

$$R = \mathbf{w}^T \mathbf{R} = \sum_{k=1}^K w_k R_k \quad (19)$$

i.e. the weighted sum of the individual total returns of the stocks. The **expected return** of the portfolio is simply

$$\mathbb{E}[R] = \sum_{k=1}^K w_k \mathbb{E}_\lambda[R_k] \quad (20)$$

We can estimate this by sampling the historical returns of the k th stock $R_{k,i}$ of the required length for each stock and then estimating its mean. That is, for each k ,

$$R_k \approx \hat{R}_k = \frac{1}{n} \sum_{i=1}^n R_{k,i} \quad (21)$$

Definition 8.3 (Portfolio Variance, Volatility)

The **covariance matrix** of the portfolio is defined as the covariance matrix of the random vector of

returns \mathbb{R} .

$$\Sigma = \text{Cov}(\mathbf{R}) := \begin{pmatrix} \text{Var}(R_1) & \dots & \text{Cov}(R_1, R_K) \\ \vdots & \ddots & \vdots \\ \text{Cov}(R_K, R_1) & \dots & \text{Var}(R_K) \end{pmatrix} \quad (22)$$

We can estimate this by sampling the one-period historical returns R_k and computing the sample variance/covariance

$$\sigma_{k_1 k_2} = \text{Cov}(\{R_{k_1, i}\}, \{R_{k_2, i}\}) \quad (23)$$

The **portfolio variance** is simply the variance of R .

$$\text{Var}(R) = \mathbf{w}^T \Sigma \mathbf{w} \approx \sum_{k_1, k_2=1}^K w_{k_1} w_{k_2} \sigma_{k_1, k_2} \quad (24)$$

and the **portfolio volatility** is simply the standard deviation of R

$$\sqrt{\text{Var}(R)} \quad (25)$$

Definition 8.4 (Alpha)

Alpha is a measure of performance. The **alpha** of a security/portfolio with return R_p compared to some market return R_m is defined as

$$\alpha := \mathbb{E}[R_p - R_m] = \mathbb{E}[R_p] - \mathbb{E}[R_m] \quad (26)$$

If we are given samples of returns $\{R_{p, i}\}$ and $\{R_{m, i}\}$, then we can estimate the α simply as

$$\hat{\alpha} = R_p - R_m = \sum_{k=1}^{j-i} R_{p, i+k} + \sum_{k=1}^{j-i} R_{m, i+k} \quad (27)$$

Definition 8.5 (Beta)

Beta is a measure of volatility. The **beta** of a stock compared to some market is defined

$$\beta := \rho_{pm} \frac{\sigma_p}{\sigma_m} = \frac{\text{Cov}(R_p, R_m)}{\text{Var}(R_m)} \quad (28)$$

where ρ_{pm} is the correlation between R_m and R_f and σ represents the volatility of returns. If we are given samples of returns $\{R_{p, i}\}$ and $\{R_{m, i}\}$, then we can estimate the beta using the sample correlation and standard deviation:

$$\hat{\beta} = \hat{\rho}_{pm} \frac{\hat{\sigma}_p}{\hat{\sigma}_m} = \rho(\{R_{p, i}\}, \{R_{m, i}\}) \frac{\sigma(\{R_{p, i}\})}{\sigma(\{R_{m, i}\})} \quad (29)$$

If a stock has beta value of 1, then its price activity is strongly correlated with the market (or the benchmark B_j). A $\beta < 1$ means that the security is less volatile than the market, and $\beta > 1$ means more volatile. A negative β indices that the stock is inversely correlated with the market.

8.1 Markowitz Portfolio Theory - Mean Variance Portfolio

Consider a one-period market with K independent securities which have identical expected returns and variances, i.e. consider $\{P_{k, 0}, P_{k, 1}\}$ for $k = 1, \dots, K$. Then, the returns

$$R_k = \log(P_{k, 1}) - \log(P_{k, 0})$$

are random variables such that $\mathbb{E}[R_k] = \mu$ and $\text{Var}(R_k) = \sigma^2$. Let w_k denote the fraction of wealth invested in the k th security. Now consider two portfolios

1. Portfolio A: 100% invested in stock 1 so that $w_1 = 1$ and $w_k = 0$ for $k = 2, \dots, K$
2. Portfolio B: An equi-weighted portfolio so that $w_k = \frac{1}{K}$ for all k .

Let R_A and R_B denote the portfolio returns of A and B. Then, we have

$$\begin{aligned}\mathbb{E}[R_A] &= \mathbb{E}[R_B] = \mu \\ \text{Var}(R_A) &= w_1 \text{Var}(R_1) = \sigma^2 \\ \text{Var}(R_B) &= \text{Var}\left(\frac{1}{K} \sum_{k=1}^K R_k\right) = \frac{1}{K^2} \sum_{k=1}^K \text{Var}(R_k) = \sigma^2/K\end{aligned}$$

which means that even though the expected returns of portfolios A and B are the same, the volatility of B is much less than that of A, making it much more advantageous. We can clearly see that the Sharpe ratio of A is μ/σ (assuming the risk-free return is 0), but the Sharpe ratio of B is $\mu\sqrt{K}/\sigma$, which means that as we increase our number of independent stocks K , the ratio goes up by a factor of \sqrt{K} .

8.1.1 Efficient Frontier without Risk-Free Asset

Given a portfolio of stocks $\{P_1, \dots, P_K\}$ with corresponding return vector $\mathbf{R} = (R_1, \dots, R_K)$, let us take its expected value $\boldsymbol{\mu} = \mathbb{E}[\mathbf{R}]$ and covariance matrix $\boldsymbol{\Sigma} = \text{Cov}(\mathbf{R})$. Its portfolio return is $R = \mathbf{w}^T \mathbf{R}$. We will assume that we must invest all of our cash into something, which will manifest in the constraint equation $\mathbf{w}^T \mathbf{1} = 1$.

