Money and Financial Markets: An Introduction*

Daniele Girardi†

These lecture notes introduce money and financial markets, in order to better understand the institutions that underpin the monetary and financial system in modern market economies. First, we will define money and discuss its functions. We will introduce the so-called 'broad money' aggregate and understand how money is created by bank lending. Second, we will introduce bonds and (nominal and real) interest rates. Third, we will discuss the Central Bank and its main functions: how it uses monetary policy to influence credit growth and economic activity; and how it helps prevent and mitigate financial crises. Fourth, we will introduce the stock market and briefly discuss what determines stock prices and their fluctuations.

1 Money

1.1 The functions of money

Unlike other terms you might encounter in your academic studies, 'money' is a very familiar word. Everyone uses money, everyday. Money's usefulness for society is pretty obvious. Without some form of money, trade and the division of labor would be severely constrained. Transactions would only take place through barter, and would thus be subject to the so-called 'double coincidence of wants'. If money did not exist, for example, a shoes maker would be able to swap shoes for food only if and when some farmer needed a new pair of shoes.

Yet, if you think about it, it is not so easy to define what money is in general terms. It is

*Lecture notes for the summer school module 'Institutions, Trade & Economic Growth '.

†Lecturer, Department of Political Economy, King's College London.

easier to make specific examples. For instance, coins and banknotes are certainly money: you can use them to buy things. This suggests that the most convenient way to define money is through the function it performs. *Money is an asset that serves simultaneously as a medium of exchange, a unit of account and a store of value.*

Let us look more closely at this definition. First of all, money is an asset: something that can be owned and that is expected to provide economic benefits. But this is not sufficient: not every asset is money. To be classified as money, an asset must play three specific social roles: it must be accepted in exchange for goods and services in the market (the *medium of exchange* role); it must be the measuring rod in terms of which the price of all goods and services is expressed (the *unit of account* role); it must be expected to retain its value – or at least a substantial part of it –in the foreseeable future (the *store of value* role).

There are strong complementarities between these three roles, meaning that it is efficient to have one single asset (money) performing all three of them simultaneously. The medium of exchange must be a good store of value, because if something loses its value very quickly with time, it will not generally be accepted as a means of payment (would you accept to receive your monthly wage in the form of perishable food?).[1] If one asset is universally adopted as the medium of exchange, it is natural to set the price of all items in terms of it; otherwise you would need to make a conversion from the unit of account to the medium of exchange everytime you considered an economic transaction.[2]

1.2 Money as a special IOU

In general, assets can either be financial or non-financial. Non-financial assets are able to produce goods or services for their owner. For example, industrial plants, works of art, houses and agricultural land are non-financial assets. Financial assets, instead, are just IOUs: claims on someone else in the economy.[3] In other words, financial assets represent debit-credit relations.

[1] This does not mean that all assets that are good stores of value are suitable as means of payment. Houses are generally good stores of value, but it would be extremely impractical to use them as means of payment.

[2] For example, imagine an economy in which the unit of account is bananas but the medium of exchange is gold. You consider buying a car, and find that it is worth 20,000 bananas. You and the seller would need to determine in some way how many ounces of gold a banana is worth, in order to understand how much gold you would have to pay for the car.

[3] An IOU (abbreviation for "I owe you") is a written document acknowledging a debt.

For example, a corporate bond is a financial asset: its holder has the right to receive a certain payment from the company that issued it. This means that a firm that issues bonds is in debt to its bondholders, while the bondholders hold a claim against the firm.

We have already stated that money is an asset. But is it a financial or a non-financial asset? The answer is that *money is a financial asset, an IOU*. Money represents a claim against its issuer. The holders of money hold a claim against the issuer of money; the issuer of money is in debt with the holders of money. Like all financial assets, thus, money represents a debit-credit relation. For example, every banknote represents a debt of the Central Bank to the holder of the banknote.

We have already mentioned the intimate relation between the two functions of medium of exchange and store of value. Everyone accepts money as a means of payment because they are very confident that it will retain some exchange value in the future. For instance, you accept money as a payment for your work because you are confident that you will be able to exchange it for goods and services in the future. *Confidence* is thus another fundamental ingredient in the definition of money: an IOU must enjoy strong confidence from the public, in order to serve as money. As economists at the Bank of England put it, "money in the modern economy is an IOU that everyone in the economy trusts" (McLeay, Radia, and Thomas, 2014, p. 7).

SUMMING UP: THE DEFINITION OF MONEY

Money is a financial asset that serves simultaneously as medium of exchange, unit of account and store of value. Like all financial assets, money represents a credit-debit relation (or IOU); but in order for a financial asset to work well as money, the public must have great confidence in its value.

1.3 Types of money

The neat definition of money that we have just presented may lead you to think that the dividing line between money and other financial assets is clear-cut, or even that there is only one type of money in the economy. In fact, there are several types of money, and the line between money and other financial assets is blurred. You should not think of ‘being money’ as a binary attribute – meaning that either an asset is money or it is not. Rather, you should think of it as an attribute that different financial assets display in different degrees. In principle, every financial asset could serve as money. But some assets are ‘more money’ (or, less awkwardly speaking, more *liquid*) than others. In this sense, there exists a hierarchy of money: some financial assets play the role of money better than others, because they are more readily and widely accepted as means of payment.

Given that money represents an IOU, everyone in principle could create money just by issuing an IOU. The problem is to get the public to have confidence in it, and thus accept it as means of payment. Suppose that I write an IOU to my favorite butcher, who has known me for years and therefore decides to accept it. She thus agrees to give me some meat in exchange for my personal IOU. However, there is little chance that she will be able to exchange my IOU with gas for her car at the gas station. And if, by some wild chance, also the gas station owner knew me very well and decided to accept my IOU, thus giving the butcher some gas in exchange for it, she will certainly not be able to spend it at the supermarket. Surely, after a short period, the butcher or the gas station owner will come asking that I repay my debt in some more liquid form of money – like banknotes, a check or a bank transfer.

To the contrary, my butcher will certainly accept to exchange meat with banknotes (which are IOUs issued by the Central Bank), and she will surely be able to exchange these banknotes for gas, and in turn the gas station owner will be able to exchange it for food at the supermarket, and so on. That’s why my personal IOU is extremely low in the hierarchy of money, while banknotes stand on the top. In other words, my IOU is so illiquid (has so low *acceptability*) that it can hardly be considered money, while banknotes are the most liquid form of money. As we will see presently, the ultimate reason for this is that banknotes and bank deposits are backed by the State, while my personal IOU is backed only by me.

1.4 The 'broad money' aggregate

At the very top in the hierarchy of money, there is government money. This takes two forms: currency (banknotes and coins) held by the public and Central Bank reserves held by commercial banks. Immediately below government money (actually almost at the same level), there are bank deposits – savings and checking accounts held by individuals and firms at commercial banks.

Government money and bank deposits are the most liquid forms of money. Although their forms and issuers are different, currency, Central Bank reserves and bank deposits are practically equivalent in terms of liquidity (or *acceptability*).

The aggregate sum of government money plus bank deposits is called 'broad money' and is currently labeled M4 in UK national statistics.[4] Following common practice, in the remainder of these notes we will use the term money to indicate the 'broad money' aggregate (M4 in the UK).[5] The broad money aggregate is our measure of the total amount of liquidity held by banks, households and firms. Let us now look more in detail at its components.

Currency First consider *currency*: banknotes and coins. These are issued by the Central Bank. In other words, they represent an IOU from the State, and in particular the Central Bank (in the UK, the Bank of England) to their holders. If you hold some currency, the Bank of England is in debt with you. More formally, banknotes and coins are a liability for the Bank of England and an asset to their holders.

But how does the Bank of England repay its debt with currency holders? In the past, in the UK as well as in many other countries, the Central Bank's debt was redeemable in gold; you could always go to the Bank of England and convert your banknotes in gold. So, back then, the Bank of England pledged to pay its debts in gold.

Today, in virtually all countries, currency is no longer convertible in gold, nor in any other commodity. Rather, the Central Bank honors its debt with money itself. You can go to the Central Bank and convert your banknotes in other banknotes of the same value, or in coins – but

[4] For more detail on what specifically is included in M4, see the following page in the Bank of England website `https://www.bankofengland.co.uk/statistics/details/further-details-about-m4-data`

[5] Strictly speaking, it is not 100% correct to identify 'broad money' with money: as we have seen, the distinction between money and other financial assets is not clear-cut. However, for practical purposes, it makes sense to refer to the M4 aggregate as simply 'money'.

not in gold. This kind of currency, that is not linked to any commodity, is called *fiat currency* or *fiat money.* Fiat money is irredeemable: it is only a claim on further fiat money. (You may now be asking: how can the value of fiat currency be maintained, if it is not guaranteed to be convertible in any useful commodity? We will answer this question in a couple of pages.)

Central Bank reserves Also *Central Bank reserves* are government money. These are reserve balances held by commercial banks at the Central Bank. Thus, in this case, the issuer (debtor) is again the Central Bank, while the holders (creditors) are commercial banks. Reserves are basically checking accounts that banks hold at the Bank of England (which is one of the reasons why the Central Bank is called 'the bank of the banks'). In practice, Central Bank reserves are just an electronic record of the amount owed by the Central bank to each individual bank.

Why would banks want to 'park' some money in their reserves at the Central Bank? First of all, holding reserves is convenient for banks for practical reasons. Everyday, banks make a great number of payments to one another. It would be extremely impractical to settle all these payments between banks using banknotes and coins. When Barclays needs to make a payment to Lloyds, it tells the Bank of England to transfer some of its reserves to Lloyds. Moreover, banks need to hold some currency in liquid form to meet deposit withdrawals. Again, it would be very cumbersome for banks to hold very large amounts of banknotes and coins for this purpose. The Central Bank guarantees that banks can swap any amount of their reserves for currency, everytime they want to. If many customers request to convert their deposits into banknotes, a commercial bank can swap some of its reserves for currency and repay those customers.

Holding some Central Bank reserves is thus necessary for commercial banks, to make payments to each other and to cover possible deposit withdrawals. In addition to this, in many countries, banks are obliged to hold Central Bank reserves by the law. The US Central Bank, the Federal Reserve, for example, requires each bank to hold an amount of reserves at least equal to 10% of their deposits. The Bank of England, instead, does not impose any reserve requirements on British banks.[6]

[6] In countries like the US where the Central Bank imposes some reserve requirements, the reserves that a bank hold above this legally required level are called *excess reserves.* For example, if a US bank has 100 million dollars in customers' deposits and holds 15 million dollars in reserves, we say that this bank has 10 million dollars of

Base money (or government money or narrow money)

	FIAT CURRENCY	CENTRAL BANK RESERVES	BANK DEPOSITS
Form	Banknotes and coins.	An electronic record of the amount owned by the C.B. to each individual bank.	An electronic record of the amount owned by a private bank to each deposit holder.
Issuer (Debtor)	Central Bank.	Central Bank.	Private banks.
Holder (Creditor)	People holding currency.	Private banks.	People and firms holding deposits.

Broad money

Table 1: The different types of money in the economy

Bank deposits The third component of 'broad money' consists of *bank deposits* – like checking accounts and savings accounts. Differently from fiat currency and Central Bank reserves, which are issued by the State, bank deposits are issued by commercial banks. So they represent a debt relation between commercial banks and their depositors: banks owe currency to depositors.

