CORVINUS
UNIVERSITY of
BUDAPEST

Budapest, 2018.

CORVINUS UNIVERSITY OF BUDAPEST
DEPARTMENT OF FINANCE

Basics of Finance

Authors

Gábor Kürthy (Chapter 1, Chapter 2)
József Varga (Chapter 3)
Tamás Pesuth (Chapter 4)
Ágnes Vidovics-Dancs (Chapter 5.1 - 5.3)
Ildikó Gelányi (Chapter 5.4)
Géza Sebestyén (Chapter 5.5)
Eszter Boros (Chapter 6)
Gábor Sztanó (Chapter 7)
Erzsébet Varga (Chapter 8)

Editor
Gábor Kürthy

Reviewers
Ágnes Vidovics-Dancs (Chapter 1)
György Surányi (Chapter 2)
Gábor Kürthy (Chapters 3, 4, 5, 6, 7, 8)

Budapest, 2018.

ISBN 978-963-503-743-8

TABLE OF CONTENTS

CHAPTER 1 *TECHNICAL INTRODUCTION*

To understand finance properly, one needs to have a solid grasp on the elemental definitions and techniques of accounting. For being able to keep track of the following chapters, we suggest the Reader studying the next few pages thoroughly.

The balance sheet is a financial statement that represents an economic agent's (a household, a company, a bank, a budgetary institution etc.) wealth by two approaches. Assets are listed on the left-hand side or asset side, resources financing the assets are listed on the right-hand side or liability-equity side. The balance sheet is always in balance, that is:

ASSETS = LIABILITIES + EQUITY

When constructing the balance sheet, assets are listed first, then liabilities. Shareholders' equity is always a residual, i.e. it is the difference between assets and liabilities. The equity of a company or a bank is frequently referred to as capital, which can lead to misunderstandings. In this context, capital is not a pile of cash that can be invested. It is only a notional entry that shows what would be left if all the company's debts were repaid. A negative book value of the capital means that the company is insolvent in the long run, i.e. it cannot repay all of its liabilities.

Example: the balance sheet of a company
Company "ABC" has 57,000 EUR worth of assets that are partially financed by long- and short-term liabilities. Long-term (or non-current) liabilities - such as bonds issued or mortgages - are due over 12 months. Short-term (or current) liabilities - such as bills or taxes payable - mature within 12 months.

ABC Company, Balance sheet, at 31-Dec-2017

Assets (EUR)		**Liabilities (EUR)**	
Land	25,000	Long-term liabilities	15,000
Machinery	20,000	Short-term liabilities	3,000
Inventories	4,000	**Equity (EUR)**	39,000
Supplies	6,000		
Cash	2,000		
Total assets:	***57,000***	***Total liabilities and equity: 57,000***	

Economic events can change the balance sheet in four ways:

- *both sides increase by the same amount*
- *both sides decrease by the same amount*
- *the asset side is restructured, i.e. some assets increase and others decrease*
- *the liability-equity side is restructured, i.e. some liabilities or equity items increase and others decrease*

The following example helps to understand the issues described above.

Example: bookkeeping

In the first month of 2018, the following events happened to ABC company:

➡ 1,000 EUR worth of inventories were bought, the company promised the supplier to pay in 60 days.

Assets: Inventories, +1,000 EUR *Liabilities: Short-term liabilities, +1,000 EUR*
Equity: no change

➡ 1,500 EUR worth of goods were sold for 1,900 EUR.

Assets: Supply -1,500 EUR *Liabilities: no change*
Cash, +1,900 EUR *Equity: + 400 EUR*

➡ The company repaid 400 EUR short-term loan and 20 EUR interest.

Assets: Cash -420 EUR *Liabilities: Short-term liabilities -400 EUR*
Equity: - 20 EUR

➡ The National Competition Authority imposed a fine of 500 EUR on the company, to be paid within 6 months.

Assets: no change *Liabilities: Short-term liabilities, +500 EUR*
Equity: -500 EUR

➡ An investment bank granted the company 5,000 EUR worth of loan with the maturity of 20 years. The company spent the proceeds on a new machinery.

Assets: Machinery, +5,000 EUR *Liabilities: Long-term liabilities, +5,000 EUR*
Equity: no change

In order to keep the balance, we followed the *rules of double entry bookkeeping*. At the end of January-2018, after booking all the events, the balance sheet looks like as follows:

ABC Company, Balance sheet, at 31-Jan-2018

Assets (EUR)		**Liabilities (EUR)**	
Land	25,000	Long-term liabilities	20,000
Machinery	25,000	Short-term liabilities	4,100
Inventories	5,000	**Equity (EUR)**	38,880
Supply	4,500		
Cash	3,480		
Total assets:	***62,980***	***Total liabilities and equity: 62,980***	

Economic events, mostly exchanges, always happen to two agents simultaneously, which leads to *quadruple entry bookkeeping* on the systemic level. If “A” does business with “B” then two changes arise in both balance sheets, which means four changes altogether. (In Latin, four is quattuor, that is where the term comes from.) Consider the following examples.

Example: quadruple entry bookkeeping

➡ Company "X" buys inventories from company "Y" and pays 4,000 EUR for them. Before the transaction, the book value of the traded inventories was 3,500 EUR.

Company "X"		**Company "Y"**	
Inventories +4,000		Supply -3,500	
Bank account -4,000		Bank account +4,000	
			Equity +500

➡ Company "Z" pays 2,000 EUR wage to Mrs. M.

Company "Z"		**Mrs. M.**	
Bank account -2,000		Bank account +2,000	
	Equity -2,000		**Equity +2,000**

➡ Mr. Q. repays 100 EUR of debt plus 5 EUR interest to Mrs. S. in cash.

Mr. Q.		**Mrs.S.**	
Cash -105	Liabilities -100	Claims -100	
	Equity -5	Cash +105	**Equity +5**

From an accounting point of view, there are two kinds of goods. ***Real economic goods*** are on *the asset side of one single balance sheet,* while ***financial goods*** appear in *two balance sheets simultaneously*: on the asset side of one balance sheet and on the liability-equity side of another one. A slice of bread or a bicycle are real economic goods; a mortgage loan, a commercial bill or a share are financial goods.

Financial goods on the liability-equity side are obligations, while those on the asset side are claims. A commercial bill with the nominal value of 500 EUR is a claim to its owner and an obligation for its issuer. The issuer is legally bound to pay 500 EUR when the bill matures. In case of default - i.e. if the issuer is not able or not willing to pay -, the owner of the bill can sue the issuer.

Shares might be regarded as financial assets too, although in this case the issuer is not legally bound to pay any money. Shares are rather economic than legal promises to pay dividends or provide capital gains through the appreciation of their price. Disappointed owners - in case of no dividend payments and no price increase - can punish the issuer company on the market by selling the shares. What happens if investors start selling the shares of some company en masse? The price of the shares plummets which leads to the simultaneous devaluation of the assets of the company. (Remember, the value of the assets always has to be equal to the summed value of the liabilities and the equity. A price drop of shares devalues equity and assets simultaneously.) As assets are collateral behind the liabilities, a significant drop in their value can lead to financial difficulties.

CHAPTER 2 *MONEY AND BANKING FROM A HISTORICAL AND THEORETICAL PERSPECTIVE*

2.1 Money in history and theory

The historical emergence of money can be related to the emergence of market-based economies. For thousands of years, communities were organised by redistributive institutions, centralised rules coordinated production, consumption, investment, etc. Ancient Egypt and Babylonia are the most typical examples of redistributive empires, but even in the feudalistic European kingdoms, markets played only a minor role for a long time. (Polányi, 1944: Chapter 4) These markets are frequently described as places where rural farmers and urban manufacturers exchanged their products directly. However, it is easy to see that barter is a very inefficient way of exchange, as the probability of double coincident of wants is low and by the growth of the number of market participants, it gets even lower. There is no historical proof that direct barter has ever played an important role in coordinating markets (Wray, 1993). Since the very advent of locally organised markets, buyers and sellers have been using a commonly accepted specific good as the medium of exchange. Primitive forms of money had been used before as store of value and standard of deferred payments, but in lack of markets, they have not functioned as the medium of exchange or the measure of value (Polányi, 1957). The most important function of ancient coins was probably the fact that the state (the king or the queen) accepted them when paying the taxes. This characteristic made primitive money generally acceptable on markets: as everyone had to pay taxes, sellers accepted coins because they knew they could (and in most cases, they should!) use them for tax payment, and they knew that other sellers from whom they would buy goods were thinking in the same way. It is tautological, but money is accepted because it is thought to be accepted.

From this short introduction it has to be clear that ***money is defined by its functions***, i.e. we call something money if it is a

- *medium of exchange,*
- *standard of deferred payments,*
- *store of value,*
- *measure of value.*

Money as a means of exchange helps market-coordination in several dimensions. As mentioned above, general acceptance simplifies trade by splitting long and hard discoverable chains of exchanges into small parts. Without money, theoretically, the economy as a whole must find the market-clearing exchanges simultaneously to avoid disequilibrium. This cannot be implemented without the knowledge of some central planner, who, of course, does not exist. Besides, the use of money spares information: in a barter economy with N kinds of goods, the number of relative prices that market participants have to remember is $\frac{N*(N-1)}{2}$, whereas in an economy with N kinds of

goods and money, the number of absolute prices is N. If $N = 100$, the size of the information set is 4,950 and 100 in the two cases respectively.

In a monetised economy, debt is easy to measure and administer with the help of money, by which the two sides of a transaction (service and payment) can be separated in time. Expressing debts in monetary terms leads to more efficient economic, social and legal procedures. For example, economic agents do not need to pay taxes by sacrificing their labour and working for the state anymore. The state can collect taxes paid in money, and then pay for a competent workforce to execute special jobs. (Without money, some John has to work on state-owned farms or must serve as a soldier. The problem is that being an excellent wool-manufacturer, John knows nothing about agriculture or warfare. If instead, he pays taxes in money, skilled farmers and soldiers can be hired.) Another example from the field of jurisdiction: in a society without money, a culprit always has to be imprisoned, executed or sentenced to penal servitude. However, with money, in a lot of cases, the sinner can literally pay for his sins.

