Quantitative Finance

Muchang Bahng

Summer 2024

# Contents

1 Interest Rates and Equity

The first thing we would like to have a model for is the value of a stock.

> **Definition 1.1 (Stock Price)**
>
> A stock is a stochastic process X_t.[a]

1.1 Supply and Demand Effects of Price

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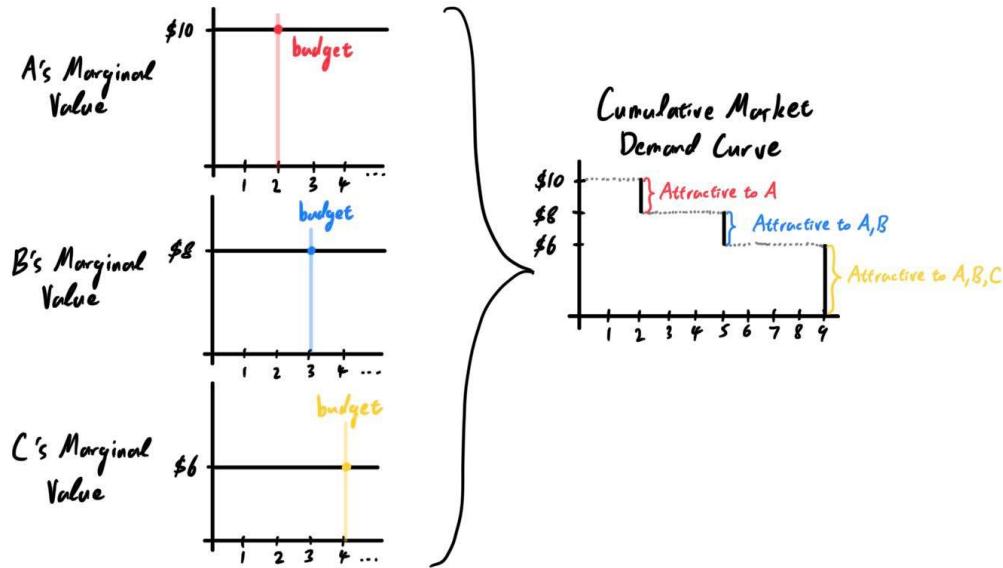
The market makers profit off of the **bid-ask spread**, which is the difference in the buy and sell price. If a seller wants to sell at $99.99 and a buyer wants to buy at $100.00, then the market maker can buy the stock from the seller for $99.99 and sell it to the buyer at $100.00, making a $0.01 profit.

We can also look at it from an economic perspective. We can draw the supply and demand curves for a certain company's stock. We can treat stocks as goods, with the company being the supplier and public investors as the consumers.

1. The demand curve says: as the price for a stock, i.e. a piece of the company, increases the quantity demanded by the market (consumers) will decrease because the costs of buying the stock will outweigh the marginal value of the share. There is a very important distinction to make here between stocks and other type of goods. Unlike consumable goods which we can clearly observe a diminishing marginal value (drinking 1 can of soda vs 2 cans vs 3 cans vs...), stocks do not seem to have any diminishing marginal value. That is, if I buy 10 shares, I expect to get precisely 10 times the profit than if I bought just 1 share. The value is not diminishing, so why is the demand curve monotonically decreasing? That is because it is not diminishing *for an individual*, but if we look at a market of many individuals, they all have different interpretations of what the marginal value of the stock is. Individual A may interpret the marginal value as $10, while B sees it as worth $8, and C sees $6. All three marginal utility curves would be horizontal, but if we take the market utility, we can see that it is decreasing (e.g. at the price of $7, a share is attractive to C but not A, B, and at $9, a share is attractive to B, C but not A). Buying an infinite number of shares is not realistic, so let us assume that each person has $25 worth to invest.

---

[a]We will work with both discrete and continuous time stochastic processes.
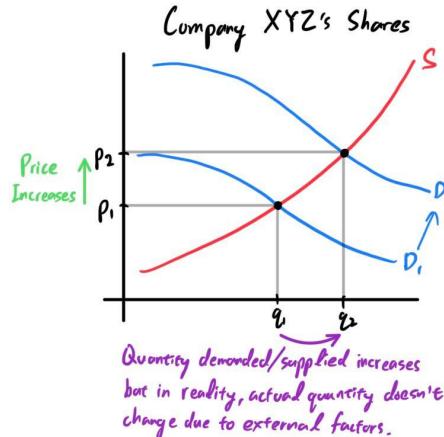
We can clearly see that the sum of these individual horizontal marginal value functions summed up produces a monotonically decreasing marginal value function and a monotonically decreasing demand curve for the total market.

2. The supply curve is simpler to interpret, since the only supplier is the issuing company (or whatever underwriting investment banks). By the law of supply, we can interpret that if the worth of the company's shares increases, then it would be logical for the company to issue more shares, all else being the same. This interpretation may be too oversimplistic, since we have to account for investor relations, dilution, and other factors, but it will do for now to justify that the supply curve of a stock is monotonically increasing.
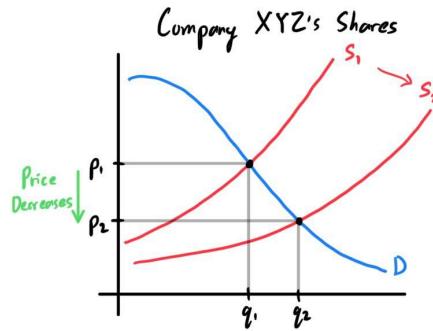
Note that in general, the demand of a company's shares is much more volatile than the supply since there are many ways for the demand to shift, while few for the supply. The demand (not quantity demanded) of a company's shares can be changed from news, world events, production, and such, which occurs often and may even violently shift the outlook of the business. Unlike certain goods that are produced continuously, a company's floating shares will not change unless there is a new issuance or a private investors sells shares at market price. Therefore, supply (not quantity supplied) does not change easily.

Now, let us have the supply and demand curves of a company's share. Assume that it is at the equilibrium position, where p_1^* is the current market price (bought and sold) and q_1^* is the quantity demanded and supplied, as balance.

1. If the demand increases, then the shift in the demand curve raises the equilibrium price to p_2^*, which raises the price of the stock. The quantity demanded/supplied also increases to q_2^*, but this desire to increase quantity cannot be met due to other factors. Therefore, we can interpret this case as such: The increase in demand can change the price accordingly, but the increase in desired quantity is offset by other factors (not being able to issue stocks freely).

2. If the supply increases, then this is much simpler to explain. Due to dilution, one stock now represents ownership of a smaller portion of the company, driving the price down. The equilibrium (optimal) quantity demanded and supplied increases due to this higher supply, which is reflected in the reality that there are more more shares issued now.

1.2 Returns and Dividends

We can work out its return as the percentage increase or decrease of its price. But given that X_t is close to X_s, we can look at the Taylor expansion of $\ln(X_t)$ to be

$$\ln(X_t) \approx \ln(X_s) + \frac{1}{X_s}(X_t - X_s) \implies \frac{X_t - X_s}{X_s} \approx \ln(X_t) - \ln(X_s) = \ln\left(\frac{X_t}{X_s}\right) \tag{1}$$

Therefore, we work with *log returns* as an appropriate approximation of the return. As we will see later, the Black Scholes, which deals with continuous compounding interest rates, require lognormal as well.

> **Definition 1.2 (Log Return)**
>
> The **log return**, or **return**, of a stock from time s to t is defined
>
> $$R_{s,t} = \frac{X_t - X_s}{X_s} \approx \log\left(\frac{X_t}{X_s}\right) = \log(X_t) - \log(X_s) \tag{2}$$
>
> However, the return is almost always annualized, so we just write the return as R.

> **Question 1.1 (Lognormal)**
>
> This is because like a bunch of other stuff, we assume normal distributions. But it has specific flaws, which the lognormal fixes.

Definition 1.3 (Lognormal Distribution)

It has its advantages:

1. It is time-additive: If we have prices X_r, X_s, X_t for $r < s < t$, we can see that

$$R_{[i,j]} + R_{[j,k]} = \log\left(\frac{P_j}{P_i}\right) + \log\left(\frac{P_k}{P_j}\right) = \log\left(\frac{P_k}{P_i}\right) = R_{[i,k]} \tag{3}$$

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 $X_t = X_s(1-k)$ 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.

Definition 1.4 (Dividend)

Given time s, t with $s < t$, the dividend that a stock X pays within time interval $[s, t]$ is the random variable $D_{s,t}$

Definition 1.5 (Total Return)

Given that you own stock X, the amount of money you would have made from time s to t is the random variable

$$X_t - X_s + D_{s,t} \tag{4}$$

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

$$\mathbf{R}_{s,t} = \frac{X_t + D_{s,t} - X_s}{X_s} \approx \ln(X_t + D_{s,t}) - \ln(P_i) \tag{5}$$

which is the price return plus the dividend rate. Again, the total return is almost always annualized.

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

Definition 1.6 (Expected Return)

The **expected return** of a stock from time s to t is simply the expected value

$$\mathbb{E}[R_{s,t}] = \mathbb{E}[\log(X_t) - \log(X_s)] \tag{6}$$

where the expectation can be over the whole σ-algebra of sequences or can be conditioned on the two random variables X_s, X_t. The expected return over a year is $\mathbb{E}[R]$.

1.3 Interest Rates

Definition 1.7 (Zero Coupon Bonds)

A convenient way to measure interest rates are through *zero-coupon bonds*, which do not pay coupon payments but are instead issued at a discount to their par value. Upon maturity, it pays the bondholder in full face value (e.g. U.S. treasury bills).

This means that the value of a ZCB is precisely determined only by the interest rate r. Therefore, we define the value of a ZCB as follows.

Definition 1.8 (ZCB)

A ZCB with maturity T that pays 1 has a value of

$$Z_r(t,T) \tag{7}$$

where the r may be omitted if the rate is assumed to be known or not calculated yet. Depending on how we model it, it may be treated as
    1. discrete, where

$$Z_r(t,T) = (1+r)^{T-t} \tag{8}$$

    2. continuous, where

$$Z_r(t,T) = e^{-r(T-t)} \tag{9}$$

While we won't directly work with ZCBs as much, the following provides a very convenient notation to determine prices.

Example 1.1 (Basic ZCB Application)

Given that you will receive \$100 in time T, the present value of this future cash flow is

$$100 \cdot Z(t,T) \tag{10}$$

If you have the decision to receive X dollars today or Y dollars in 1 year, then we can take the future cash flow Y and discount it to today and compare them. We should therefore accept Y in 1 year iff

$$Y \cdot Z(t, t+1) > X \tag{11}$$

Therefore, to calculate the present value of some future cash flow, just multiply it by Z. If you want to take some present cash flow and forward it into the future, then divide it.

1.4 Measures

While the expected returns are important, the concept of **risk** is equally if not more. This is a slippery term, used to mean different things, but it is essentially some characterization of the distribution of the returns. The variance, or standard deviation, of R was one form of risk, but there are other ones that we will introduce here.

Definition 1.9 (Sharpe Ratio)

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

$$S = \frac{R - R_f}{\sigma} \tag{12}$$

> 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.

Another property we would like to investigate is perhaps some tail bounds.

> **Definition 1.10 (Value at Risk)**
>
> The **value at risk (VaR)** over a time horizon of T at $p \in [0, 1]$ is
>
> $$\Phi^{-1}(p) \tag{13}$$
>
> which tells us the loss we could incur by the end of T with a probability of p, where $\Phi()$ is the CDF of $R_{0,T}$. p is usually taken to be 1% or 5%.

2   Secondary Markets

A secondary market is where securities exchange hands between two parties. For equity or commodities, there is a supply that circulates hands. This motivates the following definitions to describe the size of a market.

> **Definition 2.1 (Liquidity)**
>
> **Liquidity** is defined by the extent to which one can buy and sell large quantities of a security without moving the price.

> **Definition 2.2 (Volume)**
>
> **Volume** is the amount of trades that are made in a day. If party A buys from party B, the volume has just increased by 1.

When trading *contracts* such as derivatives, there is a slight nuance since there is no fixed supply. If two parties want to agree on some terms, a new contract is simply created, where one buys the contract and the other sells. If one buys a contract and then sells the same contract, it is the same as if they did not buy anything in the first place.

> **Definition 2.3 (Open Interest)**
>
> **Open Interest** is the amount of contracts outstanding. If party A sells a future to party B, the open interest just increased by 1 since a contract has just been created (out of thin air!). If B now sells to C, open interest doesn't change, but volume +1. If C sells to A, Open Interest $= 1 - 1 = 0$.

It turns out that in commodity futures markets, over 99% of all future contracts are cancelled out so that investors are net 0.

2.1 Order Types and Queue Priorities

Definition 2.4 (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.

There are other order types as well, which we will mention but are not relevant unless one starts to trade.

1. *Good For Day (GFD).* The order is canceled at the end of the day.

2. *Good Till Canceled (GTC).* The order stays on the books until canceled manually.

3. *All or None (AON).*

4. *Fill or Kill (FOK).*

5. *Immediate or Cancel (IOC).*

6. *Fill and Kill (FAK).*

7. *One Cancels the Other (OCO).*

Definition 2.5 (Short Selling)

In addition to holding a long position, we can **short sell**, or **short**, a stock.[a] Given that some shares are priced at X, we can borrow shares from some investor (probably an institutional investor), sell it on the market for X, and hope for the price to drop so we can buy it back.
1. If the price drops to $X - A$, then we buy it back at a lower price and make A in profit.
2. If the price rises to $X + A$, then we must buy it back at a higher price and lose A.

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.

---

[a]Selling something you don't own is a new concept to those new to finance, but think of it as owning -1 shares.

Definition 2.6 (Short Squeeze)

The third risk is the deadliest, and at worst, you can get **short-squeezed**. Let's explain how this works.
1. Say you and a bunch of other investors shorted GME.
2. 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.
3. 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.
4. 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.

Definition 2.7 (Queue Priority)

An important part is to know about **queue priority**, which represents the order in which different orders are filled.
1. Usually, it is ranked first by price and then by time of entry.
2. Sometimes, it is ranked by volume for specific markets.

2.2   Relevant Parties

Let's think about the people who would want to trade, beginning with people that are out of the exchange.

Definition 2.8 (Off-Floor Parties)

Off Floor Parties mean outside traders who are interacting with the exchange. They include
1. *Retail investors* like small, non-professional traders that might dabble in options.
2. *Consumers* such as Kellogs (which needs a lot of corn to make cereal) might be interested in buying contracts to hedge their underlying exposure.
3. *Producers* like corn farmers might want to buy some contract to lock in the price in which they can sell it at.
4. *Hedge Funds* who may have more information and research (e.g. drones flying over farms).
5. *Banks* who want to trade on behalf of clients or do proprietary trading.
6. *Funds* who manage large money for clients.

There are a lot of middlemen in exchanges that take on risk in return for profits (either fixed through a fee like clearing houses or variant through profits like market makers).

Example 2.1 (Risks Associated with Trades)

Let's think about what kinds of risk there are when investing a stock. 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.
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 2.9 (Counterparty)

A **counterparty** is the other party in a financial transaction. **Counterparty risk** is the likelihood or probability that one of those involved in a transaction might default on its contractual obligation. It can exist in credit, investment, and trading transactions.[a]

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 2.10 (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 *capital requirements*.[1]

Definition 2.11 (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.

Therefore, the cost of going into the trade is not just the trade itself, but also the capital requirements, since they imply an opportunity cost of not being able to invest in other securities. These requirements are generally very inflexible unless you have a good relationship with your broker and are a big trader.

---

[a]This is formalized through creditworthiness, which takes many forms such as bond ratings, corporate credit ratings, and FICO scores.

[1]This is slightly different from margin. We will get to margin later.

The clearinghouse is the first entity that should exist, but the vast majority of investors do not interact directly with the clearinghouse. Rather, there is another middleman between the investor and the clearinghouse, called the *broker*.

> **Definition 2.12 (Broker)**
>
> A **broker** is a person or firm that executes orders on behalf of a trader.[a] Their responsibilities include:
> 1. Maintaining the account of an investor who wants to trade.
> 2. The margin required by the clearinghouse is also sent to the trader, and the trader must have this margin in their account with the broker.

We can have a functional market composed of investors, brokers, and clearinghouses. However, this market would be quite illiquid with very few executions compared to the modern markets we see now. The reason we have such liquidity is because of *market makers*, also known as *liquidity providers*.

> **Definition 2.13 (Market Makers)**
>
> **Market makers** are firms that quote a bid and ask price for a security. By doing this, traders will always see someone willing to buy or sell some product, enhancing liquidity and therefore "making a market." Market making firms like Akuna have pit traders, who have headsets and tablets, that are in real-time communication with the screen traders upstairs.

The off-floor traders like us (retail, institutional, etc.) are **market takers**. Some famous market makers are Jane Street/Citadel in the equity markets, and Optiver/Akuna in the options markets.
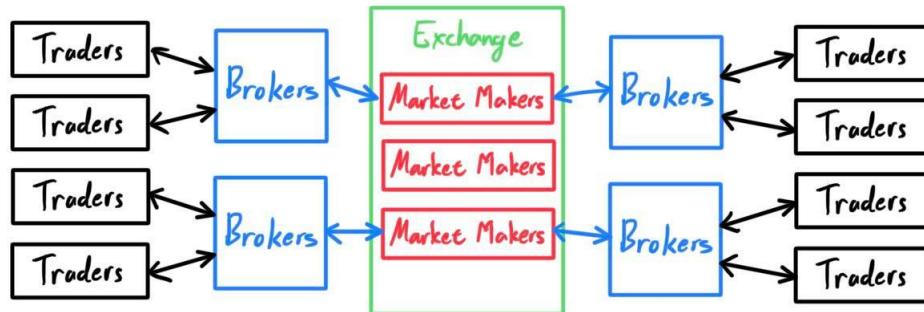
Figure 1: General structure of a trading floor/exchange.

Now let's see how this all comes together.

1. An off floor party is looking to buy shares of AAPL. He relays an *order* to the broker.

2. The broker, who is on the floor but out of the AAPL trading pit goes to the market makers of AAPL and asks them for a quote on AAPL shares.[2] The broker takes a commission for their services.

3. The market makers in the pit (along with screen traders upstairs) give the broker their quotes. Note that they usually must give a quote, i.e. both a bid and ask price, but sometimes they don't.

4. The broker takes the highest bid, say 99 and lowest offer, say 101, and relays it back to the investor. This is usually automated in exchanges with the *matching engine*, which chooses the best quotes and

---

[a]Familiar ones include Robinhood, Charles Schwab, Interactive Brokers, Fidelity.

[2]Before the rise of large market makers, *locals* were individuals who acted as small market makers. They were called locals because they were physically standing there with access to the local market. However, they don't exist due to the competitiveness of these firms.

splits the orders across market makers accordingly.[3]

5. The investor buys at price 101 and the market maker sells at 101. The trade is executed and recorded with the clearinghouse.

6. Imagine right after that someone wants to sell AAPL. The whole process happens again, and the market maker buys at 99. This trade is also executed and recorded in the clearinghouse.

Note that since market maker bought at 99 and sold at 101, they have profited on the $2 **spread**. In fact, the exchange/clearinghouse would take a fixed portion of this as fees as well, but if there are more trades and/or bigger spread, market makers can profit more. If the market maker needs to trade, then they must **cross the spread**, i.e. sell at the current market bid or buy at the current market offer.

The biggest threat to market makers are opposing traders who may make more advantageous trades than them. This can be due to information asymmetry (e.g. hedge funds, banks), larger order sizes (e.g. institutional investors can execute a block trade), or more advanced algorithms (e.g. hedge funds), but these counterparties are known as **papers**. The line for papers is a bit blurry and retail traders can be considered papers (typically a really aggressive trader is called a paper), but their threats are usually too small to be considered one. Since papers also go through brokers to trade with market makers, and brokers take commission from the papers' large orders. Therefore, the broker will always be on the side of the paper rather than the market maker (they can always go to another market maker if they lose one).

2.3 Mock Trading and Market Making

From the role of market makers, it is essential that market makers

1. buy and sell approximately equally across the spread, and

2. they don't have a heavy position on one side of the market due to directional risk.

This can be done by

1. having a good approximation on the *theoretical price*, or *theo*, of the security. This is done by having good theoretical models but can be minimized by having a directionally neutral position in the security (i.e. we don't care if the security moves up or down).

2. having fast execution on trades with efficient software and hardware. In fact, market makers often have FPGAs running automated trades near these exchanges, called **colocation**, which allows for this.

We go over basic terminology.

1. If we want to *buy*, we say *1.25 Bid for 50*.

2. If we want to *sell*, we say *Sell 10 at 1.25*.

An *order* consists of either a buy or a sell, but a *quote* consists of both a buy order and a sell order, e.g. *3.00 bid at 4.00 100 times*.

You can visualize this by imagining a market as some (futures) ladder.

---

[3]The algorithm for matching is transparent.
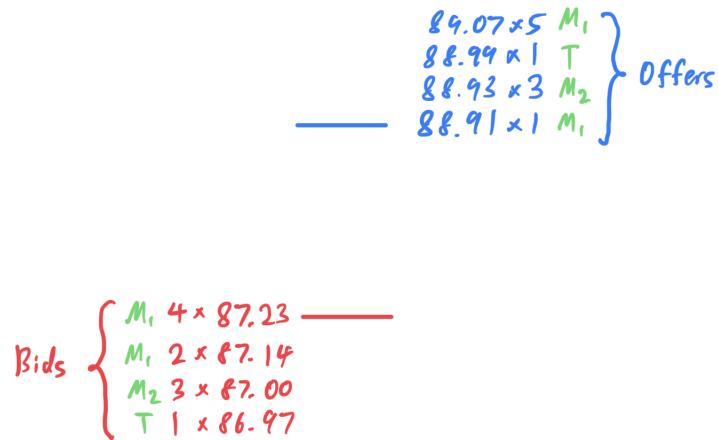
Figure 2: In a ladder, bids are placed in the bottom and offers are placed in the top. M_1, M_2 represent orders placed by market makers, but traders can also put in orders as well, marked by T.

Now when someone wants to actually execute a trade (not just quote it), they must *cross the spread*. They can

1. buy by *lifting the offer*, or

2. sell by *hitting the bid*.
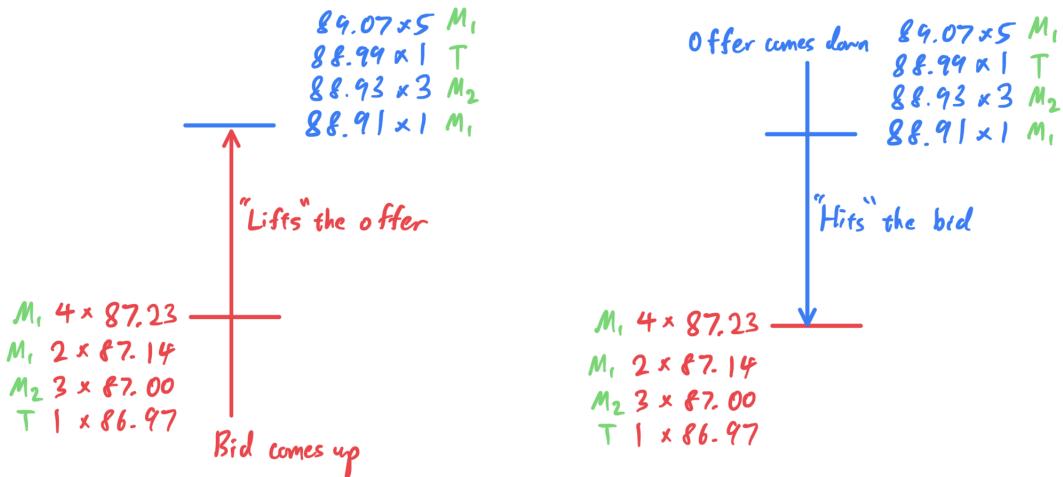
Figure 3: When a trader wants to execute a trade, they must cross the spread by either lifting the offer (buying from market) or hitting the bid (selling to market).