Table 1 summarizes the characteristics of the components of the 'broad money' aggregate.

1.5 Where does confidence in money come from?

We have mentioned that *confidence* is a necessary ingredient of money: to be accepted as money, a financial asset must be trusted by the public. Why are we so sure that banknotes, coins and checks will always be accepted as valuable means of payment?

Ultimately, confidence in government money and bank deposits arises from the backing of

required reserves and 5 million dollars of excess reserves.

the State. The State underpins these financial assets, allowing them to be the most liquid form of money, by accepting them for tax payments. In other words, currency and bank deposits are at the top of the hierarchy of money because the State guarantees their value, by pledging to accept them as tax payments. For this reason, we can be confident that these forms of money will always be in demand, which implies that they will always retain some exchange value. This allows banknotes, coins, checks, wire transfers and debit cards to be universally accepted as means of payment.

Summing up: the hierarchy of money

Money is an IOU, and any IOU could in principle serve as money. However, different IOUs have very different degrees of acceptability (or liquidity), so there is a hierarchy of money, with more liquid assets at the top. We usually call money only those financial assets that stand at the top of the hierarchy of money: fiat currency, Central Bank reserves and bank deposits. They are the most liquid forms of money because they are backed by the State, that guarantees that it will accept them as tax payment.

1.6 Money creation

Bank deposits represent the vast majority of the broad money holdings of individuals and firms. In the United Kingdom, as of July 1st 2024, there were $86,894$ million pounds of currency (banknotes and coins) in circulation and $2,221,113$ million pounds in checking and savings accounts and the like.[7] In other words, households and firms kept over 96% of their cash in the form of bank accounts, and less than 4% in the form of currency. And indeed this average is probably very much influenced by the existence of criminal organizations engaging in illegal activities, which keep most of their money in banknotes. So on average households and licit firms actually keep even more than 96% of their cash in bank accounts.

The amount of deposits held in a single individual bank depends on the choices of individuals

[7] You can download the most recent available data, as well as historical series, on aggregate currency and bank deposits at the Bank of England Website at the following link `https://www.bankofengland.co.uk/boeapps/database/index.asp?first=yes&SectionRequired=A&HideNums=-1&ExtraInfo=false&Travel=NIxSTx`

and firms, which decide how much liquidity to hold in an account at that specific bank. But at the aggregate level, it is bank lending that creates deposits. When a bank makes a loan, it does not give to the borrower a bag full of banknotes and coins. Rather, it credits the bank account of the borrower with an amount of money equal to the size of the loan, thus increasing the quantity of bank deposits (and therefore the overall quantity of money) in the system.

Imagine, for example, that you obtain a £50,000 home mortgage from Bank A. Bank A will credit your savings account with £50,000. You will then transfer this money to the bank account of the house seller (which may be held at Bank A or at another bank). As a byproduct of this process, £50,000 of new money, in the form of bank deposits, have been created (in other terms, the broad money aggregate has increased by £50,000), just through a keystroke of the bank employee. *Whenever a bank concedes a loan, it simultaneously creates a corresponding deposit, thereby increasing the quantity of broad money in circulation.* This is how money is 'created' in our economy.

The process of bank lending described above does not entail any *necessary* change in the quantity of Central Bank reserves held by banks. Faced with an increase in deposits, a bank may or may not decide to demand more Central Bank reserves. It may do so if it feels that it needs more liquidity to face possible withdrawals by its depositors.[8] If the increase in the stock of deposits causes banks to want to hold more reserves, then they will simply demand more reserves to the Central Bank. The Central Bank normally supplies all the reserves that private banks demand, and asks for some asset in exchange. So the banks will swap some asset that they own (for example Government bonds) with some Central Bank reserves. If this is the case, also the quantity of bank reserves (also called base money) will have increased as a result of the increase in bank loans.

This discussion should have made it clear that *banks are not just intermediaries that channel pre-existing funds from savers to investors.* According to a common but inaccurate view (sometimes called 'loanable funds theory'), banks simply lend out the money that depositors entrust to them. This might be a reasonable (although simplified) description of what happens at the

[8]In countries where a reserve requirement ratio is imposed, the bank might need to increase its quantity of Central Bank reserves if the increase in deposits makes the reserve/deposits ratio fall below the level required by law. This is, however, not the case in the UK.

level of an individual bank, but not at the level of the entire banking system. At the aggregate level, the amount of money that banks can lend is not constrained by the amount of deposits that have been previously opened. As we have seen, causality runs in the opposite direction: it is bank lending that creates deposits. At the aggregate level, people deposit what banks have lent to them.

Moreover, it should now be clear that *the amount of money in circulation is not fixed by the Central Bank*, but depends on the aggregate behavior of private banks and their customers. In other words, the Central Bank does not fully control the money supply. You might have encountered simplified economic models that assume that the Central Bank simply fixes the quantity of money in the system at its discretion. This can be a helpful modelling simplification when studying some problems, but it is not an accurate description of the monetary system. While the Central Bank does exert some influence on the volume of credit, and therefore the quantity of broad money, through its monetary policy, it does not fully control it.

Limits to money creation by banks The fact that private banks have the ability to create money through their lending activity does not mean that they can do so without any limit. In practice, there are several factors that limit the amount of broad money that banks can create.

First, banks are constrained in their ability to create money by the need to remain profitable. They can increase outstanding loans only to the extent that there are profitable lending opportunities. Put in a different way, banks cannot make too many risky loans without risking bankruptcy. There is only a certain number of trustworthy borrowers that a bank can lend to.

Second, the volume of new loans depends on demand for loans by households and firms. In turn, demand for new loans depends mainly on two factors: the expected rate of growth of the economy and the ongoing interest rate. The interest rate, in turn, is determined by the monetary policy enacted by the Central Bank (as we will see more in detail). The economic expectations of the public and the Central Bank's monetary policy are thus able to influence and constrain the volume of credit and money creation in a decisive way.

Third, the behavior of money holders also influences the volume of credit and the process of money creation. By repaying previous loans, firms and households 'destroy' money (i.e., they reduce the quantity of broad money in circulation). No matter how much banks increase their lending, if people use that money to repay outstanding loans the newly created money is immediately destroyed.

It is thus the demand for new loans by trustworthy borrowers at the given interest rate that determines the quantity of money that the private banking system can create. Demand for credit is decreasing in the interest rate, meaning that a lower interest rate makes firms and households want to borrow more, while a higher interest rate makes them want to borrow less. The relation between the interest rate and credit is negative because a higher interest rate makes borrowing more expensive.

Summing up: the money creation process

In modern economies, money is created by commercial banks through their lending activity. Whenever a bank concedes a loan, it simultaneously creates a corresponding deposit, thereby increasing the quantity of money in circulation. This means that two common notions are wrong: (1) banks do not lend out deposits; to the contrary, it is loans that create deposits; (2) the Central Bank does not control the money supply; it is the aggregate behavior of the commercial banking system that determines the amount of money in the economy. This does not mean that banks can create unlimited amounts of money at their will: the amount of loans that can be extended depends on (a) the existence of profitable lending opportunities; (b) the willingness of households and firms to take up debts; (c) the rate of repayment of previous loans. The Central Bank influences credit (and thus money) creation by setting the interest rate: lower interest rates stimulate lending; higher interest rates discourage lending.

2 Bonds and interest rates

2.1 Bonds

Households obtain funds to finance their housing investment or consumption mostly through bank loans. Firms, instead, have two other important ways to obtain external funds to finance their investment: selling bonds and selling stocks. In this Section we focus on bonds. Bonds are also used by Governments to borrow funds from the public.

A bond is essentially a loan that the bond buyer extends to the bond issuer. The bond buyer acquires the right to receive a certain fixed payment from the bond issuer at a given point in the future. For instance, when you buy a bond from a firm, you pay some money to the firm now, and in exchange you receive a promise to be paid a certain (higher) sum at a certain point in the future. The bond thus represents a promise of a future fixed payment.

Imagine, for example, that Microsoft decides to raise some money to fund its investments by issuing one-year bonds that pay £110 each. If you buy one of these bonds, you acquire the right to be paid £110 by Microsoft in one year. The relation between the price you pay for the bond and the amount that the bond will pay to you is directly related to the interest rate on this bond. For example, imagine that the price of this Microsoft bond is £100. This means that the rate of return that you get by buying the bond today and holding it until expiration is equal to $(110 - 100)/100 = 0.10 = 10\%$. The interest rate is 10% per year. Conversely, if we knew that the interest rate on these one-year Microsoft bonds was 10% per year and that they guaranteed a £110 payment, we could infer that their price must be $110/(1 + 0.10) = 100$.

In general, if we indicate the price of a bond as P_B and the amount that the bond will pay in one year as x, the interest rate per year is given by the formula

$$i = \frac{x - P_B}{P_B} \tag{1}$$

Rearranging equation 1 we can figure out that the price of the bond must be

$$P_B = \frac{x}{1 + i} \tag{2}$$

The price of a one-year bond at the moment it is issued is equal to the final payment divided by 1 plus the interest rate. One important implication of these formulas is that there is an inverse relation between the price of the bond and the interest rate: *the higher the interest rate, the lower the price of the bond*. When you read on the newspaper that 'bonds prices have decreased', that means that the interest rate on bonds has increased (and vice versa).

Bonds can be sold before their expiration. For example, you can buy a one-year Microsoft bond and then sell it after eight months, for a certain price agreed with the buyer. The buyer will purchase the bond from you, and she will thus be entitled to receive the 110 dollars after four months. A bond can pass from many hands before it expires. The market for bonds that have already been issued is called *secondary market*. If you buy a bond from the bond issuer at the moment of issuance, you are buying in the *primary market*. If you purchase it from another investor that has held it for some time, you are buying in the *secondary market*.

Are bonds money? Just like money, bonds represent an IOU, are priced in the unit of account, and are good stores of value. In principle, they could be considered a type of money. However, bonds are much less *liquid* (in other words, they enjoy less *acceptability*) than banknotes, coins and bank accounts. They are not universally accepted as means of payment. A supermarket will certainly accept banknotes or a debit card linked to a checking account as a payment for its articles; but it will not accept a bond. Bonds must be exchanged for more liquid forms of money before they can be used as means of payment. And the conversion of a bond in currency is not guaranteed: it may take time to find a buyer for your bond, and the price that you will be able to receive may be very uncertain. In some cases, it may even prove impossible to find someone willing to buy a given bond at a positive price. For these reasons, bonds are not considered money (they are not part of the *broad money* aggregate)

Note that some bonds are more liquid than others. For example, UK Government bonds (through which the Government borrows money from the public) are extremely liquid. The market for *gilts* – as UK government bonds are also called – is so large that you can be virtually certain to be able to promptly exchange UK Government bonds for currency at any moment (although at a market price that is not known in advance). This is even more true of US government bonds, which are the most liquid bonds globally and whose market is just huge. These bonds are thus much nearer than other bonds to fulfil the definition of money. However, conventionally, not even UK or US government bonds are counted as money.

2.2 Nominal and real interest rate

The interest rate on one-year UK government bonds (or *gilts*) was approximately 11% in January 1975.[9] For each pound you invested in one-year gilts in January 1975, you would receive around 1.11 pounds after one year. In January 2004, instead, the interest rate on one-year gilts was (approximately) a much lower 4%. For every pound invested in one-year gilts in January 2004, you would only get around 1.04 pounds after one year.