Money as a store of value secures future both in the short- and in the long-run. Between two exchanges, the purchasing power can rest in money, which creates the possibility of intertemporal optimisation. The store of value and the standard of deferred payments functions are closely connected. Surplus agents (those who spent less in a period than they earned) can finance deficit agents. The claim of surplus agents serves as the store of their postponed consumption, while the debt of deficit agents is the price of their impatience for which they will pay later.

Money measures the value of goods, services, wealth, debt, etc., which makes these things comparable. With money usage, a price can be assigned not only to exchangeable products but even to unmarketable stuff like works of art, clear environment or life. It is needless to say that the measure of value and the means of exchange functions are hand in hand with each other - exchanges are based on the knowledge of prices expressed in monetary units.

The transition from redistributive societies to market economies had been a long process, during which not only goods but labour and capital also became the subjects of exchange. This evolution was accompanied by different mutations in the monetary system. Financial innovations were partially forced by the requirements of growing markets, as money supply had to keep pace with economic development. On the other hand, the interests of the state (and not necessarily those of the public…) lead to notable alterations of money and monetary institutions as well. In the next section, we will examine the most important changes through the lenses of an accountant, i.e. we will intensively use the techniques familiarised in the first chapter of this book. However, the essence of the following is not the bookkeeping but the theoretical and practical understanding of how modern monetary systems have evolved and are operating today.

2.2 Production of money and creation of money

This tour begins with an imagined world where gold coins fulfil the functions of money. Gold coins are minted by the Royal Mint from raw gold extracted from mines or panned from rivers. Miners are obliged by law to surrender their gold to the Mint who issues stamped gold currency after the coinage. Coins are ***legal tender*** in the economy which means that producers, merchants, shop owners, etc., must accept them in exchange for their goods. On the other side, buyers are also obliged by the legal tender law to use coins as a medium of exchange. However, upon mutual agreements, parties are allowed to use alternative settlements in a given business situation. Summing up: *if the buyer wants to pay with gold coins the seller must accept them; if the seller wants to be paid in gold coins the buyer must adapt; if they both agree they can use different methods*.

Before the introduction of coinage and legal tender laws, a price system consisting of relative prices (for instance 10 kilograms of bread = 1 pair of boots) had evolved in the economy. An important vector of this system measured the values of goods in terms of gold (Figure 1).

FIGURE 1: PRICES BEFORE THE INTRODUCTION OF COINAGE.

	1000 kg of wheat
	500 liters of beer
1 ounce of gold =	…
	…
	5 acres of land

The Royal Mint minted one ounce of gold into 10 pieces of coins with the nominal value of 1 shilling stamped on them (or the equivalent of it, for example, 2 pieces of 5 shilling coins). One shilling as a measurement of mass equals one-fifth of an ounce, thus the Royal Mint issued coins that contained half of the quantity of gold of their nominal value. However, because of the legal tender law, economic agents had to accept these coins at nominal value. Thus the price system changed (Figure 2).

FIGURE 2: PRICES AFTER THE INTRODUCTION OF COINAGE.

	1000 kg of wheat
	500 liters of beer
5 shillings =	…
	…
	5 acres of land

What happened at the Mint? The miner sold 1 ounce of gold for 5 shillings, the Mint kept the rest. Suppose that the unit cost of minting 1 ounce of gold was 1 shilling. After

deducing this cost, the profit of the Mint was 4 shillings. This profit served as the revenue of the State.

By appearing on the asset side, the gold found in the mine increases the miner's balance sheet. As the miner knows that he has to surrender the gold, and eventually he will keep half of the amount mined, both his liabilities and equity increase by half of the value of the gold (Figure 3). It is worth to notice that mining is not an exchange of goods or services, thus no other balance sheet has to be involved when accounting for it.

For the sake of simplicity, the costs of mining were set aside, or put it in another way, the result of the mining reflected by Figure 3 is the net result after deducting the costs.

FIGURE 3: ACCOUNTING FOR THE MINING OF 1 OUNCE OF GOLD.

Miner

gold, +10 shillings	liabilities, +5 shillings
	equity, +5 shilling

Surrendering his gold and getting the proceeds after coinage rearranges the balance sheet of the miner: he is relieved of his gold, half of which being his liability and receives 5 shillings in gold coins. The balance sheet of the Mint (which can be identified with the State) increases by 4 shillings: 4 shillings as coins on the asset side and 4 shillings of equity. The remaining 1 shilling is partially the salary of mint-workers (plus salary in coins on the asset side, plus the same amount of equity on the other side), and partially is paid out for minting materials (for example copper) (plus the price of material on the asset side, and minus the material sold on the same side). Figure 4 accounts for these changes.

FIGURE 4: ACCOUNTING FOR THE MINTING OF 1 OUNCE OF GOLD.

Mint (State)

gold coins, + 4 shillings	equity, +4 shillings

Miner

gold, -10 shillings	liabilities, -5 shillings
gold coins, +5 shillings	

Others

gold coins, +1 shilling	equity, +0,5 shillings
supply, -0,5 shillings	

The profit from coinage can be further enhanced through debasement. This happens when the Royal Mint decreases the gold content of the coins in one of the following two ways: (1) by the power of law, it announces that from now on one ounce of gold is minted into 12 shillings, but new and old coins have equal market value; (2) every old coin is reminted, i.e. one-fifth of their gold content is extracted, and the proceeds are kept by the state.
Task: account for the changes that the debasement causes in the different balance sheets!

With the general acceptance of royal coins, the introduction of minting and the legal tender laws lead to more efficient market coordination. However, there are practical and economic problems with coin usage. For security reasons, especially in wholesale commerce, transportation and shipment costs of the gold currency can be enormous. Besides that, coins wear off with intensive circulation thus the ratio between their face value (nominal value) and real value increases. As sellers have to accept coins at their face value, buyers pay with worn ones. ***Gresham's law*** prevails: *bad quality money drives out good quality money from circulation* as the former is used as a means of exchange, while the latter is used in the function of store of value.

The most important economic problem is the relative shortage of money. This occurs when the growth rate of the real economy is higher than the growth rate of gold-mining, and the velocity of circulation cannot compensate for the difference. The ***Fisher-equation*** (1) helps to understand this phenomenon.

$$MxV = PxQ \qquad (1)$$

On the right-hand side of the equation, PxQ measures the nominal value of economic transactions in a given period of time. This is the scalar product of the ordered P and Q vectors: a (p_i, q_i) pair denotes the price and the quantity of the i-th transaction, thus the $p_i{*}q_i$ product is the nominal value of the i-th transaction. The summation of all transactions ($PxQ = \sum p_i{*}q_i$) is the nominal value of all economic affairs in the examined time period.

The left-hand side of the Fisher-equation measures the monetary side of the economy. One piece of coin can mediate more transactions in a given time period ("A" pays "B" with this coin, "B" pays "C" with the same coin, etc.) The ***velocity of money*** (V) is *the average number of transactions a given coin takes part in*. Last but not least, M is the quantity of money in the economy.

In each and every atomic transaction, money and goods with the same value flow in the opposite directions, so the Fisher-equation seems to be a mere identity. However, in order to hold this identity at the end of every time period, the dynamics of the two sides have to adapt to each other. The nominal value of transactions (PxQ) can increase only if the nominal value of money circulation (MxV) follows the growth. Out of the two factors on the monetary side, the velocity of money is constrained by institutional reasons (Bordo - Jonung, 1987).

First, the structure of the economy determines how often money should change hands. An economy where production is vertically integrated requires fewer money flows than a horizontally integrated one. In the former, big companies manage whole production lines (sowing the seeds, harvesting the wheat, baking and selling the bread) without the need of money between the successive production stages, while in the latter, smaller firms

specialise in different activities and money-flows accompany the production. Though the structure of economies is changing constantly, this evolution is a slow process, and so the velocity of circulation adjusts gradually.

Second, the development level of the payment and settlement systems determines how easily money can change hands. In a pure cash economy, parties must meet physically to hand over a payment, whereas, with electronic payment systems, this process is much faster though still bounded.

> There could be special circumstances when the velocity of circulation increases beyond these boundaries. During hyperinflation money loses its function of store of value, as economic agents expecting fast growing prices try to exchange money for goods even if they do not need those goods to consume. This, of course, reinforces the hyperinflation process. However, hyperinflation or even inflation triggered by excessive growth of money supply is unimaginable in a commodity money system. The reverse case prevails: the insufficient money supply can lead to deflating prices and as agents in the hope of further decreasing prices postpone their consumption to a slowing economy.

As the velocity of money is constrained, the growth rate of the quantity of money should adapt to the growth rate of the nominal value of transactions in the long run. In a monetary system where commodity-money is the result of production (mining and minting) and the material of it is scarce, the rate of economic growth will be constrained, sooner or later, by the growth rate of the quantity of money. This situation is called the relative shortage of money.

The practical and economic problems of coin-usage described above lead to financial innovations that eventually changed the whole monetary system. First of all, in our imagined economy, private agents founded depository institutions. Anyone could depose his gold coins in one of these institutions who issued promissory notes (or simply, notes) in exchange. A promissory note is a written promise of the issuer to pay at sight the appropriate amount of gold coins. That is, whenever someone turns in a 10 shilling note, the issuer must immediately redeem it into 10 shillings in coins. Notes are not legal tender, however, as long as market participants believe that they can be converted into coins, notes - upon mutual agreement - can be used as means of exchange.

What happens when a merchant places gold coins in a depository institution? He swaps money for notes, which is a structural change on the asset side of his balance sheet. Both sides of the depository's balance sheet increase: money appears on the asset side, while the promise of paying this money out whenever it is demanded, is a new liability (Figure 5). The note is a financial instrument: it is the asset of the merchant and simultaneously, the liability of the issuer. In its physical form, it is at the merchant, the term "*promissory note*" in the depository's balance sheet is the name of this special liability. The note is a security as well, it can be traded away on the market either for money or for some goods. The promise embodied in the note is not personal: it is not the merchant but any future owner of the note who can ask for the redemption.