Now the question comes to, how does one make a good market?

Fading. You should never fade in the direction of the incoming trade (go down when someone bought or go up when someone sold).

Say that I have an estimate of 100 on the true price of some security. I want to quote a market around it, say 90 and 110. I want to make sure that you have enough edge. Say that my edge is 5. Then I don't want to make a market that is tighter than 10, i.e. 95 at 105.

1. What happens if someone (A) buys from me? Then I sell at 110 and he **lifts** my offer. He thinks that the price is higher and buys. Therefore, I should adjust my markets *up*. There are two scenarios that can happen now.

    (a) Say I lift my market up too little.

    Say someone buys from me and I lift my offer too much, to say 120 at 130. A can then sell to me at 120 and I buy at 120. A has essentially made an arbitrage trade since he bought at 110 and sold at 120. I just lost money, but on the bright side, I got more information and reduced the risk of A lifting me again.

    (b) Replaying this scenario, A can buy from me and I lift my market to 120 at 130. He can buy from me again, essentially lifting my market even more. I can keep on selling, making directional bets. This now becomes a game of whose theos are correct since I now accumulate a short position. This is not what market makers want to do, even though this gives me more information.

2. Now someone (B) sells to me and **hits** my bid. She thinks that the price is lower and sells. Therefore, I should adjust my markets *down*.

    (a)

If someone sells to me on my market of 90 at 110, she hits my bid at 90. Her theo is lower than 90, and I should lower my market.

> **Example 2.2 (2nd Session)**
>
> Some tips. The market was 100_2 at 130_4. Then 120_2 at 150_4. Then 180_6 at 210_0.
> 1. When you have no clue on what the market on a product should be, do not make the market too tight. If there are 6 traders and you make your market 1 tick wide, the most you can arb is 3 ticks. Is this reward justified for the risk taken? 30 was a bit tight but 50 might have been better.
> 2. If everybody bought 150, then why is everybody trading at 135?
> 3. Don't cross 10 levels in the ladder and clear out all the bids or all the offers.
> 4. There were times when someone sold 140 and bought 142 right after.
> 5. If the market tightens and no one's trading, it's a good time to sell some risk.
> 6. After info drop, you made a market of 135 at 137. There's too much risk taken.

2.4 OTC Markets and 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.

3   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.

3.1   Moving Averages

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

Definition 3.1 (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 3.2 (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 he 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.

3.2   Bollinger Bands

Definition 3.3 (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.

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

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

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

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.

3.3 Relative Strength Index

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

Definition 3.4 (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.

3.4 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 3.5 (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.

4   Statistical Arbitrage

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. Let's first look at an intuitive examples with temperature betting.

> **Example 4.1 (Temperatrue Arbitrage)**
>
> Say you are given the following trades:
> 1. 65 at 66 on the temperature of San Francisco, S
> 2. 45 at 47 on the temperature in Kyiv, denoted K
> 3. 25 at 26 on the temperature of San Francisco minus Kyiv, $S - K$
>
> What do you do? Let's compare these markets. You see that buying K and $S - K$, which you can do for $47 + 26 = 73$ must be the same as buying S alone, which is 66. Therefore there is a price mismatch. You should therefore long S, short K, and short $S - K$.

To do this, take two sets of stock prices P_t and Q_t and model them with the assumption that the log returns are linearly correlated by the following relationship.[4]

$$\frac{\Delta P}{P} = \beta \frac{\Delta Q}{Q} \iff \log(P_j) - \log(P_i) = \beta\big(\log(Q_j) - \log(Q_i)\big) \tag{14}$$

This results in the model

$$\log(P) = \beta \log(Q) + \alpha + \epsilon \tag{15}$$

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

$$\Delta \log(P) = \beta \Delta \log(Q) + \Delta\epsilon \iff \frac{\Delta P}{P} = \beta \frac{\Delta Q}{Q} + \Delta\epsilon \tag{16}$$

So, if we look at

$$\{\epsilon_i\}_{i=0}^n = \{\log(P_i) - \beta \log(Q_i) - \alpha\}_{i=0}^n \tag{17}$$

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 threshhold. 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.

---

[4]What beginners assume is that if P rises by 5%, then Q will also rise by 5%. But what if companies P and Q, who use the same supply chains, exist such that P is, say twice as dependent as Q? Then, if P goes down by 5%, then Q may go down by 2.5%, and our assumption should be that it is *linearly correlated*.

4.1 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:

$$\lambda = \frac{\mathrm{MV}_Q}{\mathrm{MV}_P}$$

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

$$V = \mathrm{MV}_P + \mathrm{MV}_Q$$

where $\mathrm{MV}_P = n_P P$ and $\mathrm{MV}_Q = n_Q Q$, 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

$$\begin{aligned}
\Delta V &= \Delta\,\mathrm{MV}_P + \Delta\,\mathrm{MV}_Q \\
&= \mathrm{MV}_P \left[\frac{\Delta\,\mathrm{MV}_P}{\mathrm{MV}_P} + \frac{\mathrm{MV}_Q}{\mathrm{MV}_P}\frac{\Delta\,\mathrm{MV}_Q}{\mathrm{MV}_Q} \right] \\
&= \mathrm{MV}_P \left[\frac{\Delta P}{P} + \lambda \frac{\Delta Q}{Q} \right] \\
&= \mathrm{MV}_P \left[\beta \frac{\Delta Q}{Q} + \Delta\epsilon + \lambda \frac{\Delta Q}{Q} \right] \\
&= \mathrm{MV}_P \left[(\beta + \lambda)\frac{\Delta Q}{Q} + \Delta\epsilon \right]
\end{aligned}$$

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

$$\Delta V = \mathrm{MV}_P\, \Delta\epsilon$$

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. $\mathrm{MV}_P > 0$, then we want $\Delta\epsilon$ to also be positive. Likewise, if we are shorting P, then $\mathrm{MV}_P < 0$ and so we want $\Delta\epsilon < 0$.

4.2 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

$$\Delta\epsilon_i = \alpha\epsilon_{i-1} + \xi_i \tag{18}$$

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

$$\Delta\epsilon_i = \xi_i + \sum_{j=1}^{p} \alpha_j \epsilon_{i-j} \tag{19}$$

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

5 Portfolios

Definition 5.1 (Portfolio)

A **portfolio** is a collection of assets $X^{(1)}, X^{(2)}, \ldots, X^{(n)}$ with weights $\omega_1, \ldots, \omega_n$ s.t. $\sum \omega_i = 1$.[a]
1. The returns $\mathbf{R} = (R^{(1)}, \ldots, R^{(n)})^T$ of each component may be correlated, and the **portfolio return** is the random variable.
$$R = \boldsymbol{\omega}\mathbf{R} = \sum w_i R^{(i)} \tag{20}$$

2. The covariance matrix

$$\boldsymbol{\Sigma} = \mathrm{Cov}(\mathbf{R}) := \begin{pmatrix} \mathrm{Var}(R_1) & \ldots & \mathrm{Cov}(R_1, R_K) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}(R_K, R_1) & \ldots & \mathrm{Var}(R_K) \end{pmatrix} \tag{21}$$

can be used to calculate the **portfolio variance** and **portfolio volatility**.

$$\mathrm{Var}(R) = \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \implies \sigma_{\mathrm{port}} = \sqrt{\mathrm{Var}[R]} \tag{22}$$

Definition 5.2 (Alpha)

The **alpha** of a portfolio with return R compared to some market return R_m is defined as

$$\alpha := \mathbb{E}[R_p - R_m] = \mathbb{E}[R_p] - \mathbb{E}[R_m] \tag{23}$$

Definition 5.3 (Beta)

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

$$\beta := \rho_{pm} \frac{\sigma_p}{\sigma_m} = \frac{\mathrm{Cov}(R_p, R_m)}{\mathrm{Var}(R_m)} \tag{24}$$

where ρ_{pm} is the correlation between R_m and R_f and σ represents the volatility of returns.
1. If a stock has beta value of 1, then its price activity is strongly correlated with the market (or the benchmark B_j).
2. A $\beta < 1$ means that the security is less volatile than the market.
3. A $\beta > 1$ means more volatile.
4. A negative β indices that the stock is inversely correlated with the market.

---

[a]Since we can short stocks, the elements can be negative, but we should restrict this somehow since we don't want to allow unlimited shorting, which would require infinite margin.

5.1 Markowitz Portfolio Theory - Mean Variance Portfolio

Consider n securities $X_1, \ldots, X_n$, with returns R_1, R_2 iid. s.t. $\mathbb{E}[R_i] = \mu$ and $\mathrm{Var}[R_i] = \sigma^2$. Now within a one period market, consider the two portfolios, which have the same expected return, but have different variance.

1. Portfolio A, with 100% invested in stock 1:

$$R_A = R_1 \implies \mathrm{Var}(R_A) = w_1 \mathrm{Var}(R_1) = \sigma^2 \tag{25}$$

2. Portfolio B, an equi-weighted portfolio:

$$R_B = \sum \frac{1}{n} R_i \implies \mathrm{Var}(R_B) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} R_n\right) = \frac{1}{n^2}\sum_{i=1}^{i}\mathrm{Var}(R_i) = \sigma^2/n \tag{26}$$

The volatility of B is much better, where the Sharpe goes up by a factor of $\sqrt{n}$ as we increase our number of independent stocks. The key word is independence, which may not be a realistic assumption. Regardless, this reduces down to an optimization problem over some constrained space $\Omega \subset \mathbb{R}^n$, where we must find the optimal weights $\mathbf{w} = (w_1, \ldots, w_n)$ that maximizes expected return for some amount of risk, or equivalently gives us the lowest volatility for some around of expected return. This solution is called the **risk-return efficient portfolio**.

$$\arg\min_{\mathbf{w}\in\Omega} \mathrm{Var}(R_{\mathbf{w}}) = \arg\min_{\mathbf{w}\in\Omega} \frac{1}{2}\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \tag{27}$$

We look at variants of this optimization problem that basically includes different constraints.

**Example 5.1 (Efficient Frontier without Risk-Free Asset)**
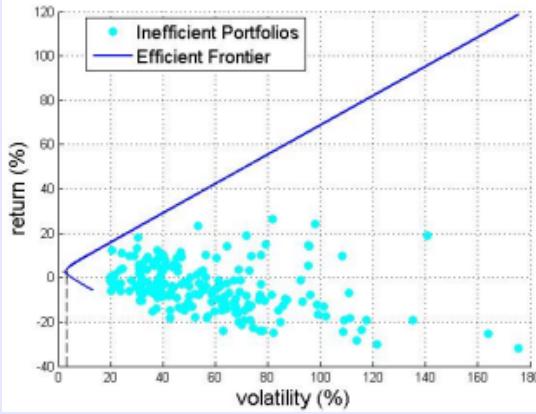
Let us put the constraints that the expected returns must be p and that we must invest all of our money. This leads to

$$\Omega = \{\mathbf{w} \in \mathbb{R}^n \,|\, \mathbf{w}^T \boldsymbol{\mu} = p, \mathbf{w}^T \mathbf{1} = 1\} \tag{28}$$

The solution to this optimization problem is called the **Markowitz efficient frontier**, which traces out a hyperbola of the form

$$\sigma_R^2 = Ap^2 + Bp + C \tag{29}$$

if we get rid of the $\boldsymbol{\omega}$ 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 $(\mathbb{E}[R], \mathrm{Var}(R))$. We can see that none of the points ever cross the frontier.

Example 5.2 (All Long Positions)

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

$$0 \leq w_k \leq 1 \text{ for } k = 1, \ldots, K \tag{30}$$
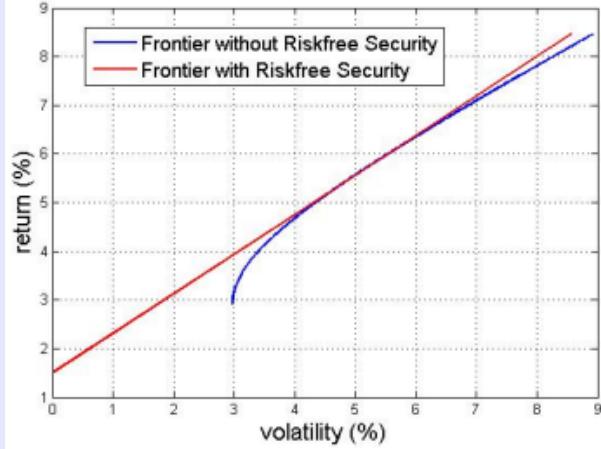
which would need to be solved numerically, possibly using nonlinear models.

Example 5.3 (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 $\{P_1, \ldots, P_K\}$, 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 the same objective subject to constraint equations

$$R_f \left(1 - \sum_{k=1}^{K} w_k \right) + \sum_{k=0}^{K} w_k R_k = R \tag{31}$$

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.

5.2   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

$$(R_p - r_f) \times (R_m - r_f) \tag{32}$$

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 5.1 (CAPM)

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

$$\overline{R} = r_f + \beta(\overline{R}_m - r_f) \tag{33}$$
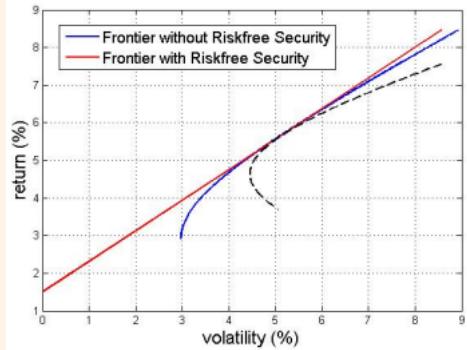
where r_f is the risk-free rate.

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

$$\mathbb{E}[R_\alpha] = \alpha\overline{R} + (1 - \alpha)\overline{R}_m$$
$$\mathrm{Var}(R_\alpha) = \alpha^2\mathrm{Var}(R) + (1 - \alpha)^2\mathrm{Var}(R_m) + 2\alpha(1 - \alpha)\mathrm{Cov}(R, R_m)$$

Note that as α varies, the mean and standard deviation $(\mathbb{E}[R_\alpha], \mathrm{Var}(R_\alpha))$ 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 $\sigma(R_\alpha) = \sqrt{\mathrm{Var}(R_\alpha)}$) is

$$\left.\frac{d\mathbb{E}[R_\alpha]}{d\sigma(R_\alpha)}\right|_{\alpha=0} = \left.\frac{d\mathbb{E}[R_\alpha]}{d\alpha} \middle/ \frac{d\sigma(R_\alpha)}{d\alpha}\right|_{\alpha=0}$$
$$= \left.\frac{\sigma(R_\alpha)(\overline{R} - \overline{R}_m)}{\alpha\sigma(R) - (1 - \alpha)\mathrm{Var}(R_m) + (1 - 2\alpha)\mathrm{Cov}(R, R_m)}\right|_{\alpha=0}$$
$$= \frac{\sigma(R_m)(\overline{R} - \overline{R}_m)}{-\mathrm{Var}(R_m) + \mathrm{Cov}(R, R_m)}$$

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

$$\frac{\sigma(R_m)(\overline{R} - \overline{R}_m)}{-\mathrm{Var}(R_m) + \mathrm{Cov}(R, R_m)} = \frac{\overline{R}_m - r_f}{\sigma(R_m)} \tag{34}$$

gives the result.

5.3 Efficient Market Hypothesis and No-Arbitrage Principle

> **Definition 5.4 (Arbitrage Portfolio)**
>
> Let us have some portfolio, which a value that is a stochastic process $V(t)$. It is considered an **arbitrage portfolio** if its current value $V(t) \leq 0$ and there exists some $T > t$ s.t.
>
> $$\mathbb{P}[V(T) \geq 0] = 1 \text{ and } \mathbb{P}[V(T) > 0] > 0 \tag{35}$$

> **Theorem 5.2 (Axiom of No Arbitrage)**
>
> There exists no arbitrage portfolios, i.e. there is no free lunch.

This is indeed what one would consider an arbitrage portfolio in the colloquial sense. If you had positive capital to invest, $V(t) > 0$, then you can invest it in risk-free Treasury bonds, and you have a guarantee increase in value. This is why you need to start with no or negative capital. If you have 0 cash, you can borrow at a risk free rate r, but if you use that to invest in treasury bonds, you are also getting returns at rate r, cancelling out the effects.

The first probability $\mathbb{P}[V(T) \geq 0] = 1$ is needed since it must be a guaranteed profit. If you have some cash on you, whether it'd be through a loan or you start off with some, you could take a coin flip bet to double your cash or lose all of it. If you win and indeed double it, this does not say anything about the theorem since the probability of this increase was not 1.

> **Theorem 5.3 (Monotonicity Theorem)**
>
> Assume no arbitrage. If we have portfolios A and B s.t. for some point in time T
>
> $$\mathbb{P}[V^A(T) \geq V^B(T)] = 1 \tag{36}$$
>
> then $V^A(t) \geq V^B(t)$ for all t. In addition, if $\mathbb{P}[V^A(T) > V^B(T)] > 0$, then the inequality is strict, i.e. $V^A(t) > V^B(t)$.

> **Corollary 5.1 (Equality)**
>
> If $\mathbb{P}[V^A(T) = V^B(T)] = 1$, then $V^A(t) = V^B(t)$ for all t.

5.4 Replication Principle and Monotonicity Theorem

> **Theorem 5.4 (Replication Principle)**
>
> If we have two portfolios that have equal value at one point in time. Then they should have equal value at every other point in time.

This leads to another

6 Derivatives

> **Definition 6.1 (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
> $$Y_t = g(X_t) \tag{37}$$
> 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 6.1 (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.

> **Example 6.2 (Weather Derivatives)**
>
> A weather derivative can have payout of $g(S_T)$, where
> $$g(S_T) = \mathbb{1}\{S_T > 50\} = \begin{cases} \$1 & \text{if } S_T > 50 \\ 0 & \text{if else} \end{cases} \tag{38}$$
> 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 $T = 1$.

The three fundamental derivative contracts we will work with are forwards/futures, swaps, and options. Let's provide a brief summary of them. Forwards are the simplest type of contract and are the easiest to value. Additionally, through them we can define the forward price of any financial asset, which is essential for more complex valuation models. Futures are exchange-traded, futures-settled versions of forwards. Swaps are contracts where two parties agree to exchange cash flows based on some underlying asset. Options are contracts that give the holder the right, but not the obligation, to buy or sell an asset at a fixed price and are the most complex. They have many interesting properties.

The most common underlying assets for derivatives are:

1. stocks
2. bonds
3. other derivatives
4. commodities
5. currencies
6. interest rates
7. market indexes

They can be traded either OTC (between 2 companies privately) or on an exchange.[5]

---

[5]OTC derivatives constitute a greater proportion of the derivatives market, but exchange-traded derivatives have many

7   Forward Contracts and Forward Prices

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

> **Definition 7.1 (Forward Contracts)**
>
> **Forward contracts** are contracts between two parties to buy or sell an asset at a future date for a price agreed upon today. For notation, let's write
> 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. $V_K(t, T)$ be the **true value** of the contract at current time $t \leq T$ of being long a forward contract with delivery price K and maturity T.
> 5. The **price** is a noisy estimate of V. Let P_t be the price of a long forward.

Note that $V_K(t, T)$ is what everyone wants to find out but is extremely difficult. However, there are some properties that it must satisfy.

> **Lemma 7.1 (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
> $$V_K(T, T) = S_T - K \tag{39}$$
> This is referred to as the **value at maturity** or the **payout** of the contract.

> **Example 7.1 (Red Sox)**
>
> If I buy the Boston Red Sox regular season wins at 100, in size one dollar per won, I have a long forward position with delivery price of 100. If the Red Sox wins 110 games I make \$10, and if they have 91 wins I lose \$9. The payout is $S_T - K$ where S_T is the number of wins.
>
>
>
> Figure 4: Payoff (not profit-loss) graphs for the buyer and sellers of the Red Sox contract.
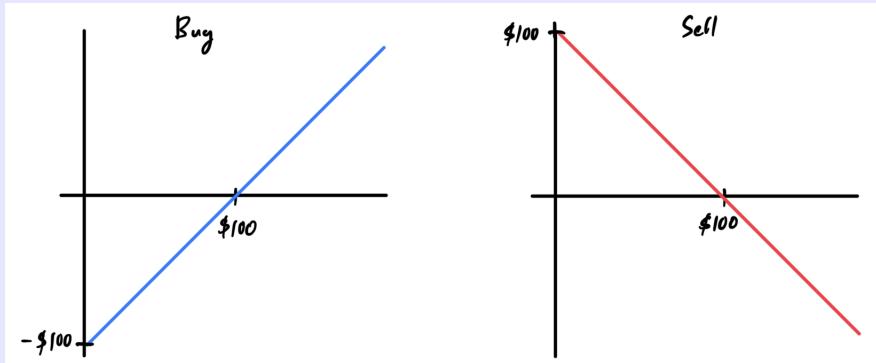
7.1   Contract Specifications

Forward contracts are in general traded OTC, with two parties agreeing to the terms of the contract. This also means that there isn't a clearinghouse to guarantee the contract, so there is a risk of default. Post-2008, there has been a push to move towards exchange-traded contracts, which are standardized and have

---

advantages, such as standardization, liquidity, and elimination of default risk. Originally, derivatives were used to ensure balanced exchange rates for goods traded internationally. With the differing values of national currencies, international traders needed a system to account for differences.

a clearinghouse to guarantee the contract. 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.

7.1.1 Settlement

Second, forwards go through **stock-type settlement** meaning that no cash flows happen until the delivery date. These can be physically settled or cash settled also.

> **Definition 7.2 (Physical Settlement)**
>
> If a contract is **physically settled**, it means that the buyer actually pays K and receives the asset at time T.

> **Definition 7.3 (Cash Settlement)**
>
> However, some forwards are **cash settled**, meaning one simply receives (pays if negative) the amount $S_T - K$ at time T.[a]

The profit/loss is the same, but there are some key differences. 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.

> **Definition 7.4 (Stock-Type Settlement)**
>
> **Stock-type settlement** means that no cash flows happen until the delivery date.

> **Exercise 7.1 (Hull 1.5)**
>
> An investor enters into a short forward contract to sell $100,000$ British Pounds for US dollars at an excahnge 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?[a]

> **Solution 7.1**
>
> 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.

7.2 Forward Price

When A and B enter into a forward contract, it should be the case that the value of this contract should be 0 when both parties agree to it. To see why, consider the following equivalent scenarios between parties A and B, with interest rates being 10%.[6]

---

[a]For example, if you look at forwards on the SP500 index, you must cash settle since you can't deliver fractional shares.
[a]A **pip** is 1/100 of a cent.
[6]Thanks to this Quora post for clarifying.

1. **A** must buy some product at \$400 from **B** in 1 year, which has a value of 0 for both parties.

2. **A** must buy some product at \$300 from **B** in 1 year *and* **A** must give **B** \$90.91 now (present value of \$100 in 1 year). The forward itself has positive value for **A** and negative value for **B**.

3. **A** must buy some product at \$500 from **B** in 1 year *and* **B** must give **A** \$90.91 now. The forward itself has positive value for **B** and negative value for **A**.

Note that the second and third points are the same as initiating a 0-value forward contract plus a separate loan. In scenario 2, **A** is loaning **B** \$90.91, which means that **A** is expected to receive \$100 and pay \$400 for the product, which is equivalent to paying \$300 for the product. This goes for scenario 3, where **B** now loans **A** \$90.01.

While this fact was obvious from the start, this scenario really emphasizes that the delivery price, along with the time period and interest rates, is a huge determinant in the value of the forward. A lower delivery price means positive value to the buyer (and negative to seller), and a higher delivery indicates positive value to the seller (and negative to buyer). [7]The specific price where this value is 0 is called the *forward price*.