Given that lending to the UK Government was equally risk-less in 1975 and 2004 (after all, the British government has never missed a single payment on its gilts in its entire history), it might seem that buying gilts was much more profitable in 1975 than in 2004.

Yearly inflation, however, was as high as 23.4% in the course of 1975. In 2004 it was much lower: consumer prices increased only around 1.7% in the course of 2004.[10] This information is very relevant, because the purchasing power of the money that the bondholder receives from the bond-issuer at the end of the year depends on the rate of inflation. If you invested 100 pounds in one-year gilts in January 1975, you would receive approximately 111 pounds at the end of the year. But in the meanwhile the purchasing power of a pound would have decreased by more than 23% because of inflation. If you made the same 100-pounds investment in one-year gilts in January 2004, you would get only around 104 dollars at the end of the year, but the purchasing power of dollars would have decreased by only 1.7% in the meanwhile. After taking into account inflation, it looks like one-year gilts were a better investment in 2004 than in 1975!

This example makes it clear that in order to calculate the *real* interest rate on a bond, it is

[9] As explained earlier, these are bonds through which the UK government borrows money from the private sector. Historical data on one-year gilt yields are surprisingly hard to find. In order to construct a series, here I'm using the annualized three-month Treasury bill rate (downloaded from the IMF Financial Statistics Database) as a proxy. While this is not perfect, it is a decent approximation to use in this example (the two are unlikely to diverge hugely).

[10] There are different ways of measuring inflation, using different price indices, and they produce different numbers. So you might see different inflation estimates for these two years. However, they all suggest that inflation was much higher in 1975 than in 2004, on a scale similar to that implied by the numbers reported here.

necessary to correct for inflation. A higher 'nominal' interest rate means that you get a higher nominal profit from your bond. But the *real* value of this profit, in terms of purchasing power, is influenced by the rate of inflation.

To be more precise, we can define the *nominal interest rate* as the interest rate expressed in terms of current pounds (or dollars, or euros, or whatever currency the bond is denominated in); the *real interest rate*, instead, is the interest rate expressed in terms of real purchasing power. The real interest rate is approximately equal to the nominal interest rate *minus* the rate of inflation. Indicating the real interest rate as r, the nominal interest rate as i and the inflation rate as π, we have[11]

$$r_t \approx i_t - \pi_{t+1} \tag{3}$$

This important formula is called 'Fisher equation' (after the early 20th century economist Irving Fisher). Intuitively, it says that the real rate of return on a bond is equal to the amount of purchasing power that you gain (by receiving more dollars than you initially invested), minus the purchasing power that you lose (because of inflation). The sign $\approx$ means that the real interest rate is *approximately* (but not exactly) equal to the interest rate minus the inflation rate. You can find a formal derivation of the Fisher equation – that is, a demonstration that equation 3 is true – in Appendix A

Note that the relevant nominal interest rate is that of period t: this is the period when the bond is issued. It is the nominal interest rate at the moment of issuance that determines the price of the bond. The relevant inflation rate, instead, represents the growth rate of the price level between the period of issuance and the period of expiration of the bond. This determines the extent to which money loses value during the period in which the bond is held. We call this the inflation rate at time $t+1$ (indicated as π_{t+1}).

The formula in equation 3 raises a question: when investors buy bonds (in period t) they are interested in the real interest rate that the bond pays; but how on earth can they know it, if it depends on the future inflation rate (the inflation rate at period $t+1$), that they cannot observe?

[11] Remember that the inflation rate is the rate of increase of the price level. Formally, $\pi_t = (P_t - P_{t-1})/P_{t-1}$, where P_t is the price level, measured through either the GDP deflator or the Consumer Price Index.

Indeed in period t investors can observe the nominal interest rate i_t, but not the future inflation rate π_{t+1}. All they can do is try to forecast the inflation rate. They thus plug in the formula the inflation rate that they *expect* to materialize at time $t+1$. We can adjust equation 3 to take this fact into account: from the point of view of investors, the real interest rate is determined by the nominal interest rate minus the *expected* inflation rate. Let us denote the expected inflation rate as π^e_{t+1}. The formula thus becomes as follows

$$r_t \approx i_t - \pi^e_{t+1} \tag{4}$$

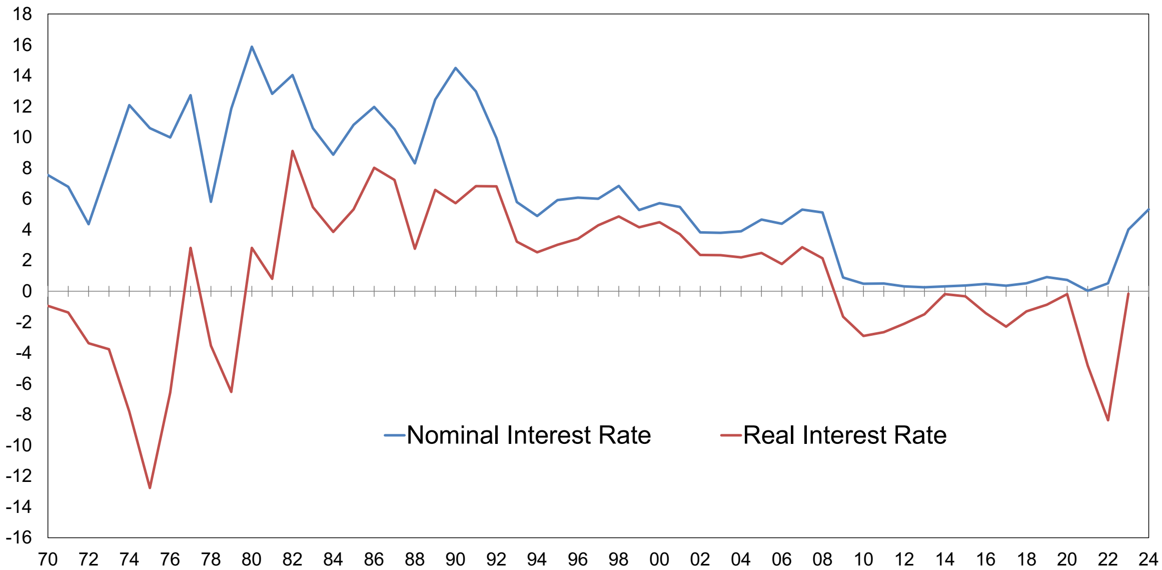
The real interest rate formulas in equations 3 and 4 are very important: make sure you remember them and understand the difference between them. Equation 3 describes the *realized* real interest rate (also called 'ex-post real interest rate'). The realized real interest rate is the real rate of return that investors will actually earn on the bond at its expiration. Equation 4, instead, describes the *ex-ante* real interest rate. Ex-ante means 'before the fact', and indeed the ex-ante real interest rate is the real interest rate as calculated before the rate of inflation is known with certainty.

The ex-ante real interest rate is thus the real interest rate that investors expect to earn when they buy the bond, while the realized interest rate is the one that they actually get. Only if investors' expectations turn out to be correct, the ex-ante and the realized real interest rates will be equal.

Nominal and real interest rates in the United Kingdom Figure 1 displays nominal and real short-term interest rates on UK government bonds between 1970 and 2024. Interest rates are measured in January of each year, and the real interest rate is computed ex-post, as the difference between the nominal interest rate and a measure of actual realized inflation over the year. Note that there is no data point for the ex-post real interest rate for 2024, as this would require data on inflation over the course of 2024, which is still unknown at the time of writing.

While nominal interest rates were pretty high in the 1970s, inflation was even higher: the real interest rate was *negative* for most of the decade. In the 1980s nominal interest rates were

Figure 1: Short-term interest rate on UK government bonds



Source: Computed based on interest rates data from IMF International Financial Statistics and inflation data from OECD Statistics

even higher, while inflation slowed down, resulting in a much higher real interest rate. Real interest rates remained positive (although lower than in the 1980s) in the 1990s and until the 2008-09 Great Recession. After 2008, real interest rates on short-term government bonds went negative again, this time not because of high inflation (to the contrary, inflation remained low and stable between 1992 and 2021) but because of a low nominal rate, as the Central Bank was trying to boost economic activity amid a slow recovery. Inflation and nominal interest rates both increased again in 2022-2023.

Consider again our previous example about interest rates in 1975 and 2004. Figure 1 shows that, although the *nominal* interest rate was much higher in 1975 than in 2004, the opposite is true of the *real* interest rate. Correcting for inflation can be important!

2.3 The interest rate and the premium for risk

What determines the real interest rate on a given bond? The answer is that the real interest rate on a particular bond depends on a mix of economy-wide factors and factors related to the specific characteristics of the bond.

The specific characteristics of a bond which matter for the determination of its interest rate are *maturity* and *riskiness*.

The maturity of a bond is the length of time over which the bond is repaid. Bonds can have different maturity. A one-year bond is fully repaid after one year, while a 10-years bond is fully repaid after 10 years.

In normal times, other things being equal, investors require a higher average annual interest rate to purchase bonds with longer maturity. This is because they need to be compensated for locking in their money for a longer period of time, and because the real rate of return on longer maturity bonds is more uncertain, as inflation in the distant future is harder to forecast.

Bonds also differ in terms of their riskiness, in particular their *default risk*. Default risk arises because the bond issuer may not actually be able (or willing) to pay the due amount when the bond reaches maturity. As always happens when money is lent, the lender (the bondholder) bears some risk because the borrower (the bond issuer) may not be able or willing to pay back. Some bonds (for example UK or US government bonds) are seen as virtually risk-less because it is considered certain that the issuer will honour its obligations. Other bonds, instead, are risky, because there is a non-negligible probability that the issuer will not pay.

Let us now set aside the role of maturity, in order to focus only on default risk. In the following discussion, we thus assume for simplicity that all bonds have the same maturity (say one year) and focus on the role of default risk in producing different interest rates on different bonds.

We can see the interest rate on a given bond as the sum of two components: the risk-less interest rate plus a 'risk premium'. The risk-less interest rate is the interest rate that would be required by investors on bonds that are sure to be repaid. It is the same in the whole economy and it depends on the monetary policy of the Central Bank. As we will see in the next section, the Central Bank sets the risk-less interest rate on the basis of economy-wide factors. The risk premium, instead, is different for each bond and depends on its default risk. The interest rate on a given bond is thus equal to the risk-less interest rate plus a bond-specific risk premium.

Let us indicate the risk-less interest rate as i, the interest rate on a risky bond as i_R and the bond-specific risk premium as x_R. The interest rate on a risky bond is then equal to:

$$i_R = i + x_R \tag{5}$$

The risk premium on a risky bond (x_R) is determined by two factors:

1. The first factor is the *probability of default*: the probability that the bond issuer will not pay the due amount when the bond expires. *The higher the probability of default, the higher the risk-premium, and so the higher the interest rate.*

   The explanation for that is fairly simple. If a very risky bond paid the same interest as a less risky one, no one would buy it! The interest rate on a riskier bond must be higher to compensate investors for a higher risk of default.