Definition 8.6 (Efficient Frontier without Risk-Free Asset)

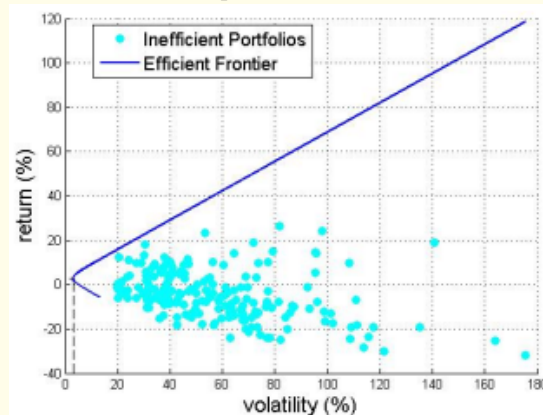
We would like to construct a risk-return efficient portfolio (by determining the w_k 's) so that it has the highest return for the same amount of risk, or the lowest risk for some amount of return. That is, letting $R_{\mathbf{w}}$ be the return of a portfolio with weight \mathbf{w} , we would like to find

$$\arg \min_{\mathbf{w} \in \mathbb{R}^K} \text{Var}(R_{\mathbf{w}}) = \arg \min_{\mathbf{w} \in \mathbb{R}^K} \frac{1}{2} \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$$

subject to the constraint equations $\mathbf{w}^T \boldsymbol{\mu} = p$ and $\mathbf{w}^T \mathbf{1} = 1$, where p is our target portfolio returns. The solution to this optimization problem is called the **Markowitz efficient frontier**, which traces out a hyperbola of the form

$$\sigma_R^2 = Ap^2 + Bp + C$$

if we get rid of the ω parameter. Here we randomly sample a bunch of \mathbf{w} 's from $[0, 1]^K$ (subject to the constraints, of course), construct the return of the portfolio $R_{\mathbf{w}}$, and then plot the points $(\mathbb{E}[R], \text{Var}(R))$. We can see that none of the points ever cross the frontier.



Note that this efficient portfolio allows us to have unlimited short positions, which may or may not be realistic. If we wanted to work only with all-long portfolios, then we would impose the nonlinear restrictions

$$0 \leq w_k \leq 1 \text{ for } k = 1, \dots, K$$

which would need to be solved numerically, possibly using nonlinear models.

8.1.2 Efficient Frontier with Risk-Free Asset

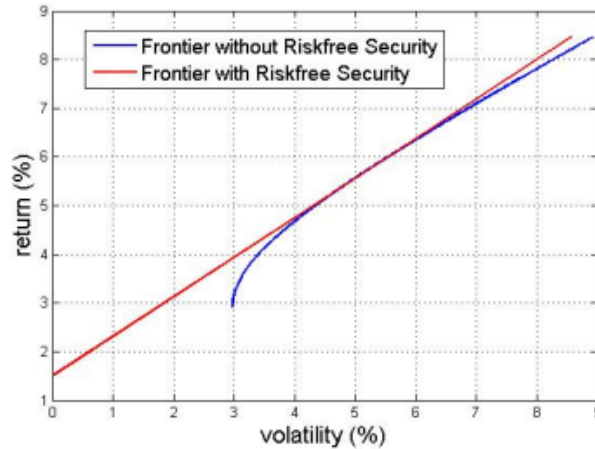
If we include a risk-free asset such as cash C with risk-free return R_f to the portfolio $\{P_1, \dots, P_K\}$, then we slightly modify our equations. By definition, the risk-free asset has volatility $\sigma_0 = 0$ and its weight must be equal to $w_f = 1 - \sum_{i=1}^K w_k$, so we still want to minimize

$$\sigma^2 = \frac{1}{2} \mathbf{w}^T \Sigma \mathbf{w}$$

subject to constraint equations

$$R_f \left(1 - \sum_{k=1}^K w_k\right) + \sum_{k=0}^K w_k R_k = R$$

When we allow our portfolio to include the risk-free security, the efficient frontier becomes a straight line that is tangential to the risky efficient frontier and with a y -intercept equal to the risk-free rate.



We can include other linear portfolio constraints, such as no-borrowing, no-short sales, or certain sector constraints. While analytic solutions are generally no longer available, the resulting problems are easy to solve numerically.

8.2 Capital Asset Pricing Model

Let us have some stock P with random variable of return $R_p = R_{P,[i,j]}$ and the market M with random variable of return $R_m = R_{M,[i,j]}$, both within period $[i, j]$. There may be some sort of risk-free return r_f available (e.g. U.S. treasury bonds), so we can observe the returns of these two assets past the risk-free return by considering the joint distribution

$$(R_p - r_f) \times (R_m - r_f) \quad (30)$$

This may or may not be correlated, but the capital asset pricing model shows that there exists a linear relationship between the expected values of these two distributions. The central insight of the CAPM is that in equilibrium the riskiness of an asset is not measured by the standard deviation of its return but by its beta.

Theorem 8.1 (CAPM)

Now let $\bar{R}_m = \mathbb{E}[R_m]$ denote the expected return of the market, and $\bar{R} = \mathbb{E}[R]$ denote the expected return of a security or portfolio. Then, the **capital asset pricing model (CAPM)** asserts that there exists a linear relationship

$$\bar{R} = r_f + \beta(\bar{R}_m - r_f) \quad (31)$$

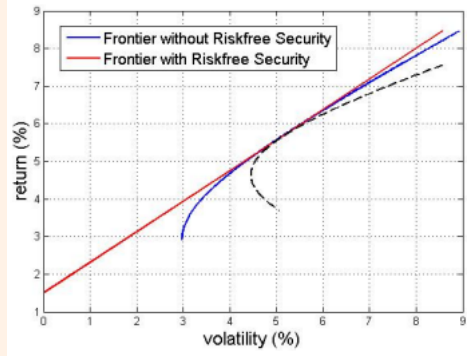
where r_f is the risk-free rate.

Proof.

Let us consider a portfolio of weights α and $1 - \alpha$ on the risky security and market portfolio, respectively. Let R_α denote the random return of this portfolio as a function of α . We then have

$$\begin{aligned} \mathbb{E}[R_\alpha] &= \alpha\bar{R} + (1 - \alpha)\bar{R}_m \\ \text{Var}(R_\alpha) &= \alpha^2\text{Var}(R) + (1 - \alpha)^2\text{Var}(R_m) + 2\alpha(1 - \alpha)\text{Cov}(R, R_m) \end{aligned}$$

Note that as α varies, the mean and standard deviation $(\mathbb{E}[R_\alpha], \text{Var}(R_\alpha))$ trace out a curve in \mathbb{R}^2 that cannot cross the efficient frontier, as shown in the dotted line.