> **Definition 7.5 (Forward Price)**
>
> The **forward price** $F(t,T)$ at the current time $t \leq T$ is the delivery price K such that
>
> $$V_K(t,T) = 0 \implies V_{F(t,T)}(t,T) = 0 \tag{40}$$
>
> That is, it is the price that makes the contract have zero value at the current time.

So we can think of the forward price as the fair price of an asset in the future, which is determined by the current spot price and the interest rate.

> **Lemma 7.2 (Forward Price at Expiry)**
>
> 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.

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. So if you find that the strike is greater than the forward price, then the person who longs that contract would have to buy at a high price, meaning that the value is negative. Otherwise it is positive.

> **Example 7.2 (Constant Stock)**
>
> 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.

To try valuing forward contracts through some formula, we should have an analogous formula for forward prices. Essentially, the fair value of our forward price should be captured in this formula.

$$\text{forward price } = \text{current spot price of underlying} + \text{benefits of buying later}$$
$$- \text{costs of buying later}$$

---

[7]Depending on the buyer and seller's estimates of the interest rates, these may differ, but for now we assume one true rate.

Let's think about the benefits of buying later, which is equivalent to the costs of buying now.

1. *Interest on Cash.* You can get interest on the cash you have now.

2. *Inventory Costs.* There are inventory costs that you would incur from storing products, e.g. commodities. There are also transportation and insurance costs.

The benefits of buying now, or the costs of buying later, is as such:

1. *Convenience Yield.* The buyer may urgently need the product now to capitalize on some immediate business opportunity.

2. *Fixed Income Payments.* By not owning a dividend-paying stock or bond now, you will miss out on fixed-income payments like dividends or coupon payments.

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.

Now we want to put this formula in more mathematical terms. Let's do this for underlying assets that gradually increase in complexity.

Theorem 7.1 (Forward Price of No Income Asset)

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

$$F(t,T) = S_t \cdot Z_r(t,T)^{-1} \tag{41}$$

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.*

$$S_t = V_K(t,T) + Ke^{-r(T-t)} \implies V_K(t,T) = S_t - Ke^{-r(T-t)} \tag{42}$$

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

$$0 = S_t - Ke^{-r(T-t)} \implies K = S_t e^{r(T-t)} \tag{43}$$

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 $F(t,T) < S_t e^{r(T-t)}$. 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 $S_t e^{r(T-t)}$ 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 $S_t e^{r(T-t)} - F(t,T) > 0$ by assumption, and you have an arbitrage opportunity.
2. Assume that $F(t,T) > S_t e^{r(T-t)}$. 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 $S_t e^{r(T-t)}$. Therefore, this leaves me with a net gain of $F(t, T) - S_t e^{r(T-t)} > 0$ by assumption, and I have an arbitrage opportunity.
> Therefore, since I have an arbitrage opportunity, it must be the case that $F(t, T) = S_t e^{r(T-t)}$.

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).

Exercise 7.2 (Blyth Chapter 2 Exercise 2)

At time t you own one stock that pays no dividends, and observe that

$$F(t, T) < \frac{S_t}{Z(t, T)} \tag{44}$$

What arbitrage opportunity is available to you, assuming you can only trade the stock, ZCB, and the forward contract?

Solution 7.2

Note that if we rearrange the equation, it becomes

$$F(t, T) \, Z(t, T) < S_t \tag{45}$$

This should really be an equality, so it tells us that at least the LHS is undervalued, or the RHS is overvalued. So we want to take long positions on the forward price and the ZCB, while we take a short position on S_t.
1. We can first short-sell the stock at S_t and buy a forward contract for no cost.[a]
2. We now have S_t in cash, which we loan it out at the ZCB rate.[b]
3. At time T, you have $S_t/Z(t, T)$ in cash. You use this to buy back the stock at the delivery price $F(t, T)$ (not the spot price S_T) from your forward contract, and you return this stock back to the lender who you shorted-sold. Therefore, you are left with

$$\frac{S_t}{Z(t, T)} - F(t, T) > 0 \tag{46}$$

in cash.

Exercise 7.3 (Blyth Chapter 2 Exercise 1)

Let S_t be the current price of a stock that pays no dividends.
1. Let r_{bid} be the interest rate at which one can invest/lend money, and let r_{off} be the interest rate at which one can borrow money, $r_{\text{bid}} \leq r_{\text{off}}$. Both rates are continuously compounded. Using arbitrage arguments, find the upper and lower bounds for the forward price of the stock for a forward contract with maturity $T > t$.
2. How does your answer change if the stock itself has bid price $S_{t,\text{bid}}$ and offer price $S_{t,\text{off}}$?

---

[a]If you're betting that the forward price will go up, then this means that it is advantageous to buy the contract so you can lock in a lower forward price to buy the underlying later.
[b]If we are betting that the ZCB value, i.e. the present value of \$1, will go up, then this means that interest rates will decrease. If interest rates decrease, you would like to loan *now* rather than later to make a profit.

Solution 7.3

Let's go through two cases. Assume you start with nothing.
1. You can short-sell a stock at S_t, loan out the S_t in cash at rate r_{off}, and buy a forward contract at some forward price $F(t,T)$. At time T, you receive $S_t e^{r_{\text{off}}(T-t)}$ and can buy the stock at some forward price to return. In this case, we would have arbitrage if the money I am left with

$$S_t e^{r_{\text{off}}(T-t)} - F(t,T) > 0 \tag{47}$$

so this can't be the case.
2. You can borrow S_t in cash at the rate r_{bid}, immediately buy a stock at S_t, and sell a forward at some forward price $F(t,T)$. At time T, I can sell the stock at price $F(t,T)$ and then use this money to pay back my loan. We would have arbitrage if the money I am left with

$$F(t,T) - S_t e^{r_{\text{bid}}(T-t)} > 0 \tag{48}$$

Neither of these can occur by the no arbitrage principle, so we have the bounds

$$S_t e^{r_{\text{bid}}(T-t)} < F(t,T) < S_t e^{r_{\text{off}}(T-t)} \tag{49}$$

If the stock had a bid-ask price, then we can slightly adjust our formula. In the first scenario, we short sell a stock at the bid price, and in the second, we buy a stock at the ask price. Therefore, we have

$$S_{t,\text{ask}} e^{r_{\text{bid}}(T-t)} < F(t,T) < S_{t,\text{bid}} e^{r_{\text{off}}(T-t)} \tag{50}$$

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 7.2 (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,

$$F(t,T) = (S_t - I)e^{r(T-t)} \tag{51}$$

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.

$$S_T + Ie^{r(T-t)} = V_K(T,T) + K + Ie^{r(T-t)} \implies S_t = V_K(t,T) + Ke^{-r(T-t)} + I \tag{52}$$

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

$$S_t = Ke^{-r(T-t)} + I \implies F(t,T) = K = (S_t - I)e^{r(T-t)} \tag{53}$$

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

$$S_t e^{r(T-t)} - F(t,T) - Ie^{r(T-t)} > 0 \tag{54}$$

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 $S_t e^{r(T-t)}$, and execute the contract to sell it at $F(t,T)$. Our net profit is

$$F(t,T) + Ie^{r(T-t)} - S_t e^{r(T-t)} > 0 \tag{55}$$

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 7.3 (Forward Price of Compounded Income Asset)

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

$$F(t,T) = S_t e^{(r-q)(T-t)} \tag{56}$$

Theorem 7.4 (Forward Price of Forex)

The forward price of one unit of foreign currency is given by

$$F(t,T) = X_t e^{(r_\$ - r_f)(T-t)} \tag{57}$$

where X_t is the price at time t of one foreign currency and T is the maturity of the forward contract.

Proof.

We use replication. Consider the two portfolios.

1. You start off with X_t dollars, which you immediately convert to 1 foreign currency and put it in interest to get $e^{r_f(T-t)}$ foreign currency at time T.
2. You start off with X_t dollars. You immediately buy a forex forward at $F(t,T)$ for no cost and put the dollars in interest. At time T, we have $X_t e^{r_\$(T-t)}$ in cash. We execute the forward and convert our dollars to the foreign currency, leaving us with $\frac{X_t}{F(t,T)} e^{r_\$(T-t)}$ foreign currency.

By replication these two portfolios must equal in value, and we have

$$\frac{X_t}{F(t,T)} e^{r_\$(T-t)} = e^{r_f(T-t)} \implies F(t,T) = X_t e^{(r_\$ - r_f)(T-t)} \tag{58}$$

At this point, we can take any financial security and interpret the forward price of it as the fair price of it at some predetermined point in the future.

Example 7.3 (Forward Price of Commodities)

The forward price of a commodity is

$$F(t,T) = S_t(1 + r(T-t)) + s(T-t) + i(T-t) \tag{59}$$

where s, i are the annual storage, insurance costs per commodity unit.[a] If these values are known, you can use them to predict other variables like the potential convenience yield.

Example 7.4 (Forward Price of Dividend-Paying Stock)

Let
1. S_t be the stock price
2. r be the interest rate over the life of the forward contract
3. d_i is each dividend payment expected prior to maturity of the forward contract
4. t_i is the time remaining to maturity after each dividend payment.
5. r_i the applicable interest rate (i.e. *forward rate*) from each dividend payment to maturity of the forward contract.

Then, the forward price of a dividend-paying stock can be written as the cash plus all the dividend payments, plus their interest.

$$F(t,T) = S(1 + r(T-t)) - \sum_i d_i(1 + r_i t_i) \tag{60}$$

Example 7.5 (Forward Price of Bond)

Let
1. r be the interest rate over the life of the forward contract
2. c_i be each coupon expected prior to the maturity of the forward contract
3. t_i be the time remaining to maturity after each coupon payment
4. r_i be the applicable interest rate from each coupon payment to maturity of the forward contract

Then, the forward price of a bond can be written as

$$F(t,T) = B(1 + r(T-t)) - \sum_i c_i(1 + r_i t_i) \tag{61}$$

Exercise 7.4 (Blyth Chapter 2 Exercise 4)

FX forwards are among the most liquid derivative contracts in the world and often reveal more about the health of money markets (markets for borrowing or lending cash) than published short-term interest rates themselves.

1. On 3 October 2008, the euro dollar FX rate was trading at €1=$1.3772, and the forward price for a 3 April 2009 forward contract was $1.3891. Assuming six-month euro interest rates were 5.415%, what is the implied six-month dollar rate? Both interest rates are quoted with act/360 daycount and semi-annual compounding. There are 182 days between 3 October 2008 and 3 April 2009.
2. Published six-month dollar rates were actually 4.13125%. What arbitrage opportunity existed? What does a potential arbitrageur need to transact to exploit this opportunity?
3. During the financial crisis, several European commercial banks badly needed to borrow dollar cash, but their only source of funds was euro cash from the European Central Bank (ECB). These banks would: borrow euro cash for six months from the ECB; sell euros/buy dollars

---
[a]This assumes no compound interest (?)

in the spot FX market; and sell dollars/buy euros six months forward (to neutralize the FX risk on their euro liability). Explain briefly how these actions may have created the arbitrage opportunity in (2), which existed for several months in late 2008.

Solution 7.4

We will ignore the daycount conventions. We can substitute these values in the forward prices for forex to get

$$1.3891 = 1.3772 \cdot e^{(r_\$ - r_f) \cdot 0.5} \implies r_\$ = 0.07138 \tag{62}$$

Therefore, the general idea is that we want to long domestic interest rates and short foreign interest rates, which means that we want to buy domestic bonds and sell foreign bonds.

7.3 Valuation

Theorem 7.5 (Valuation of Forward Contracts)

The value of a forward contract on an asset satisfies

$$V_K(t, T) = \big(F(t, T) - K\big) Z_r(t, T) \tag{63}$$

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

Proof.

No Arbitrage. We assume the two scenarios.

Note that for fixed t_0, T with varying $t_0 \leq t \leq T$,

$$V_{F(t_0, T)}(t, T) = 0 \text{ for } t = t_0 \tag{64}$$

and may not be 0 as t changes. However,

$$V_{F(t, T)}(t, T) \tag{65}$$

is constantly 0.

7.4 Forward Rates

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 7.6 (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 f_{12} (agreed at t) at which one can borrow or lend money from T_1 to T_2.
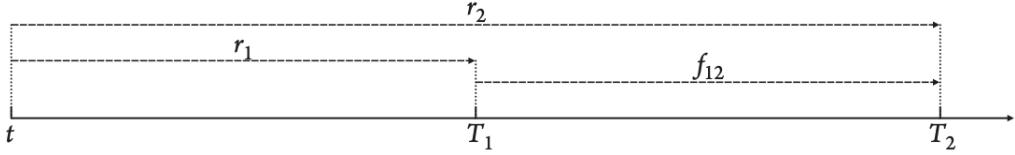
Figure 5: Common visual of forward rate, from Blyth.

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

Theorem 7.6 (Forward Rate)

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

$$Z_{r_1}(t, T_1) Z_{f_{12}}(T_1, T_2) = Z_{r_2}(t, T_2) \tag{66}$$

where r_k is the zero interest rate to borrow from t to T_k.

Proof.

This is trivial using the replication argument.

This is great theoretically, but how do we actually trade something like this? What if we wanted to bet that this forward rate was going to be less than some percentage? We can do this by trading forward contracts with the underlying being ZCBs! As stated in the theorem below, the forward price of zero coupon bonds is precisely the same as the value of a zero coupon bond with the forward rate, which we can think of as locking in a future interest rate from T_1 to T_2.

Theorem 7.7 (Forward Price equals Value of ZCB at Forward Rate)

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

$$F(t, T_1, T_2) = Z_{f_{12}}(T_1, T_2) = \frac{Z_{r_1}(t, T_2)}{Z_{r_2}(t, T_1)} \tag{67}$$

Therefore, if we model the interest as simply compounded, we have

$$F(t, T_1, T_2) = (1 + f_{12})^{-(T_2 - T_1)} \tag{68}$$

and continuously compounded gives

$$F(t, T_1, T_2) = e^{-f_{12}(T_2 - T_1)} \tag{69}$$

Proof.

We can directly calculate the forward price of the underlying ZCB. It's a one-time payment of 1 at time T_2, so the only cost to consider here is the time value of money. Therefore, $F(t, T_1, T_2) = Z_{f_{12}}(T_1, T_2)$, and the second equality follows from the forward rate identity above.[a]

---

[a]Note that this is not trivial as reducing them to exponentials! This was just proved in the previous theorem.

Example 7.6 (Forward Libor Rate)

Exercise 7.5 (Blyth 3.1)

Suppose $t \leq T_1 \leq T_2 \leq T_3$, where t is current time, and $\varepsilon > 0$. Recall that $Z(T_1, T_2)$ is the price at time T_1 of a ZCB with maturity T_2, and $F(T_1, T_2, T_3)$ is the forward price at time T_1 for a forward contract with maturity T_2 on a ZCB with maturity T_3.
1. For each of the pairs of A and B in Table 3.1, choose the most appropriate relation-ship out of $\geq, \leq, =$ and ?, where ? means the answer is indeterminate. Give brief reasoning.
2. What can you say about interest rates between T1 and T2 if
    (a) $Z(t, T1) = Z(t, T2)$?
    (b) $Z(t, T1) > 0$ and $Z(t, T2) = 0$?

Solution 7.5

For (1), the answers are listed.
1. $Z(t, T_1) \leq 1$. Assuming positive interest rates, you discount 1 to present value.
2. $Z(T_1, T_1) = 1$. You receive the money now so it's the same as 1.
3. $Z(t, T_2) \geq Z(t, T_3)$. You lock up your money for longer until T_3, so it should be worth less.
4. $Z(T_1, T_2) \geq Z(T_1, T_3)$. Same as above.
5. $Z(T_1, T_3) \leq Z(T_2, T_3)$. You lock up your money for longer at T_1, so it's worth less.
6. $Z(T_1, T_1 + \Delta) ? Z(T_2, T_2 + \Delta)$. Indeterminate since you don't know what prevailing interest rates are.
7. $F(t, T_1, T_2) \geq F(t, T_1, T_3)$. Reduce them to the ZCB values from T_1 to T_2 and T_1 to T_3. You are now comparing $Z(t, T_2)$ and $Z(t, T_3)$, which from above, $Z(t, T_3)$ is worth less.
8. $F(t, T_1, T_3) \leq F(t, T_2, T_3)$. Same as above, but now you are comparing $Z(t, T_1)^{-1}$ and $Z(t, T_2)^{-1}$.
9. $\lim_{T \to \infty} Z(t, T) = 0$. If you are lending money for an infinite time, it's as if you don't have any money. Interest rates r I suppose would rise though (?).

For (2),
1. Either $T_1 = T_2$ or $f_{12} = 0$.
2. $f_{12} < 0$.

Exercise 7.6 (Blyth 3.2)

Forward Rates.
1. The one-year and two-year zero rates are 1% and 2% respectively. What is the one-year forward rate (that is, f_{11})? Assume all rates are annually compounded.
2. If the two-year forward one-year rate (f_{21}) is 3%, what is the three-year zero rate?

8 Swaps

Definition 8.1 (Swap)

Swaps are contracts through which two parties exchange cash flows or liabilities from two different financial instruments. One cash flow is generally fixed, while the other is variable and is based on a benchmark interest rate, floating currency exchange rate, or index price.
1. T_0 is the start date, T_n the maturity date, and T_i the payment dates for $i = 1, 2, \ldots, n$. In vanilla swaps we will have $T_i + \alpha = T_{i+1}$ for fixed α.

2. The **fixed leg** of the swap consists of payments αK at each payment date for some fixed **delivery price/rate** K.
3. The **floating leg** of the swap consists of payments αP_i for each payment date, where P_i may be dependent on the variable benchmarks.

8.1 Contract Specification

The most common examples of swaps can be about:

1. interest rates

2. mortgage bonds

3. exchange rates

Swaps don't trade on exchanges, and retail investors do not generally engage in swaps. Rather, swaps are over the counter contracts primarily between businesses that are customized to the needs of both parties.

Definition 8.2 (Interest Rate Swaps)

In **interest rate swaps**, we have
$$P_i = L_{T_{i-1}}[T_{i-1}, T_i] \tag{70}$$
where the variable payments are dependent on the Libor rate fixing at T_{i-1} for the period T_{i-1} to $T_i = T_{i-1} + \alpha$, paid at T_i.
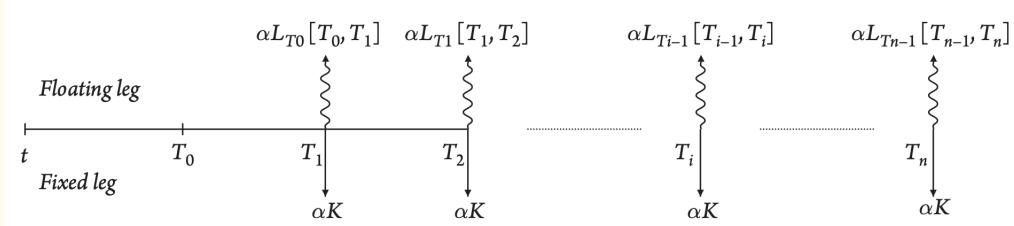
Figure 6: Visual of an interest rate swap.

Definition 8.3 (Forex Swaps)

8.2 Forward Swap Rate and Swap Value

8.3 Swaps as a Difference Between Bonds

9 Future Contracts

Definition 9.1 (Futures)

Like forward contracts, **future contracts** are contracts between two parties to buy or sell and asset at a future ate for a price agreed upon today. For notation, let's write
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. $V_K(t, T)$ be the **true value** of the contract at current time $t \leq T$ of being long a future contract

with delivery price K and maturity T.

5. The **price** is a noisy estimate of V. Let P_t be the price of a long future.

The difference between futures and forwards is that first, futures are exchange traded and thus have standards to follow, and second, futures are settled daily. Future contracts are forwards contracts, but these extra factors slightly change the properties of the contract.

9.1 Contract Specifications

We want to talk about

1. What the **underlying** asset is.

2. The **contract size**. 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 big ask spread.

3. The **(expected) grade**, or quality, of the underlying, if applicable (e.g. commodities). This allows for a range of grades depending on the actual grade of the contract.

4. Delivery date and location, which may be slightly different than when the contract expires.

5. The minimum and 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.

6. The maximum number of contracts that a trader can hold. This is to prevent parties from cornering a market.

9.1.1 Bond Futures

Let's start off with bonds. U.S. Treasury futures and options contracts are available for each of the U.S. Treasury benchmark tenors: 2-year, 5-year, 10-year, and 30-year. Additionally, CME Group offers Ultra 10-Year Note and Ultra T-Bond futures which offer greater precision for trading the 10-year and 30-year maturity points on the yield curve, respectively. Filling out the curve further, 3-Year Note futures (reintroduced in 2020) bring finer hedging granularity to the front of the curve, while 20-Year Bond futures (launched in 2022) provide a precise instrument for hedging 20-Year bond exposure.

Each of the bond and note futures contracts has an associated delivery bond basket that defines the range of bonds by maturity that can be delivered by the seller to the buyer in the delivery month. For example, the 5-year contract delivers into any U.S. government fixed coupon bond that has a remaining maturity of longer than 4 years and 2 months and an original maturity of no more than 5 years and 3 months. The delivery mechanism ensures the integrity of futures prices by ensuring that they are very closely tied to the prices of U.S. government bonds and their yields (interest rates).

When calculating the minimum tick size, they tend to say something like "In percent of par to one-eighth of 1/32nd of 1% of par," which means $0.01 \cdot 1/8 \cdot 1/32$ of the total contract size. Furthermore, the quotes are represented as a percentage of the face value, e.g. 103.23 would mean 103,230 for a \$100,000 contract.

Underlying	Size	Increments	Deliverable Maturities
2-Year T-Note	$200,000	$\frac{1}{8} \cdot \frac{1}{32}$ c ($7.8125)	1 3/4 to 2 years
3-Year T-Note	$200,000	$\frac{1}{8} \cdot \frac{1}{32}$ c ($7.8125)	2 9/12 to 3 years
5-Year T-Note	$100,000	$\frac{1}{4} \cdot \frac{1}{32}$ c ($7.8125)	4 1/6 to 5 1/4 years
10-Year T-Note	$100,000	$\frac{1}{2} \cdot \frac{1}{32}$ c ($15.625)	6 1/2 to 8 years
Ultra 10-Year T-Note	$100,000	$\frac{1}{2} \cdot \frac{1}{32}$ c ($15.625)	9 5/12 to 10 Years
T-Bond	$100,000	$\frac{1}{32}$ c ($31.25)	15 years up to 25 years
20-Year T-Bond	$100,000	$\frac{1}{32}$ c ($31.25)	19 2/12 to 19 11/12 years
Ultra T-Bond	$100,000	$\frac{1}{32}$ c ($31.25)	25 years to 30 years

Table 1: Standards of Bonds Futures Contracts

In practice, most participants trade U.S. Treasury futures contracts with the intent of either closing out the futures position or rolling them into longer expiry futures contracts. U.S. Treasury futures are listed on the March, June, September, and December quarterly cycles. More specifically, the futures expire in March, June, September, and December 24 of each year. This standard is consistent with the U.S. Treasury auctions, which take place in February, May, August, and November.

> **Example 9.1 (Bond Futures Standards)**
>
> Mathematically, this means that if the current time is t, the delivery date is T_1, and the bond maturity date is T_2, then for a futures on a 10-year T-Note,
> 1. T_1 must be in one of the days of March, June, September, December 24.
> 2. $T_2 - T_1 \approx 10$.

For bond futures, there is the concept of a **continuous contract**, which is a price of a certain type of contract that accounts for all the expiration dates.

> **Question 9.1 ()**
>
> Why do the prices of these futures increase as T_1 is further?

9.1.2 Equity Index Futures

Futures on stocks, called **single-stock futures**, are generally not very accessible for retail investors (e.g. OneChicago).[8] Rather we use **stock index futures**, which we list some specs down in decreasing order of popularity.