   It is possible to quantify the minimum risk premium implied by a given probability of default. Let us denote the probability of default as p. For example $p = 0.06$ means that there is a 6% probability that the bond will not be repaid. For a risky bond to provide the same *expected* rate of return as a risk-less one, after accounting for the probability of default, it must pay the following risk premium:

$$x_R = (1+i)\frac{p}{1-p} \tag{6}$$

   The risk premium in equation 6 is the one that *just compensates* investors for the probability of default, in the sense that it makes the expected rate of return on the risky bond equal to that on a riskless one.

   For example, consider a bond that has a probability of default of 2%, in an economy where the Central Bank has set the riskless interest rate at 3%. The risk premium required for this bond to give the same expected rate of return as a riskless bond is $x_R = 1.03\frac{0.02}{0.98} = 0.021 = 2.1\%$.

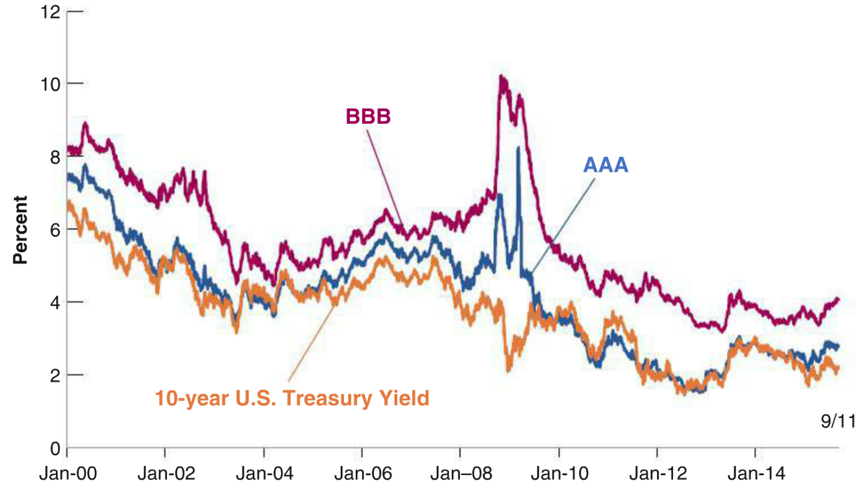
   The derivation of equation 6 can be found in Appendix B.

2. The second factor is the *degree of risk aversion* of bond holders. Even if the expected return on the risky bond was the same as on the riskless bond, the risk itself would make them reluctant to hold the risky bond. In general, people are risk-averse. This means that they tend to prefer a bond which gives 100$ with certainty, rather than a bond with an uncertain return with an expected value of 100$. Thus, they will ask for an even higher risk premium on the risky bond, with respect to the one that just compensates for the probability of default (given by equation 6). How higher will depend on their degree of risk aversion. If investors were not risk averse at all, the expected return on a bond would just need to be the same as the expected return on a riskless bond (as in formula 6). With some risk aversion, the expected return on the risky bond must be higher. This means that if for some reason investors become more risk averse, the risk premium will go up (and so the interest rate on risky bonds will go up), even if the probability of default has not changed.

Risk premiums in actual economies Figure 2 plots the interest rates on three different types of bonds from January 2000 to September 2014 in the United States (the figure would look broadly similar in the UK). The yellow line refers to US Government bonds, which are considered practically risk-less. The blue and red lines refer to corporate bonds rated respectively as safe (AAA bonds) and less safe (BBB bonds) by ratings agencies.

Figure 2 provides three main insights. First, even the bonds of the safest firms (AAA corporate bonds) pay an higher interest rate with respect to Government bonds. The US Government can normally borrow at cheaper rates than US corporations. Second, the rate on lower rated (BBB) corporate bonds is higher than the rate on the most highly rated (AAA) corporate bonds by a premium that has usually been around 5% in the period depicted. Third, look at what happened when the financial crisis exploded in 2008. Although the rate on Government bonds decreased, reflecting the decision of the US Central Bank (the Federal Reserve) to slash the policy rate, the interest rate on lower-rated bonds increased sharply, reaching 10% at the height of the crisis. In other words, despite the fact that the Fed was lowering the policy rate (and thus the riskless interest rate was decreasing), the rate at which lower rated firms could borrow became much higher. This made it difficult for firms to borrow, and thus to invest. The increase in the

Figure 2: Yields on 10-year Government bonds, AAA corporate bonds and BBB corporate bonds in the USA (Jan 2000 - Sep 2014)



Source: Blanchard, 2016

interest rate on corporate bonds during the crisis was due to the fact that, while the Fed had decreased the riskless interest rate, the risk premium had increased greatly. In normal times, however, the average risk premium is stable and the different interest rates tend to move in step, so that when the Fed decreases the riskless rate, also all the other interest rates decrease, and it becomes cheaper for firms to borrow.

Summing up: bonds

A bond is a financial asset that guarantees a fixed payment after a certain period of time. Firms, Governments and other organizations issue bonds as a way to borrow money. The price of a bond is inversely related to the interest rate on that bond: the higher the interest rate, the lower the price of the bond.

What is really relevant for investors is not the nominal but the real interest rate. The real interest rate is the interest rate in terms of purchasing power. The real interest rate on a bond is (approximately) equal to the nominal interest rate minus the inflation rate.

The interest rate on a given bond can be seen as the sum of the risk-less interest rate plus a risk premium. The risk-less interest interest rate is set by the Central Bank on the basis of economy-wide factors. The risk premium depends on bond-specific characteristics (in particular the probability of default) and on the degree of risk-aversion of bond-holders. The higher the probability of default, the higher the risk premium. The higher the risk aversion of investors, the higher the risk premium.

Throughout this section, we have maintained that the risk-less interest rate is determined by the monetary policy of the Central Bank. The next section explains what the Central Bank does– in particular how and why it controls the risk-less interest rate.

3 The Central Bank and monetary policy

This section provides a brief systematic overview of the functioning of the Central Bank and monetary policy, with special reference to the United Kingdom.

3.1 The Central Bank balance sheet

We start by looking at (a simplified version of) the balance sheet of the Central Bank. By analyzing the elements in the Central Bank balance sheet, we can get a more precise idea of its functioning. Indeed each component of the Central Bank's balance sheet plays a significant role in the functioning of the Central Bank and of the whole economy.

In simplified fashion, the balance sheet of a Central Bank looks as in Figure 3. The *liabilities* are the debts of the Central Bank. In other words, they are claims held by economic actors (banks, families, foreigners...) against the Central Bank. The *assets* are credits that the Central Bank holds. They are claims held by the Central Bank on other economic actors. The overall amount of liabilities of the Central Bank must always and necessarily equal the amount of its assets, thanks to a residual item called 'capital' that bridges the differences between the two.

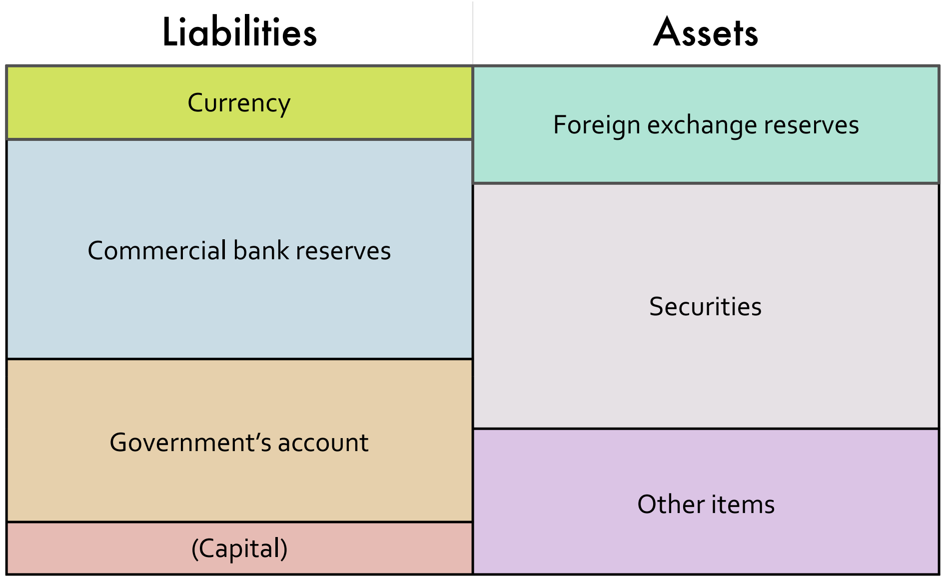


Figure 3: Stylised Central Bank balance sheet

3.1.1 The liabilities of the Central Bank

Currency The banknotes and coins in circulation (held by households in their wallets and in their homes, and by commercial banks in vaults or ATMs) are a liability for the Central Bank: as explained in Section 1, they represent debts of the Central Bank.

Banknotes and coins enter in circulation through commercial banks. Households obtain banknotes from banks by making withdrawals from their own bank accounts. In turn, commercial banks obtain banknotes from the Central Bank in exchange for their reserves. As we have seen in Section 1, the Central Bank guarantees that banks can swap any amount of their reserves for currency, any time they want to.

Commercial banks' reserves We have already discussed commercial banks' reserves extensively in Section 1. To summarize: Reserves are balances that commercial banks hold in their accounts at the Central Bank. As such, they represent claims on the Central Bank, held by commercial banks. Reserves are (together with banknotes) the most liquid asset in the economy. They also represent, together with banknotes and bank accounts, one of the components of the broad money aggregate. Reserves can be seen as current account balances held by commercial banks at the Central Bank, in pretty much the same way in which checking accounts are held by households and firms at commercial banks.

As you will remember from Section 1, reserves are used to settle payments between banks. When HSBC needs to make a payment to NatWest, it transfers some of its Central Bank reserves to NatWest.

To smooth these transactions, banks routinely lend reserves to one another overnight ('overnight' meaning that these loans are very short-term). The interest rate on these loans is called *overnight inter-bank interest rate*. In the UK it is also called Sterling Overnight Index Average, or SONIA.[12] Given that lending between banks overnight is considered really safe in normal times (the exception being a financial crisis so severe that some banks might go bankrupt by the end of the day), SONIA is seen as a good measure of the risk-less interest rate, just like short-term government bonds.

Just like a commercial bank can pay interest on its customers' saving and checking accounts, the Bank of England pays interest on the reserves held by commercial banks. The interest rate that the Bank of England pays on reserves is called the Bank Rate, and, as we will see in the next section, is a key instrument of monetary policy.

Government's account The Central Bank acts as the banker of the Government. For example, the UK Government (specifically the Treasury Department) has its account at the Bank of England, called the Consolidated Fund. When you pay your UK taxes, you send the money to HM Revenue and Customs (HMRC), which deposits it into the Consolidated Fund. When the Government spends money, it pays it out of the Consolidated Fund. In the UK, payments from the government bank account must be authorised in advance by Parliament. The government submits a Consolidated Fund Bill – a request to spend some of the money in the Consolidated Fund – to the House of Commons, which then votes on it.

If over time the Government consistently spends more than its revenues, the Government account at the Central Bank can get empty. In this case, typically, the Central Bank will

[12] To learn more, you can consult the Bank of England webpage on SONIA: `https://www.bankofengland.co.uk/markets/sonia-benchmark`.

lend money to the Government to fill the gap. So the Government balance at the Central Bank can become negative, but the Government can still pay its expenses with loans from the Central Bank. The practice of having the Central Bank directly finance Government deficits is called 'monetization' of Government debt. When this happens, (a part of) the Government deficit is financed by newly created money instead of borrowing existing money from households. Monetization of Government debt has been common practice among industrialized economies in the Post-World War II period (broadly speaking, until the 1980s). However, today advanced economies tend to avoid it (in some countries it is even banned by law), in fear that it could be abused by the Government, leading to overheating of the economy and very high inflation.