It is worth to emphasise that promissory notes were convertible to gold coins by anyone and anytime at face value. That is, a 1 shilling note equals a 1 shilling coin as long as

convertibility is secured and economic agents trust the promise of the issuer institution. As we will see, this promise proved to be the most important disciplinary force in the monetary system.

FIGURE 5: PLACING MONEY IN A DEPOSITORY INSTITUTION.*

Depository Institution		Merchant	
gold coins, +10	promissory notes, +10	gold coins, -10	
		promissory notes, +10	

**** as everything is measured in shillings, we do not use the term on the figures anymore***

Because of the practical problems of coin-usage, promissory notes became enormously popular. As depository institutions proved to be trustworthy, note owners did not worry about the security of convertibility. Practically, private businesses began to replace coins with notes. However, as note issuance did not change the total amount of circulating media of exchange, the problem of relative money shortage was still to be solved.

It is worth to notice that funding and managing a depository institution described above is not a profitable activity. Clerks at the offices change coins for notes and vice versa, but unless there is some cut on the deposits (say, the depository issues 99 shillings in notes for 100 shillings in coins), the costs of salaries and security arrangements will lead to massive losses in the income statement.

> **Task**: account for the changes in the balance sheets if the depository institution applies a 2 percent cut on deposited coins!

In order to gain some profit in their business, depository institutions started to lend out money. Shortly after the introduction of this new activity, managers observed a very interesting phenomenon: the bulk of borrowers, immediately after signing the contract, without even leaving the building, deposited their borrowed coins, i.e. they changed their gold coins for promissory notes. This seemingly weird act is quite understandable: debtors needed purchasing power, they wanted to buy something on the market, and for practical reasons, by that time notes have become more popular than coins. Promissory notes substituted gold coins perfectly in their functions of payment and medium of exchange, besides, their physical characteristics, such as weight and portability made them easier to use than coins. As a consequence of this observation, managers made an efficiency modification to the lending business: by signing the credit-contract, borrowers received promissory notes directly. Of course, anyone could redeem the notes into coins, but only the minority of costumers did so.

What happens to the balance sheets of the different parties? Consider first the case when costumers borrow gold coins directly (step 1), then convert them into promissory notes (step 2). Lending out gold coins is a structural change on the asset side of the creditor: coins are swapped for the promise of repayment. The latter is a new liability of the

costumer, while the coins received increase the asset side. Converting coins into notes restructures the asset side of the borrower and increases the balance sheet of the lender (Figure 6B).

FIGURE 6A: THE BALANCE SHEET OF THE DEPOSITORY BEFORE PROVIDING THE LOAN.

Bank / Depository	
gold X	promissory notes X

FIGURE 6B: ACCOUNTING FOR A 5 SHILLING LOAN IN TWO STEPS.

Bank / Depository		Costumer	
gold coins, -5 loan, +5		gold coins, +5	loan, +5
gold coins, +5	promissory notes, +5	gold coins, -5 promissory notes, +5	

Consolidating these two steps (Figure 6C) leads to the second case, which shows the result of lending out notes directly, instead of coins. Now the borrower has the notes amongst her assets and is liable to pay back the loan; the lender has the promise of the borrower as an asset and is liable to redeem the new notes into coins whenever it is demanded. The borrower and the lender swapped their IOUs: the borrower promised to pay back money according to the conditions of the credit contract, the lender promised to pay at sight the coins, upon the request of any future owner of the notes.

FIGURE 6C: ACCOUNTING FOR A 5 SHILLING LOAN IN ONE STEP.

Bank / Depository		Costumer	
loan, +5	promissory notes, +5	promissory notes, +5	loan, +5

FIGURE 6D: THE BALANCE SHEET OF THE DEPOSITORY AFTER PROVIDING A 5 SHILLING LOAN.

Bank / Depository	
gold X loan 5	promissory notes X + 5

By this act, purchasing power is *created* out of nowhere, which is very similar to the money creation in the modern banking system (see later). Initially the depository had gold coins, and after providing the loan it still has the same amount (compare Figures 6A and 6D!). Thus, technically gold is not needed to provide a loan. However, as the created notes are convertible to gold, business wise it is crucial to have enough reserves.

The introduction of lending leads to many changes. First of all, lenders are not only depository institutions anymore, we can call them banks henceforth. Second, lending increases the quantity of money circulating in the economy, which could solve the problem of relative money shortage. On the microeconomic (institutional) level, more complex managing skills are needed. Before starting to lend out money, the convertibility of the total amount of the notes was technically secured, as the amount of gold coins on the depositories' asset side equalled the amount of notes amongst their liabilities. Lending creates new liabilities (new notes) that are backed by the promise of repayment. As there is no difference between notes originating from depositing coins, and those created by lending, the whole quantity of promissory notes are now partially backed by gold coin ***reserves*** of banks, and partially by loans they granted. Reserves are generally the *highest forms of money or ultimate liquidity, for which economic agents can convert their claims on the members of the banking system*.

Assuming that all debts will be repaid, only a liquidity problem prevails: in the short run, the convertibility of notes is constrained by the level of reserves. However, there are always some debtors who default on their debt, which can lead to solvency problems in the long run. To minimise the probability of failing to keep the promise of convertibility, banks must properly examine their debtors' financial status, and price the loans accordingly.

Example: pricing loans

Let p denote the probability of default on a loan. If p is 3% then expectedly 3 shillings out of 100 shillings lent will not be repaid. Of course, banks try to minimise this probability, but even with the most sophisticated screening techniques, it cannot be beaten down to zero.

Let i_L denote the interest rate on loans. If i_L is 6% then a debtor has to pay 6 shillings of interest on a 100 shilling debt annually. If the maturity of the loan is less than one year, the debtor pays less interest (for example 3 shillings for half a year). If the debt matures over one year, this 6 percent must be paid accordingly.

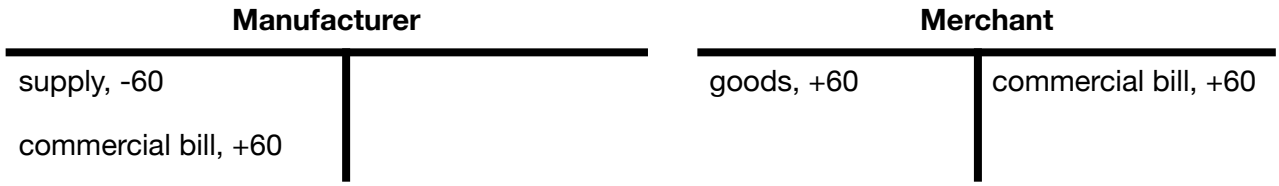
Let c_L denote the unit cost of managing a loan. If c_L is 1% then the cost of lending 100 shillings is 1 shilling. This cost involves elements such as screening and monitoring the debtor, writing the contract, etc.

Finally, r_d denotes the cost of the liability created by lending. Although banks do not borrow funds to lend, still, the liabilities (promissory notes, later deposits) incur costs: they have to be managed, sometimes interest is paid on them etc.

Considering the above data the expected profit of the bank on a 100 shilling loan is:
$100 * (1 - p) * r_L - 100 * p - 100 * c_L - 100 * r_d$.
The first term is the income of the bank: $100*(1-p)$ shillings will be repaid with interest. The second term ($-100*p$) is the credit-loss from defaults, the next term is the cost of lending 100 shillings, the last term is the cost of the created liability. Assuming perfect competition, this profit is zero if $r_L = \frac{p + c_L + r_d}{1 - p}$. If the probability of default is 3 percent, the unit cost of lending and the cost of liability are both 1 percent then the minimal interest rate on loans should be 5.15 percent.

Besides bank loans, private credit started to flourish in the economy in the form of ***commercial bill***s. A commercial bill *is a security embodying* the *short-term debt* of a merchant who is unable to pay upon delivery. Instead, he promises to pay later by signing the bill. If the supplier does not want his business to fail, he accepts the bill that he can trade away later, i.e. he can use the IOU of the original issuer (the merchant) to pay for some shipment. (See Figure 7 for accounting details.) Commercial bills are not legal tender, thus, as mentioned before, a mutual agreement between the trading partners is needed to settle business with them. The liquidity of bills depends on the trustworthiness of the issuer. If "A" wants to pay "B" with the bill issued by "C", "B" will accept this form of payment only if he believes either that "C" is able to repay when the bill matures or some "D" will accept the bill in some later transaction. The more reputable a private agent is, the more liquid his IOUs will be. However, only close business partners can rate the reputation of an agent, which makes the liquidity of commercial bills pretty constrained.

FIGURE 7: ISSUING A COMMERCIAL BILL.



Task: account for the changes in the balance sheets if the sold supply is worth 50 shillings, i.e. the manufacturer earns 10 shillings profit by selling it!

Banks can help this problem by either accepting commercial bills or buying them by issuing their own bills. The two cases differ only technically, as in the first case the bank merely signs or stamps the commercial bill (this is the "bank's acceptance"), which guarantees payment if the issuer defaults, while in the second case, the bank writes its own IOU and swaps the commercial bill for it. However, the two cases are essentially the same: the bank takes the credit risk of the issuer on its own balance sheet, replacing it with its own credit risk, which is, of course, significantly lower than the original issuer's.

It is even better if the bank buys the commercial bill by paying out gold coins or by issuing promissory notes. The former can be ruled out on the grounds discussed before: the seller of the bill will convert coins into notes immediately for practical reasons. The difference between the nominal value of the commercial bill and the discounted amount paid for it generates profit for the bank.

Historically it was a long process til the banks started to discount commercial bills by issuing promissory notes. At the first stage of this process the banks bought commercial bills by issuing their own bill, the so called banker's bill. As it was a short term security embodying the bank's debt with interest, it did not create an immediate liquidity risk. The banks managed the discounting business by matching the cash (gold coin) outflow caused by the maturing banker's bills with the cash inflow from maturing commercial bills. A minor modification to this business was when the banks started to issue the banker's

bills in standardised denomination (e.g. instead of issuing a bill with the face value of 47 shillings, the bank issued two 20 shillings bills, one 5 shillings bill, and two 1 shilling bills). Major modifications came later, when the banks issued bills without maturity and interest. This changed the nature of the above mentioned management. A banker's bill without maturity is essentially a promissory note: it is convertible to gold on demand. Thus, now the bank has to match cash outflows triggered by redemption with cash inflows from maturing commercial bills. This business could be managed by changing the discount rate, i.e. the rate by which incoming commercial bills were valued. If the net cashflow of the bank turned negative, the bank increased the discount rate by which some bill-owners were discouraged from discounting their bills.