ADV[9]	Underlying	Size	Increments	Ex. Quote
1,697,303	E-Mini S&P 500	50	0.25 ($12.50)	5253.50
231,570	E-Mini Nasdaq 100	20	0.25 ($5.00)	18675.75
160,145	E-Mini Dow	5	1 ($5.00)	38191
19,487	E-Mini S&P MidCap 400	100	0.1 ($10)	2918.60
12,599	Nikkei 225 (USD)	5	5 ($25)	
7839	S&P 500	250	0.25 ($62.50)	5412.75
	E-Mini Russell 2000	50	0.10 ($5.00)	
	Micro E-Mini S&P 500	5	0.25 ($1.25)	5253.75
	Micro E-Mini Nasdaq 100	2	0.25 ($0.50)	
	Micro E-Mini Dow	0.5	1 ($0.50)	
	Micro E-Mini Russell 2000	5	0.10 ($0.50)	

Table 2: Standards of Equity Index Futures Contracts

---

[8]If you Google search "Apple futures" you don't really get anything.

These contracts, like bonds, trade during the months of March, June, September, and December. Trading terminates at 9:30 a.m. ET on the third Friday of the contract month.

9.1.3 Oil and Natural Gas Futures

The two main fuel commodities traded are crude oil and natural gas. Let's compare the two.

1. Crude oil is used for vehicle fuel while natural gas is for stoves and electric lights. This difference in purpose and thus demand likely is the cause of the prices.

2. Crude oil needs to go under a heavy refining process while natural gas is minimal.

3. Crude oil is liquid at room temperature while natural gas is a gas.

Definition 9.2 (Crude Oil)

Crude oil is a naturally occurring, unrefined petroleum product composed of hydrocarbon deposits and other organic materials. It is primarily used to produce gasoline, diesel fuel, and other essential products, being liquid in room temperature.
Crude oil needs to undergo a refining process to convert into gasoline, and thus there are different grades of oil.
    1. **Light crude oil** is a type of crude oil that has a low density, measured by the American Petroleum Institute (API) gravity, and flows freely at room temperature.[a] **Heavy crude oil** has higher density and is harder to refine.
    2. **Sweet** oil refers to the sulfur content of the oil, with sweet oil having less than 0.5% sulfur. **Sour** oil has more than 0.5% sulfur.
The main premium oils that are traded are light sweet oils, with the two benchmarks being **West Texas Intermediate (WTI)**, which is extracted from fields located in Texas, North Dakota, and Louisiana of the United States, and **North Sea Brent**, which is extracted from the North Sea near Europe; its oil fields include Brent, Ekofisk, Forties, and Oseberg.

The difference between WTI and Brent is that WTI is slightly lighter and sweeter than Brent. Also, the WTI Crude Oil futures delivery point is Cushing, Oklahoma, a landlocked area that's centrally located around the major oil fields but far from where most oil and refined products—such as gasoline and petrochemicals—are ultimately destined, making it more difficult to transport. Brent is extracted at sea, making it easier to transport.

Underlying	Size	Increments	Ex. Quote
WTI Crude Oil (CL)	1000 barrels	0.01 ($10.00)	$86.73
E-Mini Crude Oil	500 barrels	0.01 ($5.00)	$86.73
Micro WTI Crude Oil	100 barrels	0.01 ($1.00)	$86.73
Brent Crude Oil	1000 barrels	0.01 ($10.00)	$78.94
Henry Hub Natural Gas	10,000 MMBtu	0.001 ($10.00)	$2.843
E-Mini Natural Gas	2500 MMBtu	0.005 ($12.50)	$2.845
Micro Henry Hub Natural Gas	1000 MMBtu	0.001 ($1.00)	$2.843

Table 3: Standards of Oil and Natural Gas Futures Contracts

Crude Oil futures trade on every month, going up to the next 10 calendar years.

For oil, natural gas, and copper, the delivery month is the month after the expiration month and delivers on all months.

---

[a]API levels fall between 10 and 70, with higher numbers being lighter. Even light enough to float in water.

9.1.4   Grain Commodity Futures

Let's start with wheat, which has many variants.

1. **Hard red winter wheat** is a type of wheat that is high in protein and is often used for bread.

2. **Soft red winter wheat** is a type of wheat that is low in protein and is often used for pastries.

3. **Hard red spring wheat** is a type of wheat that is high in protein and is often used for bread.

4. **Durum wheat** is a type of wheat that is high in protein and is often used for pasta.

There aren't as many variants of underlying corn in corn futures. They are delivered to Toledo, Ohio.

Finally, soybeans also do not have many variants.

Underlying	Size	Increments	Ex. Quote
Soft Red Winter (SRW)	5000 Bushels	$\frac{1}{4}$ c ($12.50)	$568\frac{1}{4}$ c
Mini Soft Red Winter	1000 Bushels	$\frac{1}{8}$ c ($1.25)	$568\frac{1}{4}$ c
Corn (ZC)	5000 Bushels	$\frac{1}{4}$ c ($12.50)	$400\frac{3}{8}$ c
Mini Corn (XC)	1000 Bushels	$\frac{1}{8}$ c ($1.25)	$400\frac{3}{8}$ c
Soybean	5000 Bushels	$\frac{1}{4}$ c ($12.50)	$1655\frac{1}{2}$ c
Mini Soybean	1000 Bushels	$\frac{1}{8}$ c ($1.25)	$1655\frac{1}{2}$ c

Table 4: Standards of Grain Futures Contracts

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.

The month of contract, which is the month that the commodity will be delivered and is not necessarily the same as the expiration date. It turns out that this discrepancy is important in analyzing the convergence of the future price to the spot price and determining the risk.

> **Example 9.2 (Kansas City vs Chicago Wheat Spread)**
>
> The following article here tells us a good story about the divergence between prices in these two markets.

9.1.5   Livestock Commodity Futures

Feeder cattle are weaned calves just sent to the feedlots (about 6-10 months old), and live cattle are cattle which have attained a desirable weight (1150-1375 pounds for heifers, and 1200-1500 pounds for steers), to be sold to a packer.

It may be a bit odd that poultry isn't mentioned here, but this article explains that during the three attempts to offer chicken futures, traders were not interested enough. Apparently, three things are needed for a futures market to succeed: the volume traded and commodity value must be high to attract sufficient interest, large buyers must be necessary, and price volatility is needed.

Underlying	Size	Increments	Ex. Quote
Live Cattle	40,000 pounds	0.025 c ($10.00)	123.25 c
Lean Hog	40,000 pounds	0.025 c ($10.00)	123.25 c
Feeder Cattle	50,000 pounds	0.025 c ($12.50)	123.25 c

Table 5: Standards of Livestock Futures Contracts

The months of delivery are February, April, June, August, October, and December.

9.1.6   Metals Commodity

These metals are mined and refined, and are used in a variety of industries.

Underlying	Size	Increments	Ex. Quote
COMEX Gold	100 Troy Ounces	0.10 ($10.00)	$2,364.90
Micro Gold	10 Troy Ounces	0.10 ($1.00)	$2,364.90
COMEX Silver	5000 Troy Ounces	0.005 ($25.00)	$31.365
Micro Silver	1000 Troy Ounces	0.005 ($5.00)	$31.350
Platinum	50 Troy Ounces	0.10 ($5.00)	$1,033.6
Micro Platinum	10 Troy Ounces	0.10 ($1.00)	$1,033.6
Copper	25,000 pounds	0.05 c ($12.50)	4.6625 c
E-Mini Copper	12,500 pounds	0.2 c ($25.00)	4.6624 c
Zinc	25 metric tons	0.50 ($12.50)	3,000.00

Table 6: Standards of Metals Futures Contracts

9.1.7   Settlement Standards

For cash-settled contracts, remember[10] that forward contracts have stock-type settlement, meaning that at time $t = 0$, you can enter into a forward contract at no cost and you would get a one-time profit of $S_T - K = F(T,T) - F(0,T)$ upon delivery. However, in future contracts, it is futures-settled, meaning that at the end of every day, the difference in forward prices are settled.

> **Definition 9.3 (Futures-Type Settlement)**
>
> **Futures-type settlement** is the following type of settlement. Given current time $t = 0$ and T, let us label each market day between the two as $t_0 = 0, t_1, t_2, \ldots, t_n = T$. If you buy a futures contract at t_0, then at the end of each day t_i, the difference between the forward price at t_i and the forward price at t_{i-1} is settled. That is, after day t_i $(i = 1, \ldots n)$, you receive (pay if negative)
>
> $$\Phi(t_i, T) - \Phi(t_{i-1}, T) \tag{71}$$

> **Definition 9.4 (Variation)**
>
> The money that is settled daily between the two parties of a futures contract is called the **variation**.

Note that the variation is simply the difference between the forward prices at the end of the day. While it is technically cash that you need to keep into your account, this is different from the margin, which is the amount of money you need to have in your account to trade the contract.

> **Example 9.3 (Futures Type Settlement)**
>
> Say that **A** buys a COMEX Gold futures contract from **B** that expires in say 5 business days. We can see that this is a futures-type settlement because the difference between the forward prices are settled at the end of each day.

---

[10]For first readers, you can skip ahead to read the definition of the futures price before reading ahead.

t	S_t	$\Phi(t,T)$	ΔA	ΔB	A	B
t_0	2332.02	2359.70	0	0	0	0
t_1	2334.20	2356.90	-2.80	+2.80	-2.80	2.80
t_2	2352.51	2379.30	+22.40	-22.40	+19.60	-19.60
t_3	2360.12	2364.10	-15.20	+15.20	+4.40	-4.40
t_4	2362.50	2362.50	-1.60	+1.60	+2.80	-2.80

Table 7: Cash Flows Between parties **A** and **B** upon entry into 4 day futures contract. In total, **A** makes a profit.

Exercise 9.1 (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 9.1

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.

9.2 Futures Price

This means that due to the time value of money, the futures price is not the same as the forward price, and so in this scenario the futures price is considered the "fair price" of the asset in the future. In many cases, these two terms are used interchangeably, but it is good to know the details once.

Definition 9.5 (Futures Price)

The **futures price** $\Phi(t,T)$ at the current time $t \leq T$ is the delivery price that satisfies

$$V_{F(t,T)}(t,T) = 0 \tag{72}$$

That is, it is the price that makes the contract have zero value at the current time.

Definition 9.6 (Cotango, Backwardation Market)

A futures market is said to be in
1. **cotango** if futures prices are above spot prices: $\Phi(t,T) > S_t$.
2. **backwardation** if futures prices are below spot prices: $\Phi(t,T) < S_t$.

The difference between these two prices is known as the **basis gap/risk**, which must converge as $t \to T$. Cotango markets are known to be the default, while backwardation is a bit rarer.
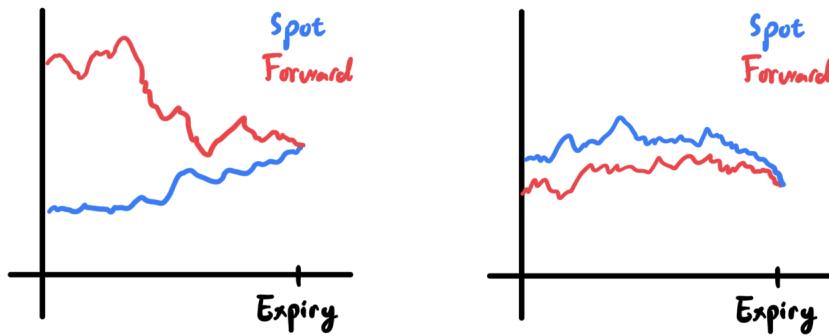
Figure 7: Spot and future prices must converge upon expiry. If this wasn't the case, then we can construct an arbitrage pairs trade at expiry with the underlying and its forward contract.

Example 9.4 (Convergence of Basis Gap)

It actually isn't the case that the delivery is.

We are still left a little unsatisfied with the futures vs forward price. To better compare them, we introduce the theorem.

Theorem 9.1 (Constant Interest Rate)

If interest rates are constant, then
$$\Phi(t, T) = F(t, T) \tag{73}$$

Proof.

Now what happens if r is not constant?

Definition 9.7 (Futures Convexity Correction)

If S_T is positively correlated with the interest rate r (i.e. tends to increase as r increases). If we are long a futures contract, then we receive mark-to-market gains earlier than the forward in environments when interest rates are high, and thus the gains can be invested at a higher rate. Similarly, losses from the long position have to be paid early when rates are low (and thus the future value of the losses is lower). Thus in this case, we would prefer to hold a long futures position relative to a long forward position. Therefore, we have

$$\Phi(t, T) - F(t, T) > 0 \tag{74}$$

which we call the **futures convexity correction**.

What is the forward price of a futures contract? A futures contract is a forward contract, and it itself should have some fair price.

Theorem 9.2 (Forward Price of Futures Contract)

The forward price of a futures contract is the futures price.

> **Exercise 9.2 (Blyth 5.2)**
>
> Futures on bonds and stocks.
> 1. Is the futures price of a fixed rate bond likely to be higher, lower, or the same as its forward price? Give brief reasoning.
> 2. Is the futures price (for a futures contract with maturity $T < T_0$) of a floating rate bond likely to be higher, lower, or the same as its forward price? How does your answer change if $T = T_0$?
> 3. Is the futures price of a stock likely to be higher, lower, or the same as its forward price.

9.3 Futures on Interest Rates

10 Option Contracts

> **Definition 10.1 (Options)**
>
> An option is a contract that gives the owner the right, but not the obligation, to buy or sell a specific asset.
> 1. T is the time until delivery, usually in years. This is fixed per contract and is called the expiration date.
> 2. K is the strike or exercise price.
> 3. S_t is the spot price of the asset at time $t \in [0, T]$.
> Since the buyer of the option gains the *right*, not the obligation, the buyer must give some premium. Therefore, a premium changes hands.

Clearly, this option will have some value, which is estimated by this premium just mentioned, aka the price.

> **Definition 10.2 (Value)**
>
> Let $V_K(t, T)$ be the **true value** of the contract at current time $t \leq T$ of being long an options contract with strike K and maturity T. This will vary over time.
> The **price** is a noisy estimate of $V_K(t, T)$. Let C_t be the price of a call option and P_t be the price of a put option.

> **Example 10.1 (Option Chain)**
>
> An **option chain** is a chart that depicts information regarding a class of options.
> 1. The top has a list of delivery times: Jun 21, Jul 19, Aug 16, Sep 20, etc.
> 2. For a specific delivery time we have a list of strike prices in the middle, with calls on the left and puts on the right.
> 3. For each strike price and option type we have its market, consisting of a bid/ask prices and their respective sizes, along with the total volume.

Figure 8: Option chain for TSLA viewed on May 26, 2024 at Interactive Broker's Trader Workstation (TWS).

To accurately value this, the trader must have a good feel for both the directional movement of the market *and* the timing. A equity trader just needs directional.

10.1   Contract Specifications

10.1.1   Underlying Contracts

Every options contract has a certain underlying security that comes with its own multiplier M.

Example 10.2 (Stocks and ETFs)

Stock options have an underlying of $M = 100$ contracts and the S&P Mini has an underlying of $M = 50$ ETFs.

Example 10.3 (Future Options)

For commmodities, you are really dealing with future options, i.e. options that have an underlying asset of a future contract. The underlying is always 1 future, but conventionally, we say the multiplier in terms of the underlying of the future. Therefore, a future option on wheat is 5000 bushels, since there are 5000 in 1 wheat future.

Example 10.4 (Company Options)

Company ABC's shares trade at $60, and a call writer is looking to sell calls at $65 with a one-month expiration. If the share price stays below $65 and the options expire, the call writer keeps the shares and can collect another premium by writing calls again.

If the share price appreciates to a price above $65, referred to as being in-the-money, the buyer calls the shares from the seller, purchasing them at $65. The call-buyer can also sell the options if purchasing the shares is not the desired outcome.

10.1.2   Settlement

Options can either be **physically settled** (settles into the right but not the obligation to buy/sell the actual underlying) or **cash settled** (cash deposited the difference).

1. A contract settles to a certain price, say 100 (the underlying) at 1:00pm. This is used to determine the daily profit or loss on an option and to calculate margin requirements. However, the underlying is still trading.

2. From this moment to perhaps 5 minutes later at 1:05pm, there is a **settlement window** that we can choose to execute optionality of a trade. At this moment, the underlying can still change from 100 to some other, meaning that it can change between ITM or OTM. Usually, the option will exercise automatically if it is very ITM (there are thresholds for it).

3. Once we decide to execute (or not), the **last executed trade price** of the underlying will determine our true PnL.

> **Example 10.5 (Contrary Exercise)**
>
> Say that you have long 200 180-strike calls and short 100 futures as well to delta hedge.
> 1. If the future settles to 160, it is very OTM (underrealized), and it is automatically not executed.
> 2. If the future settles to 200, and it is very ITM (outrealized), and so it is automatically executed.
> 3. If the future settles to 179.9, it is close to ATM, so during the settlement window, we can choose to execute it or not. Say that during this window the future realizes to 180.5. Then, we gain $0.5 \cdot 200 = 100$ on the calls and we lose $-0.6 * 100 = -60$ on the futures. We can choose to execute this call.
>
> This is sent to the exchange, and you get your underlyings by the end of day, usually at 4:00pm.

> **Example 10.6 (ES Options)**
>
> Let's talk about the most widely traded option, the S&P. Note that SPX options settle to cash while the SPY options settle to shares of the SPY ETF. The **ES (E-Mini S&P)** has different types, each with different settlements.
> 1. **Regulars** settle into the E-Mini S&P 500 Index futures, which are cash settled.
> 2. **Quarterlies** are directly cash settled.

10.1.3 Exercise

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. Generally, stock and future options tend to be American, and options on indexes tend to be European.

Exercise prices are quantized.

> **Example 10.7 (Digital Options)**
>
> Options that have a payout of 1 if a certain event is met and 0 if else are called **digital options**. Therefore, it has payout at T
> $$\begin{cases} 1 & \text{if } S_T \geq K \\ 0 & \text{if } S_T < K \end{cases} \tag{75}$$

Exercise 10.1 (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 10.1

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 $S_T > 26$.
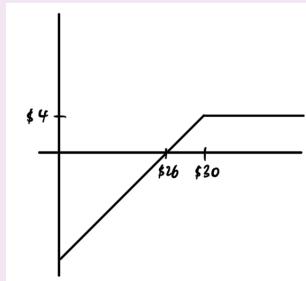
Figure 9: The payoff graph for the put option from the sellers side.

Exercise 10.2 (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 10.2

You can simply draw the profit-loss diagram.
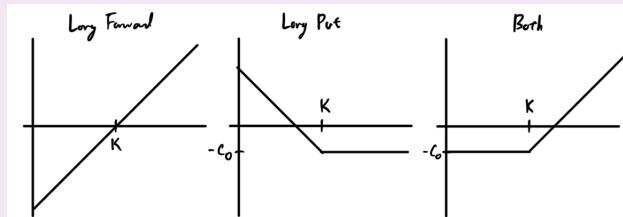
Figure 10: The payoff graph for the long forward and long put option at the same strike price.

Exercise 10.3 (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

$$1000 - \max\left\{0, 1000\left(\frac{169}{S_T} - 1\right)\right\} \tag{76}$$

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 10.3

This seems a bit easy since we can draw the payoff diagram as such.
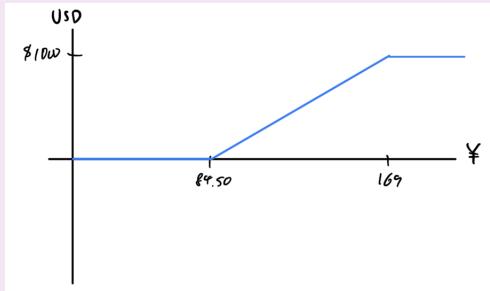
Figure 11

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.
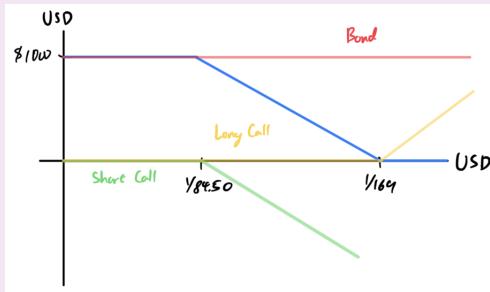
Figure 12

10.2 Intrinsic Value and Time Premium

Let's try to construct some model for this price. The most obvious factor is the difference between the actual price and the strike price, which tells us if our option is currently profitable or not (if we exercised it now).

Definition 10.3 (Intrinsic Value)

The **intrinsic value** is simply what the holder of the option contract would make if they chose to exercise it. That is, for a long call option, it is

$$\text{IV}_K(t,T) = \max\{0, S_t - K\} \tag{77}$$

and for a long put option, it is

$$\text{IV}_K(t,T) = \max\{0, K - S_t\} \tag{78}$$

At some time t, if the value is positive, then the contract is **in-the-money (ITM)** or else it is **out-of-money (OTM)**. If $K = S_t$, then it is **at-the-money (ATM)**.[a] However, since the strike prices are quantized (e.g. $65, $70, $75, etc.), the contract with strike price closest to the underlying spot is known as the ATM.

With this, we can already draw *payoff diagrams*, which show the intrinsic value of the option (i.e. the value of the option *at expiry*).

Example 10.8 (4 Basic Options: Payoff Diagrams at Expiry)

The 4 basic sides of an option contract are listed below.

1. If party **A** buys a call option from party **B**, then **A** is in a **long call** position and has the right to buy at the strike price. **B** is in a **short call** position and has the possible obligation to sell at the strike price.
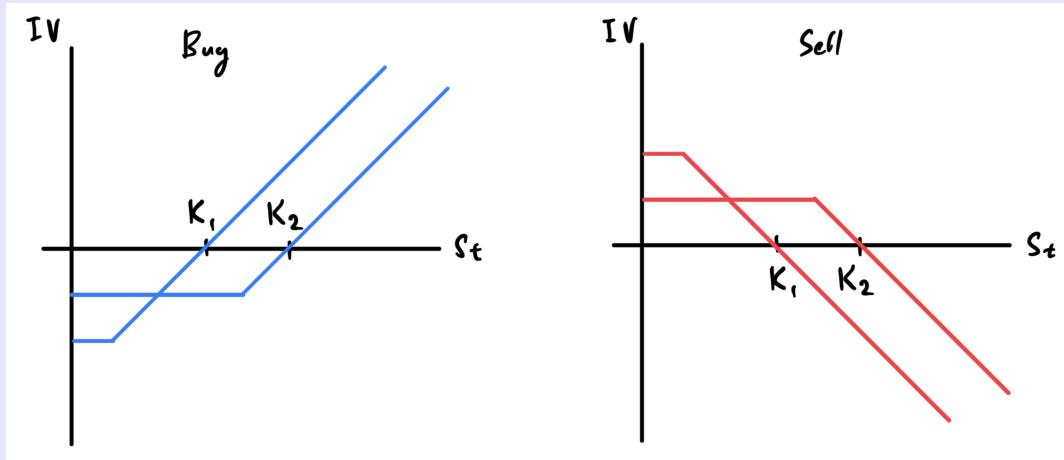
Figure 13: For calls, the premium is higher as the strike price gets lower, since you can buy at lower prices. Note that the buyer has limited loss and has unlimited gain with the long put, while the seller has limited gain and unlimited loss with the short put.