Capital Capital is a residual item in the balance sheet: it is just the difference between the total value of assets and liabilities. (Capital = Assets − Liabilities.)

For a firm or a commercial bank, this magnitude is very important, because it represents the net worth of the firm – the value that the firm would remain with after paying all its liabilities. A firm or a commercial bank having negative capital – meaning that its liabilities are worth more than its assets – is a big problem. It means that, even if it sold all its assets at their market value, it would not be able to repay its debts. To avoid the risk of failure, a firm or commercial bank should maintain a high enough level of capital, so that some losses could be absorbed without immediately resulting in a negative net worth. For firms and commercial banks, the amount of capital is thus a very meaningful magnitude. For Central Banks, things are different. Even if a Central Bank's capital decreased to the point of becoming negative, there would be no risk of default. The Central Bank cannot default on its debts, given that it can repay all its debts by issuing currency. So, differently from firms and commercial banks, the amount of capital held by the Central Bank is generally not considered a very important variable.

The presence of capital in the balance sheet of the Central Bank (and of any financial or non-financial firm, for that matter) explains why the total value of assets and liabilities must always necessarily be equal. If, for example, there is a decrease in the value of assets without any corresponding decrease in liabilities, the capital of the Central Bank would decrease by the same amount, so the overall discrepancy between liabilities and assets would remain zero. For this reason, we say that the equality between the assets and the liabilities of the Central Bank is an *accounting identity*: it is always true because of the accounting definitions (in particular because of the definition of capital).

3.1.2 The assets of the Central Bank

Foreign exchange reserves All Central Banks hold some foreign exchange reserves (also called 'forex reserves' or 'FX reserves'). These are quantities of foreign currency. For example, the Bank of England owns a certain quantity of US Dollars, Euros, Yens, and other foreign currencies.

The Central Bank wants to hold some foreign exchange reserves for many reasons. The most important is to guarantee that the country can meet its external obligations. For example, every day UK firms, households and the Government obtain a certain quantity of US Dollars by exporting goods and services to the USA, and by borrowing from US banks and individuals. After earning these US dollars, they typically go to the Bank of England and exchange them with UK Sterlings. At the same time, UK economic agents need to pay a certain amount of US dollars to import goods and services or repay previous debts. These agents will go to the Bank of England and exchange some Sterlings with US Dollars.

If at all points in time, the overall amount of US Dollars obtained by UK economic agents is equal to the amount that they need to pay, the Central Bank can just give to the importers the US Dollars brought in by the exporters. But of course in some days (or years) the amount of US Dollars that UK economic agents must pay exceeds the amount that they earn. When this happens, the Central Bank consumes some of its dollar reserves to fill the gap. By holding foreign exchange reserves, the Central Bank makes sure that the country is able to pay its foreign obligations.

The Central Bank also uses foreign exchange reserves to influence the exchange rate. For example, imagine that the Bank of England was concerned that the value of the Sterling in terms of the US dollar is falling too fast. The Bank of England would therefore like to stop the Sterling from depreciating relative to the US dollar (or at least mitigate the depreciation). To do so, the

Bank can use its US dollar reserves to buy Sterlings in international markets. In this way, the Bank of England is able to increase demand for, and therefore the market value of, the British currency. Similarly, if the Bank of England is concerned that the Sterling is appreciating too much, it can purchase more foreign exchange reserves, paying them with Sterlings, thus deflating the market value of the Sterling in terms of other currencies.

The fact that the Central Bank holds foreign exchange reserves does not mean that it hoards somewhere large quantities of foreign banknotes. To maintain foreign exchange reserves, Central Banks hold very liquid foreign assets, for example short-term Government bonds. For instance, the US dollar reserves held by the Bank of England can take the form of one-year US Government bonds. Given that US Government bonds are very liquid, the Bank of England can quickly sell them for US dollars, every time it needs to.

Of course most of the foreign exchange reserves of a Central Bank are constituted by currencies that are relevant for trade: the currencies of the countries with which the country in question trades most, and the currencies that are used in international trades. Given that many international trades are settled using US dollars (for example commodities like oil are traded in dollars in international markets), Central Banks, including the Bank of England, want to hold most of their foreign exchange reserves in the form of US dollars (more precisely, short-term US Treasury bonds).

Securities The Central Bank holds a certain amount of domestic securities (meaning domestic financial assets). These are mostly Government bonds, which are virtually free of default risk.

Unlike in the past, today the Bank of England does not buy newly issued Government bonds directly from the Treasury. Rather, it buys them in secondary markets, meaning that it buys previously issued government bonds from banks and other financial actors. The reason is that buying newly issued government bonds directly from the government would constitute a form of 'monetization' – the financing of the government deficit with newly created money. Monetization is seen as anathema today because, if taken to extremes, it can flood the economy with money and eventually lead to runaway inflation.

When the Central Bank buys or sells securities like government bonds, we say that it performs *open market operations*. Open market operations affect the quantity of banks' reserves in the system. When the Central Bank buys securities from banks, it pays with reserves. This increases the amount of bank reserves. When the Central Bank sells securities to banks, banks use their reserves to pay. This reduces the amount of reserves. Hence, Central Bank purchases of securities increase bank reserves; Central Bank sales of securities reduce bank reserves.

During and after the 2008 financial crisis and during the 2020 Covid-induced recession (as well as during other crises in the past), the Bank of England and other Central Banks around the world purchased extremely large quantities of Government bonds. Moreover, they also started acquiring a variety of more risky assets from banks (mostly, bonds issued by firms and corporations). This large-scale purchase of bonds by the Bank of England and other Central Banks since 2009 has been dubbed *quantitative easing*, a term you might have heard before. As a result of quantitative easing, the amount of securities held by the Bank of England (as well as other Central Banks) increased dramatically. The other side of the coin is that also the amount of Central Bank reserves held by commercial banks increased dramatically. In other words, quantitative easing greatly increased the size of the Central Bank balance sheet, inflating both the liability side (through increases in commercial bank reserves) and the assets side (because of increases in securities held).

Quantitative easing had two main motivations. First, it helped banks improve their balance sheets during these crises, by providing them liquidity in exchange for less liquid assets. Second, it aimed to keep interest rates on risky and long-term loans lower than they would otherwise have been, in order to support credit creation and economic activity in the aftermath of these crises.

Other items (net) Other items include various elements that are not included in the classification above. For example Central Banks normally have physical assets like property or gold bars. Other liabilities, instead, can be accounts of other institutions (besides the Government), for example foreign Central Banks.

Summing up: the Central Bank's balance sheet

The balance sheet of the Central Bank lists all the assets and liabilities that the Central Bank holds at a given point in time. The total amount of assets must always and necessarily be equal to the total amount of liabilities. The most important liabilities are currency (banknotes), commercial banks' reserves and the Government account. The most important assets are foreign exchange reserves and domestic securities (mainly Government bonds bought on the secondary market). To provide all the reserves that the banking system needs and to contribute to its control of interest rates, the Central Bank performs open market operations. In particular, the Central Bank sells securities when it wants to reduce the amount of reserves, while it buys assets when it wants to provide more reserves. Since 2009, Central Banks around the world embarked on an unprecedented program of large-scale asset purchases, called *quantitative easing*, which dramatically enlarged both the asset side and the liability side of the Central Bank balance sheet.

3.2 Monetary policy: objectives and instruments

The main and most important function of the Central Bank is to conduct monetary policy. Broadly speaking, the objective of monetary policy is to influence macroeconomic conditions, in particular inflation and unemployment. The main instrument that monetary policy uses to pursue its macroeconomic objectives is the interest rate.

3.2.1 The objectives of monetary policy

Generally, the objectives of monetary policy are (a) to keep inflation low and stable, and (b) to keep economic activity high enough to avoid high unemployment. In short, the Central Bank wants low inflation and low unemployment. However, there are differences across different Central Banks in the relative emphasis put on these two objectives.

On one end of the spectrum, the US Federal Reserve is required by its statute to take into consideration both unemployment and inflation when setting interest rates, both with the same importance. By its statute, keeping inflation low and maintaining full employment are two parallel objectives that the US Central Bank must pursue. The US Federal Reserve, therefore, has a *dual* mandate.

At the opposite end of the spectrum, the European Central Bank (ECB) is required to consider only the rate of inflation when making its decisions: its only mandate is to keep inflation low and stable. The ECB, therefore, has a *single* mandate.

The Bank of England is somewhat in between these two models, although probably closer to the European Central Bank. Its mandate mentions both price stability and economic activity, but there is a clear hierarchy between the two, with price stability being the primary objective and economic activity only a secondary one. Meeting the inflation target – currently set at 2% – is the primary and paramount aim. Only conditional on having secured low inflation, the Bank of England should contribute to growth and employment.[13]

3.2.2 The instruments of monetary policy

In normal times, conducting monetary policy means setting the interest rate. The Central Bank leverages its ability to control the short-term interest rate in order to promote low inflation and low unemployment.

More precisely, Central Banks aim to control the overnight inter-bank rate. As we have seen in Section 3.1.1 when discussing Central Bank reserves, the overnight inter-bank rate is the interest rate that banks charge each other for short-term reserve loans, and in the UK is called Sterling Overnight Index Average, or SONIA.

Given that lending some reserves overnight to another bank is normally considered completely safe (except in the midst of a serious financial crisis), the overnight interbank-rate determines the short-term risk-less interest rate that prevails in the economy. By controlling the SONIA, the Bank of England controls the risk-less interest rate in the economy.

The interest rate influences unemployment and inflation through different channels. In general, a lower interest rate stimulates housing investment and the expansion of credit (because it

[13] Specifically, the Bank of England monetary policy mandate as written in its statute is *"to maintain price stability, and subject to that, to support the economic policy of Her Majesty's Government, including its objectives for growth and employment."*

makes it cheaper to borrow money). In some situations in which firms would like to invest more to expand productive capacity, but are kept from doing so by an high cost of credit, a lower interest rate can also stimulate business investment. In addition to this, lower interest rates tend to increase the prices of financial assets (like stocks and bonds) and houses, thus increasing the wealth of households and inducing them to spend more. For these reasons, a lower interest rate boosts aggregate demand and GDP growth. Higher GDP growth decreases unemployment and increases inflation.

To the contrary, raising the interest rate has a negative effect on housing investment and credit availability (and possibly business investment), and reduces the wealth of families by depressing asset prices; so it decreases aggregate demand and GDP, in turn increasing unemployment and lowering inflation.

In short, the Central Bank lowers the interest rate when it wants to increase GDP growth and reduce unemployment. It increases the interest rate when it wants to cool down the economy and reduce inflation.

3.2.3 The nuts and bolts of monetary policy

How does the Central Bank control the short-term interest rate? The nuts and bolts of interest rate control vary between different Central Banks. We will focus here on the Bank of England, but also describe the mechanisms used by the US Federal Reserve and the European Central Bank by way of comparison.