Example: commercial bill
Suppose that manufacturer "X" has a commercial bill with the nominal value of 60 shillings issued by merchant "Y". "X" has to pay for a shipment next week but he has no money, and the bill matures only in 90 days. As he is not sure that he can pay for the shipment by trading away the bill, he decides to sell it to a bank. The bank buys the bill for 59 shillings paid out in promissory notes. The 1 shilling difference is the price of two services. First, promissory notes are more liquid than the commercial bill, so the bank provides a liquidity service. Second, the bank relieves the costumer of the credit risk. The 90-days discount rate the bank applied was $\frac{1}{60} = 1.67\,\%$.

The bank buys the commercial bill by expanding its balance sheet. The asset side is increased by the discounted value of the bill (59 shillings in the above example), while the liability side is increased by the newly created promissory notes. Upon maturity, the debtor will pay the nominal value of the bill in notes, which removes the commercial bill from the asset side and simultaneously, destroys the same amount of notes from the other side. The difference between the nominal value and the discounted value increases the capital of the bank. If the issuer of the bill fails to serve his debt then this will be a minus bill from the assets (at a discounted value) and a negative item in the capital.

FIGURE 8: DISCOUNTING A COMMERCIAL BILL.*

Manufacturer		**Bank**	
commercial bill, -60		commercial bill, +50	promissory notes, +59
promissory notes, +59	equity, -1		

**** the bank books the bill at its discounted value (59) instead of its nominal value (60). The capital gain of 1 shilling will enter the balance sheet upon repayment.***

It is worth to notice that there are two different series of events that influence the same balance-sheet structure. If "A" borrows money from a bank and pays "B" with this money for some shipment, then "A" has some goods on his asset side and a debt on the liability side; "B" swapped goods for promissory notes; the bank granted a loan (assets) by creating notes (liability). If "A" buys some goods from "B" and pays by issuing a commercial bill, after which "B" sells this bill to the bank who buys it by issuing promissory notes, the result is the same regarding the balance sheets of the three parties.

An important difference is, however, that in the latter case the private debt the bank has amongst its assets is a tradable security, while in the former case it is a contracted debt that cannot be sold easily.

2.3 Fiat money

Despite its powerful means, such as taxing private agents, determining monopoly rights over mine-products, or debasing the currency to raise revenues, the state has been the biggest debtor of the economy for a long time. Sometimes violent or innovative solutions were needed to successfully manage high levels of budget deficit or debt. Before the modern era of market-based economies, kings and queens frequently went into war against their creditors or executed them as the most efficient way of decreasing debt. These ancient methods of debt management, however, faded away over time. (For further exciting features of sovereign defaults, see Vidovics-Dancs (2014).) If debt neither can be repaid nor served in some other way, a bank is needed who willingly grants a new loan upon the maturity of the old one.

The financial affairs of our imagined world gave new ideas to the experts of the crown, which eventually led to the profound transformation of the monetary system. The state convinced some wealthy businessmen to found a new bank (henceforth: the central bank) and lend the raised capital to the state. In exchange, the state endowed this new bank with the right to issue legal tender in the form of banknotes. Banknotes are the liabilities of the central bank.

The significance of this change is worth to emphasise. Before the founding of the central bank, gold coins and promissory notes issued by different banks were circulating in the economy, but only the coins were legal tender. The acceptance of the notes issued by any bank depended on trust, i.e. on the belief that the bank is ready and able to convert them into coins. The acceptance of the central bank's banknotes is, however, forced by the law, which makes these notes and the coins equivalent. Banknotes are called ***fiat money*** because they *do not have intrinsic value* end yet *must be accepted* by the rules of the state. Fiat money can be distinguished from *fiduciary money,* that does not have intrinsic value either, but its acceptance depends on trust and not on power (Mises, 1912).

Besides the power of the law, economic incentives were introduced to create demand for the new currency. First, the crown and the central bank promised to convert banknotes into gold coins on demand. Second, the state made it possible to pay taxes with banknotes. Two schools of thoughts are connected to these incentives: *Metallists* think that money is valuable as long as it has gold (or silver, or copper, etc.) content or it is convertible to gold at a fixed rate; *Chartallists* think that the power of the state gives value to the money: if it is not only possible but obligatory to pay taxes in banknotes, then it creates demand for the new currency (for details, see Bell, 2001).

The perceptive reader can observe the sentence "*I promise to pay the Bearer on Demand the sum of Ten shillings."* on the English 10 shilling banknote (Figure 9). Once it meant the

mutual guarantee of the Queen / King and the Governor of the Bank of England to redeem this note at sight to 10 shillings worth of gold. The nature of this promise has changed over time: today it means that the State and the Central Bank guarantee to safeguard the purchasing power of money.

FIGURE 9: THE ENGLISH 10 SHILLING BANKNOTE.*

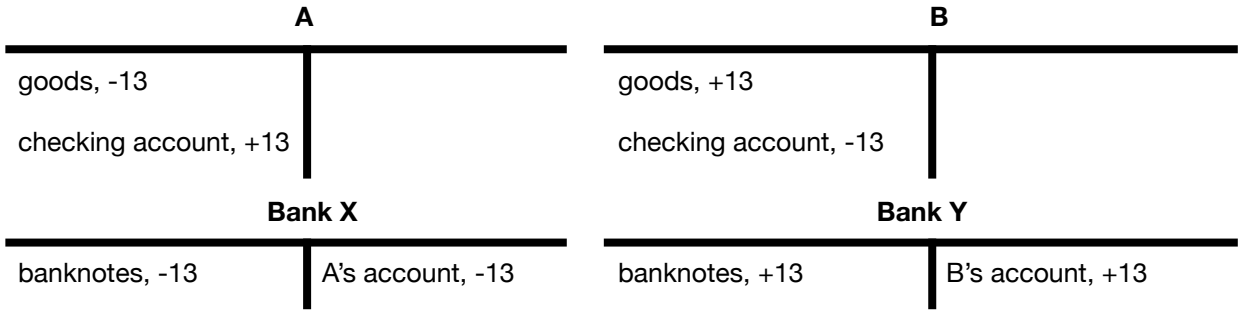
**** Source: worldbanknotescoins.com***

The introduction of the new paper currency initiated significant changes within the banking system. Ordinary banks gradually substituted their gold coin reserves with banknotes issued by the central bank, which was rationalised by the convertibility of banknotes to gold coins. As economic agents (and banks as well) could redeem paper currency to gold at the central bank, it was enough for the banks to hold the central bank's banknotes as reserves. If costumers wanted to convert their promissory notes issued by some bank to gold, they could do it in two steps: first, they applied for redemption to the issuer of the promissory note, then they could convert the given banknotes to gold at the central bank.

For the recommendation of the central bank, members of the banking system made their costumers use checking accounts instead of promissory notes, thus promissory notes gradually vanished from circulation. A checking account is a liability of a bank that can be used for withdrawing cash (central bank notes in our case) or, with the help of a checkbook, for buying goods. In the latter case, the account holder gives a signed preprinted check to the seller of the good. Later, the seller can turn in the check to his own bank who deposits the amount written on the check on the seller's account.

Example: checking accounts 1
"A" has a checking account at bank "X", and "B" has a checking account at bank "Y". "B" buys goods from "A" for 13 shillings and pays with a check. "A" turns in the check to bank "X", who deposits 13 shillings on "A"'s account then contacts bank "Y". "Y" debits "B"'s account with 13 shillings and pays bank "X" in banknotes. See Figure 10 for accounting.

FIGURE 10: SETTLING BUSINESS WITH CHECKING ACCOUNTS.



Example: checking accounts 2
Suppose that there are 5 banks in the economy, X1, X2, X3, X4, and X5. By the end of some business day, the banks have collected checks from their costumers according to the table below (Table 1).

TABLE 1: BILATERAL GROSS POSITIONS OF BANKS BEFORE SETTLEMENT.

	X1	X2	X3	X4	X5
X1	3 912	6 021	4 700	1 976	5 728
X2	5 110	6 230	2 918	7 504	1 635
X3	1 132	907	8 702	6 482	7 203
X4	3 374	8 320	439	3 800	854
X5	8 827	2 835	3 325	5 585	1 677

Checks written by the customers of a given bank are in the rows of the table, while those given to them are in the columns. The amounts in the diagonal are checks written by and given to costumers of the same bank, while other cells are containing checks written by and given to customers of two different banks. The amount 1,976 in the fourth cell of the first row thus means that at the end of the day, bank X1 is due to pay 1,976 shillings in banknotes to X4 (or X4 is due from X1 1,976 shillings). The sum of all the numbers in the table is the total nominal value of the real economic transactions settled by checks. This amount is 109,196 shillings. Instead of paying the gross amounts, the banks first net out their dues and pay according to their net positions (see Table 2).

The due to bank X1 (22,355 shillings) is the sum of the first column of Table 1, the due from X1 (22,337 shillings) is the sum of the first row. X1's net position is 18 shillings, which is a claim on other banks. X3's net position is -4,342 shillings, which is a debt payable to other banks. The sum of the net positions is zero, as someone's revenue is always someone else's expenditure. The sum of net cash flows is 9,494 shillings which is less by a factor of 11 than the amount of gross cash flows (109,196) would have been. (9,494 is the absolute value of debts [4,342+5,152] or claims [18+916+8,560].)

TABLE 2: GROSS AND NET POSITIONS OF THE BANKS.

	Due to	Due from	Net position
X1	22 355	22 337	***18***
X2	24 313	23 397	***916***
X3	20 084	24 426	***-4 342***
X4	25 347	16 787	***8 560***
X5	17 097	22 249	***-5 152***

The final settlement among the banks can be arranged in two ways. Either debtors (X3 and X5) put the proper amounts in a box from where creditors (X1, X2, and X4) can take their claims, or they can ask the central bank to debit and credit their accounts. Having accounts at the central bank is another financial innovation that helps payments between members of the banking system.