2. If party **A** buys a put option from party **B**, then **A** is in a **long put** position and has the right to sell at the strike price. **B** is in a **short put** position and has the possible obligation to buy at the strike price.
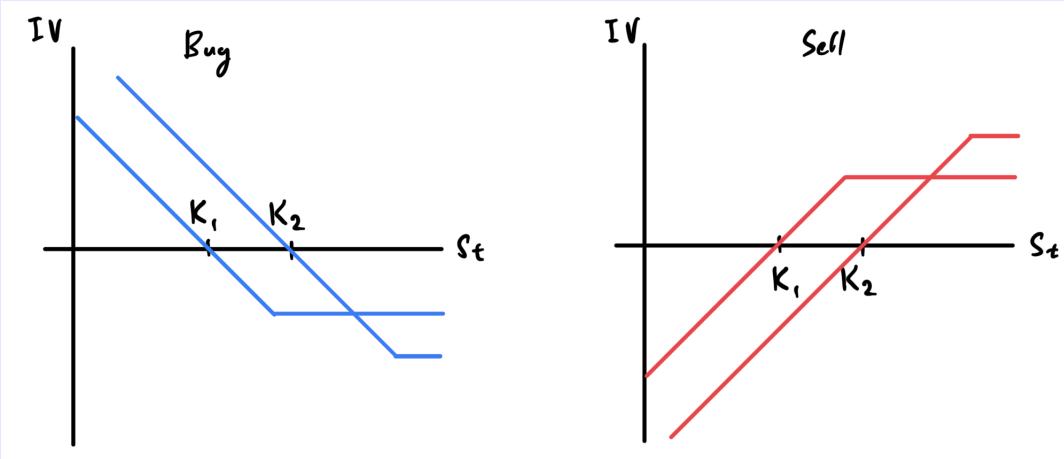
Figure 14: For puts, the premium is higher as the strike price is higher, since you can sell at a higher price. Note that the buyer has limited loss and limited gain, and the seller also has limited loss and limited gain.

---

[a]Technically, ATM options are the same at OTM, but as we will see later, ATM options have very specific and desirable characteristics, and such options tend to be the most actively traded.

There are already a few interesting properties. First, note that the long call and short put (like a double negative) take long positions in the underlying, and the other two short positions.

With just the intrinsic value, two options with the same strike price, but with one expiring in 1 month and the other expiring in 6 months, must be worth the same. This is usually not what we see in the market, and the 6-month contract is worth more. Therefore, there is some extrinsic component as well.

Definition 10.4 (Extrinsic Value)

The **extrinsic value** of an option contract $\mathrm{EV}_K(t,t)$ is affected by two components. Traders are willing to pay this additional amount due to some protective characteristics offered by options than an outright long or short equity position.
  1. The *time to expiry*. There is the probability of the underlying asset's price moving in or against our favor during the time remaining. This gives us more *optionality*, i.e. the chance to exercise our option if we choose.
  2. The *volatility*. This also has to do with the *volatility* of this underlying spot price, which we will talk about later.
  3. The *distance to ATM*. Having a larger distance to the ATM option means that there is less extrinsic value since the probability of crossing from ITM to OTM or OTM to ITM is smaller. If the option becomes more ITM, then intrinsic value will rise, but if it is more OTM, then the intrinsic value is still 0 and extrinsic also decreases.

It must clearly satisfy
$$\lim_{t \to T} \mathrm{TV}_K(t,T) = 0 \tag{79}$$

However, there are times when the time value can be 0, at which the contract is said to be trading **at parity**.

Theorem 10.1 (Components of Options Premium)

The premium paid for an option can be divided into two components, the intrinsic value and the time value.
$$V_K(t,T) = \mathrm{IV}_K(t,T) + \mathrm{TV}_K(t,T) \tag{80}$$

This must satisfy the following.
  1. The time premium must monotonically vanish at $t = T$, i.e. is a *wasting asset*.
$$\lim_{t \to T} \mathrm{TV}_K(t,T) = 0 \implies \lim_{t \to T} V_K(t,T) = \mathrm{IV}_K(t,T) \tag{81}$$

  2. The value of an option with underlying value of 0 always is also 0. That is, if $S_T = 0$ (almost surely), then
$$V_K(t,T) = 0 \tag{82}$$

To construct a naive model of an option's value that takes into account both the intrinsic value and the time premium, we can first take the expected intrinsic value at expiry.[11] Then, by buying the option for the duration, you are forgoing the opportunity to invest it with interest rate r, leading to the *theoretical value*, or the present discounted expected value, of the contract.

$$V_K(t,T) = Z_r(t,T)\mathbb{E}[\mathrm{IV}_K(T,T)] = \begin{cases} Z_r(t,T)\mathbb{E}_{S_T}[\max\{S_T - K, 0\}] & \text{for calls} \\ Z_r(t,T)\mathbb{E}_{S_T}[\max\{K - S_T, 0\}] & \text{for puts} \end{cases} \tag{83}$$

This makes sense, but it has a severe flaw: it violates the no-arbitrage principle. Let's see how.

---
[11] This is what Natenberg does in the theoretical pricing models chapter before introducing Black Scholes.

Example 10.9 (Arbitrage Opportunity?)

Suppose that $S_0 = 50$ and we have

$$\mathbb{P}(S_1 = 40) = \mathbb{P}(S_1 = 70) = \mathbb{P}(S_1 = 110) = \frac{1}{3} \tag{84}$$

Then, the expected payoff of an option with strike price 60 that expires at $T = 1$ is

$$\mathbb{E}[\text{IV}_{60}(1, 1)] = \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot 10 + \frac{1}{3} \cdot 50 = 20 \tag{85}$$

Discount this at a rate of say $r = 0.05$ and we have a valuation of $19.
To see how we can construct an arbitrage, let us sell an option for $19, then borrow an extra $31, and finally buy the underlying asset. At time $T = 1$, let us consider the following scenarios.
  1. The asset price is $S_1 = 40$. Then, the buyer won't want to exercise. You have the stock, which you sell for $40, and you pay back the loan of $31/0.95$, leaving you with $40 - 31/0.95 = 7.37$.
  2. The asset price is $S_1 = 70$. Then, the buyer wants to exercise, and you sell the stock for $60. You are left with $60 - 31/0.95 = 27.37$.
  3. The asset price is $S_1 = 110$. Then, the buyer wants to exercise, and you sell the stock for $60. You are left with $60 - 31/0.95 = 27.37$.
Therefore, this is an arbitrage portfolio since you get riskless profit from nothing. The first reader may find it difficult to find this arbitrage as if it was pulled from thin air. Later on, we will explore how this arbitrage opportunity is even found, and it turns out that the error in this derivation comes from the assumption of the probability distribution itself; it is not a risk neutral measure.

10.3   Put Call Parity

Before we start developing more options theory, let us talk about a simple property of options. This states that one long call and one short put equals one long forward, all at the same strike K and expiry T. The naive argument is that if you look at the payoff graphs, you can see that it is true.
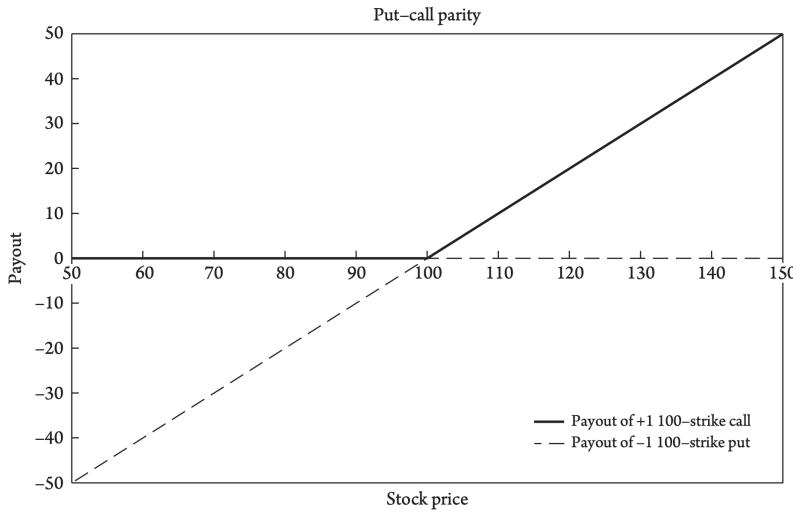
Figure 15: Adding the payoff graphs of the two options position at expiry shows that the position is equal to a forwards position at expiry.

However, this is not the correct argument, since this result is only true *at expiry*. Fortunately, we have the (corollary of the) monotonicity theorem that tells us that if the portfolios are equal with probability 1 at

some point in time, it is equal at all points in time. This is the general argument we will use here.[12]

> **Lemma 10.1 (Put-Call Parity)**
>
> The prices of the put and call option of the same strike price are related to the value of a forward option by
> $$C_K(t,T) - P_K(t,T) = V_K(t,T) \tag{86}$$
> $$= (F(t,T) - K)Z(t,T) \tag{87}$$
>
> In the specific case where interest rates are 0, then we can simplify notation and write[a]
> $$C - P = S_t - K \iff C + K = S_t + P \tag{88}$$

> **Proof.**
>
> Say that you are long call and short put with common strike K and expiry T. Then, this is the same thing as having a forward contract. Therefore,
> $$C_K(t,T) - P_K(t,T) = V_K(t,T) \tag{89}$$

An immediate consequence of this is that is the delivery price $K = F(t,T)$ we have by definition $V_K(t,T) = 0$ and therefore
$$C_{F(t,T)}(t,T) = P_{F(t,T)}(t,T) \tag{90}$$
That is, these **at the money forward (ATMF)** options have the same price, which holds despite any distributional assumptions.

> **Theorem 10.2 ()**
>
> An ITM call option or put should never be worth less than its intrinsic value.

> **Proof.**
>
> We have
> $$C_t = S_t - Ke^{-rt} + P_t > S_t - Ke^{-rt} > S_t - K \tag{91}$$

> **Theorem 10.3 (Bounds on Call Prices)**
>
> We have
> $$\max\{0, S_t - KZ(t,T)\} \leq C_K(t,T) \leq S_t \tag{92}$$

> **Corollary 10.1 ()**
>
> The price of an American call and European call on a non-dividend paying stock are equal.

10.4   Calls are Puts and Puts are Calls

There is a duality between calls and puts. In fact, it is possible to turn one into the other by either longing or shorting the underlying. There may be terminology differentiating a covered call and a protective call,

---
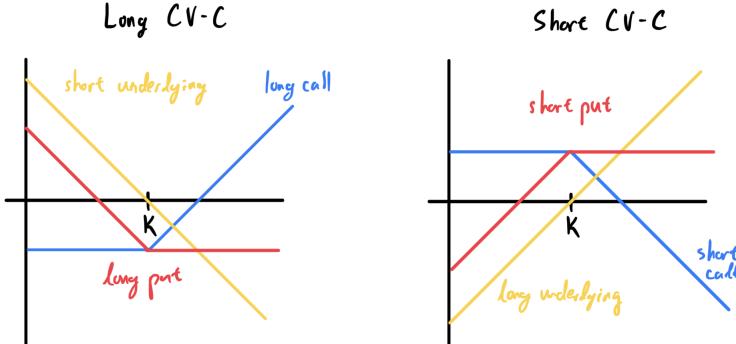
[12]Without even knowing the formula for the pricing of these options, like Black Scholes, it is amazing that we can still prove this relationship.

[a]I like to use the mnemonic of the S&P to realize that S_t and P must be on the same side.

but there are really two instruments. When we say we are *longing* a covered call/put, we are longing the option and then hedging with the appropriate underlying.
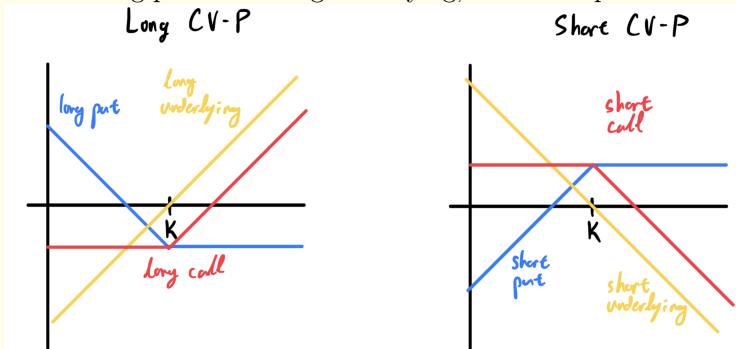

Definition 10.5 (Covered Call, CV-C)

A long **covered call** is a long call and a short underlying, which is equivalent to a long put.

Definition 10.6 (Covered Put, CV-P)

A long **covered put** is a long put and a long underlying, which is equivalent to a long call.
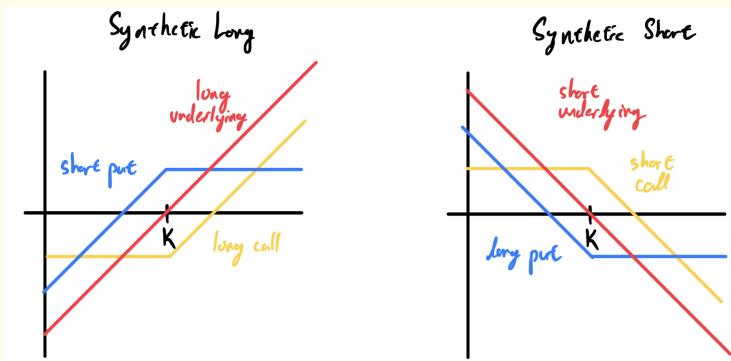
While it is worthwhile to draw the payoff diagrams at first, it is good to remember that if we own a covered call, we are taking an opposite position of a put (so must be a long put). Therefore, you should be indifferent to trading one or the other (same strike) since you can always turn calls to puts or puts to calls by longing/shorting an underlying.

This means that we can also take two options at the same strike and delivery and use them to make a pure long or pure short position in the underlying. These are called *synthetic* long and short positions.

Definition 10.7 (Synthetic Long/Short)

A **synthetic long** is to buy a (long call + short put) position. A **synthetic short** is to buy a (short call + long put) position.

Clearly, selling a synthetic long is buying a synthetic put, and the two are counterparties to each other.

If they are so equivalent, then what is the actual difference? There are two main points that we consider:

1. The exchange fees of trading two securities is obviously greater than one, so covered options and synthetic longs/shorts are more expensive. However, they are also sold together in a package as *combinations*, or *combos*, so this can be avoided.

2. Trading these combos is safer, since you are guaranteed to make both trades simultaneously. On the other hand, if you attempt to buy a call and a sell an underlying, the price that you want it executed may not go through during the second that you attempt to buy the second security. Therefore you must buy and sell combos together always.[13]

3. As a market maker accrues positions in various options and underlyings, it may be simpler to analyze our positions using these combos rather than all at once. If we own a short put and a covered call, it is simpler to think of this as a short put and a long put, which has a flat payoff diagram (assuming same strike and expiry).

4. You pay less to cross the spread (e.g. 2 cents for put and 2c for call vs 1c for spread)

10.5   Straddles and Strangles
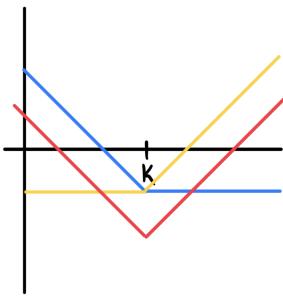
Definition 10.8 (Straddles, Y)

A **long straddle** is achieved when you buy a call option and a put option at the same strike price K and expiration date, and it is a long volatility position. If you bought the call for price C_t and bought the put at price P_t, then we have the following properties:
 1. *Value.* This is worth $C_t + P_t$.
 2. *Profit Lower Bound.* The most you can lose is $C_t + P_t$, the value of the straddle.
 3. *Positive Profit.* You profit if the stock moves at least $C_t + P_t$ away from the strike. If $S_T \notin [K - (C_t + P_t), K + (C_t + P_t)]$, then we profit.
 4. This is generally not a directional bet so you would buy ATM options, where $C_t = P_t$, but this does not always need to be the case.
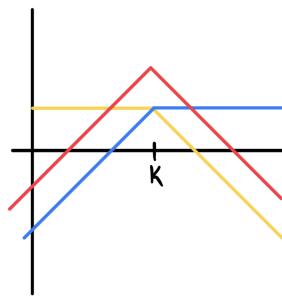
---

[13]For example, consider that we bought a straddle and the market collapsed, we might want to take off the call and let the put "ride". What usually happens is that then the market will rally and you lose on both the call and the put.

Theorem 10.4 (ATM Straddle Price)

The ATM straddle price is
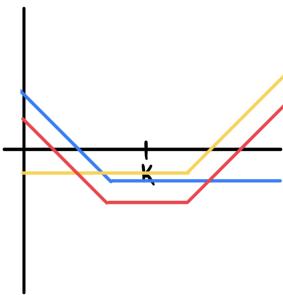
$$C_t + P_t = S_t \sigma \sqrt{\frac{t}{2\pi}} \tag{93}$$

Definition 10.9 (Strangles)
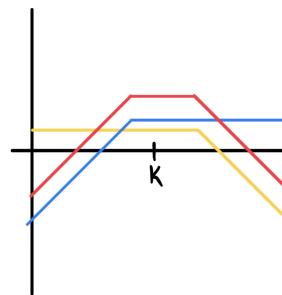
A **long strangle** is achieved when you buy a call and a put option at *different* strike prices K_c and K_p (usually both out-of-money) and the same expiration date, and it is a long volatility position. If you bought a call option for price C_t and bought the put at price P_t, then we have the following properties:

1. *Value.* This is worth $C_t + P_t$.
2. *Profit Lower Bound.* The most you can lose is again $C_t + P_t$, the value of the strangle. However, since they are both out of the money, you pay less premium than you would do on a straddle.
3. *Positive Profit.* You profit if the stock moves at least $C_t + P_t$ to the left of the put strike K_p or moves at least $C_t + P_t$ to the right of the call strike.
4. This is generally not a directional bet so you would buy ATM options, where $C_t = P_t$, but this does not always need to be the case.

The following fact quantifies the differences between the straddle and the strangle.

Theorem 10.5 (Gammas of Straddles and Strangles)
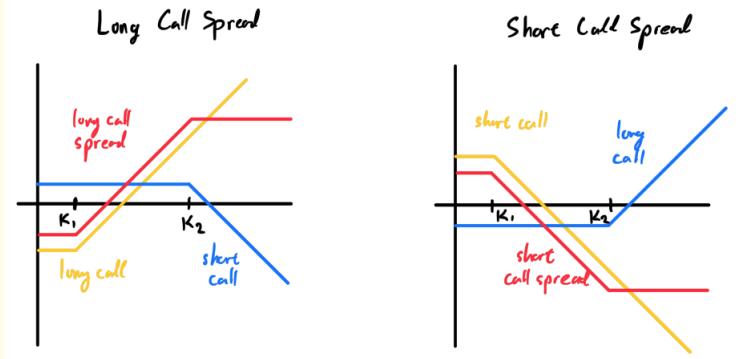
The straddle will always have a higher gamma than the strangle and therefore will be more appropriate for traders who are betting on an increase in the actual volatility of the market.

10.6   Call, Put, and Box Spreads

Definition 10.10 (Call Spread, CS)

A **call spread** consists of a long call with smaller strike K_1 and a short call on a larger strike of K_2.[a]
It has the following properties.
  1. *Value.* The K_1 strike option will be worth more since it has a lower strike, so the total value of this is $C_1 - C_2 > 0$.
  2. *Profit Lower Bound.* The most you can lose is if they are both OTM, meaning that you just lose what you paid for, which is $C_2 - C_1$.
  3. *Profit Upper Bound.* The most you gain is if they are both ITM, meaning you gain $-(C_1 - C_2) + (K_2 - K_1)$.
  4. The longer the spread $(K_2 - K_1)$ is, the more you stand to gain, but the premium on the K_2 call will be more smaller compared to that of the K_1, meaning you have more to lose.
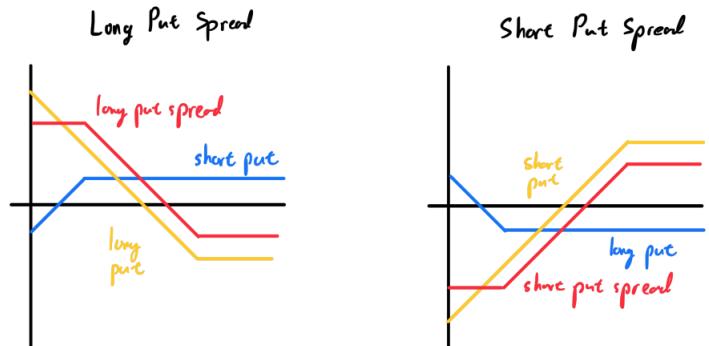
Definition 10.11 (Put Spread, PS)

A **put spread** consists of a short put with smaller strike K_1 and a long put with a larger strike K_2.[a]
It has the following properties.
  1. *Value.* The K_2 strike option will be worth more since it has a higher strike, so the total value of this is $P_2 - P_1 > 0$.
  2. *Profit Lower Bound.* The most you can lose is if they are both OTM, meaning that you just lose what you paid for, which is $P_2 - P_1 > 0$
  3. *Profit Upper Bound.* The most you can gain is if they are both ITM, meaning you gain $-(P_2 - P_1) + (K_2 - K_1)$
Often, the theo of a put spread is shown as positive, but

---

[a]It is also called a bull call spread since you are betting directionally that it will go up.
[a]It is also called a bear put spread since you are betting directionally that the underlying will go down.

Theorem 10.6 (Upper Bound on Value Call Spread)

A call spread can never be worth more than the difference between the strikes.

Proof.

Say $K_2 > K_1 \implies P_2 > P_1$. Then

$$\begin{cases} C_1 & = S_t - Ke^{-rt} + P_1 \\ C_2 & = S_t - Ke^{-rt} + P_2 \end{cases} \implies C_1 - C_2 = (K_2 - K_1)e^{-rt} + (P_1 - P_2) \tag{94}$$

and since the larger strike put is worth more, the expression above is less than $(K_2 - K_1)e^{-rt} < K_2 - K_1$.

With this, we can construct an approximation of the value of a call spread right when the underlying price is at the midpoint of the two strikes.

Corollary 10.2 (Call Spread Value at Midpoint)

Given a K_1/K_2 call spread, if the underlying $S_t = (K_1 + K_2)/2$, then the value of the call spread is

$$C_1 - C_2 = \frac{K_2 - K_1}{2} \tag{95}$$

Example 10.10 (C-SP Market Making)

To make a market in a X/Y C-SP $(X < Y)$, you simply calculate the theo, which must be positive.

$$C_X - C_Y \tag{96}$$

Example 10.11 (P-SP Market Making)

To make a market in a X/Y P-SP $(X < Y)$, you simply calculate the theo, which must be negative, and turn it positive.

$$-(-P_Y + P_X) = P_Y - P_X \tag{97}$$

Using the theorem above, along with put call parity, we can use this to approximate options prices on an option chain, a process known as **rough opening**. For example, as a trader, if you see that futures are at limit (during the trading day) and there are orders sitting in the book. You want to quickly do rough openings to look for arbitrage.

**Example 10.12 (Rough Opening with Bounds on Spreads)**
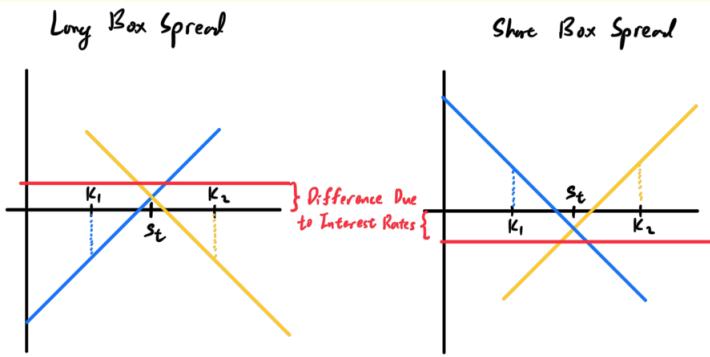
Say that the underlying price is 1460 and the prices of a 1450-strike call and put is 60 and 50. Then, let's try to estimate the call and put prices of a 1500-strike call and put.
1. If we buy a 1450/1500 call spread, you are paying $60 - C$, it should be worth half the difference in the strikes when the underlying price is halfway, i.e. 1475. So if the underlying is 1475, the CS is worth 25.
2. However, the underlying price is 1460, which is slightly less, so the CS should be worth a little less, say 20. Then we solve $60 - C = 20 \implies C = 40$ and by put call parity $P = 80$.