Bank of England: setting the Bank Rate The Bank of England, and in particular its Monetary Policy Committee, sets a key interest rate, called the Bank Rate. As we have already mentioned when discussing the Central Bank liabilities (Section 3.1.1), the Bank Rate is the interest rate that the Bank of England pays on banks' reserves. Moreover, the Bank of England has a lending facility (called Short-Term Repo, or STR) where financial institutions can borrow reserves from the Bank of England in the short-term, paying an interest equal to the Bank Rate. The Bank Rate is thus the interest rate that the Bank of England pays and charges on reserves.

The Bank Rate, in turn, influences all other interest rates in the UK economy.

Most directly, the Bank Rate determines the the over-night inter-bank interest rate. As you will recall, this is the interest rate that banks pay to borrow reserves from one another, and in the UK it is called SONIA. Most of the time (an exceptionally severe financial crisis being the exception), SONIA is bound to be very close to the Bank Rate. Indeed, SONIA cannot go significantly above the Bank Rate: if a bank tried to charge another bank an interest rate substantially higher than the Bank Rate, the borrowing bank would be better off borrowing from the Bank of England lending facility. But SONIA cannot go significantly below the Bank Rate either: if that happened, the lending bank would prefer keeping its reserves at the Bank of England and earning the Bank Rate on them, rather than lending them at a lower rate.

In turn, the rate that banks charge to one another influences the interest rates that they pay on the savings accounts of their customers, and the rates they charge to people applying for mortgages or other loans. So the Bank of England choice of the Bank Rate affects the rates paid more widely by individuals and firms in the economy, including the interest rate on the bonds issued by firms and the government.

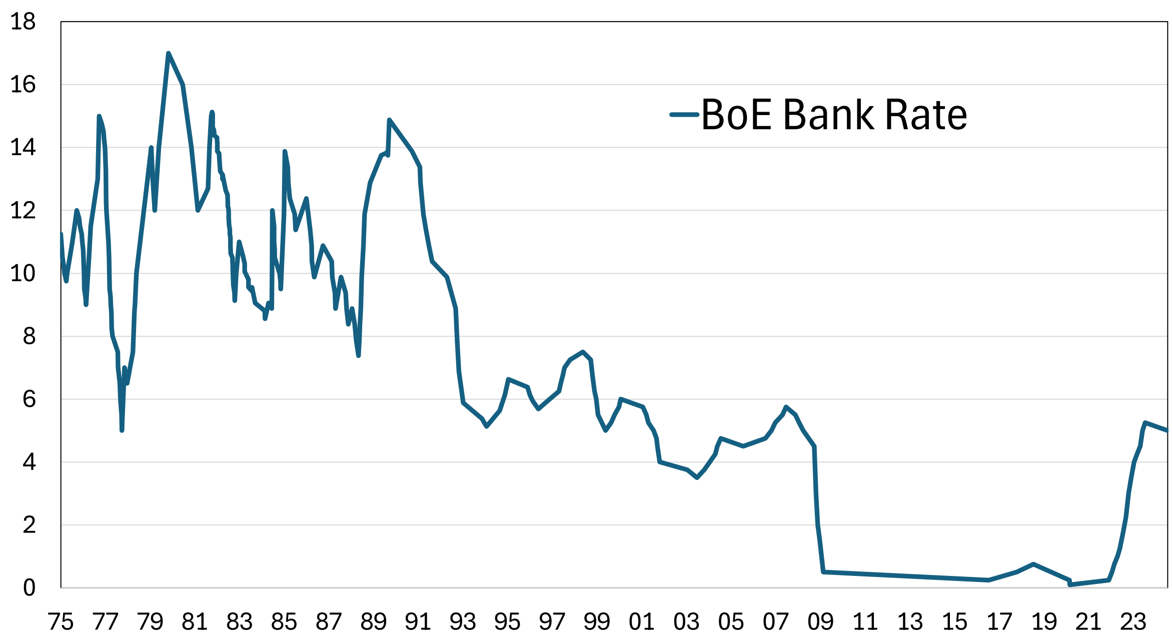
To understand the connection between the Bank Rate and all other (riskier) interest rates in the economy, recall equation 5. This says that that the interest rate on a risky loan (or equivalently a risky bond) equals the risk-less interest rate plus a risk premium. This is why, by setting the risk-less interest rate, the Bank of England also affects the interest rates on risky loans and bonds.

Figure 4 displays the Bank Rate set by the Bank of England from January 1975 to August 2024. The Bank of England kept the Bank Rate very high in the 1980s, in order to fight high inflation. Inflation was much lower and stable during the 1990s and early 2000s, allowing the Bank of England to gradually decrease the policy rate. The Bank Rate was slashed during the 2008-09 Great Recession in order to avoid an economic depression, and kept close to zero during the subsequent slow recovery. It was suddenly raised again since 2022, in reaction to the resurgence of inflation.

Note that the dynamics of the Bank Rate shown in Figure 4 are pretty similary to those of the nominal short-term rate on UK government bonds displayed in Figure 1. This is because the

Bank Rate strongly influences all other interest rates, including on UK government bonds.

Figure 4: The Bank of England Bank Rate (Jan 1975 to Aug 2024)



Source: Bank of England Official Bank Rate Statistics (`https://www.bankofengland.co.uk/boeapps/database/Bank-Rate.asp`)

US Federal Reserve: open market operations The Federal Reserve uses open market operations to control the overnight inter-bank rate, which in the USA is called Federal Funds Rate. As we have seen, open market operations essentially consist in selling assets to banks in exchange for reserves, or buying assets from banks paying with reserves. The Federal Funds Rate is the price of reserves. So the Federal Reserve uses open market operations to increase or decrease its supply of reserves, until their price reaches the desired value. If for example the Fed wants to increase the interest rate, it will decrease the supply of reserves by selling assets. When the Federal Reserve wants to decrease the interest rate, it increases the supply of reserves by buying financial assets.

European Central Bank: The Corridor System The European Central Bank (ECB) runs a system similar to that used by the Bank of England: its main monetary policy tool is the interest rate that it pays and charges on bank reserves. The main difference is that the interest rate paid and the interest rate charged are different. Therefore the European Central bank does not set a single interest rate like the Bank of England, but rather a corridor, with a lower bound and an upper bound.

Specifically, the ECB offers to commercial banks a lending facility and a deposit facility. In the lending facility, it lends reserves to commercial banks, charging them a given interest rate, called the *lending facility rate*. The ECB deposit facility is where banks hold their reserves and earn a given interest rate on them, called the *deposit facility rate*.

By setting the lending facility rate and the deposit facility rate, the Central Bank controls the overnight inter-bank interest rate (which in the Eurozone is called EONIA). Specifically, the lending facility rate is the ceiling that the overnight interest rate cannot exceed, while the deposit facility is the floor under which the overnight interest rate cannot fall. The logic is the same as in the UK Bank Rate system: the interest rate that banks charge to one another cannot go beyond the lending facility rate – no bank will borrow from you, if they can borrow from the Central Bank at a lower rate; similarly, the overnight inter-bank interest rate cannot go below the deposit facility rate – because no bank will lend to you, if you offer to pay an interest rate lower than the one it would earn in the Central Bank deposit facility.

That is why this is called a corridor system: the interest rate that banks charge to one another has to stay inside the corridor formed by the deposit facility rate (the floor) and the lending facility rate (the ceiling). By setting the floor and the ceiling, the European Central Bank controls the short-term risk-less interest rate in the Eurozone.[14]

Note that the Bank of England system is effectively just a version of the corridor system, in which the floor and the ceiling are equal. Similar corridor systems are used by the Bank of Canada, the Reserve Bank of Australia, and the Reserve Bank of New Zealand.

Figure 5 shows the working of the corridor system operated by the European Central Bank

[14] For example, suppose that the lending facility rate is set by the ECB at 3.1%, while the deposit facility rate is set at 2.9%. Imagine that BNP Paribas wants to lend reserves, while Banco Santander wants to borrow reserves. BNP Paribas will never accept an interest rate lower than 2.9%, otherwise it would just put the reserves in the deposit facility rather than lending them to Banco Santander. And Banco Santander will never accept to pay an interest rate higher than 3.1%, otherwise it could just lend the money from the lending facility. So the loan from BNP Paribas to Banco Santander will have an interest rate comprised between 2.9% and 3.1%.

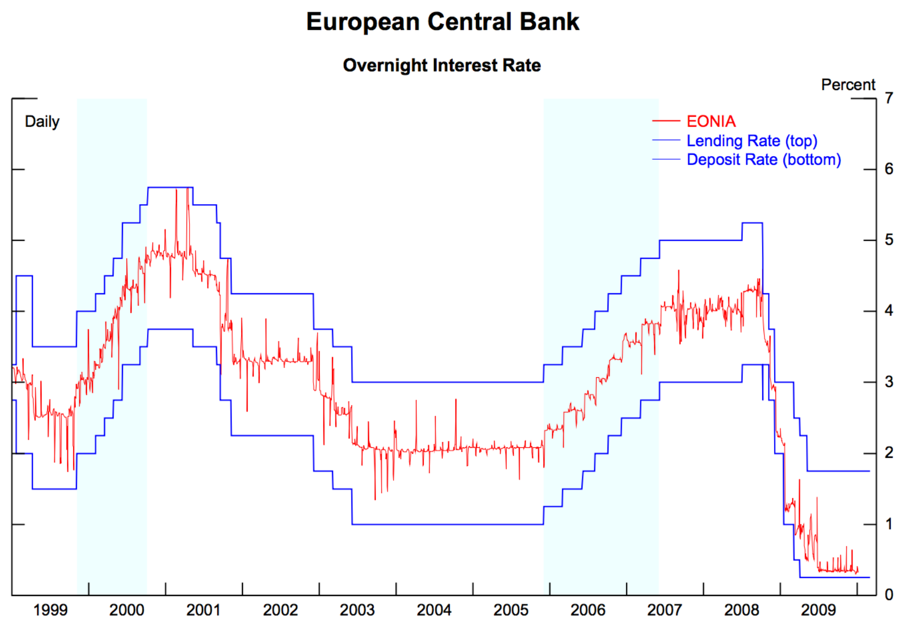


Figure 5: The European Central Bank interest-rate corridor

Source: Bowman, Gagnon and Leahy (2010)

between 1999 and 2009. The blue lines represent the lending facility rate (on top) and the deposit facility rate (at the bottom), while the red line represents the average interest rate on overnight loans between banks (called EONIA - Euro Overnight Index Average). The short-term interest rate tends to stay around the center of the corridor most of the time.

3.2.4 Quantitative easing: an unconventional monetary policy tool

In addition to setting the Bank Rate, the Bank of England (and other Central Banks) can also use bond purchases to further influence interest rates. Typically, the Bank of England uses large-scale bond purchases to further reduce risky and long-term interest rates when the Bank Rate is already at (or close to) zero and therefore cannot be reduced further.

The use of large-scale bond purchases to further reduce interest rates when the Bank Rate is already at zero has been called *Quantitative Easing* (QE). This is how the Bank of England summarizes this practice in its website: *QE is one of two tools we can use to lower interest rates, particularly when Bank Rate is very low and there is limited scope to lower it further. We first* *began using QE in March 2009 in response to the Global Financial Crisis. At that time Bank Rate was already very low. In fact, it couldn't be lowered any further at that point. So we needed another way to lower interest rates, encourage spending in the economy, and meet our inflation target. QE involves us buying bonds to push up their prices and bring down long-term interest rates.*[15]

QE is generally seen as an unconventional monetary policy tool, which was introduced during exceptional events, namely the 2008-09 financial crash and subsequent sharp economic downturn. Other large rounds of QE were performed during and in the aftermath of the Covid-19 crisis in 2020-21. It is not clear whether QE will be gradually rolled down, or whether it is here to stay as a permanent feature of monetary policy in the 21st Century.