At $\alpha = 0$, the slope of this curve must equal the slope of the capital market line. The slope of the α -curve (where $\sigma(R_\alpha) = \sqrt{\text{Var}(R_\alpha)}$) is

$$\begin{aligned} \left. \frac{d\mathbb{E}[R_\alpha]}{d\sigma(R_\alpha)} \right|_{\alpha=0} &= \frac{d\mathbb{E}[R_\alpha]}{d\alpha} \bigg/ \frac{d\sigma(R_\alpha)}{d\alpha} \bigg|_{\alpha=0} \\ &= \frac{\sigma(R_\alpha)(\bar{R} - \bar{R}_m)}{\alpha\sigma(R) - (1 - \alpha)\text{Var}(R_m) + (1 - 2\alpha)\text{Cov}(R, R_m)} \bigg|_{\alpha=0} \\ &= \frac{\sigma(R_m)(\bar{R} - \bar{R}_m)}{-\text{Var}(R_m) + \text{Cov}(R, R_m)} \end{aligned}$$

The slope of the capital market line is $(\bar{R}_m - r_f)/\sigma(R_m)$, and equating the two

$$\frac{\sigma(R_m)(\bar{R} - \bar{R}_m)}{-\text{Var}(R_m) + \text{Cov}(R, R_m)} = \frac{\bar{R}_m - r_f}{\sigma(R_m)} \quad (32)$$

gives the result.

9 Forward Contracts

People trade stocks and commodities, but derivatives give more flexibility. To see why, consider the case below.

Example 9.1 (Apples and Pies)

Suppose you have an apple orchard and you make pies. You can sell apples and pies at a fixed price, but you can also sell them at a future price. For example, you can sell apples at a future price of \$1 per apple.

Definition 9.1 (Derivatives)

A **derivative** is a financial instrument whose value depends on (is *derived* from) the value of some other, more basic, underlying variable. Essentially, given some stochastic process X_t describing some variable, the derivative is some function

$$Y_t = g(X_t) \quad (33)$$

The reason that we say an underlying *variable* is that it includes real assets (e.g. property), financial assets (e.g. stocks), indices (stock, inflation, or housing price index), or an event (e.g. weather, or the amount of rainfall in a given season to hedge against a bad winter). Therefore, these variables can be anything, like the number of people attending a fair, which may be an indicator of the profit of the fair or some catastrophic event, and we can derive value out of that event. Therefore, the variable does not necessarily have to be an asset.

Example 9.2 (Stock Derivative)

For example, a stock option is derived from the value of the stock, which itself is derived from the value of the underlying company. We will see more specific examples of this later.

Example 9.3 (Weather Derivatives)

A weather derivative can have payout of $g(S_T)$, where

$$g(S_T) = \mathbb{1}\{S_T > 50\} = \begin{cases} \$1 & \text{if } S_T > 50 \\ 0 & \text{if else} \end{cases} \quad (34)$$

where S_T is the total snowfall in inches during the year up to $T = 1$. Note that both S_T and $g(S_T)$ are random variables whose value is unknown until $T = 1$.

Definition 9.2 (Payoff Graph, Profit-Loss Diagram)

To see the dependency between the underlying and the derivative, it helps to plot g , known as the **payoff graph**. The x -axis tells you the value of the underlying asset, and the y -axis tells you the payoff that a certain buyer or seller will make given the price of the underlying asset. A **profit-loss diagram** is the same thing but it accounts for the premium paid on a contract. This is important for options since the buyer pays a premium to enter the contract, but doesn't make a difference for forward or future contracts.

Now we will see the most common type of derivative: the forward contract.

Definition 9.3 (Forward Contracts)

Forward contracts are simply contracts between two parties to buy or sell an asset at a future date for a price agreed upon today.

1. T is the time until delivery, usually in years. This is fixed per contract.
2. K is the delivery price. This is also fixed per contract.

3. S_t is the spot price of the asset at time $t \in [0, T]$.
4. $V_K(t, T)$ is the value of the contract at current time $t \leq T$ of being long a forward contract with delivery price K and maturity T . This will vary over time.
5. $P(t, T)$ is the price of the forward contract at time t with maturity T . This is the price that you pay to enter the contract.

Example 9.4 (Red Sox)

If I buy the Boston Red Sox regular season wins at 90, in size one dollar per won, I have a long forward position with delivery price of 90. If the Red Sox wins 100 games I make \$10, and if they have 81 wins I lose \$9. The payout is $S_T - K$ where S_T is the number of wins.

9.1 Primary Markets and Types of Forwards

Example 9.5 ()

We have a forward contract that agrees to buy/sell an asset for $K = \$100$ in $T = 3$ months. Then, the buyer and seller have the following payoff graphs.

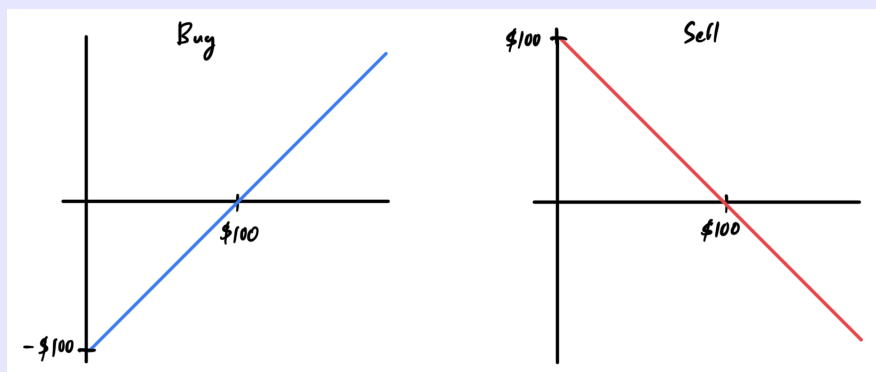


Figure 9

We can see from the following that this is a zero-sum game. Therefore, *no wealth* is created or destroyed in a forward contract, it is only transferred. Since there was no cost to enter the contract, the profit is really a linear relationship between the price of the underlying asset. This is unlike options, where there is a cost to enter since the buyer or seller has the *right*, but not the obligation to buy/sell the asset. This right puts an extra premium on the option that you must pay for.