Definition 10.12 (Box Spread)

A **long box spread** consists of a synthetic long at a lower strike K_1 and a synthetic short at a higher strike K_2. It has the following properties.
1. If we assume that $K_1 < S_t < K_2$, then we must pay a positive theo for the synthetic long (for lower strikes, long call is worth more than short put) and we must also pay a positive theo for the synthetic put (for higher strikes, long put is worth more than short call).
2. *Buyer lends money.* The buyer will pay a fixed $(C_1 - P_1) + (P_2 - C_2)$. At expiry, they will receive the same amount plus the risk free interest rate (by no-arbitrage).
3. *Seller borrows money.* The seller will receive fixed $(C_1 - P_1) + (P_2 - C_2)$. At expiry, they will pay the same amount plus the risk free interest rate (by no-arbitrage).
4. *Longer box spreads.* Longer (in difference of strikes) box spreads represent a greater amount of money borrowed.

The box is worth a little less than the difference of the strikes, so if I buy a box, I am flat positive. It is a ZCB (zero coupon bond).

One strategy for market makers is that they can buy boxes with borrowed money from the clearing firm. Unfortunately, the interest rate r^* of the clearinghouse will be greater than the interest rate implied in the box spread, so there will be no arbitrage opportunity here as well.
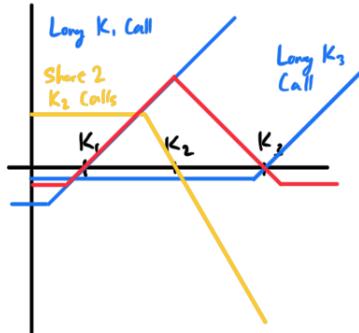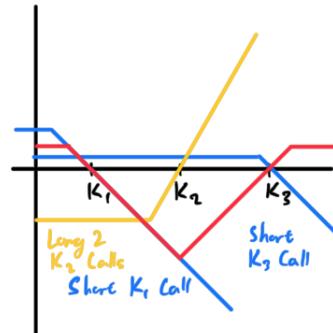
10.7   Butterfly Spreads

Conservative traders who do not like the unlimited risk associated with shorting options can use a butterfly spread to limit their loss potentials. It is a short volatility strategy that makes money from volatility decreasing but mostly from time decay. We can construct butterfly spreads with long calls, long puts, short calls, and short puts, but we will only introduce long calls here.

Definition 10.13 (Call Butterfly Spread)

A **long call butterfly spread** consists of a long call with strike K_1, 2 short calls with strike K_2, and long call with strike K_3, where $K_1 < K_2 < K_3$. It can also be considered a short straddle (shorting volatility) and a long strangle (hedging large movements). It is a short volatility bet with the following properties:
1. *Value.* The value of this spread is $C_1 - 2C_2 + C_3$.
2. *Profit Lower Bound.* The most you can lose is the price that you paid for the spread, which is $C_1 - 2C_2 + C_3$.
3. *Profit Upper Bound.* The most you can earn is if S_T is at K_2, which is $K_2 - K_1$. We must subtract the cost of this spread as well.
4. We usually have the middle call be ATM so that we are direction neutral.
5. We usually employ a *symmetrical butterfly* where $K_2 - K_1 = K_3 - K_2$ due to some nice properties mentioned later.

Theorem 10.7 (Call and Put Butterflies)

Notice that a symmetrical call butterfly spread is the same as a put butterfly spread. This should be quite intuitive since the payoff diagrams of the butterflies are symmetrical across the Y-axis. However, we can actually prove it using put-call parity.

Proof.

Let C_i and P_i have the same expiry and strike K_i. Then, by PCP,

$$C_1 - 2C_2 + C_3 = (P_1 + S_t - K_1) - 2(P_2 + S_t - K_2) + (P_3 + S_t - K_3)$$
$$= P_1 - 2P_2 + P_3 - K_1 + 2K_2 - K_3$$

and since it is symmetrical, the strikes vanish.

Example 10.13 (BS Market Making)

To make a market in a $X/Y/Z$ fly spread ($X < Y < Z$) whether they are calls or puts, we can simply compute

$$C_X - 2C_Y + C_Z \text{ or } P_X - 2P_Y + P_Z \tag{98}$$

Theorem 10.8 (Value of Fly)

The value of a symmetrical fly spread can never be less than 0.
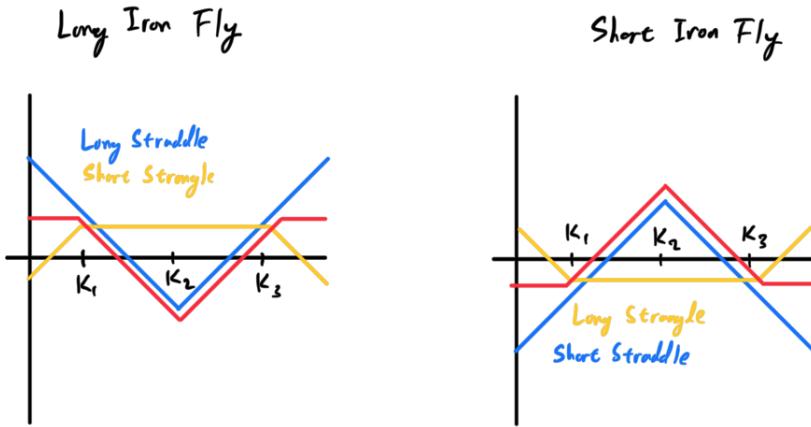
Proof.

The most you can lose is $C_1 - 2C_2 + C_3 = 0$, so it is bounded below by 0 and it must be greater.

They are also brilliant from a time decay standpoint, especially in the last 30 days. They decay at an extraordinary rapid pace in the last 30 days. So these trades are ideal to reap the rewards of heavy time decay without having to assume an unreasonable amount of risk.

Definition 10.14 (Iron Butterfly Spread)

An **iron butterfly spread** consists of a long straddle a short strangle.

From looking at the payoff diagrams, it may seem like the long call fly and the long iron fly are opposing positions, but this is not true since the long iron fly consists of 2 puts. The greeks of a long call fly and a short iron fly are the same since they are translationally invariant, but their theos are not.

Theorem 10.9 ()

The theos of iron and call butterflies sum to the differences of their strikes.
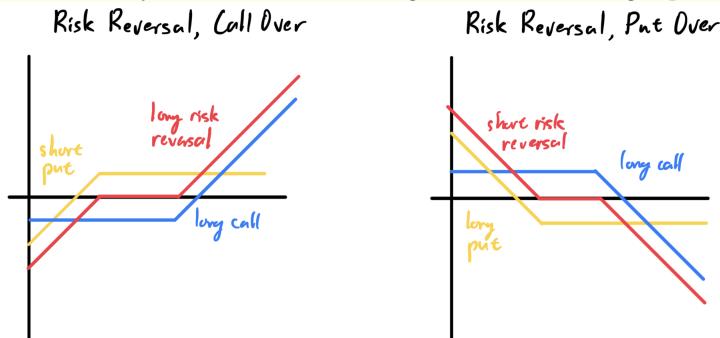
10.8 Reversals

When we hedge against a short call option by buying the underlying, it may be too expensive to do this. A cheaper way is by buying a risk reversal (call over). Additionally, if you are shorting an underlying, you can buy a RR (call over) to reverse your risk and hedge it (note that this is equivalent to a put spread).

Definition 10.15 (Risk Reversal, RR)

A **risk reversal** consists of buying and selling OTM options.[a] Whether you are call or put over determines which one you buy, and these two are counterparties to each other.
1. A **RR, call over** means you long a call at a higher strike and short a put at a lower strike.
2. A **RR, put over** means you short a call at a higher strike and long a put at a lower strike.

Note that these risk reversals can have either positive, 0, or negative theos, but when we usually quote them, we make sure to quote the one that has a positive theo (e.g. is call price is higher than put price, then we quote a call over RR).

---

[a]This does not have to be OTM, but buying an ITM option will cost you more premium.

To make markets on risk reversals, you take theos of lower strike puts and higher strike calls.

> **Example 10.14 (RR Market Making)**
>
> To make a market in a X/Y RR, $(X < Y)$ the X represents the strike of the put and Y that of a call. Which one you long and short depends on the following.
> 1. If it is call over, then you have X for -1 and Y for $+1$, i.e. $X(-1)/Y(1)$.
> 2. If it is put over, then you have X for $+1$ and Y for -1, i.e. $X(1)/Y(-1)$.
>
> Then, to price it, we have two scenarios. Let's take a `RR (-1) 178 (1) 185` (which is call over).
> 1. If the theo of the call is 3.13 and that of the put is 2.18, then the total value of this is 0.95. You want an edge of 0.05, so you make a market for 0.90 at 1.00, 1 up.
> 2. If the theo of the put is 3.13 and that of the put is 2.18, then the total value of this is -0.95. Therefore, a trader who wants to "buy" this RR receives 0.95. But buying this call over `RR (-1) 178 (1) 185` for 0.95 is the same thing as selling a put over `RR (1) 178 (-1) 185` for 0.95. Therefore you are really putting the theo of the put over RR at 0.95.

This means that if a call over RR is a negative value, you're really making a market on the put over RR for a positive value. Same for put over RR. Ultimately, when you have a X/Y RR, you really want to look for the absolute value in differences in the values, and whichever theo is greater is the one you're longing.

1. If the call theo is greater, then you are longing the call and it is automatically a call over.

2. If the put theo is greater, then you are longing the put and it is automatically a put over.

> **Example 10.15 (Estimating Underlying Price from Risk Reversal)**
>
> Say that risk reversal `RR (-1) 178 (1) 185, 78` (short 178-put long 185-call, call over, 78 delta) has a price of -1.873. You are trying to estimate the price of the underlying. Let's see how we would do this.
> 1. If the future is below 178, say $178 - \epsilon$, then the put will be ITM and call OTM. Also the put will have more extrinsic value (since it's closer to ATM). So the price of the risk reversal must fall between $(-\infty, -\epsilon]$. [a] If the future is worth $185 + \epsilon$, then the RR price must fall between $(\epsilon, +\infty)$. So, we can see that it must be either between the strikes, or it can be at least $178 - 1.873 = 176.167$. However, we should account for a nontrivial difference in extrinsic value (say 1) so it must fall somewhere between say $[177.167, 184]$.
> 2. Assume that the future is at the midpoint, i.e. at 181.5. Then, the distance to ATM is the same, and the options both have the same time to expiry and volatility of the underlying. This means that their extrinsic value must be the same. They both do not have intrinsic value so the value of this contract should be 0. However, it is negative, meaning that it is slightly closer to the put strike than the call strike. Therefore, we can estimate it to be 179.5.
> 3. We know that when the underlying price is $\pm\infty$, then the delta of the RR is $+1$ (for both cases). It bottoms out in between[b] so we can see that the delta should be lowest at 181.5, which is less than 78.
> 4. We can make an even tighter bound. The value of the RR is 0 at $S_t = 181.5$. If we shift the price to $S_t = 181.5 - x$, then since the delta is lower than 78, say 70, the value (assuming 0 gamma) must move to $0 - 0.78x = -1.873$. But since we know from before that the delta is increasing (positive gamma), let's estimate the average delta over this interval to be slightly higher, say 0.85. Then, we solve $x = 1.873/0.85 \approx 2.20$. Therefore, we can estimate the price to be $181.5 - 2.20 = 179.3$.
>
> The answer is 179.55.
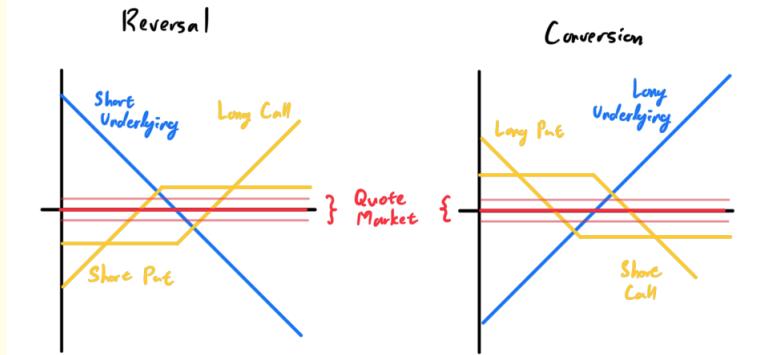
---

[a]On the other hand, if the risk reversal is worth less than $-\epsilon$, then the future is at least $178 - \epsilon$.
[b]Look at the function $f(x) = e^x + e^{-x}$.

Definition 10.16 (Conversion, Reversal)

A **reversal** (reversing the underlying so short it) consists of a short underlying, a long call, and a short put. A **conversion** is simply the opposing position: a long underlying, a short call, and a short put.

The reason we have these positions is that a reversal is really just a synthetic long put (short underlying and long call) plus a short put. This should equal out to 0, but by trading these three instruments, either separately or as a combo, a MM can quote a market around this theo to gain more edge by cancelling out a position with a synthetic option.[a]

Example 10.16 (Reversal Convesion Market Making)

To make a market in the X Rev/Con, you take the difference between the underlying price and the synthetic underlying price (constructed from the put and the call). Therefore, by put call parity, the market is

$$\hat{S} = C_X - P_X + K \tag{99}$$

10.9 Ratio Spreads and Ladders

Definition 10.17 (1x2 Call Spread)

A **1 by 2** consists of buying one K_1 call and selling two K_2 calls (is a negative price because the 2 is probably worth more).

Definition 10.18 (Call, Put Ladder)

A **put ladder** consists of selling a K_1 put, selling a K_2 put, and buying a K_3 put where $K_1 < K_2 < K_3$.

11 Black Scholes

11.1 Risk Neutral Probability

Let's consider a simple model on a non-dividend paying stock at price \$100 at time $t = 0$. Assume annually compounded interest rates are 10% and that

$$\mathbb{P}(S_1 = 120) = p, \;\; \mathbb{P}(S_1 = 90) = 1 - p \tag{100}$$

---

[a]However, in an no-arbitrage scenario, this is impossible.

That is, it either increases by 20% or decreases by 10%, these events labeled A and B. Let's try to value these options using 2 different ways.

1. Assume that we long an option. Then, the payoff at time T would be

$$\max\{S_T - 110, 0\} = \begin{cases} 10 & \text{with prob.} = p \\ 0 & \text{with prob.} = 1 - p \end{cases} \tag{101}$$

Therefore, our expected payoff is $10p$, which we discount to get $10pZ(0,1)$. Note from before that this method of valuation is inherently faulty.

2. Assume that we long an option and short Δ stock. Then, the payoff at time T would be

$$\max\{S_T - 110, 0\} - S_T\Delta = \begin{cases} 10 - 120\Delta & \text{with prob. } = p \\ -90\Delta & \text{with prob. } = 1 - p \end{cases} \tag{102}$$

Let's minimize the variance of this portfolio by setting the two equal to each other.

$$10 - 120\Delta = -90\Delta \implies \Delta = \frac{1}{3} \tag{103}$$

Therefore, let's consider this portfolio of long one option and short $1/3$ stock. Then, our profit is

What would our exposure be if we held various options? Say that we bought a call at strike 110. Then, our payout would be

Say that we shorted a call option at strike 110. Then, our payout would be

$$-\max\{S_T - 110, 0\} = \begin{cases} -10 & \text{with prob.} = p \\ 0 & \text{with prob.} = 1 - p \end{cases} \tag{104}$$

This gives us negative exposure[14] and to hedge this, we may consider holding a position in the underlying stock. Let's say we want to hold Δ positions of the stock. Then, our payout at time T is

$$S_T\Delta - \max\{S_T - 110, 0\} = \begin{cases} 120\Delta - 10 & \text{with prob. } = p \\ 90\Delta & \text{with prob. } = 1 - p \end{cases} \tag{105}$$

If we wanted to minimize the variance of this portfolio, then we can just solve for

$$120\Delta - 10 = 90\Delta \implies \Delta = \frac{1}{3} \tag{106}$$

This is called a *covered call*.[15] We can develop these for hedges against the four basic options.

Recall our initial attempt at pricing options using the discounted expected payoff. In fact, let's redo this with the current example to gain some insight on how one would find this arbitrage. Assume that the asset price is S_0 and that we have a one-step binomial model with

$$\mathbb{P}(S_1 = S_0(1 + u)) = p \text{ and } \mathbb{P}(S_1 = S_0(1 - d)) = 1 - p \tag{107}$$

Let the strike price be K, which should be

11.2 Continuous Time Limit and Derivation

The binomial tree is a good initial model, but clearly the values that a stock can take are not a set of discrete numbers. It is more mathematically convenient to take the limit of this tree to the reals. We can also think of this as the limiting case of the percent change being normal.

---

[14]There would be a premium that we would receive when we sold the option, but for now we assume that we just hold it.

[15]It is nice to remember that for calls, you take the opposite position in the underlying (e.g. buy calls, sell underlying) and for puts, you take the same position (e.g. buy puts, buy underlying).

Question 11.1 (TBD)

Show that the risk neutral probability converges to the log normal distribution.

The problems are modeled through the Black-Scholes model comes in, which is stated with the following PDE.

Definition 11.1 (Black Scholes PDE)

The **Black-Scholes PDE** is the following.

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{108}$$

We can model the price of a stock and want to construct a measure that follows this stochastic differential equation.

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{109}$$

under some real world measure $\mathbb{P}$. This is equivalent to

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \text{ where } dt = (dW_t)^2 \tag{110}$$

We first start off with the intuitive pricing of a call option as a discounted expected payoff.

$$C_0 = \mathbb{E}[e^{-rT}\max\{S_T - K, 0\}] \tag{111}$$

In Black Scholes, there is a formula like this, but not under the real world measure $\mathbb{P}$, but this is true under another measure $\mathbb{Q}$. This is formalized in the following theorem.

Theorem 11.1 (Girsanov Theorem)

There exists another measure $\mathbb{Q} \sim \mathbb{P}$ s.t.

$$dS_t = rS_t dt + \sigma S_t \tilde{W}_t \tag{112}$$

where $\tilde{W}_t$ is another Brownian motion.

Let's go back to the main point. We want to solve this SDE to get geometric Brownian motion.

Theorem 11.2 (Solution to SDE)

It turns out that the solution is

$$S_t = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma\tilde{W}_t\right\} \tag{113}$$

Proof.

Using Ito's lemma, we have

$$d(\ln S_t) = \frac{1}{S_t}dS_t - \frac{1}{S_t^2} \cdot \frac{1}{2}(dS_t)^2 \tag{114}$$

$$= rdt + \sigma d\tilde{W}_t - \frac{1}{2}\sigma^2 dt \tag{115}$$

$$= \left(r - \frac{1}{2}\sigma^2\right)dt + \sigma d\tilde{W}_t \tag{116}$$

and so

$$[\ln S_T]_0^T = \int_0^T \left(r - \frac{1}{2}\sigma^2\right) dt + \int_0^T \sigma d\tilde{W}_t \implies \ln\left(\frac{S_T}{S_0}\right) = \left(r - \frac{1}{2}\sigma^2\right)T + \sigma(\tilde{W}_T - \tilde{W}_0) \quad (117)$$

and since $\tilde{W}_0 = 0$ since it is Brownian motion, we have the result.

Now we just need to calculate $C_t = \mathbb{E}[e^{-rT} \max\{S_T - K, 0\}]$, where only S_T is random. Since we just found a solution to S_T above, which has the only random component $\tilde{W}_T$, this distribution is related to Brownian motion $\tilde{W}_T$. Note that

$$\ln\left(\frac{S_T}{S_0}\right) = \underbrace{\left(r - \frac{1}{2}\sigma^2\right)T}_{\text{mean or normal}} + \underbrace{\sigma}_{\text{std. dev.}} \tilde{W}_T \implies \ln\left(\frac{S_T}{S_0}\right) \sim \mathcal{N}\left(\left(r - \frac{1}{2}\sigma^2\right)T, \sigma^2\right) \quad (118)$$

since $\tilde{W}_T$ already is a normal with mean 0 and variance T. Now we can calculate the integral/expectation. The max term is annoying to deal with, so we divide it

$$C_T = \mathbb{E}_\mathbb{Q}[e^{-rT}(S_T - K)\, \mathbb{1}(S_T > K)] \quad (119)$$

$$= e^{-rT}\{\mathbb{E}_\mathbb{Q}[S_T\, \mathbb{1}(S_T > K)] - \mathbb{E}_\mathbb{Q}[K\, \mathbb{1}(S_T > K)]\} \quad (120)$$

$$= e^{-rT}\{\mathbb{E}_\mathbb{Q}[S_T\, \mathbb{1}(S_T > K)] - K\mathbb{P}(S_T > K)\} \quad (121)$$

We can already see that each of these expectations evaluate to each term in the Black Scholes solution. Let's solve the simpler second term first. We have

$$\mathbb{P}(S_T > K) = \mathbb{P}\big(\ln(S_T/S_0) > \ln(K/S_0)\big) \quad (122)$$

$$= \mathbb{P}\left(\underbrace{\frac{\ln(S_T/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}}_{N(0,1)} > \frac{\ln(K/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \quad (123)$$

$$= \mathbb{P}\left(\underbrace{\frac{\ln(S_T/S_0) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}}_{N(0,1)} < \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \quad (124)$$

$$= \Phi\left(\frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \quad (125)$$

As for the second expectation, since $S_T = S_0 e^{\ln(S_T/S_0)}$, we can substitute $X = \ln(S_T/S_0)$ and write

$$\mathbb{E}_\mathbb{Q}[S_T\, \mathbb{1}(S_T > K)] = \mathbb{E}[S_0 e^X\, \mathbb{1}(S_0 e^X > K)] \quad (126)$$

$$= S_0 \mathbb{E}[e^X\, \mathbb{1}(X > \ln(K/S_0))] \quad (127)$$

$$= S_0 \int_{\ln(K/S_0)}^\infty e^x \frac{1}{\sqrt{2\pi}b} e^{-\frac{(x-a)^2}{2b^2}}\, dx \quad (128)$$

$$= S_0 \int_{\ln(K/S_0)}^\infty e^{rT}\, N(a + b^2, b^2)\, dx \quad (129)$$

$$= S_0 e^{rT}\, \Phi\left(\frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \quad (130)$$

where $a = \left(r - \frac{1}{2}\sigma^2\right)T$ and $b = \sigma\sqrt{T}$. Therefore, by substituting these into the original expectation, we have

$$C_T = e^{-rT}\left\{S_0 e^{rT}\, \Phi\left(\frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) - K\Phi\left(\frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right)\right\} \quad (131)$$

$$= \Phi\left(\frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) S_0 - \Phi\left(\frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) K e^{-rT} \quad (132)$$

It has the following solution, which we will state now but will elaborate later.

> **Definition 11.2 (Black-Scholes Option Pricing Formula)**
>
> Given a call option, the **Black-Scholes** pricing formula states that
>   1. For a European call option, its price is
>
> $$C_K(t,T) = \Phi(d_1)S_t - \Phi(d_2)Ke^{-r(T-t)} \tag{133}$$
>
>   2. For a European put option, its price is
>
> $$P_K(t,T) = -\Phi(-d_1)S_t + \Phi(-d_2)Ke^{-r(T-t)} \tag{134}$$
>
> where r is some risk-free applicable interest rate, Φ is the normal CDF, σ is the volatility of the underlying, and
>
> $$d_1 = \frac{\ln(S_t/K) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \text{ and } d_2 = d_1 - \sigma\sqrt{T-t} \tag{135}$$

Let's explore a little bit of what each term means in a colloquial sense.
   1. The $\ln(S_t/K)$ represents the **log-moneyness**[16]

   2. The d_1 represents the normalized measure of the log-moneyness, considering the time remaining under expiration and the stock's volatility. It is also called the **z-score in the d_1 space**.

   3. The $\Phi(d_1)$ represents the probability of the option being ITM at expiry.

   4. The d_2 can be thought of as a measure of the expected value of the option at expiration, considering the probability of expiring ITM and the potential impact of volatility on the option's price. It is also called the **z-score in the d_2 space**.