Summing up: monetary policy

The Central Bank conducts monetary policy by setting the interest rate. Specifically, the interest rate that the Central Bank controls is the overnight inter-bank rate: the rate that banks charge to one another on short-term loans (called SONIA in the UK, Federal Funds Rate in the US, and EONIA in the Eurozone). Given that short-term lending between banks is considered completely safe, SONIA determines the short-term risk-less interest rate that prevails in the UK economy.

There are two main ways for Central Banks to control the overnight inter-bank interest rate. The first is to conduct *open market operations*. This means changing the supply of reserves to influence their price (the interest rate). The second is to employ a 'corridor system. In a corridor system, the Central Bank sets the rate at which it lends money to banks (lending facility rate) and the rate that it pays on banks' reserves (deposit facility rate). The lending facility rate is the ceiling that the inter-bank interest rate cannot overstep, while the deposit facility rate provides a floor under which the inter-bank interest rate cannot go. The Federal Reserve uses open market operations to set the interest rate, while the European Central Bank uses a corridor system. The Bank of England uses a particular version of the

[15] https://www.bankofengland.co.uk/monetary-policy/quantitative-easing, accessed on Aug 22, 2024.

corridor system, where the lending facility rate and the deposit facility rate are set at the same level (called the Bank Rate).

Central Banks lower the interest rate when they want to increase GDP growth in order to reduce unemployment. They raise the interest rate when they want to slow down the economy in order to reduce inflation.

When the short-term interest rate is at or close to zero and cannot be lowered further, Central Banks can use quantitative easing (QE) to further decrease interest rates and stimulate economic activity. QE consists in buying large quantities of medium and long-term bonds.

3.3 The Central Bank as a lender of the last resort during financial crises

Before modern Central Banks were founded, financial crises with bank failures were much more common. In particular, 'bank runs' would happen very often.

A bank run is a situation where the (justified or unjustified) fear that a bank may fail turns into a self-fulfilling prophecy: depositors suddenly withdraw their money in large numbers at the same time, which can cause the bank to default. In other words, fear that a bank may fail can cause it to *actually* fail, even if its initial situation was in fact sound.

For example, imagine that Bank A is in good financial conditions. It made loans that are very likely to be repaid, and it invested in financial assets that are doing well. However, rumors start to circulate, according to which Bank A is not doing well and some loans will not be repaid. These rumors are false, but people are unsure whether there is some truth to them. To be safe, many people that deposited money at Bank A prefer to withdraw their accounts and move them to another Bank, or convert them in banknotes. If enough people do so, the bank will run out of funds and experience a real crisis. Given that the loans made by Bank A cannot easily be called back before the due date, Bank A will not be able to reimburse all its depositors, and it will have to close.

Before the creation of modern Central Banks, a typical cause of bank runs was 'contagion'. In a contagion, one bank fails because it made bad loans, and this causes depositors at other banks

Figure 6: A bank run in Germany during the Great Depression
Source: Bundesarchiv, Bild 102-12023 / Georg Pahl / CC-BY-SA 3.0

(that are actually in good financial conditions) to become worried and withdraw their deposits, forcing also some of these 'good' banks to close. In other words, the failure of one single bank often had a domino effect, which would bring down other banks too.

A major factor that contributes to make bank runs possible is the so-called *liquidity mismatch*: the loans that banks make take some time to be repaid, while deposits can be withdrawn by customers at short notice.

Consider this simplified example. In year 1860, Bank A has 100 millions worth of assets, composed of 70 millions of loans made, that will be repaid in 5 years, and 30 millions of cash in its vaults. Bank A has also 80 millions worth of liabilities, corresponding to the deposits of customers. It is certainly a very healthy bank: eventually, when the loans will be repaid, it will have sufficient money to repay all customers and make a fat profit. However, if some badly run bank fails in the same region of Bank A, depositors may become nervous. If they all withdraw their deposits, Bank A will not be able to repay all of them. The cash that it has in its vaults (30 millions) are not sufficient to repay all deposits (80 millions). The 70 millions that the Bank lent out will be recovered only after the 5 years will have expired. This is liquidity mismatch: your debts can be recalled on short notice, but your credits cannot. So, even if your overall credits are much higher than your debts, if panic pushes your creditors to withdraw all their money at the same time, you may not be able to repay.

The example above illustrates the difference between a *solvency crisis* and a *liquidity crisis*. If the value of your debts is higher than the value of your assets, your are bankrupt or insolvent: you face an insolvency crisis. If your assets are worth more than your debts, but you fall short of liquidity during a panic, your are solvent but lack liquid funds: you face a liquidity crisis.

A particularly devastating episode happened in the USA in 1907, when generalized panic led to a series of bank runs, that caused many bank failures and pushed the entire national financial system on the verge of collapse. This episode is often called 'The Panic of 1907'. Banks needed a source of emergency funds to prevent the panics and the resulting runs from driving them out of business. After the Panic of 1907, the United States decided to establish a Central Bank, the Federal Reserve, precisely to provide this source of liquidity in times of financial panic.

The US Federal Reserve was designed to act as a 'lender of last resort': should an otherwise healthy bank suffer a loss of confidence and a bank run, the Fed would lend it all the money that is necessary to repay all customers that want to withdraw their money. This would reassure all depositors that their money is not at risk. Indeed the very existence of a Central Bank ready to be a lender of the last resort is usually sufficient to prevent bank runs: if people know that the Federal Reserve will act in this way, they know that they will always be able to withdraw their money and thus they don't panic in the first place. The existence of a lender of the last resort is usually sufficient to prevent panics, in such a way that the emergency funds are seldom actually needed.

When the Central Bank acts as a lender of the last resort, it loans money to banks against the value of the assets of the bank. This means that the assets of the bank are used as collateral in these loans: should a bank not be able to repay the money, the Central Bank will seize some of the assets posted as collateral, until the debt is repaid.

The historical evidence suggests that also the Bank of England (which was founded much earlier, in 1694, but only gradually acquired the features and functions of a modern Central

Bank) has often played the role of lender of last resort in its long history, thus mitigating the incidence of bank runs. This has happened as recently as during the 2007/2008 financial crash, when the Bank of England stood ready to provide emergency loans to financial institutions in order to sedate financial panic.

One potential problem with having the Central Bank acting as a lender of the last resort is that large financial institutions may act irresponsibly, take too much risks, knowing that the Central Bank will always step in to save them. Economists call this *moral hazard*: knowing that you will be rescued in any case, you act irresponsibly. In theory, there is a simple solution to this: the Central Bank should provide emergency funds only to institutions that are healthy and solvent but temporarily lack liquidity, not to banks that are insolvent. In other words, the Central Bank should intervene if the bank faces a liquidity crisis, but not if it faces a solvency crisis. If a bank has made very bad or too risky loans and investments, it should not be eligible for receiving emergency funds from the Central Bank.

In practice, however, it can be very difficult to distinguish between a liquidity crisis and a solvency crisis. During a crisis the price of many assets (houses, stocks, bonds) falls, and this changes the value of assets and liabilities of each institution. A solvent institution can become temporarily insolvent because its assets temporarily lose value during the panic. Assessing the true value of the assets of a bank, and thus deciding whether it is solvent or not, and how much money can be lent to it, can become difficult during a financial crisis. This can also create uncertainty for investors and depositors as to whether their bank is solvent or not.

For these and other reasons, the existence of a Central Bank is not sufficient to eliminate bank runs altogether (although it certainly reduced their frequency).

A further measure that can drastically reduce the occurrence of bank runs is the institution of some kind of public deposit insurance. This means that the Government insures depositors: should your bank go bankrupt, the Government will reimburse your deposit (up to a certain amount). The United Kingdom introduced deposit insurance in 1979, under the name of Deposit Insurance Scheme. Since then, the UK government ensures checking and savings account up to a ceiling, which is currently £85,000. As a result, there is no reason for depositors to run to the bank to withdraw their money (at least, not amounts below $85,000$ pounds).[16]

Although deposit insurance can also lead to moral hazard, it is considered necessary to avoid bank runs. However, as the 2007-2008 crisis demonstrated, deposit insurance is no longer enough. There are today many ways for banks to borrow short-term money, other than just through deposits. For example there are so-called 'wholesale' markets in which banks borrow overnight from other banks and investors. These other funds are not insured and indeed during the crisis there were runs on many banks, not from the traditional depositors but through wholesale lenders.

Summing up: the Central Bank as a lender of the last resort Before the advent of modern Central Banks, financial crises were more common. Often these crises led to "panics" in which people raced to their banks to withdraw their deposits. The Central Bank typically acts as a lender of the last resort during financial crises. This means that it provides emergency loans to banks when they face a crisis of confidence. To receive the emergency loans, a bank must usually post its assets as collateral. In theory the Central Bank should provide emergency loans only to banks that are fundamentally healthy but are facing a temporary liquidity crisis, not to banks that are fundamentally bankrupt. In practice, it is often difficult to distinguish between the two.

4 The Stock Market

In Section 2 we focused on bonds. Bonds are very important debt instruments. For example, Governments borrow money from the public only through bonds. However, for firms bonds are not the only source of finance. Another way for firms to obtain external funding is to issue stocks (also called shares).

Unlike a bond, a corporate stock does not guarantee a predetermined amount of money. Rather, a stock gives you the right to receive a share of the profits of the firm. These variable

[16] By way of comparison, the United States introduced *federal deposit insurance* in 1934. It was one of the measures undertaken by F.D. Roosevelt to fight the Great Depression. As of 2024, the US Government insures bank deposits account up to a ceiling of $\$250,000$.

payments are called *dividends*. When you buy some stocks of a company, you become one of the owners of the firm. The quantity of stocks that you own, relative to the total stocks issued, determines the share of the firm that you own.

Not all the profits of the firm are paid to stock holders as dividends. The firm retains some of its profits in order to finance its new investments. Only profits that are not retained for investment are distributed to stockholders in the form of dividends. Being a share of profits, dividends tend to be very correlated with them: the higher the profits of a firm, the higher the dividends it pays; when profits increase, so do dividends.

As in the case of bonds, stock prices are influenced both by economy-wide factors (macroeconomic conditions and macroeconomic policy) and factors specific to the firm in question. In order to assess the overall movements of the stock market, we look at stock-market indexes. The most used stock market index in the United Kingdom is the *FTSE 100* index. This index is calculated as the weighted average of the prices of the stocks of 100 large companies listed in the London Stock Exchange, with weights proportional to firm size. As such, it captures the overall movements of the UK stock market. Similar indexes exist in virtually all World countries. For example, the *Standard & Poor's 500 Composite Index*, in short *S&P 500*, is the main index of US stock prices, and it is computed as a weighted average of the stock prices of 500 large firms listed in the US.