Definition 9.4 (Settlement)

If a contract is **physically settled**, it means that the buyer actually pays K and receives the asset at time T . However, some forwards are **cash settled**, meaning one simply receives (pays if negative) the amount $S_T - K$ at time T . The profit/loss is the same, but there are some key differences.

1. A cash settled forward has no further exposure to the asset price since you're just receiving the difference in cash. However, a physically settled contract implies that the buyer now owns the asset at T , which continues to have exposure to asset price movements.

9.2 Forward Contract Valuation

Now let's talk about how you would value a forward.

Definition 9.5 (Value at Maturity, Payout)

Since the party that longs the forward contract must pay K at T to buy an asset which is worth S_T , the value of the contract at T is

$$V_K(T, T) = S_T - K \quad (35)$$

This is referred to as the **value at maturity** or the **payout** of the contract.

Definition 9.6 (Forward Price)

The **forward price** $F(t, T)$ at the current time $t \leq T$ is the delivery price K such that

$$V_K(t, T) = 0 \implies V_{F(t, T)}(t, T) = 0 \quad (36)$$

That is, it is the price that makes the contract have zero value at the current time. By definition, it must be the case that $F(T, T) = S_T$ since to buy an asset at T for the price $F(T, T)$ is the same as buying the asset at the spot price S_T .

Since the forward price tells you what the value of the underlying asset K should be at time T for the contract to have no value, it is a good indicator of the actual value of the contract. When the forward is created, it has no intrinsic value, but as time goes on, the value of the contract changes as the spot price changes.

Example 9.6 ()

Suppose that a stock which pays no dividends always has price 100 and interest rates are always 0. Then

1. $F(t, T) = 100$ since if you had the right to buy the stock at 100 at any time T , then this would be the same as buying the stock at the spot market, but it is always 100 in the spot market. So the value of this contract must be 0 for all T .
2. $V_K(t, T) = 100 - K$ since if you have the right to buy the stock at K in time T , the spot price will always be worth 100 so you're really gaining $100 - K$. Therefore the value of this contract must be $100 - K$ for all T .

9.3 Forward Prices**Theorem 9.1 (Forward Price of No Income Asset)**

For an asset paying no income (e.g. a stock with no dividends), the forward price is

$$F(t, T) = S_t e^{r(T-t)} \implies F(0, T) = S_0 e^{rT} \quad (37)$$

Proof.

Replication. At current time t , let's take two portfolios.

1. A: Consists of one unit of the underlying stock.
2. B: One long forward contract with delivery price K , plus $Ke^{-r(T-t)}$ of cash.

At time T , B can invest the cash at the risk free rate r and get K cash. Then, it executes the forward contract at K to get one stock. Therefore, at time T , both portfolios are worth S_T and the values of both portfolios at time t must also be the same. This is called the *replication proof*.

$$S_t = V_K(t, T) + Ke^{-r(T-t)} \implies V_K(t, T) = S_t - Ke^{-r(T-t)} \quad (38)$$

We want to find the value of K such that the value of the contract is 0. Therefore,

$$0 = S_t - Ke^{-r(T-t)} \implies K = S_te^{r(T-t)} \quad (39)$$

Proof.

No Arbitrage. Another way to prove this is with the **No-arbitrage principle**. What we do is assume that the forward value is above or below our stated value, and prove that under these assumptions there is an arbitrage opportunity.

1. Assume that $F(t, T) < S_te^{r(T-t)}$. Then, you can buy a forward contract at delivery price $K = F(t, T)$ for free. You can also short the underlying stock at S_t and invest the proceeds at the interest rate r . At time T , you can take your money, which is now worth $S_te^{r(T-t)}$ and execute the contract by buying one share of the stock for $K = F(t, T)$. This leaves you with no stocks and money of $S_te^{r(T-t)} - F(t, T) > 0$ by assumption, and you have an arbitrage opportunity.
2. Assume that $F(t, T) > S_te^{r(T-t)}$. Then, you can sell a forward contract at delivery price $K = F(t, T)$ for free. I borrow S_t money at the interest rate r and buy a stock. At time T , I sell the stock at $F(t, T)$ and also must pay back from who I borrowed at a rate of $S_te^{r(T-t)}$. Therefore, this leaves me with a net gain of $F(t, T) - S_te^{r(T-t)} > 0$ by assumption, and I have an arbitrage opportunity.

Therefore, since I have an arbitrage opportunity, it must be the case that $F(t, T) = S_te^{r(T-t)}$.

Intuitively, with the no arbitrage strategy, we can see that if the forward price is too high relative to the stock price S_t , we buy the stock now and sell the contract (sell the stock forward at the higher price). If the forward price is too low, then we sell the stock now and buy the contract (buy the stock forward at the cheaper price).

Corollary 9.1 (Dependence on Forward Price)

The forward price is only dependent on the current underlying price S_t , the interest rate r , and the time to maturity $T - t$. Counterintuitively, it does not depend on the growth rate, the standard deviation, or any distributional assumptions of S_T . Furthermore, any two assets which pay no income and which have the same spot price S_t will have the same forward price regardless of any views about their future movements.

9.4 Forward Price of Income Assets

However, this is unrealistic in a few things. It first does not account for dividends, the cost of storage, or the cost of carry.

Theorem 9.2 (Forward Price of Income Asset)

Suppose an asset pays a known amount of income (e.g. dividends, coupons, rent) during the life of the forward contract and the present value at t of the income is I . Then,

$$F(t, T) = (S_t - I)e^{r(T-t)} \quad (40)$$

Proof.

Replication. Let us have two portfolios.