   5. $\Phi(d_2)$ represents the probability that the option will be exercised under risk-neutral conditions, assuming that the option is held until expiration.

   6. Therefore, the term $\Phi(d_1)S_t$ represents the expected value of receiving the stock at expiration.

   7. The $\Phi(d_2)Ke^{-(T-t)}$ represents the present expected value of paying the strike price at expiration.

11.3   Properties

The Black Scholes formula is the foundation of how other options behave. We can derive almost everything from this formula. Therefore, we should check a few properties.

> **Theorem 11.3 (PCP)**
>
> These formulas make sense with PCP, since we can use the identity for a normal $\Phi(x) = 1 - \Phi(-x)$[a] to take the difference and see that
>
> $$C_K(t,T) = \Phi(d_1)S_t - \Phi(d_2)Ke^{-r(T-t)}$$
> $$P_K(t,T) = -\Phi(-d_1)S_t + \Phi(-d_2)Ke^{-r(T-t)}$$
> $$C_K - P_k = S_t - Ke^{-r(T-t)}$$
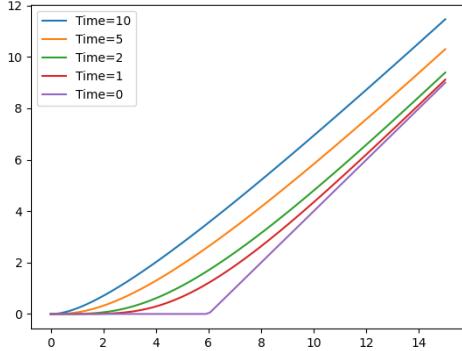
Let's take a look at how the value of this behaves with changes in some of the parameters. We can do this by looking at payoff diagrams for time t that is not necessarily at expiry, known as *risk diagrams*.

---

[16]The **moneyness** is used to represent how ITM or OTM is for an option, which is $K - S_t$ for a call and $S_t - K$ for a put. To symmetrize this, we use the log-moneyness.
[a]Derived from $\mathbb{P}(X \geq x) = \mathbb{P}(X \leq -x) = 1 - \mathbb{P}(X \geq -x)$.
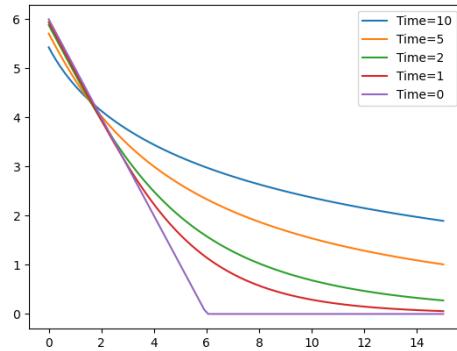
Theorem 11.4 (BS wrt Time)

As time decreases, the value converges to the payoff diagram at expiry.

(a) The value of a call option across the underlying space.

(b) The value of a call option across the underlying space.
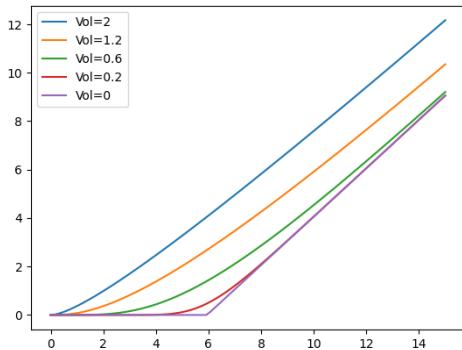
Proof.

We can interpret this multiple ways.
1. Mathematically, we can see that at the time increases, the present value of the strike is decreased further.
2. The extrinsic value is lost as time goes on, so all there's left is the intrinsic value.
3. If we have 0 time, we are at expiry and so the value must be the payoff diagram.
4. If we have infinite time, we will have a probability of 1 of being ITM.
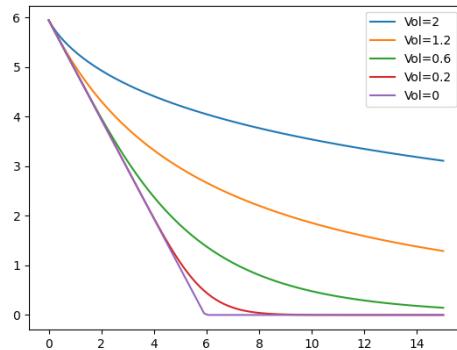
Theorem 11.5 (BS wrt Vol)

As vol decreases, the value converges to the payoff diagram at expiry.

(a) The value of a call option across the underlying space.

(b) The value of a call option across the underlying space.

Proof.
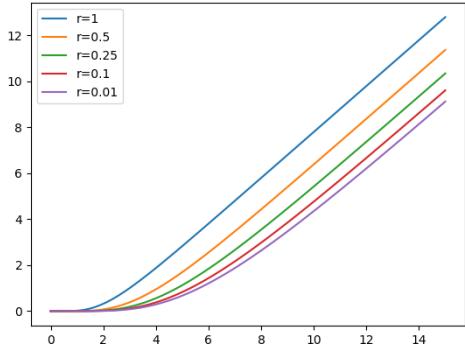
We can interpret this in multiple ways.
1. As volatility decreases, the extrinsic value decreases since the probability of it being ITM is lower.
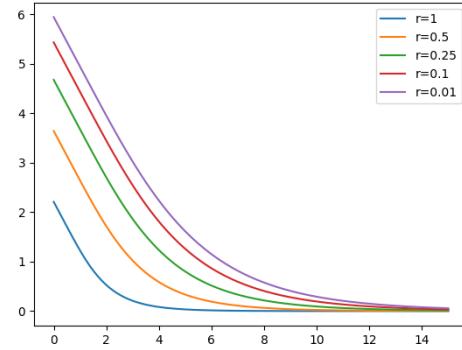
2. If vol is 0, the underlying value will never change so it must be the payoff diagram.
3. If vol is infinity, the probability of the option being ITM is 1.

Theorem 11.6 (BS wrt Vol)

As interest rates decrease, the value converges to the payoff diagram at expiry.

(a) The value of a call option across the underlying space.   (b) The value of a call option across the underlying space.

Proof.

We can interpret this in multiple ways.
1. If the interest rate is high, then the discounted strike price will diminish, leading to a rise in calls and drop in puts.

11.4 Bachelier Model

The key point of the Black-Scholes model is that it assumes a log-normal distribution of prices of $S_T \mid S_t$. This makes sense since it doesn't account for negative prices, but there are times when the prices does indeed become negative. The *Bachelier model* accounts for this by simply assuming that the distribution of the price is normal.

Definition 11.3 (Bachelier Formula)

The **Bachlier formula** for a call and put option is

$$C_t = (S_t - Ke^{-r(T-t)})\Phi(d) + \sigma\sqrt{T-t}\,\phi(d)$$
$$P_t = -(S_t - Ke^{-r(T-t)})\Phi(-d) + \sigma\sqrt{T-t}\,\phi(-d)$$

where

$$d = \frac{S_t - Ke^{-r(T-t)}}{\sigma\sqrt{(1 - e^{-2r(T-t)})/2r}} \tag{136}$$

We can check that this also satisfies PCP.

12   Primary Greeks

The intrinsic value is really always fixed, so when we are looking at Greeks, we are focusing on measuring the change in extrinsic value. As said before, this is dependent on volatility, distance to ATM, and time to expiry. The delta and gamma actually measures this change in intrinsic value (by changing the underlying value), which also changes the extrinsic value since the distance to ATM also changes. Therefore,

1. Delta and Gamma are concerned with the change in both intrinsic and extrinsic value. Vanna is $dDel/dVol$. Charm is $dDel/dt$.

2. Theta and Vega are concerned with the change in extrinsic value.

3. Rho is concerend with the change in interest rate, which is negligible in most cases.
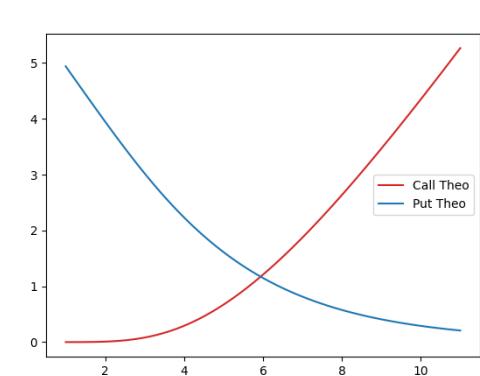
12.1   Delta
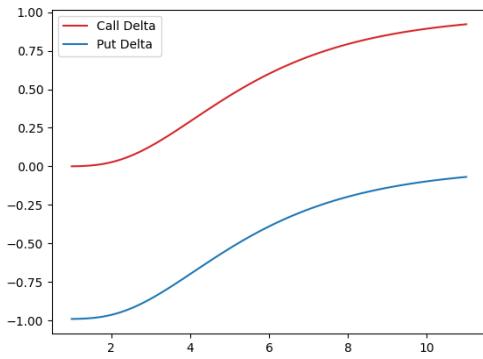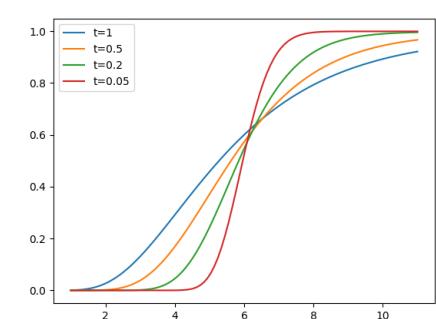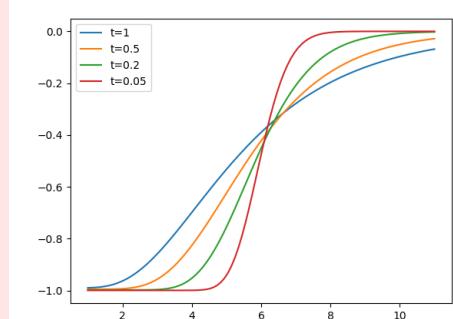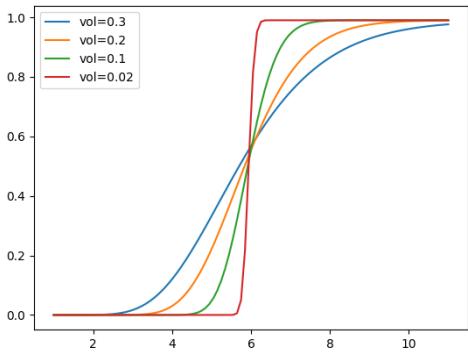
> **Definition 12.1 (Delta)**
>
> The **delta** of an option is defined
>
> $$\Delta_t := \begin{cases} \frac{\partial C_t}{\partial S_t} & = e^{-r(T-t)}\Phi(d_1) \\ \frac{\partial P_t}{\partial S_t} & = -e^{-r(T-t)}\Phi(-d_1) \end{cases} \tag{137}$$
>
> meaning that for every \$1 increase in stock price, the option price increases roughly by \$$\Delta$, with all else being equal. It indicates how many options contracts are needed to hedge a long or short position in the underlying asset. Note that the delta for a call must be in $[0, 1]$ and of a put must be in $[-1, 0]$, but traders tend to describe it in $[0, 100]$ and $[-100, 0]$.
>
>
>
> (a) The theos from BS for reference.
>
>
>
> (b) The deltas are only vertically translated. The delta of an ATM call is always 50 and an ATM put is always -50.
>
> A third, often useful, interpretation is that it is the *approximate* (not exact) probability of the option finishing ITM.

> **Theorem 12.1 (Delta wrt Time)**
>
> As time passes, the delta will be more concentrated. It will be either ± 100 for ITM calls and puts, and 0 for OTM calls and puts.

(a) Deltas get more squished.   (b) The behavior for put deltas is the same for call deltas.

It is interesting to see what happens to the delta if the option is exactly ATM. In this case, the delta can really be what you want. When options expire on some day, owners still have the optinality to exercise for the rest of the day. Expiry is before delivery (a few days). If you buy 100 ATM calls and sell 50 futures, and the futures are 100 at expiry, then Δ is undefined.

Theorem 12.2 (Delta Across Volatility)

Since volatility also behaves similar to time (increase in time = increase in volatility), we also see that as volatility decreases, the delta becomes more sensitive ATM.
1. ITM calls/puts $\Delta \to \pm 100$.
2. OTM calls/puts: $\Delta \to 0$.

(a) Delta becomes more concentrated.   (b) Delta becomes more concentrated.

Definition 12.2 (Delta Hedging)

Delta hedging is the process of dynamically hedging the option position with the underlying stocks. Given the nonlinearity, the hedge ratio has to be adjusted dynamically. For example, to hedge a short call option, we need to buy Δ amount for stock for each option. The resulting portfolio

$$V_t = -C_t + \Delta S_t \tag{138}$$

has the first derivative 0, which is called **delta neutral**.

Figure 22: Delta hedged portfolio.

However, as S_t changes, the Δ_t changes by the amount of $\Gamma_T \Delta_t S_t$. That means
1. when the price goes up, we need to buy more shares, and
2. when the prices goes down, we need to sell more shares.

A common misconception of hedging is that it will increase the expected PnL. This is not true and has no effect. Rather, on average, the true purpose of it is to *decrease variance*.

12.2 Gamma

Definition 12.3 (Gamma)

The **gamma** is the sensitivity of Δ w.r.t. S, i.e. the double derivative. The gamma is also referred to as the change in delta for a 1 point move.

$$\Gamma_t := \frac{\partial^2 C_t}{\partial S_t^2} = \frac{\partial^2 P_t}{\partial S_t^2} = e^{-r(T-t)} \frac{\phi(d_1)}{S_t \sigma \sqrt{T-t}} = K e^{-r(T-t)} \frac{\phi(d_2)}{S_t^2 \sigma \sqrt{T-t}} \tag{139}$$

If you buy an option, you have positive gamma, and if you sell an option, you have negative gamma.[a]


Figure 23: Gamma across strike for both calls and puts. Gamma peaks around ATM since this is where delta is most sensitive.

Now let's talk about a trading called **gamma scalping**. Since all long options have positive gamma, a net long portfolio will be **long gamma**. Say that you have 100 30-delta call option with a gamma of 0.02. You

---

[a]Note that there is a skew in gamma. We explore why that happens later.

must short 30 futures to get delta-neutral (the futures have 0 gamma since there's no curvature).

1. Now if the future moves up 1 point, we gain 2 delta, and so we must sell 2 more futures to be delta neutral.

2. If the futures move down 1 point, we lose 2 delta, and so we must buy 2 more futures to be delta neutral.

Therefore, we will profit no matter what. This process is called **hedging gamma**, which means you are hedging accumulated deltas due to gamma.[17] Essentially, whether a trader is long or short gamma determines a trader's bet on the *realized volatility* (as opposed to implied volatility as in vega, more on this later) of the underlying

We sell high and buy low, and the rate at which we transact is the **path**. We can choose to buy or sell every day, every week, or when the price reaches a certain threshold. Our profits will depend on this path that we take, called **path dependency**. Clearly, the higher volatility is, the more opportunities we have to gamma scalp. We have just talked about a riskless profit, so is this arbitrage? No, because this balances out with our theta. Generally, new traders should avoid large gamma positions due to the speed at which the delta, i.e. the directional risk, can change.

Lemma 12.1 (Integral of Gamma)

The integral of gamma is always 1.

$$\int_{-\infty}^{+\infty} \Gamma \, dS_t = 1 \text{ for all } t \tag{140}$$

Proof.

By the fundamental theorem of calculus.

$$\Delta(+\infty) = \Delta(-\infty) + \int_{-\infty}^{+\infty} \Gamma \, dS_t \tag{141}$$

where $\Delta(+\infty) = 1, 0$ and $\Delta(-\infty) = 0, -1$ for a call and put, respectively.

Theorem 12.3 (Gamma wrt Time)

As time goes on, we can again take the derivative and see that the gamma becomes more an more peaked around ATM. This happens since the delta is more squished.

---

[17]Hedging is a 0 EV decision, but it still costs to hedge (since we have to cross the spread). If we can hedge and break even, we are very happy.

Figure 24: Gamma is more concentrated as time passes.

You can see that gamma increases when ATM but will eventually decrease for everywhere else. The following plot gives a better representation of this.


Figure 25: Gamma does not always increase. For every option that is not ATM, it will eventually go to 0 while peaking around the ATM.

Theorem 12.4 (Gamma wrt Volatility)

Since decreasing time is also like decreasing vol, the gamma also is more concentrated. As volatility decreases, the delta becomes more sensitive ATM.

Figure 26: Gamma is more concentrated as volatility decreases.

Proof.

This makes sense since if there's more volatility, we can also gamma scalp at a larger rate earning more profits.

12.3   Cash Delta, Gamma and Portfolios

Let $C(S)$ denote the price of a call option w.r.t. the underlying price S. Say that the spot price is S_0 now and that later, the spot price changes to $S_0 + S'$. Then, the second order Taylor approximation of our PnL is

$$C(S_0 + S') - C(S_0) = C'(S_0)S' + \frac{1}{2}C''(S_0)(S')^2 \tag{142}$$

$$= \Delta(S_0)S' + \frac{1}{2}\Gamma(S_0)(S')^2 \tag{143}$$

To calculate the *portfolio price*, you must multiply this by the multiplier M and the number of options you own N. Thus your portfolio price is $P(S) = NMC(S)$ and your PnL is:

$$P(S_0 + S') - P(S_0) = P'(S_0)S' + \frac{1}{2}P''(S_0)(S')^2 \tag{144}$$

$$= NM\Delta(S_0)S' + \frac{1}{2}NM\Gamma(S_0)(S')^2 \tag{145}$$

Definition 12.4 (Portfolio Delta, Gamma)

The **portfolio delta** and **portfolio gamma** are the delta and gamma of the contract, scaled up by MN.

While delta and gamma are good ways to compute the directional and realized volatility exposure of an option, it is not good for comparing different options or portfolios of options. This is because a 1 point move in a \$200 underlying is different from a 1 point move in a \$4 underlying. We must *normalize* this in terms of percent change. This leads us to the concepts of cash delta.

> **Definition 12.5 (Cash Delta)**
>
> The **cash delta** of a portfolio represents the change in the portfolio's value w.r.t. a 100% percent change, i.e. the PnL potential at this time.[a]
>
> $$\tilde{\Delta} = \Delta(S_0) M N S_0 \tag{146}$$
>
> which represents
> 1. The first order change in PnL with a 100% increase in the underlying. Therefore, a $x\%$ increase in our underlying S_0 means a
> $$\text{Change in PnL} = \frac{x}{100} \cdot \tilde{\Delta} \tag{147}$$
> 2. The cash needed to delta hedge the underlying contracts of this portfolio. The delta of this portfolio is $\Delta(S_0) M N$ and you would want to buy/sell this many underlying contracts at price S_t.

> **Example 12.1 (Portfolio vs Cash Delta)**
>
> Take a look at two portfolios
> 1. $\mathcal{P}_1$ of one oil future trading $65 per barrel, multiplier 1000.
> 2. $\mathcal{P}_2$ of two corn futures trading at $3.50, multiplier 5000.
> The portfolio delta are
>
> $$\Delta_{\mathcal{P}_1} = 1 \cdot 1000 \cdot 1 = 1000 \tag{148}$$
> $$\Delta_{\mathcal{P}_2} = 1 \cdot 5000 \cdot 2 = 10,000 \tag{149}$$
> $$\tag{150}$$
>
> This looks like the corn future has much more exposure, since a $1 move in the underlying can change the value of the portfolio by $10,000$, while it can only change the oil by $1,000$. However, note that a $1 change from 3.50 to 4.50 is a 28% change, while 65 to 66 is a 1.5% change. Therefore, it is better to compare the change in the portfolios with the *percent* change in the underlying. The cash deltas are
>
> $$\Delta_{\mathcal{P}_1} = 1 \cdot 1000 \cdot 1 \cdot 65 = 65,000 \tag{151}$$
> $$\Delta_{\mathcal{P}_2} = 1 \cdot 5000 \cdot 2 \cdot 3.50 = 35,000 \tag{152}$$
> $$\tag{153}$$
>
> This means that the oil portfolio is riskier, since a 1% change in underlying will shift the portfolio by $650 compared to $350 for the corn.

12.4   Theta

The risk diagram of an option shows that the price of options falls to their intrinsic value.

---

[a]You can think of this as the local linear approximation of the PnL after a change of basis into percent change space.

Figure 27: A surface that shows the evolution of the risk diagram over time.

This decrease in value is really just from the extrinsic value, so let's remove the intrinsic value.


Figure 28: The extrinsic value peaks around ATM options and is roughly symmetric w.r.t. the moneyness. For ATM options, it seems that the value is decaying at a faster rate (concave down) while for more ITM or OTM options the curve flattens out to 0.

Definition 12.6 (Theta)

The **theta** measures the change of the option value with respect to time, i.e. it is the derivative of the extrinsic value.

$$\Theta_t := \begin{cases} \frac{\partial C_t}{\partial t} & = -\frac{S \cdot \sigma \cdot \phi(d_1)}{2\sqrt{T}} - r \cdot K \cdot e^{-r(T-t)} \cdot \Phi(d_2) \\ \frac{\partial P_t}{\partial t} & = -\frac{S \cdot \sigma \cdot \phi(d_1)}{2\sqrt{T}} + r \cdot K \cdot e^{-r(T-t)} \cdot \Phi(-d_2) \end{cases} \tag{154}$$

Theta is also referred to as the change in the value of an option over a 24-hour period (1 business day). It is always negative but we write positive. We don't say long or short theta. We say if we bleed/lose theta or gain theta. Theta is the same for both calls and puts at the same strike K.

We can think of theta as a counteracting force that balances out the effects of gamma. Therefore, we must distinguish its effects over ATM options and non-ATM.

Theorem 12.5 (Theta wrt Time)

Theta will increase (exponentially) as time decreases for ATM options, and will decrease to 0 for ITM/OTM options.


Figure 29: Theta decay is most prominent for ATM options. So it is most advantageous to sell ATM options if you want to reap the benefits of theta decay.

Proof.

This makes sense, since ATM there is the greatest distance from the value and the payoff diagram (at expiry). So we must cover more distance to reach it.

Theorem 12.6 (Theta wrt Vol)

Theta will increase (exponentially) as vol decreases for ATM options, and will decrease to 0 for ITM/OTM options.

Example 12.2 (Opposite Gamma and Theta Positions)

With a put skew, it is possible to have a short gamma and short theta position.
  1. We long a lot of OTM puts with long expiry, which will give us a lot of theta
  2. We short ATM options which will give us large negative gamma

12.5　Vega

Definition 12.7 (Vega)

The **vega** measures the impact of a change in the implied volatility of the underlying. The vega for a call and a put option on the same strike are the same.

$$\nu_t := S_t e^{-r(T-t)} \phi(d_1)\sqrt{T-t} = K e^{-r(T-t)} \phi(d_2)\sqrt{T-t} \tag{155}$$

Typicaly, it is represented as the price move per 1-point (percent) change in volatility.

Figure 30: Vega across underlying price.

Theorem 12.7 (Vega wrt Time)

The entire vega will monotonically go down as time passes.

Figure 31: Change in vega as time passes.

Theorem 12.8 (Vega wrt IV)

As IV decreases, the vega also decreases but stays the same for ATM options.


Figure 32: Vega decreases as volatility decreases, but now at a faster exponential rate.

If we perform delta hedging strategy continuously, we would have

$$\text{PnL} = \int_0^T \frac{S_t^2}{2} \Gamma_t (\sigma_{\text{implied}}^2 - \sigma_{\text{realized}}^2) \, dt \tag{156}$$

A practical note. Vega can refer to the vega of a position/portfolio of options, e.g. I'm long 4 vegas. (Similar to portfolio delta, where you multiply the vega of each option by the number of options and the multiplier).