FIGURE 11: FINAL SETTLEMENT USING CENTRAL BANK ACCOUNTS.

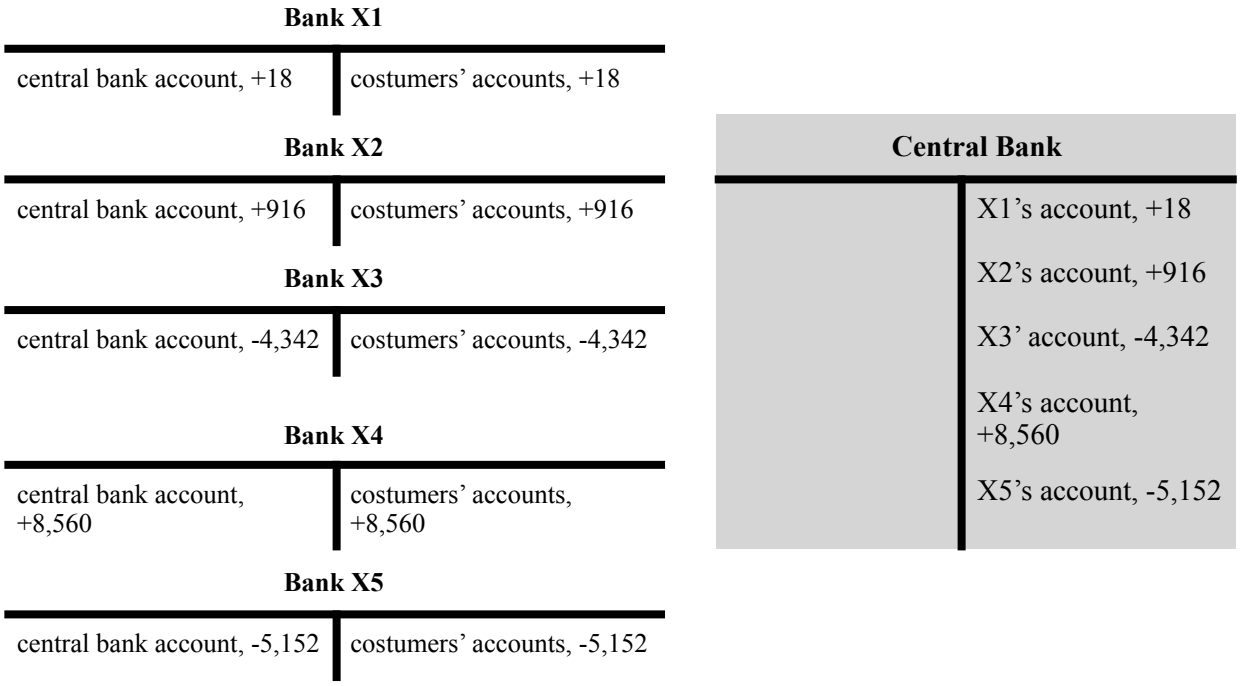
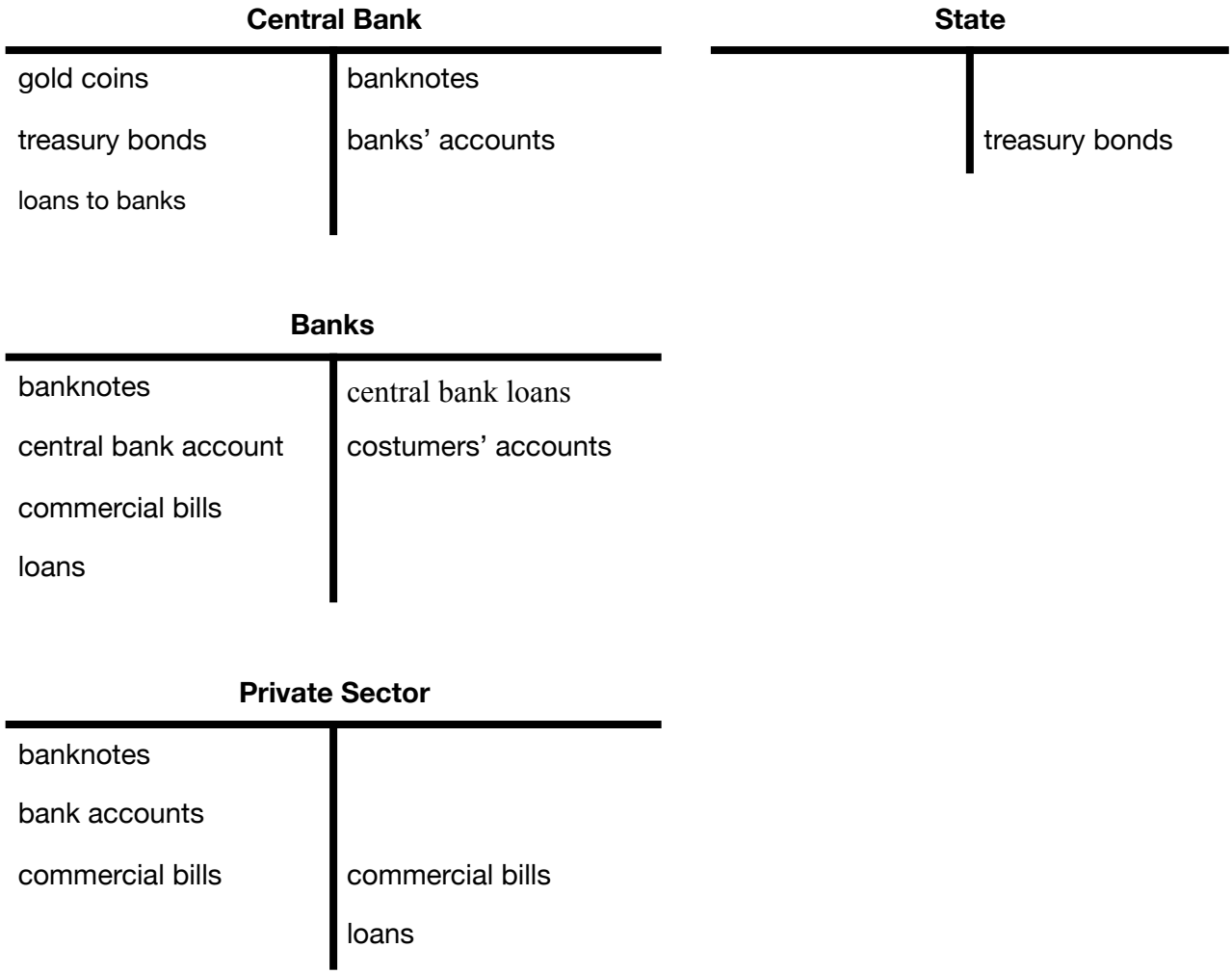


Figure 11 shows the net result of payments and settlements. At the banking level, individual costumers' accounts are credited and debited according to the checks turned in.

With the foundation of the central bank, a hierarchical monetary system evolved (Mehrling, 2013). On the top level of this hierarchy, the central bank serves as the bank of the crown (the state) and of other banks. Gold currency is still the highest form of money, but now it has started to lose its function of medium of exchange, as economic agents use either banknotes or checking accounts for the payments. The liabilities of the central bank (banknotes and current accounts of other banks) are the second highest forms of money, while checking accounts at banks are the third, commercial bills are the fourth in

this row. The second level of the institutional hierarchy is composed of the banks who keep banknotes and central bank accounts as reserves, and their liabilities (checking accounts) fulfil the role of medium of exchange in the economy. Though in a constrained manner, certain liabilities (commercial bills) of real economic agents can substitute higher forms of money in the function of exchange. (See Figure 12 for accounting details.)

FIGURE 12: THE HIERARCHICAL MONETARY SYSTEM.



The ultimate phase in the evolution of this system was the suspension of convertibility of banknotes to gold, which was triggered by a series of monetised budget deficits. As mentioned at the beginning of this section, the crown founded the central bank in order to create an institution at hand who willingly finances current deficits and maturing debts. The method of financing is the purchase of treasury bonds issued by the state. The central bank buys these bonds by printing banknotes, by which the state gains purchasing power to spend. As long as the additional money supply does not exceed the money demand generated by economic growth, this can be considered a proper and rightful procedure (Bánfi-Hagelmayer, 1989). It is worth to notice that whenever a private agent borrows from a bank, who grants the loan by issuing promissory notes or depositing the checking account, technically the same thing happens: some deficit is financed by newly created money. Considering the special relationship between the state

and the central bank, however, the two cases are not identical. Private agents face strict rules, they are screened, monitored, asked for collateral, and their loans are priced properly, whereas financing the state by the central bank does not involve such scrutiny. If a private debt matures, in most cases, it has to be repaid, whereas a maturing public debt can easily be rolled over.

The differences might be explained by the monopoly right to issue legal tender, a privilege the crown has given to the central bank. Holding the highest level of liabilities in the monetary hierarchy generates profitable business for the central bank, which is lending money to other banks facing liquidity shortage. A liquidity problem arises at the second level of the banking system when too many customers want to withdraw money by either having commercial bills discounted, or redeeming promissory notes or checking accounts into central bank notes, or when the bank is unable to settle at the end of the clearing process. In all of these cases, a short-term liquidity loan from the central bank can solve the bank's problem. As the central bank is the monopolistic supplier of legal tender notes, it can charge high interest rates for these loans. (Remember, the central bank was founded by private businessmen, so originally and for a long time thereafter, it was a profit-seeking entity.)

Public deficit financed by the central bank threatens this business by increasing the quantity of banknotes in circulation and can lead to growing prices in the economy. Inflation is partially caused by the direct spending of the state, and partially by a credit bubble generated by the loosening monetary conditions. As we have seen in the example about the checking accounts and the clearing system, 1 unit of central bank money can generate transactions in multiple units of value. Consequently, an additional unit of banknote created by the central bank, spent by the state and deposited by an economic agent in a bank, can generate multiple units of money on checking accounts. On the basis of the additional reserve, banks can lend more without having to worry about their own liquidity. The macroeconomic demand is increased through two channels: both the state and private agents can finance their excessive spending by debt. The state spends central bank money directly, private agents spend the money created by members of the banking system on the basis of the additional liquidity spent by the state.