1. A: Consists of one unit of the underlying stock.
2. B: Consists of one long forward contract with delivery price K , plus $Ke^{-r(T-t)} + I$ of cash.

Then, B will invest the cash at the risk free rate r and get $K + Ie^{r(T-t)}$ cash. Then it uses K to buy the underlying stock, ending up with $Ie^{r(T-t)}$ in cash and one unit of stock. A, by owning the stock throughout the span, also receives I of cash (present value), which is equal to $Ie^{r(T-t)}$ worth at time T . Therefore, the two portfolios are equal at time T and therefore must be equal at time t .

$$S_T + Ie^{r(T-t)} = V_K(T, T) + K + Ie^{r(T-t)} \implies S_t = V_K(t, T) + Ke^{-r(T-t)} + I \quad (41)$$

Therefore, we can solve for the forward price by setting $V_K(t, T) = 0$.

$$S_t = Ke^{-r(T-t)} + I \implies F(t, T) = K = (S_t - I)e^{r(T-t)} \quad (42)$$

Proof.

No Arbitrage. We assume the two scenarios.

1. $F(t, T) < (S_t - I)e^{r(T-t)}$. Then, we buy the forward contract for free at $K = F(t, T)$ and short the stock at S_t . We take the proceeds of the short and invest it at the risk free rate r . At time T , we buy back the stock at $F(t, T)$, and give back the stock to the lender, and additional pay them the dividends they would have received, which is $Ie^{r(T-t)}$. Therefore, our net profit is

$$S_te^{r(T-t)} - F(t, T) - Ie^{r(T-t)} > 0 \quad (43)$$

2. $F(t, T) > (S_t - I)e^{r(T-t)}$. Then, we buy the stock at S_t and short the forward contract at $K = F(t, T)$ for free. At time T , we get our income $Ie^{r(T-t)}$, pay back the lender at $S_te^{r(T-t)}$, and execute the contract to sell it at $F(t, T)$. Our net profit is

$$F(t, T) + Ie^{r(T-t)} - S_te^{r(T-t)} > 0 \quad (44)$$

Therefore, we have an arbitrage opportunity in both cases, and therefore it must be the case that $F(t, T) = (S_t - I)e^{r(T-t)}$.

If the stock pays some income at a compounded rate, then we have the following result on the forward price.

Theorem 9.3 ()

If the stock pays a continuous dividend yield q (e.g. $q = 0.02$ means that the stock pays 2% of its value in dividends each year), then the forward price is

$$F(t, T) = S_te^{(r-q)(T-t)} \quad (45)$$

9.5 Valuation of Forward Contracts

Recall from our example of the asset that pays no income. The replication argument gave

$$S_t = V_K(t, T) + Ke^{-r(T-t)} \quad (46)$$

If we substitute $F(t, T) = S_te^{r(T-t)}$ (that is, the forward price is the spot price compounded at the risk free rate), then we get

$$V_K(t, T) = (F(t, T) - K)e^{-r(T-t)} \quad (47)$$

which is the difference between the forward price and the delivery price, discounted back to today.

Theorem 9.4 (Valuation of Forward Contracts)

The value of a forward contract on an asset satisfies

$$V_K(t, T) = (F(t, T) - K)e^{-r(T-t)} \quad (48)$$

Proof.

No Arbitrage. We assume the two scenarios.

9.6 Forward Interest Rates**10 Forward Rates and Libor**

One might want to borrow money at a future time, say from T_1 to T_2 , but they are not sure what the interest rate will be. Therefore, they can buy a forward contract that expires at T_1 , with an underlying zero coupon bond that has maturity of T_2 . This essentially allows them to lock in the interest rate r_{T_1, T_2} for a future loan between T_1 and T_2 . This is called the *forward rate*.

Definition 10.1 (Forward Rate)

The **forward rate** at current time t for period T_1 to T_2 ($t \leq T_1 \leq T_2$) is the interest rate (agreed at t) at which one can borrow or lend money from T_1 to T_2 , i.e. the interest rate r_{T_1, T_2} .

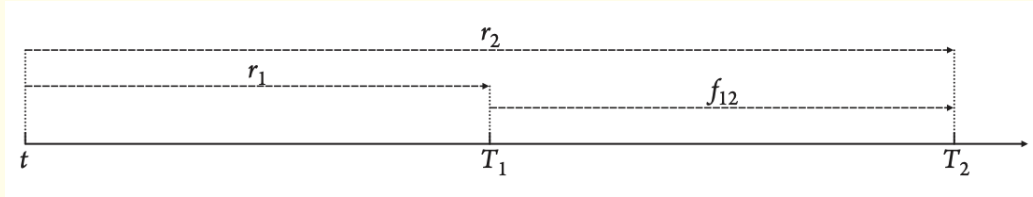


Figure 10: Letting r_1 be the interest rate from t to T_1 and r_2 be the interest rate from t to T_2 , we can see that the forward rate $f_{12} = r_{T_1, T_2}$ is the interest rate from T_1 to T_2 .

A simple replication argument gives the following result about the forward rate.

Theorem 10.1 (Forward Rate)

At time t , the fair value of the forward rate f_{12} is given by

$$f_{12} = \frac{(r_2(T_2 - t) - r_1(T_1 - t))}{T_2 - T_1} \quad (49)$$

if the rates are continuously compounded and

$$f_{12} = \left(\frac{(1 + r_2)^{T_2 - t}}{(1 + r_1)^{T_1 - t}} \right)^{\frac{1}{T_2 - T_1}} - 1 \quad (50)$$

Proof.

By the replication argument, consider two portfolios A and B with the same money at time t . A first borrows money at the rate r_1 and then at the forward rate f_{12} , and B borrows at the rate of r_2 .