12.6   Rho

The rho is the least important, since interest rates barely change. However, let's try to reason through how interest rates would affect the price of call and put options. If interest rates rise, then the forward price of the underlying will rise, or equivalently, the present value of the strike price will fall, and by PCP, the call option must rise and puts to fall. At this same time, higher interest rates will cause both options to fall. For the call option, because the price of a stock is always greater than that of an option, the rise in forward price will always outweigh the fall in option price. Thus, a call options will rise in value and puts will fall. This brings us to the formula.

> **Definition 12.8 (Rho)**
>
> The **rho** measures the impact of a change in interest rates.
>
> $$\rho_t := \begin{cases} \frac{C_t}{\partial r} & = K \cdot (T - t) \cdot e^{-r(T-t)} \cdot \Phi(d_2) \\ \frac{P_t}{\partial r} & = -K \cdot (T - t) \cdot e^{-r(T-t)} \cdot \Phi(-d_2) \end{cases} \tag{157}$$
>
> It is expressed as the change in option price for a 1 point (1%) change in interest rates.

Therefore, an option with a theo of 0.93 and a rho of 0.02 will increase in price to 0.95 if the interest rates increased from 3.5% to 4.5%.

The rho is presented as a change in option price per 1% change in rates. We also talk about the rho of a portfolio by multiplying it by the option multiplier and the number of options. Note that we don't input in a flat rate to every expiry date, but rather we parameterize the rate over time r_T and input the expected rho at time T into our calculation of $C_K(t, T)$.

Rho differs for stock options vs futures options, and we first define an intuitive way of thinking about it and deliver it with mathematical rigor. Interest rates essentially determine what you can do with money. If you buy an option, you are essentially short rho since you have tied up that money. So if interest rates increased, you can invest the cash at higher returns.

1. Borrowing money is short rho

2. Lending money is long rho

Option values with a change in rho depend on the product. The key point is that stock are settled right away. Futures require much less capital. Usually, stocks are worth much more than options, so if you buy/sell stock, you borrow/lend money and are short/long rho (assume floating rates). Same with options if you buy/sell them. For futures, however, if it is close to the forward price, then very little money changes hands to buy/sell them.

1. If you buy a stock call, you want to sell a stock as a hedge and lend the gained money. You are short rho on the call but the long rho position on the stock dominates, so you are long rho.

2. if you buy a stock put, you want to borrow money to buy a stock as a hedge. You are short rho on the put and short rho on the stock, so you are short rho.

3. For stocks, we can also interpret the difference in the PCP equation. For stock options, if rates go up, then the future price will go up and therefore $S_t - K$ will be greater, meaning that $C - P$ will be greater and thus C is higher and P is lower.

4. If you buy a future option, you want to either buy or short the future, which has negligible cost. Therefore, you are short rho almost always from the option.

The rho is largest for deep ITM options.

Theorem 12.9 (Rho Positions on Stock Options)

If interest rates rise,
1. Stock call prices will rise. Originally, the call price is

$$e^{-r(T-t)}\mathbb{E}[\max\{e^{r(T-t)}S_t - K, 0\}] = \mathbb{E}[\max\{S_t - Ke^{-r(T-t)}\}] \qquad (158)$$

Therefore, if interest rates increase $r \to r'$, then the strike would get discounted further, leading to a higher call price.
2. Stock put prices will fall. Originally, the put price is

$$e^{-r(T-t)}\mathbb{E}[\max\{K - e^{r(T-t)}S_t, 0\}] = \mathbb{E}[\max\{Ke^{-r(T-t)} - S_t\}] \qquad (159)$$

Therefore, if interest rates increase $r \to r'$, then the strike would get discounted further, leading to a lower put price.

Theorem 12.10 (Rho Positions on Futures Options)

If interest rates rise, futures options (both calls and puts) will fall. This is because the futures price will rise, but since the future is the underlying, this is a delta move. The rho now only accounts for the discount factor, which will increase, leading to a decrease in options.

Example 12.3 (Box Spread)

A box spread has no greek exposure except for rho. If we buy a box, we're lending money so we short rho (interest rates).

Rhos can be summed up across products and managed at the firm level, since interest rates are universal across all products.

13 Volatility

13.1 Realized Volatility

Given some spot price of the underlying S_t, we assume that the mean of S_{t+1} is the forward price.

$$\mathbb{E}[S_{t+1} \mid S_t] = F(t, t+1) = S_t e^r \qquad (160)$$

The standard deviation of this is the **volatility**. But to get the actual form of this distribution, we need some underlying assumptions. Now here comes the key assumption: Let's assume that the percent change in price is normally distributed, i.e.

$$\frac{S_{t+1} - S_t}{S_t} \approx \ln(S_{t+1}/S_t) \qquad (161)$$

is Gaussian.[18] Then, if we take a look at the log price $\ln(S_t)$ (constant random variable), then we can add the returns together to get $\ln(S_{t+1})$, which must be normal. This also means that S_{t+1} must be log-normal,

---

[18] Remember we did this before by taking the first order Taylor expansion.

and so the price distribution themselves must be log-normal.

$$\ln(S_{t+1}) \mid \ln(S_t) \sim N(\mu = \ln(S_t e^r), \sigma^2) \tag{162}$$

Since we can completely characterize the distribution, we can talk about the z-scores of this distribution. In fact, the z score is related to the volatility by the transformation

$$z = \frac{\ln(S_{t+1}) - r \ln(S_t)}{\sigma} \tag{163}$$

Figure 33: Assuming that we know the price at S_t, the S_T is a log normal distribution with the forward price as the mean. The volatility refers to the standard deviation of this distribution.

Definition 13.1 (Realized Vol)

There are two types of realized vol.
1. The **historical realized vol** is the annualized volatility measured with historical data. This is usually easy to compute with data, and can be computed on different time granularities (e.g. minute, hour, day, week, month, and then scaling it up by $\sqrt{n}$ where there are n periods in a year).
2. The **future realized vol** is the annualized future volatility. This is the *true* volatility of the future and is unknown to everyone.

Now just because the future realized volatility is unknown, all hope is not lost. Volatility tends to have some nice properties that allow us to decently predict future vol with historical vol.

Theorem 13.1 (Properties of Realized Vol)

Realized vol has a few properties.
1. It tends to be mean reverting, meaning that a temporarily high or low vol will tend to go back down or up.
2.

Another useful unit of measurement that is often confused with the volatility is the expected move.

> **Definition 13.2 (Expected Move)**
>
> The **expected move** of S_T is
>
> $$\mathbb{E}[|S_t - \mathbb{E}[S_T]|] = \sqrt{\frac{2}{\pi}}\sigma \approx 0.8\sigma \tag{164}$$
>
> Note that this is in absolute terms, and there is no squared term like the variance.

> **Example 13.1 (Cash Delta)**
>
> Let's look at a practical case when we have two portfolios A and B with the same cash delta of $10,000,000$. This looks like a lot of cash delta since a 1% change in the underlying can give you a $\$100,000$ difference in PnL! However, we find out that
> 1. Portfolio A has a vol of 48%, leading to a daily vol of 3% and therefore an expected move of 2.4%. This means that Portfolio A is expected to win or lose 240,000 in a day.
> 2. Portfolio B has a vol of 1.6%, leading to a daily vol of 0.1% and therefore an expected move of only 8 bps! This means that Portfolio B is expected to win or lose 80,000 in a day.
>
> Clearly, the same portfolios with the same cash delta has very different risks due to the historical vol of the underlying. Thus it is essential to consider the vol when managing risk.

13.2   Implied Volatility

To evaluate the Black Scholes formula, you really just need $K, T - t, S_t, r$, and σ. The first four arguments are easy to find and has no ambiguity. The final volatility is a bit more slippery because it turns out that if you take the historical volatility and plug it into the Black-Scholes equation, you get a value that isn't too close to the actual current price. This can be attributed to many things, one being that the Black-Scholes model is not correct. While this is certainly a possibility, another one is to delve a bit deeper on what this difference means. The function BS is monotonically increasing w.r.t. σ, and so it is invertible. So, rather than looking at the difference in the output pricing, we can look at the difference in the volatility.

> **Definition 13.3 (Implied Volatility)**
>
> The **implied volatility** of an option is the theoretical volatility σ of the underlying s.t. the Black-Scholes pricing model equals the current market price of the option. That is, taken S_t, K, T, r to be constant, it is the value of σ s.t.
> $$C_t = \mathcal{B}(K, T - t, S_t, r, \sigma) \tag{165}$$

Essentially, the implied vol is the public's general estimate of what the future realized vol will be. Therefore, the IV is the determinant of the price and the future realized vol is the determinant of the value. Therefore, the *premium*, which refers to the option's price, and the *implied volatility*, are often used interchangeably. Therefore traders often talk about the IV as the actual price of an option, and it is the case that

$$\text{Implied Vol} \iff \text{Price}$$
$$\text{Realized Vol} \iff \text{Value}$$

The implied volatility tends to be a better indicator of the future realized volatility than the historical volatility. Therefore,

1. if the implied volatility is greater than the historical volatility, this means that current prices are overestimating the volatility of the underlying. This overestimation of σ means that the option premium is overvalued, and we might likely see a drop in option prices.

$$V_K(t, T) = \mathcal{B}(K, T - t, S_t, r, \sigma) < C_t \tag{166}$$

2. if it is less, then current prices are underestimating it, meaning that the option is undervalued.

$$V_K(t,T) = \mathcal{B}(K, T - t, S_t, r, \sigma) > C_t \tag{167}$$

However, the IV is still far from perfect as an indicator of the future realized vol. In fact, it turns out that most of the time, the IV is greater than historical volatility, a phenomenon we will explain now.

Theorem 13.2 (Volatility Risk Premium)

The **volatility risk premium** refers to the following two phenomena, which are equivalent due to monotonicity of $\mathcal{B}$ w.r.t. σ.
1. Implied volatility tends to exceed realized volatility of the same underlying asset over time.
2. The market price C_t of the option tends to exceed the Black-Scholes pricing $\mathcal{B}(K, T-t, S_t, r, \hat{\sigma})$, where $\hat{\sigma}$ is the historical volatility.

The cause for this is still up for debate, but a study from Yale suggests that market participants tend to overestimate the probability of a significant crash, leading to higher demand for options. This heightened perception of risk may lead to higher willingness to pay for these options to hedge a portfolio, which is realized in the additional premium $C_t - \mathcal{B}(K, T - t, S_t, r, \hat{\sigma})$.

This is quite weird, but it gets even weirder. Let us have two options with the same expiry T, same underlying asset S_t, same interest rate (of course, since this is dependent on the environment), but different strike prices K^1 and K^2. The volatility should be the same between the two options since they have the same underlying. Therefore, the difference in their prices should be accounted purely from the difference in their strikes.

$$C_t^1 \neq C_t^2 \implies K^1 \neq K^2 \tag{168}$$

Great, so for a given moment in time t, let's take a look at all the options of a certain common underlying and their strike prices K. Their prices C_t^i should be different, and if we calculate the implied volatility for each strike K, we *should* get the same estimate of volatility.

Figure 34: What we expect our strike price vs implied volatility graph to look like for a collection of options.

Definition 13.4 (Volatility Smile)

Empirically, this is not the case, and we see a **volatility smile** centered around ATM. This means that ATM options have lower implied volatility than OTM or ITM options. Therefore, as the strike price moves away from the current stock price, the price of the options are in fact higher than what is accounted for in the Black-Scholes model, meaning that they are overvalued.

Figure 35: Volatility smile

However, this may not be symmetrical.

Definition 13.5 (Volatility Skew)

The volatility skew refers to the convex shape of the IV curve, centered near ATM. This means that ATM options have lowered implied vol than OTM or ITM options. Therefore, as the strike price moves away from the current spot, the price of the options are in fact higher than what is accounted for in the Black-Scholes model, meaning that they are overvalued. There are three general types of volatility skews.

1. *Normal/Put Skew.* In most cases, people are generally long the market because they buy stocks, ETFs, and indices. Therefore, the tendency to purchase OTM puts aas protection against sudden downside movement leads to greater demand for puts than for calls.
2. *Symmetric/Smile Skew.*
3. *Positive/Call Skew.* This can occur in stocks with a high chance of takeover or potentially lucky events that will cause profits to increase. For example, a pharm company can stay silent for a few years before making a new drug, at which the price will shoot up. There may be increased demand for OTM calls for companies such as these.

Exercise 13.1 (Talking with Friends)

What does the price of a call option go to as the volatility of the underlying goes to infinity? Describe it mathematically and heuristically.[a]

Solution 13.1

We can simply take this limit of the Black Scholes formula as $\sigma \to \infty$. Note that

$$\lim_{\sigma \to \infty} d_1 = +\infty \implies \Phi(d_1) = 1 \tag{169}$$

$$\lim_{\sigma \to \infty} d_2 = -\infty \implies \Phi(d_2) = 0 \tag{170}$$

and by plugging it in we get

$$\lim_{\sigma \to \infty} V_K(t, T) = S_t \tag{171}$$

---

[a]A discussion can be found here.

This means that it would converge onto the spot price of the stock. Let's now describe this more heuristically. It is clear that the value of a call option will increase as the upside of the option is greater if the stock is more volatile. The downside is always floored at zero so this does not change. One would assume that the value of the call option would go to infinity by considering the following case. If $S_0 = 100$ and has infinite variance, then since we are floored at 0, the expectation of $\max\{S_T, 0\}$ would go to infinity and therefore $\max\{S_T, 0\} - K$ would also tend to infinity. This means that the payout at time T is infinite, and the present value would also be infinite.

A better way to look at it is through arbitrage conditions. If the value $V_K(t, T) > S_t$, then you can always sell a call option, receive $V_K(t, T)$, and buy the stock at S_t. Now your portfolio is a short call, a long stock, and $V_K(t, T) - S_t > 0$ in cash. You invest the cash at the risk free rate to get

$$\left(V_K(t, T) - S_t \right) / Z(t, T) > 0 \tag{172}$$

During expiry, you will get $\min\{K, S_T\} > 0$ in cash more, and you have made free money.

Now rather than across strikes, let's fix a strike and talk about the IV across expiries. Sometimes, we may expect the IV to go up and other times go down depending on the circumstances, and the way we model this is called the *volatility term structure*. The last few periods will generally impact the IV levels of short-term options more than those further out.

Definition 13.6 (Volatility Term Structure)

1. If the long term historical vol is higher than the short term, then we would expect the vol to go up[a], creating a upward sloping vol structure. We would want to input incresing IVs into our models for further expiries.

2. If the long term historical vol is lower than the short term, then we would expect the vol to go down, creating a downward sloping vol structure. We would want to input decreasing IVs into our models for futher expiries.

13.3 Implied vs Future Realized Vol

We should clarify a couple relationships between the HRV, FRV, and IV. Both HRV and IV are used to predict the FRV, but IV is forward-looking while HRV is backward-looking. For example, if there is an big event such as an election happening in a week, IV will take that into account while HRV will not. So how does the market determine this IV? Recall gamma scalping vs theta decay.

Theorem 13.3 (IV is Breakeven Vol)

If the future realized volatility equals the IV, then the expected profits from gamma scalping will cancel out with theta decay, resulting in a net 0 profit. Therefore, the IV is also known as the **breakeven vol**.

Furthermore, the difference between IV and FRV is analogous to the difference between vega and gamma.

1. Long gamma is really a bet on FRV, since if more volatility actually realizes, then we are more profitable from gamma scalping, even if IV decreases.

2. Long vega is really a bet on IV, since if implied volatility increases, then we profit even if FRV decreases.

All three are correlated, yet distinct measures. The vega bets solely on the implied vol (not directly on the FRV) and the gamma bets solely on the FRV with respect to the IV. Let's look at scenarios where we would want to long one but not the other.

---

[a]We are assuming that volatility is mean reverting.

Example 13.2 (Gamma Scapling)

The gamma bets on the FRV with respect to the IV. By looking at whether the FRV is above (below) the IV, we can determine whether there is (isn't) enough volatility to justify gamma scalping.
1. If you're betting that FRV > IV, then you long gamma, which will give you more profits than bleeding theta.
2. If you're betting that IV > FRV, then you short gamma, where you reap theta.

Example 13.3 (Long Short Vega)

Say that there is a difference between the HRV and the IV. If IV > HRV, then the market thinks that the realized voltility will go up in relation to the historical. There are two scenarios.
1. HRV is mispriced.[a] We can't bet on HRV here, so we do nothing.
2. IV is mispriced.[b] We want to short this IV and thus short vega.

If HRV > IV, then the market thinks that the realized volatility will go down in relation to the historical. There are two scenarios.
1. HRV is mispriced.[c] We can't bet on HRV here, so we do nothing.
2. IV is mispriced.[d] We want to long this IV and thus long vega.

Every trader will hope that the implied vol will reevaluate asap to their vol target. It not only enables him to realize his profits more quickly, but it eliminates the risk of holding a position for an extended period of time. The longer a position is held, the greater the possibility of error.

It may seem a bit odd that gamma and vega are traded separately. In most cases, if you long gamma you also want to long vega. This is because since FRV is going up, the IV must follow the FRV and rise with it. Longing gamma but shorting vega must mean that you are betting that the market will be even more wrong about its estimation of the FRV, which may not make sense. However, if you long vega, then you may think that the IV in the future will surpass the FRV, meaning that you want to short gamma. These predictions also depends on which information you have, so let's consider a few cases where you would want opposing gamma/vega positions.

Example 13.4 (IV Crush)

An IV crush happens when the IV of an option drops abruptly. For example, leading up to an earnings announcement, options prices often rise due to increased uncertainty about the stock's future price movement. After the earnings release, uncertainty typically dissipate, triggering a decline in implied vol. Therefore, you want to short vega, but since the stock moves heavily, this is indeed realized vol and you want to long gamma.

In fact, buying an option gives us long gamma and long vega, so how would you construct a portfolio with opposing gamma and vega positions?

Example 13.5 (Opposite Gamma and Vega Positions)

We construct a diagonal (across strike and expiry) put spread.
1. Buy an ATM put with a week left to expiry. This has high positive gamma and little positive vega.
2. Sell a OTM put with a year left to expiry. This has low negative gamma and high negative vega.

---

[a] This can be due to elections or some other event that will bolster the volatility in the near future, but the HRV hasn't priced this in yet.

[b] The market may think that something big will happen, but it doesn't happen and volatility stays at consistent levels.

[c] We could have just gotten out of some event that caused a lot of volatility.

[d] The market may think that we are out of volatile times, but turns out we aren't.

> We have just created a long gamma short vega position.

13.4   Vol Edge

Since IV and prices are proportional, really they are one and the same.

Now let's talk about how we would determine the offsets of a quote. An offset of 0.05 on a 40-strike option may not carry the same risks and rewards as the 80-strike option. In fact, we can divide the two main factors to be

1. The **delta risk** refers to high deltas resulting in high directional risk.

2. The **vega risk** refers to high vegas. A small change in volatility can dramatically shift the price of the option, leading to large moves in our theo and therefore a potential loss.

Therefore, we should account for both delta and vega to standardize our offset. Clearly, an offset of 0.05 for an option with low vega and low delta is a much wider quote than that of an option with high vega and high delta. For now, let's focus on the delta position.

Definition 13.7 (Vega Edge)

The **vega edge** is defined

$$\text{vegaedge} = \frac{\text{edge}}{\text{vega}} \tag{173}$$

By converting edge to vega edge, we have done a change of basis from price to vol. This vega edge is in term of percentage, and so we can calculate everything in terms of vol.

Example 13.6 (Vega Edge Breakeven point)

Say that you have an option with a theo of 3.00 and you sell it at 3.10. The vega is 0.5, and so the vega edge is $0.10/0.5 = 0.20\%$, which is 20 bps. Therefore, if the implied vol is 21.50%, then you will lose this edge if the IV goes up to 21.70%.

Since the vega is highest in ATM options, we would expect our offsets to be highest. For ITM options, we have delta risk so we also quote high. For OTM options, however, it has low delta and low vega, so we can have smaller offsets. Ultimately, we would see the raw offsets increase as we go from OTM to ATM to ITM options.

13.5   Volatility Splines/Surfaces

Now we delve into how the actual theos are computed. This is really through our *volatility spline*.

Definition 13.8 (Volatility Spline)

Given a certain underlying, we pick about 11 deltas across the strikes, usually centered at ATM (50 delta) and 5 on each side $(\pm 0.2, \pm 0.4, \pm 0.45, \pm 0.48)$.[a] Let them be denoted $d_{-5}, \ldots, d_0, \ldots, d_5$. We take their corresponding strikes $K_{-5}, \ldots, K_0, \ldots, K_5$ and try to fit a **volatility spline**[b] predicting the implied volatilities IV_i. We can view the spline in multiple axis:
  1. *Strike Space.* The strikes to the right represent OTM calls. The strikes to the left represent OTM puts.
  2. *Log Moneyness Space.* This is similar to what we have done before.
  3. *Probability Space.* This is again analogous (but not exactly) to delta space, but we can view each node as the location where the probability of the option will finish past the strike.


Figure 36: Best cubic fit of our vol spline.

It is worth going through how this derivation is fit. The general idea is that between each points, we want to fit a cubic function s.t. the whole piecewise cubic function is C^2, i.e. has coinciding continuous 2nd derivatives. This, along with the fact that it has to interpolate the nodes, and finally the fact that the second derivative (concavity) at the end nodes must be 0, gives us a unique cubic spline that fits the points.

Definition 13.9 (Skew)

Depending on the shape of this spline, we can determine if there is a skew.
1. A **put skew** indicates that there is a right skew to the chart, i.e. the ATM vol is slightly to the left of the minimum. This indicates that there is a higher premium for put options.[a]
2. A **call skew** indicates that there is a left skew to the chart, i.e. the ATM vol is slightly to the right of the minimum. This indicates that there is a higher premium for call options.[b]
3. A **symmetric skew** indicates that there is a smile, i.e. the ATM vol is the minimum. This indicates that there isn't a preference for OTM calls or puts.[c]

The reason we want a vol spline is that it allows us to visually see where the market is predicting its theos. When we usually enter a new market, we want to make sure that our vol spline stays approximately in between the spreads of the market.

---

[a]Depending on the product, there may be more or less nodes on each side.
[b]More specifically, a natural cubic spline that is a C^2 piecewise function that interpolates the nodes.
[a]This usually happens in indices or stocks, since they usually go up slowly but crash down suddenly.
[b]This usually happens for commodities, where a sudden supply crash can send prices soaring.
[c]This usually happens for currencies, since slow/fast rises in one currency correspond to slow/fast falls in another, so it is symmetric by definition.

14 Second Order Greeks

14.1 Vanna

Definition 14.1 (Vanna)

Vanna is defined to be the derivative of the delta w.r.t. vol.

$$\text{vanna} := e^{-r(T-t)}\phi(d_1)\frac{d_2}{\sigma} = \frac{\nu}{S_t}\left[1 - \frac{d_1}{\sigma\sqrt{T-t}}\right] \tag{174}$$

In practicality, we use finite differences.
1. If we talk about it, we talk about the change of delta with a 1% increase in vol.
2. If we are tuning it on our vol spline, we refer to the **DdeldNode**, which is the change of delta with a 1 tick increase in vol (specified by the tick matrix).


Figure 37: Vanna

14.2 Weezu/Vomma

Definition 14.2 (Weezu/Vomma)

Weezu is the derivative of the vol w.r.t. vol.

$$\text{weezu} := S_t e^{-r(T-t)}\phi(d_1)\sqrt{T-t}\frac{d_1 d_2}{\sigma} = \nu\frac{d_1 d_2}{\sigma} \tag{175}$$

Raising weezu means making the vol spline more convex, i.e. raising the 30 70 nodes.

14.3 Charm

Definition 14.3 (Charm)

Charm is the derivative of the vol w.r.t. time.

$$\text{Charm} = \frac{\partial\Delta}{\partial t} = -\frac{\partial^2 V}{\partial S\partial t} \tag{176}$$