Given that the US is the main global stock market (as well as the world largest economy), it is interesting to look at it as an example. Figure 4 plots the US S&P 500 stock market index from 1955 to mid-2016. Being an index number, its absolute value has no meaning: what matters is its dynamics – its movements. In particular, in the Figure we set the initial period (January 1955) equal to 1.0. Moreover, to express stock price movements in terms of real purchasing power instead of nominal money, I divided the S&P index by the US monthly CPI index. Clearly, in the period 1955-1980 stock prices were lower and quite stable, while after 1980 they have displayed a strong increasing trend and much more volatility (booms followed by sharp slumps). According to many observers, a major factor behind this change in behavior has been financial deregulation, which increased the profitability of financial institutions but also opened the way for speculative

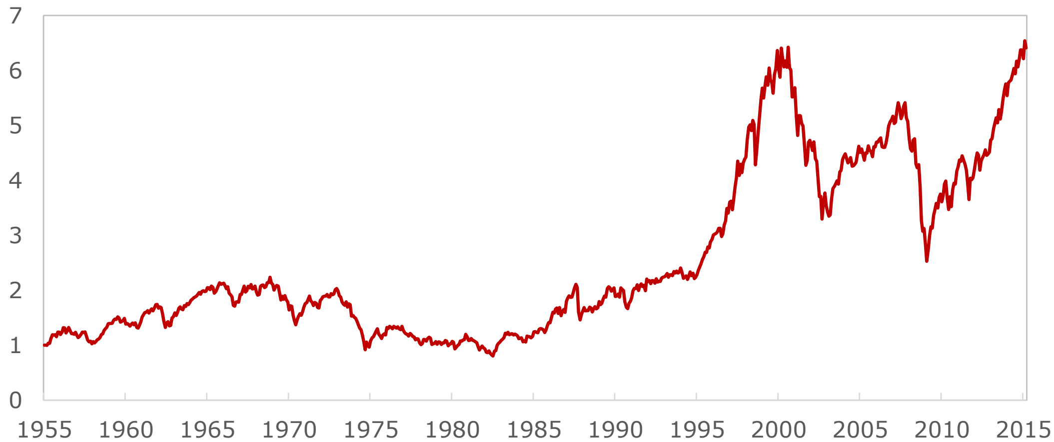


Figure 7: S&P 500 stock price index in real terms
(Index number; 1970=1.0; monthly data from Jan 1955 to Aug 2016)

bubbles. According to recent research, another main force behind the rise of stock prices in the USA has been a decrease in the bargaining power of workers, which reduced the share of firms' net income paid to workers, thus making firms more profitable for their owners.

A rigorous discussion of what determines the price of a single stock and the overall movements of the stock market is outside the scope of this course. For our purposes, it is sufficient to say that the value of a stock reflects the present value of all future expected dividend payments. As a consequence, the larger the profits that a firm is expected to obtain, the higher the price of its stocks. In the whole economy, the higher expected total profits, the higher the average value of stocks (as measured for example by the *SP 500* index in the US or the *FTSE 100* index in the UK).

As a consequence, the price of a firm's stocks increase when expectations regarding the firm's profits are revised upwards. For example if a firm presents a brand new product, that investors think will be very successful, the value of its stocks will increase. By the same token, if expectations regarding GDP growth or regarding the share of GDP absorbed by profits (as opposed to wages) are revised upwards, the *FTSE 100* index will go up. Also expectations of a decrease in taxes on corporations would cause the *FTSE 100* index to increase.

Another factor that influences share prices is the interest rate. In general, a lower interest rate will increase the value of all stocks (and thus indexes like the $FTSE100$ will go up). This happens for two reasons: (1) the value of stocks is related to dividend payments that will happen in the future. When investors calculate the present value (the value today) of money that will be received in the future, they apply a discount rate. The discount rate measures how much dollars today is worth one dollar tomorrow. The discount rate takes into account the opportunity-cost of money – the interest that could have been earned on the money if they were borrowed instead of invested in the stock – so it depends on the interest rate. The higher the interest rate, the higher the discount rate, the lower the present value of future payments, and so the lower the price of stocks; (2) with a lower interest rate, it is cheaper to borrow, and so it can be more convenient to borrow money to invest in stocks, which increases demand for stocks and their price.

Stock prices are thus positively related to expected (after-tax) profits, GDP growth, and the profit share of GDP; they are negatively related to the interest rate, the corporate tax rate, and the wage share of GDP.

Summing up: the stock market

By buying a stock, you acquire a fraction of the ownership of a firm. This gives you the right to receive a share of those profits of the firm that are not reinvested. These variable payments are called dividends. To assess overall movements in the stock market, we look at stock-market indexes (like the *FTSE 100* index in the UK or the *S&P 500* index in the US). The value of a stock reflects the present value of all expected future dividend payments. As a consequence, stock prices are positively related to expected profits and GDP growth, and negatively related to the interest rate and the wage share of GDP.

5 Further readings

- 'Money in the modern economy: an introduction ' by Michael McLeay, Amar Radia and Ryland Thomas, Bank of England Quarterly Bulletin, 2014 Q1, available at this LINK

- 'Money creation in the modern economy' by Michael McLeay, Amar Radia and Ryland Thomas, Bank of England Quarterly Bulletin, 2014 Q1, available at this link

- 'Understanding the central bank balance sheet ' by Garreth Rule, Bank of England Center for Central Bank Studies, Handbook n.32, available at this link

References

Blanchard, Olivier (2016). *Macroeconomics, 7th Edition*. Pearson.

McLeay, M., A. Radia, and R. Thomas (2014). "Money in the modern economy: an introduction". In: *Bank of England Quarterly Bulletin* 54.1. URL: `https://www.bankofengland.co.uk/quarterly-bulletin/2014/q1/money-in-the-modern-economy-an-introduction`.

Appendix

A Derivation of the Fisher equation

This appendix demonstrates formally how the important formulas of equations 3 and 4 are derived.

Imagine that there is only one good in the economy, apples. In this simple situation, the price level is just equal to the price of apples, and the inflation rate is just the rate of change of the price of apples.

In year t, the nominal interest rate is i_t. This means that, if you invest y dollars in year t, you will get $y(1+i_t)$ dollars in year $t+1$. But you don't care about dollars in themselves. Rather, you care about the apples that you can buy with these dollars. The price of apples is P_t in year t and P_{t+1} in year $t+1$. This means that y dollars buy $\frac{y}{P_t}$ apples in year t and $\frac{y}{P_{t+1}}$ apples in year $t+1$. Therefore, the real value (in terms of apples) of each dollar you invest at time t is equal to $\frac{1}{P_t}$ apples, while the real value of each single dollar you obtain at time $t+1$ is equal to $\frac{1}{P_{t+1}}$ apples.

So, if you invest x dollars in the bond, you pay the equivalent of $\frac{y}{P_t}$ apples at year t, and you obtain in return the equivalent of $(y+i_t)/P_{t+1}$ apples next year. By definition, the real interest rate on the bond (the interest rate in terms of apples) is equal to the difference between the real amount that you get at year $t+1$ and the real amount that you invested initially, divided by the real amount that you invested initially. Formally,

$$r_t = \frac{\frac{y(1+i_t)}{P_{t+1}} - \frac{y}{P_t}}{\frac{y}{P_t}} \tag{7}$$

The term y (the amount of dollars invested) can be eliminated from the formula, because it is present in all terms of the numerator as well as in the denominator. So we can write

$$r_t = \frac{\frac{(1+i)}{P_{t+1}} - \frac{1}{P_t}}{\frac{1}{P_t}} \tag{8}$$

The formula can be further simplified through some rearrangements:

$$r_t = \frac{\frac{(1+i_t)}{P_{t+1}} - \frac{1}{P_t}}{\frac{1}{P_t}} = (\frac{1+i_t}{P_{t+1}} - \frac{1}{P_t})P_t = (1+i_t)\frac{P_t}{P_{t+1}} - 1 \tag{9}$$

Moving the 1 on the left side of the equation, we get

$$(1+r_t) = (1+i_t)\frac{P_t}{P_{t+1}} \tag{10}$$

This is a much more compact formula. Some further rearrangements will make it even more appealing.

First, remember that π_{t+1} (the rate of inflation at time $t+1$) is equal to the rate of change in the level of prices. We have assumed that there is only one good in the economy, so the change in the level of prices is just the change in the price of apples. So we have

$$\pi_{t+1} = \frac{P_{t+1} - P_t}{P_t} \tag{11}$$

Applying some manipulations to equation 11, we obtain

$$\pi_{t+1} = \frac{P_{t+1} - P_t}{P_t} \Rightarrow \pi_{t+1} + 1 = \frac{P_{t+1}}{P_t} \Rightarrow \frac{1}{\pi_{t+1} + 1} = \frac{P_t}{P_{t+1}} \tag{12}$$

We can substitute the last equivalence in eq. 12 into equation 9, and obtain

$$(1+r_t) = \frac{1+i_t}{1+\pi_{t+1}} \tag{13}$$

This is the exact relation between the real and the nominal interest rate: (one plus) the real interest rate equals the ratio between (one plus) the nominal interest rate and (one plus) the rate of inflation. This formula is fairly simple, but the intuition behind it is not really immediate to get. Further manipulations can yield an even more intuitive and useful formula.

First, note that equation 13 implies the following

$$1 + i_t = (1+r_t)(1+\pi_{t+1}) = 1 + \pi_{t+1} + r + r\pi_{t+1} \tag{14}$$

Therefore we have

$$i_t = \pi_{t+1} + r_t + r_t\pi_{t+1} \tag{15}$$

In turn, this implies that the real interest rate is equal to

$$r_t = i_t - \pi_{t+1} - r_t\pi_{t+1} \tag{16}$$

When both the interest rate and the inflation rate are not too high – say, less than 20% – the term $r_t\pi_{t+1}$ is very small with respect to the other terms in the equation.[17] We can thus neglect the term $r_t\pi_{t+1}$, and say that the real interest rate is *approximately* equal to the nominal interest rate minus the inflation rate. Formally, we can write

$$r_t \approx i_t - \pi_{t+1} \tag{17}$$

We have thus derived the Fisher equation. While you must not commit all the passages of the derivation to memory (but you should try to derive it yourself at least once), it is very important that you keep formula 17 in mind. *The real interest rate is (approximately) equal to the nominal interest rate minus the inflation rate.*

B Derivation of Equation 6

As mentioned in the main text, the riskiness of a bond is determined by its probability of default. Only when the difference between the interest rates on the risky bond and the riskless interest rate is high enough to compensate investors for the risk of default, someone will be willing to hold the risky bond. The interest rate on a risky bond must be high enough to compensate investors for the risk of default. To be more precise, in order for investors to be willing to hold the risky bond, the expected return on the two bonds must be *at least* the same.

Indicating the probability of default as p, the expected rate of return on a risky bond can be written as follows

$$(1-p)(1+i+x_R) + (p)(0) \tag{18}$$

[17] For example if the interest rate is 5% and the inflation rate is 3%, the term $r_t\pi_{t+1}$ is equal to 0.0015.

With probability $(1-p)$, there will be no default, the bond will be paid regularly, and so the bond holder will get a return of $(1+i+x_R)$ on its initial investment. With probability p, there will be default, and the bond will pay zero.

Thus, in order to get the same expected return on a risky bond as on a riskless bond, the following relation must hold:

$$(1+i) = (1-p)(1+i+x_R) + (p)(0) \tag{19}$$

The left-side of equation 19 gives the return on a riskless bond. The right-hand side gives the expected return on a risky bond that has a probability of default equal to p. Reorganizing the equation gives

$$x_R = (1+i)\frac{p}{1-p}$$