Then, the value of the two portfolios must be the same at time T_2 , and so we have

$$e_{r_1(T_1-t)} e_{f_{12}(T_2-T_1)} = e_{r_2(T_2-t)} \implies f_{12} = \frac{(r_2(T_2-t) - r_1(T_1-t))}{T_2 - T_1} \quad (51)$$

and if they are discretely compounded we have

$$(1 + r_1)^{T_1-t} (1 + f_{12})^{T_2-T_1} = (1 + r_2)^{T_2-t} \implies f_{12} = \left(\frac{(1 + r_2)^{T_2-t}}{(1 + r_1)^{T_1-t}} \right)^{\frac{1}{T_2-T_1}} - 1 \quad (52)$$

Theorem 10.2 (Forward Price of Zero Coupon Bonds)

For $T_1 \leq T_2$, consider a forward contract with maturity T_1 on a ZCB with maturity T_2 . That is, the underlying asset price is $S_t = Z(t, T_2)$. Then, the forward price, i.e. the price where one can (for no cost) agree to buy a forward contract on the T_2 -ZCB with a expiry date of T_1 , is

$$F(t, T_1, T_2) = \frac{Z(t, T_2)}{Z(t, T_1)} \quad (53)$$

Proof.

Let

1. Portfolio A be one ZCB maturing at T_2 .
2. Portfolio B be one long forward with delivery price K , and K ZCBs maturing at T_1 .

Corollary 10.1 ()

Since the value of a ZCB with annually compounded rate is

$$Z(t, T_i) = \frac{1}{(1 + r_i)^{T_i-t}} \text{ for } i = 1, 2 \quad (54)$$

we can substitute to see that the forward price of a ZCB is

$$F(t, T_1, T_2) = \frac{Z(t, T_2)}{Z(t, T_1)} = \frac{(1 + r_1)^{T_1-t}}{(1 + r_2)^{T_2-t}} = \frac{1}{(1 + f_{12})^{T_2-T_1}} \quad (55)$$

11 Swaps

12 Futures

Note that when trading futures, it is exchange traded so every trade between a buyer and a seller must go through a clearing house. Let's talk about the properties of a future contract.

Definition 12.1 (Components of a Future Contract)

A future contract must specify the following components.

1. **Asset:** What is the underlying asset?
2. **Grade:** If the asset is a commodity, we want to specify the quality of the asset allowed at some price. Alternatively, we can talk about the **expected grade**, which allows for a range of grades depending on the actual grade of the contract.

3. **Contract Size:** The quantity of the asset. How much of it is being traded? The price of the contract shown is usually for one unit of the underlying commodity, so the actual value of the contract is the price times the contract size. These sizes are determined by the exchange (should be large enough for institutions to be interested in but not too big to be unaffordable). Furthermore, the mini contracts are less liquid and have a higher bid ask spread.
4. **Delivery Location:** Where is the location? This is also standardized by the exchange, and the buyer must personally deliver the asset from these locations to wherever the asset is needed.
5. **Delivery Time:** The month of contract, which is the month that the commodity will be delivered and is not necessarily the same as the expiration date.
6. **Price Increments:** The minimum price change allowed.
7. **Daily Price Limits:** The maximum price change allowed in a day, and any price change beyond this limit will halt trading (called *limit up* or *limit down*). However, trading can still resume if it goes down.
8. **Position Limits:** The maximum number of contracts that a trader can hold. This is to prevent parties from cornering a market.

Example 12.1 (Futures Pricing and Size Standards)

Here is an overview of the pricing and size standards of many futures contracts set by the exchange.

Underlying	Size	Increments	Ex. Quote
Crude Oil (CL)	1000 barrels	1 c (\$10.00)	\$86.73
Crude Oil (CL) Mini	500 barrels	1 c (\$5.00)	\$86.73
Natural Gas (NG)	10,000 mmBtu	$\frac{1}{10}$ c (\$10.00)	\$2.843
Corn (ZC)	5000 Bushels	$\frac{1}{8}$ c (\$12.50)	$400\frac{3}{8}$ c
Wheat (ZW)	5000 Bushels	$\frac{1}{8}$ c (\$12.50)	$568\frac{1}{4}$ c
Soybean (ZS)	5000 Bushels	$\frac{1}{8}$ c (\$12.50)	$1655\frac{1}{2}$ c
Gold (CS)	100 Troy Ounces	10 c (\$10.00)	\$2349.10
Silver (SI) Full	5000 Troy Ounces	$\frac{1}{10}$ c (\$5.00)	\$27.247
Silver (SI) Micro	1000 Troy Ounces	$\frac{1}{10}$ c (\$1.00)	
Copper (NG)	25,000 Pounds	$\frac{1}{200}$ c (\$12.50)	
E-Mini S&P 500	50	0.25 c (\$12.50)	
E-Mini Nasdaq 100	20	0.25 c (\$5.00)	
E-Mini Dow	5	1 c (\$5.00)	
E-Mini Russell 2000	50	0.10 c (\$5.00)	
Micro E-Mini S&P 500	5	0.25 c (\$1.25)	
Micro E-Mini Nasdaq 100	2	0.25 c (\$0.50)	
Micro E-Mini Dow	0.5	1 c (\$0.50)	
Micro E-Mini Russell 2000	5	0.10 c (\$0.50)	
Treasury Bonds	\$100,000	$\frac{1}{32}$ c (\$31.25)	
Treasury Bonds	\$200,000	$\frac{1}{32}$ c (\$62.50)	

Table 2: Standards of Futures Contracts

1. **Corn** is delivered to Toledo, Ohio.
2. **Oil** (WTI) is delivered to Cushing, Oklahoma.
1. For oil, natural gas, and copper, the delivery month is the month after the expiration month and delivers on all months.
2. For corn, wheat, and soybean, we only deliver on the months of March, May, July, September, and December.

Example 12.2 (Harvest and Delivery Standards)

While most futures deliver throughout the year, for grain (corn, wheat, soybean) futures, they deliver on the months of March, May, July, September, and December. This is because these are the months of the harvest.

It turns out that this discrepancy is important in analyzing the convergence of the future price to the spot price and determining the risk.

13 Options

Definition 13.1 (Long Options)

An option is a contract that gives the owner the right, but not the obligation, to buy or sell a specific asset (the **underlying**) at a specific price K (**strike** or **exercise price**) at a specific date T (**expiration date**). When you buy (long)

1. a **call option** for $\$X$, then you have the *right* to buy the asset and the seller has the *obligation* to sell.