If macroeconomic supply cannot adapt to macroeconomic demand, prices start to increase. What private agents see in this situation is that the purchasing power of their checking accounts or banknotes is falling. As checking accounts are redeemable to banknotes, and banknotes are still convertible to gold coins, and the nominal gold content of banknotes, so the price of gold is unchanged, more and more agents will turn their balances at banks into banknotes, and banknotes at the central bank into gold coins. As banknotes are created, the possibility of converting deposits (checking accounts) into banknotes is (with the help of the central bank) infinite, however, as the gold reserves of the central bank are finite, convertibility on the highest level of the hierarchy sooner or later must be suspended.

Historically, excessive public deficit and debt led to the suspension of convertibility. However regarding that the money demand of the growing economy can only be met with money-creation, sooner or later the gap between the gold-reserves of the banking system and the amount of the outstanding money-of-account would have been so huge, that convertibility could not have been hold.

2.4 Modern monetary systems

The monetary system evolved by the end of this (imagined) story is similar to the monetary systems operating in modern economies. There are some differences, of course. Central banks cannot finance the state directly anymore, although they can trade with treasury bonds on the secondary market, i.e. they can buy and sell treasuries that have been already issued. Instead of gold, the most important monetary reserves held by central banks are key currencies, such as the US Dollar, the Euro or the British Pound. Convertibility of national currencies is the possibility of exchanging them to other currencies, especially to those listed above.

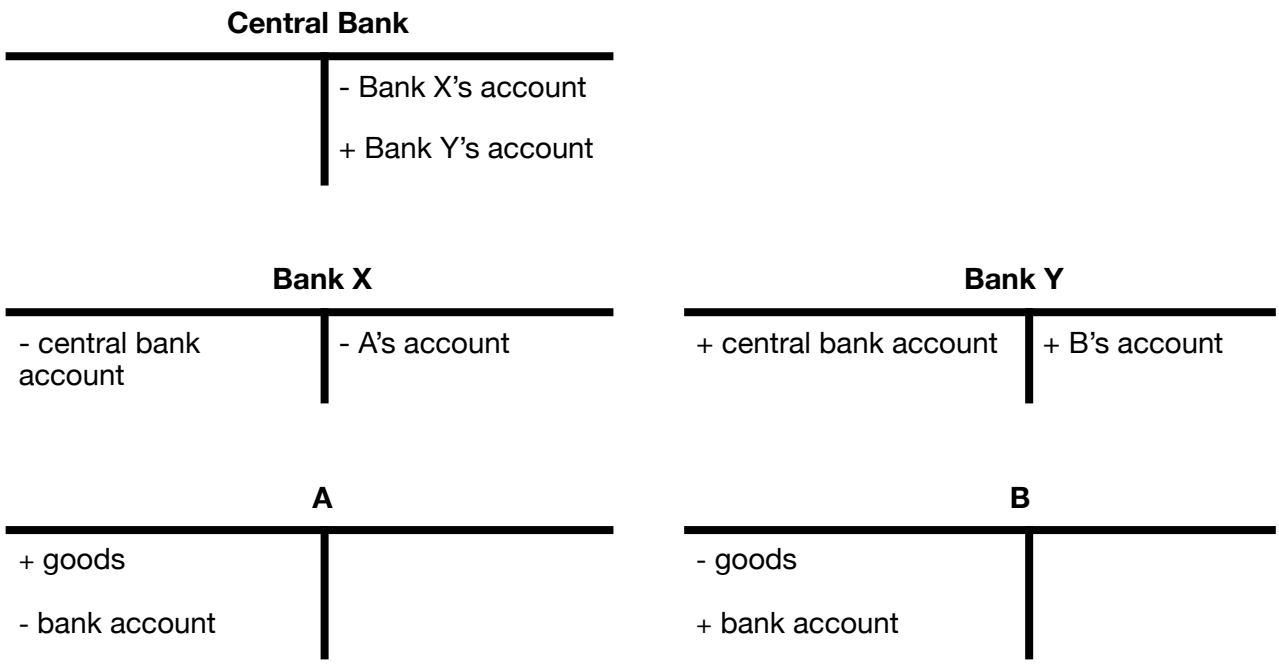
The convertibility of the home currency to foreign currencies constrains the ease with which the public deficit and debt can be (re)financed per se. Although technically the central bank can still help serving the debt by creating money, as this money is convertible to foreign currency, the excessive debt can easily lead to a currency crises.

Industrial and technological development has changed the payments and settlements systems. Economic agents holding deposit accounts in banks are using plastic cards instead of checkbooks. However, the essence of a payment between trading partners is still the ***parallel settlement***. *If "A" pays "B"* - except for the case when they are the costumers of the same bank - *two payments are generated simultaneously*. The account of "A" is debited by his bank, and the account of "B" is credited by her bank; and parallel with this payment, central bank money flows between the accounts of the two banks (Figure 13).

The transaction between "A" and "B" is successful if two conditions are fulfilled. First, the budget constraint is held, that is, "A" needs to have enough money on his account to pay for the goods or services provided by "B". Second, the liquidity constraint is held, that is, "A"'s bank needs to have enough central bank money on its account to pay "B"'s bank.

> On the international level, a single payment generally needs more than one liquidity conditions to hold. If a Mexican supplier sells tequila to a Hungarian liquor-store, there are three parallel payments: (1) the store pays the supplier, (2) the Hungarian commercial bank pays the Mexican commercial bank; (3) the Hungarian central bank pays the Mexican central bank. There are other possibilities to settle this business, but discussing them is beyond the scope of this chapter.
> **Task:** account for the international payment discussed above.

Figure 13: Parallel settlement.

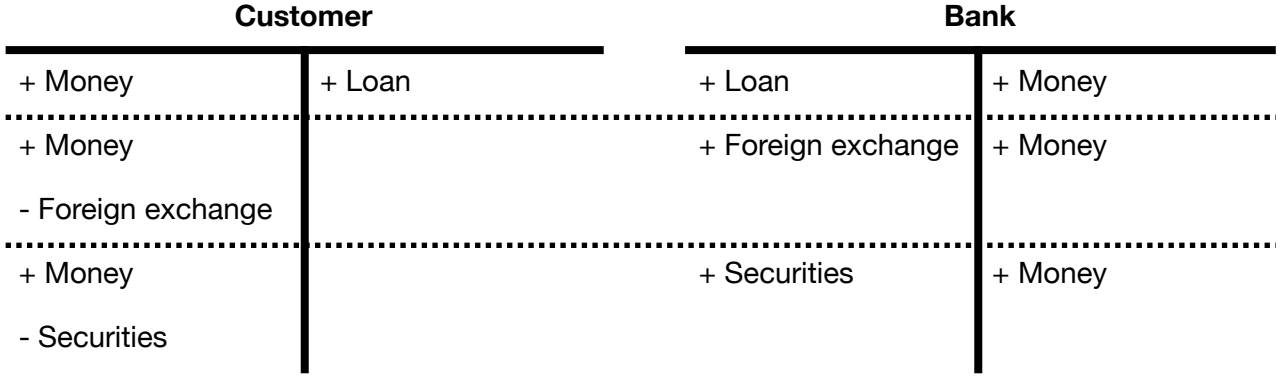


Money is created by the members of the banking system in three ways:

- *by granting loans either in a non-securitised or in a securitised form,*
- *by buying foreign exchange,*
- *by buying securities on the secondary market (i.e. issued securities).*

In all of these cases, the bank credits the current account of its costumer with the proper amount (i.e., with the amount of the loan, with the price of the foreign exchange or the securities bought; see Figure 14). The customer might be a household or a company. In these cases, a commercial bank from the second level of the banking system creates money. However, the costumer might be a commercial bank as well, in this case, it is the central bank who creates money. Out of the three methods of money creation, buying foreign exchange and buying securities are secondary. Foreign exchange is primarily created by foreign banks, by granting a loan, securities represent the debts of some economic agents. Ultimately, every unit of money is created through the indebtedness of someone.

Figure 14: The three ways of money creation.



When the reverse of these actions happens, i.e. a customer repays a bank loan, buys foreign exchange or securities from a bank, money is destroyed.
Task: account for the reverse case!

Commercial banks can create only "money of account", whereas the central bank can create cash (banknotes and coins) as well. It is rare, however, that a bank applies to the central bank for a loan granted in cash. Instead, the central bank credits the account of the bank that can be converted later into cash. Cash and central bank accounts of banks differ from each other only in their physical form, as they are both the liabilities of the central bank.

Creating money is not constrained technically, however, it generates liquidity risk, as additional deposits might be withdrawn at any time. Commercial banks manage this risk by different protective measures. First of all, they hold cash reserves and deposits at the central bank. These are the highest forms of liquidity as costumers can redeem their deposits into cash directly or into central bank accounts indirectly through the payments system. Commercial banks can, and in most cases do, hold short-term treasury bills. T-bills are liquid securities that can easily be sold on the interbank market or used as a collateral for an interbank loan or for a central bank loan.

Task: account for a liquidity loan on the interbank market!