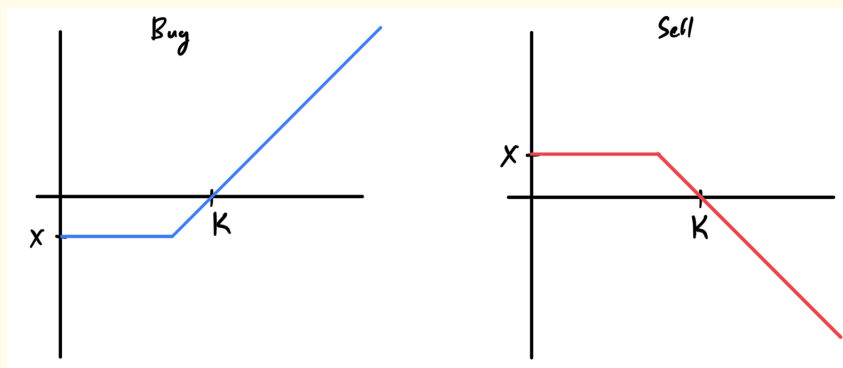


Figure 11: Note that the buyer has limited loss and has unlimited gain with the long put, while the seller has limited gain and unlimited loss with the short put.

2. a **put option** for $\$X$, then you have the *right* to sell the asset and the seller has the *obligation* to buy.

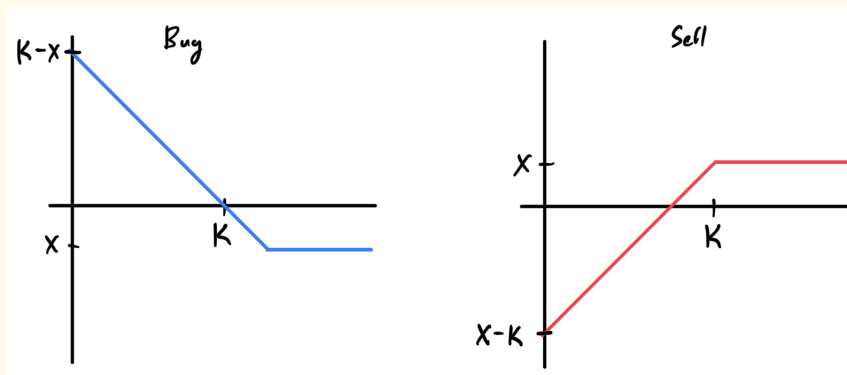


Figure 12: Note that the buyer has limited loss and limited gain, and the seller also has limited loss and limited gain.

There are three types of options.

1. **American style** options can be exercised anytime during the life of the option.
2. **European style** options can only be exercised at the expiration date.
3. **Bermuda style** options can be exercised at specific dates during the life of the option.

This has nothing to do with which continent the options are trading in. It is just a name. Most exchange traded options are American style and most OTC options are European style.

Note that unlike forwards or futures, the premium changes hands, i.e. the buyer pays the seller to enter the contract.

14 Exercises

Exercise 14.1 (Hull 1.3)

What is the difference between entering in to a long forward contract when the forward price is \$50 and taking a long position in a call option with a strike price of \$50?

Solution 14.1

The answer is clear when we look at the payoff graph for each one.

1. The first difference is that the forward contract renders both parties obligated to buy or sell the underlying at a future time, while the option only gives the seller the obligation to sell, if the buyer requests it.
2. The second is that money is paid to enter into the options contract by the buyer, while the forward contract does not.
3. Both profits/losses are unlimited in the forward contract, while in the options contract, the buyer has limited loss and the seller has unlimited loss.
4. By having the safety of a limited downside for the call options buyer, they get a slightly lower payoff $S_T - C_0$ compared to the forward contract, which is S_T . Therefore, the options buyer doesn't actually make a profit until the price of the underlying asset is at least $C_0 + K$.
5. If $S_T > K$, we would wish we were taking the forward, and if $S_T < K$, then we wish we were taking the call option.

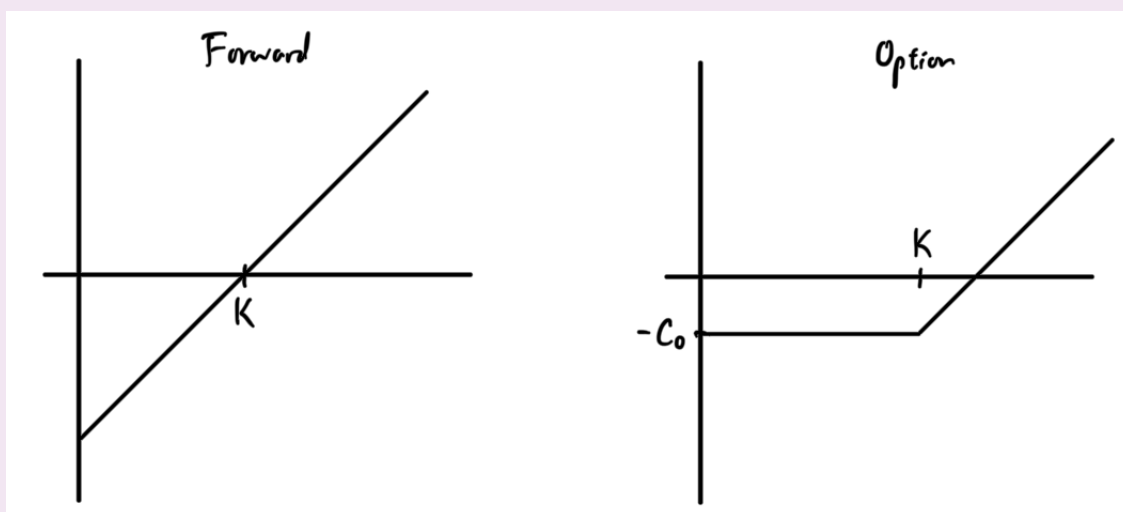


Figure 13: The payoff graph for the forward contract and the call option from the buyers side.

Definition 14.1 (Pips)

For the next problem, it helps to know that a **pip** is $1/100$ of a cent.