There is an important difference between borrowing money from another bank and requiring liquidity from the central bank. Banks can provide liquidity only to the extent of their excess reserves, whereas the ability of the central bank to create reserves is theoretically infinite. Because of this, the central bank is the ultimate liquidity backstop in the banking system, frequently called the lender of last resort. To avoid moral hazard problems – e. g. when banks do not care about their liquidity because the central bank can always help –, this role must be regulated. According to ***Bagehot's Rules*** (Bagehot, 1873), named after Walter Bagehot, an English businessman of the 19th century, *the central bank should:*

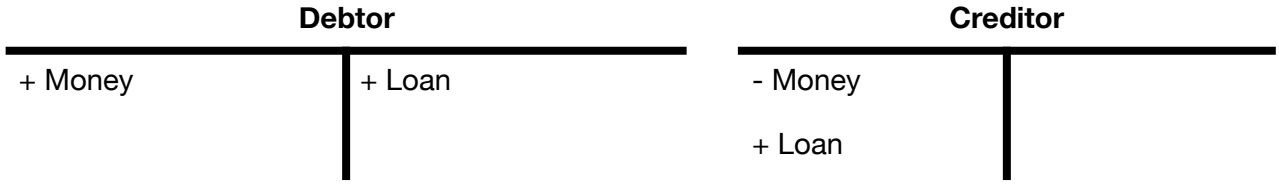
- *provide liquidity freely (without hesitation) to solvent banks,*
- *at sound collateral,*
- *at a high interest rate.*

Liquidity provision is important because a failure in the payments system could lead to a chain reaction and to a full-fledged systemic crisis. Banks and economic agents are interlocked through their balance sheets, so if "A" cannot pay to "B" because of a liquidity shortage at "A"'s bank, then "B" cannot pay "C", "C" cannot pay "D", and so on. However, it must be controlled for whether the bank has only mismanaged its short-term cash flows, i.e. it could not match cash (reserve) outflows with inflows temporarily, or it has an outstanding problem in the long run, i.e. it is insolvent. In the latter case, the bank will apply for liquidity again and again because its assets are worth less than its liabilities. The "sound collateral", in most cases, consists of some government security, however, in a wider sense any good quality asset can be referred to by this term. High interest rate on

the liquidity loan is required to have the bank feel the consequence of its mismanagement.

Every economic agent can provide money-redistributing loans, i.e. lend out money accumulated as an asset (Figure 15). This is - typically - a loan between the members of the same level of the monetary system, where the money lent had been created by an institution of some higher level before. A household, for example, can lend cash created by the central bank or money from its bank account (created by a bank) to another household or to a company. A bank can lend money from its central bank account to another bank. Money-creating loans are provided between levels of the monetary system, i.e. a bank can provide a loan by creating a bank deposit to a household or a company, or the central bank can grant a loan by creating a central bank deposit (or cash) to a bank. Money-redistributing and money-creating loans differ technically: the former restructures the asset side of the creditor whereas the latter increases both the assets and the liabilities of the creditor (Riesz, 1980). Another difference is that providing money-redistributing loans is constrained by the amount of money the creditor has, while money-creating loans can be (technically) granted infinitely.

FIGURE 15: PROVIDING A MONEY-REDISTRIBUTING LOAN.



Markets to provide liquidity are created on several levels of the monetary system, by deficit and surplus agents. Amounts paid for shipments in wholesale trade can be large enough not to let them rest on suppliers' accounts but to be tied up in the form of short-term deposits or short-term securities. Banks must be ready to provide solutions for surplus agents' problems (i.e. excess liquidity) in the short run because, on the other hand, deficit agents (for instance, firms who have to finance their production before they could realise revenues from sales) apply for liquidity loans. When a loan is granted, money is created and serves as the liquidity of the debtor. Thus, the bank does not lend out its own reserves, but when the debtor spends this money, reserves are pumped out of the bank. This liquidity, as mentioned above, can be re-borrowed on the interbank market, however, it is better if it comes back indirectly or even does not leave the bank. If some costumer of Bank A deposits excess money balances in Bank B then a reserve flow between the two banks is generated. In this wise, banks can compete for liquidity by creating attractive saving instruments. It is worth to emphasise that the subject of this competition is not the deposit itself. Deposits are liabilities of banks, and who wants to be liable? It is the additional liquidity – generated by the excess deposits – that the banks compete for, so that they can lend more on its ground.

Task: account for a bank loan in the next two cases:
- ➡ first, the business partner of the debtor is the costumer of another bank;
- ➡ second, the business partner of the debtor is the costumer of the same bank.

Considering these two cases, how can the bank decrease liquidity outflows generated by lending?

At the given level of available financial techniques and management skills the banks can estimate the secure ratio between their liquidity reserves and deposits. This ratio is not static, it can be influenced by the business cycle or by events in the calendar year. In recessions, for example, economic agents keep more cash, thus banks need to increase their reserves. Or around Christmas cash-usage is more frequent, making the banks once again keep more liquidity. Let alone these considerations, the supposed secure ratio of liquidity reserves determines how much money banks can create by lending. This ratio is well below 1, which leads to ***money multiplication***: *if the central bank creates one unit of central bank money* (which serves as the reserve of banks), *then banks can create multiple units of deposits according to the secure reserve ratio*. The following examples shed light on money multiplication.

Suppose that Bank X calculates that the secure ratio between its reserves and deposits is 10%. If at the moment the bank has 100 units of reserves and 800 units of deposits, then it can create by lending 200 more units of deposits without having to worry about liquidity problems.

Suppose that in the banking system the average of secure reserve ratios calculated by individual banks is 10%. If at the moment the banking system has 10,000 units of reserves and 70,000 units of deposits, then on the systemic level 30,000 more units of deposits can be created by lending.

Suppose that in the banking system the average of secure reserve ratios calculated by individual banks is 10%. At the moment the banking system has 10,000 units of reserves and 100,000 units of deposits, thus at the given level of reserves no more deposit creation is possible on the systemic level. However, if the central bank creates 500 more units of reserves by lending central bank money to the members of the banking system, it allows the creation of 5,000 more units of deposits.

The driving force behind the evolution of monetary systems has always been some problem that was solved by a formal, a procedural or an institutional innovation of profit seeking individuals (Table 3) or by the state. But even the state pursued selfish goals when introduced changes without the intention to manage some systemic imperfection. Still, the interplay of all these microlevel (re)solutions effected a system that fulfils two macroeconomic functions, namely providing the possibility of payments and settlements and connecting savings with investments.

TABLE 3: EVOLUTION OF MONETARY SYSTEMS.

monetary system	problems	solutions		
		formal	institutional	procedural
barter*	inefficient, probability of double coincident of wants is low	generally accepted means of exchange		
commodity money (gold coins)	practical: wears off, insecure, heavy	promissory notes in exchange for gold coins	depository institutions	promissory note - gold coin conversion
	economic: relative money shortage	promissory notes in exchange for the promise of repayment	banks	creation of promissory notes by lending
gold coins promissory notes (fiduciary money) created by banks	public deficit and debt	banknotes	central bank	financing the state by issuing banknotes
	practical: promissory notes and banknotes are seemingly identical	checking accounts (replacing promissory notes)	clearing system	clearing
		central bank accounts of banks		
	(possible) liquidity shortage of banks			central bank lending according to the Bagehot rules
gold coins checking accounts created by banks banknotes created by the central bank	inflation caused by overrissuance of banknotes	suspension of banknotes - gold coins convertibility	independent central bank (no deficit financing)	
two level banking system central bank money bank deposits	asset price bubbles caused by loose credit conditions *procyclical banking behaviour* too big and too important to fail problems**		uniting central banks and prudential authorities *Bank Union (EU)* international supervisory bodies	introduction of Basel III rules

* never existed

** discussing these issues is beyond the scope of this book

The financial position of an agent can be defined as the difference between its incomes and expenditures in a given time period. As someone's income is simultaneously someone else's expenditure, the aggregated financial position of all economic units is always zero ex post, which gives rise to perfect financial intermediation, i.e. deficits can be financed by surpluses completely. However, the preferences of surplus and deficit agents differ by more dimensions. The typical deficit agent is a private business who wants to fund investments in physical capital. The typical surplus agent is a household, who saves some part of its monthly income. Capital investments pay off expectedly in the long run, sometimes with non negligible risk of default. Households save for the short run, and they are typically risk-averse, i.e. they fear loosing even a tiny part of their savings. All these differences make it almost impossible to connect savers with investors directly.

The loanable funds theory states that financial intermediaries collect savings and fund investments thereby connecting surplus agents with deficit agents. It is easy to see, that there is a logical error in this claim if we want to apply it on the macroeconomic level. To fund deficit agents, the intermediaries - according to this theory - must collect the surpluses first. However, the surpluses are the results of excessive spending of deficit agents, who cannot spend until they borrowed these surpluses. The intermediary cannot lend something now that can be only collected later.

The banking system that creates money provides solution both to the problem of connecting savers with investors and to this logical error. It is enough to take a look at the balance sheet of a bank to see how the funding problem is solved. The bulk of the loans granted by the bank are long term investment loans, whereas the bulk of the deposits are sight deposits and time deposits maturing at most in one or two years. The bank separates the credit risk of the debtors from the liquidity risk of the depositors by holding capital reserves and liquidity reserves on the two sides of its balance sheet. In normal circumstances the expected liquidity loss and the expected capital loss are low, thus the bank can keep the liquidity reserves to deposit ratio and the capital reserves to loans ratio around 10-20% (see Figure 15).

FIGURE 15: THE BALANCE SHEET OF A TYPICAL BANK IN NORMAL CIRCUMSTANCES.

Bank

Liquidity reserves	10	Deposits	85
Loans	90	Capital reserves	15

The dynamics of this balance sheet are as follows.

- *money creation by lending: loans and deposits increase*
- *destroying money when the loan is repaid: loans and deposits decrease*
- *realisation of credit risk: loans and capital decrease*
- *realisation of liquidity risk: deposits and liquidity decrease*

Figures 16A and 16B help to understand how the banking system connects deficit agents with surplus agents through money creation. First, the banking system grants loans to surplus agents by creating deposits (money). Next, deficit agents spend this money and buy goods from surplus agents. This spending is the income of surplus agents, which until spent is by definition their savings. At the end the investments (loans) of deficit agents are financed by the savings (deposits) of surplus agents, the banking system intermediates between them, however the deposits were not collected but created.

FIGURE 16A: FINANCIAL INTERMEDIATION IN THE BANKING SYSTEM.

Deficit Agents		Banking system		Surplus Agents	
+Deposit	+ Loan	+ Loan	+Deposit		
-Money			-Deposit	-Goods	
+Goods			+Deposit	+Deposit	

FIGURE 16B: THE RESULT OF THE FINANCIAL INTERMEDIATION.