Exercise 14.2 (Hull 1.5)

An investor enters into a short forward contract to sell 100,000 British Pounds for US dollars at an exchange rate of 1.5 dollars per pound. How much does the investor gain or lose if the exchange rate at the expiration date is 1.4900 dollars per pound? What if it is 1.5200 dollars per pound?

Solution 14.2

Note that we must specify the currency of the profit or loss. It turns out that the payoff will always be in the currency that the investor is selling. The investor is selling the forward, so they are obligated to sell it at 1.5000 dollars per pound.

1. If the rate is 1.4900, then the investor sells the pounds for 150,000 USD and can buy a little bit more than 100,000 pounds. In fact, they are making a profit of $100,000 \cdot 0.0100 = 1,000$ dollars.
2. If the rate is 1.52, then the investors sells the pounds for 150,000 USD and can buy a little less than 100,000 pounds, so they lose $100,000 \cdot 0.0200 = 2,000$ dollars.

Exercise 14.3 (Hull 1.6)

A trader enters into a short cotton futures contract when the futures price is 50 cents per pound. The contract is for the delivery of 50,000 pounds. At the end of the contract, the price of cotton is 48.20 or 51.30 per pound. What is the trader's gain or loss?

Solution 14.3

The trader is selling the cotton, and the buyer is obligated to buy. Therefore,

1. the trader makes $(50 - 48.20) \cdot 50,000 = 1.8 \cdot 50,000 = 90,000$ cents, or \$900 if the price is 48.20.
2. the trader loses $(51.30 - 50) \cdot 50,000 = 1.3 \cdot 50,000 = 65,000$ cents, or \$650 if the price is 51.30.

Exercise 14.4 (Hull 1.16)

A trader writes (sells) a December put option with a strike price of \$30. The price of the option is \$4. Under what circumstances does the trader make a profit?

Solution 14.4

Even without a payoff graph, we can see that since I'm selling the option, I'm getting \$4. Now, I'm obligated to buy the asset (since it's a put) at \$30, and therefore I will profit if the price of the asset is above \$30. However, since I already have the cushion of \$4, I'm really profiting if the asset price $S_T > 26$.

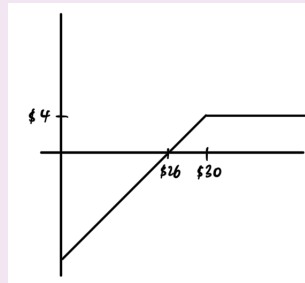


Figure 14: The payoff graph for the put option from the sellers side.

Exercise 14.5 (Hull 1.22)

I am in a long forward contract and a long put option at the same strike price K . What is the payoff of this position at expiration?

Solution 14.5

You can simply draw the profit-loss diagram.

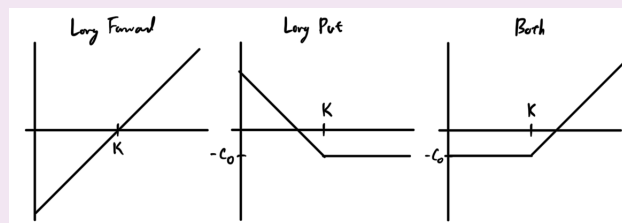


Figure 15: The payoff graph for the long forward and long put option at the same strike price.

Exercise 14.6 (Hull 1.23)

We have a **Index Currency Option Note (ICON)** which states the following:

1. If the exchange rate $USD.JPY < 169$, then the bond pays 1000 USD.
2. If $84.50 \leq USD.JPY \leq 169$, then the bond pays

$$1000 - \max \left\{ 0, 1000 \left(\frac{169}{S_T} - 1 \right) \right\} \quad (56)$$

3. If $USD.JPY < 84.50$, then the bond pays nothing.

Show that this ICON is a combination of a bond and two options.

Solution 14.6

This seems a bit easy since we can draw the payoff diagram as such.

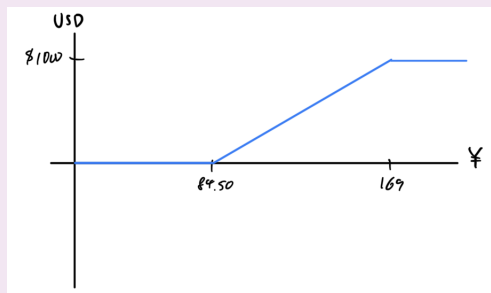


Figure 16

However, the underlying asset is shown in yen and the payoff in USD, which complicates calculations. You should rather change the underlying to be USD by changing USD.JPY to JPY.USD. Then, we can do some quick sketches to find that you should have a

1. Bond at \$1000.
2. Short call at $1/169$ for \$1000.
3. Long call at $1/84.50$ for \$1000.

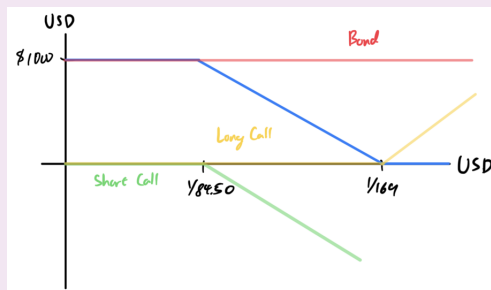


Figure 17

15 Appendix

Futures are mostly cash settled (e.g. look at futures on SP500 index. you can't deliver fractional shares.)

Margin calls with futures.

The difference between the spot and the futures price is called the basis gap/risk. They tend to converge as we approach the expiration date.

Order types (market, limit - default, stop), which exists for both long and short futures.

Regulations are done by the CFTC (Dodd-Frank Act expands CFTC oversight to OTC markets due to 2008). Accounting (does unrealized gains count as taxable income?) and taxes can be country specific.

Talk about derivatives in general - stuff like payoff graphs - definitions of derivatives

Forwards & Futures - how they work, with trading, delivery, quality of goods, settlement (delivery or cash settled). Does delivery actually happen? You must actually request to have it delivered to a warehouse you set up. This could have inventory/delivery fees as well. - basis risk (spot price of asset A is S_A , and the futures is F_A). (8) - hedging strategies with futures (short & long futures) hedge ratio, with cross-hedging, and minimizing the variance of our position. (is this delta-hedging?) - stock index futures (you need CAPM to evaluate this)