Deficit Agents		Banking system		Surplus Agents	
+Goods	+ Loan	+ Loan	+Deposit	+Deposit	
				-Goods	

Bibliography

Bagehot, Walter (1873): "*Lombard Street: A Description of the Money Market.*" Henry S. King & Co., London.

Bánfi, Tamás - Hagelmayer, István (1989): "*Pénzelmélet és pénzügypolitika*" ("*Theory of Money and Monetary Policy*"). Közgazdasági és Jogi Könyvikadó, Budapest.

Bell, Stephanie (2001): "*The role of the state and the hierarchy of money.*" The Cambridge Journal of Economics, 25, pp.: 149-163

Bordo, Michael D. - Jonung, Lars (1987): "*The long-run behavior of the velocity of circulation: The international evidence*", Cambridge University Press, Cambridge.

Mehrling, Perry (2013): *"The inherent hierarchy of money."* In: Social fairness and economics: economic essays in the spirit of Duncan Foley. London, Routledge, pp.: 394-404

Mises, Ludwig von (1912): "*The Theory of Money and Credit.*" Available at: https://mises-media.s3.amazonaws.com/Theory%20of%20Money%20and%20Credit.pdf?file=1&type=document

Polányi, Károly (1944): "*The Great Transformation. The Political and Economic Origins of Our Time*". Beacon Press, Boston.

Polányi, Károly (1957): "*The semantics of money uses.*" In: Dalton, G. (ed): "Primitive, archaic and modern economies: essays of Karl Polanyi", Garden City, Doubleday. pp.: 175-190

Riesz, Miklós (1980): "*Pénzforgalom és hitel*" ("*Money Circulation and Credit*"). Tankönyvkiadó, Budapest

Vidovics-Dancs, Ágnes (2014): *"Az államcsőd költségei régen és ma" ("Costs of sovereign defaults now and long ago")*. Közgazdasági Szemle, LXI., March, pp.: 262-278.

Wray, Randall L. (1993): "*The Origins of Money and the Development of the Modern Financial System*". Levy Institute, Working Paper No. 86

CHAPTER 3 *BANKING OPERATIONS*

3.1 Passive banking oparations

The expression of ***passive banking*** transactions derives from the early stage of bank history, when their main activity was custody services, for which a fee was charged. At that time banks were really condemned to passivity, so this name became widespread. Even today, the *operations that modify the extent and composition of liabilities* – regardless of the fact that they express debts and liabilities – are called "passive"

Deposits

Deposits create the suitable liabilities for banks, which the credit and other income-generating activities are based on. The credit institutions' own funds perform only security functions.

Deposits are distinguished according to the right of disposal:

- *bearer deposit,*
- *registered deposits.*

Bearer deposit are *unanimous*, *the holder of the proper documents can dispose over the deposited amount*. In case of ***registered deposit****s* only *the depositor and his/her representative may have disposal of registered deposits*.

According to the currency of the deposit:

- *in domestic currency,*
- *in foreign exchange, currency.*

In terms of repayment period:

- *short (shorter than a year) term,*
- *medium- (between 1 and 5 years) term,*
- *long- (longer than 5 years) term.*

According to interests:

- *fixed,*
- *variable,*
- *convertible rate,*
- *formally non-interest bearing deposits.*

We talk about ***fixed rates*** if during the repayment period *the size of the interest rate is defined in advance*. In case of ***variable-rate deposits*** *the initial interest rate is bound to some kind of reference rate* (e.g. key interest rate) and the initial interest rate is changed during maturity only in case of its change. In case of ***variable-rate deposits*** *the bank may change the initial interest rate unilaterally under certain conditions*. Formally non-interest bearing deposits are the discount liabilities.

According to the maturity of deposits:

- *demand deposit,*
- *deposits redeemable at notice,*
- *term deposit.*

In case of ***demand deposits***, *the depositor disposes of the account without any restrictions at any time without loss of interest*. The transfer of this amount is uncertain for the credit institution; therefore very low interest rate is paid on them. ***Deposits redeemable at notice*** are transition between demand deposits and term deposits. *The client may terminate the deposit at any time but the credit institution has to affect the order only after a certain time* (usually three days) depending on the contract. ***Term deposits*** *cannot be withdrawn maturity, or only at the price of losing interest*.

The payment account or simply bank account is used for the execution of payment transactions. Payment accounts are on the liabilities side of the bank balance, just as deposits. However, while in case of the deposit, the amount of the deposit is only for the disposal of funds (due to this fact the turnover of the deposit is small), the payment account is used for the transactions of the account holder.

If this account the bank accomplishes overpayment in addition to the balance, i.e. grants credit to the client, the account is called current account. The size of the available credit which can be taken automatically, without a separate application, is usually a certain percentage of the account balance (relative limit) or a specific amount limit (absolute limit).

The bank account management does not only mean booking incomes and expenses of the account merely, but a range of other related services are also included. For example, management of general ledger accounts, holding the disposal of the account holder credits granted on the current account, the so-called overdrafts, check and bill of exchange transactions, management of separate accounts, statistical report management for the monetary authorities, etc.

So if the account management is not too promising to the banks, the expected opportunity of connected transactions makes this banking transaction attractive for them after all. The interest benefit is added to this if the bank does not pay interest on the account balance. It mainly occurs in countries with low inflation rate.

Issue of securities
Besides taking deposits, the emission of the bank's own securities is a growing way of funds raising. Issuing of short-term securities generally has a liquidity motive, therefore their majority has a maturity of 90, 180 and 270 days. Many varieties of short-term securities are known, such as savings certificates, bills, etc. In general, they are fixed rate ones, but rarely flexible rate ones also occur. Interest payments are usually made at maturity, however, issuing on discount price is also known. Certificates of deposit bridge the gap between the short-term and long-term securities. The aim is the acquisition of sources aiming at financing permanent current assets with the emission of middle-term securities.

The purpose of issuing long-term securities is financing long-term investments. Its most widespread forms are bonds, savings certificates and letter of hypothecation.

The savings certificate can be bearer or registered, fixed or floating rate, or with bound denomination. In most cases, interest is paid at maturity, as compound interest. The savings certificate can be found in discount variations, as well. It is not possible to withdraw money from the note principal prior to the expiration.

The letter of hypothecation is fixed-rate security, of which coverage is created by lien on property. It makes possible to meet the lasting demand for credit of the agricultural and construction industries. Behind the letter of hypothecation there is therefore a land and real estate collateral, so the owner's claim is insured not only by the repayment promise of the credit institution but also by the collateral.

The great benefit of securities emission is that if the investor needs his capital earlier than it was excepted, it is not needed to be redeemed at the issuing credit institution, but it can be sold in the secondary securities market. It can provide long-term resources for the credit institutions without having to be afraid of repayment prematurely.

Central bank loans
If short- or long-term lack of resources arises at credit institutions, they can turn to the central bank for a refinancing loan. Banks can obtain collateralised loans only by depositing eligible securities at the central bank. The owner of the securities is still the commercial bank, so the amount of credit received for it increases the balance sheet total. Central bank money enters the asset side and there central bank credit (liabilities) enters the liability side.

The so-called ***repo*** transactions can be regarded as collateralised loans as well. Their essence is that the business bank *sells eligible securities to the central bank and and agrees to repurchase them at a predetermined price at a predetermined future date*. Thus the repo is the combination of a spot selling and a forward buying of the same security. The difference between the spot and the future price reflects the repo rate. In case of a ***reverse repo*** *the bank buys spot and sell forward the same securities*.

In case of ***rediscounting*** the *central bank purchases bills discounted by banks earlier*.

In case of loans offered beside foreign exchange coverage, the business bank takes out a loan in domestic cash equivalents to the amount of its foreign exchange deposit (which is its collateral at the central bank). ***Foreign exchange-swap*** transaction can be mentioned here, when the *business bank sells a kind of foreign exchange to the central bank for another particular foreign exchange, and the business bank repurchases it simultaneously for a specific period, on a predetermined rate*. So, in practice, it means a spot purchase and a forward selling, or the other way round, spot selling and forward purchase. T

Interbank borrowing
Compared to the corporate sector, the short-term liquidity in the banking sector should meet much more severe requirements. Companies and individuals can be some days late with their dues without a relatively greater risk, illiquidity of a bank can lead to a panic attack among depositors.

Therefore, banks often have to take intraday loans from each other in order to ensure their liquidity. Interbank credits have some basic specialties compared to the corporate credit market.
In general, the vast majority of interbank loans is short-term:

- ***overnight loans****, where the conclusion of the contract and taking out the loan happen on the same day, the repayment of the loan is a working day later.*
- ***tomnext*** *(tomorrow-next), when the borrowing takes place one day after the conclusion, and the loan repayment the day after*
- ***spotnext****, where the granting of credit happens two days after the contract has been made.*

Of course, the participants of the interbank market may reach an agreement about any kinds of constructions. Thus, interbank loans actually have no time limit, but due to the high interest rates, they are mainly used in case of liquidity problems, for a few days duration.

The second feature lies in the technique of contracting. While in case of corporate loans, literacy has an almost exclusive role, here oral agreements dominate, and technically it is followed by contracting written on-line, or by fax or other channels.

The third particularity is that there is no insurance behind the interbank transactions. It is explained by the short duration of transactions and the confidence of credit institutions in one another. Banks defend themselves against the risk by creating inner credit limits, in as much the bank offering the credit sets up limits, according to maturity types, based on the size of the partner asking for credit, its reliability so far, its creditability, banking regulation, etc. The limits can be different on the different levels of bank management.

Finally, a special feature of the interbank market is that trends in interest rates are much more erratic than in the corporate market. In case of tight liquidity market situation, since the borrower has to keep its liquidity in any case, a really high interest rate level may occur (of course its daily interest burden is bearable due to the short duration), while in broad liquidity periods the level of interbank interest rates may fall very low.

The best-known international interbank interest rate is LIBOR (London Interbank Offered Rate), which is the interest rate provided for first-class banks in the London interbank market. Its deposit equivalent is LIBID. International loans are provided in LIBOR + risk premium type of

constructions, where the risk premium covers the risk between the given borrower and a first-class bank in London.

3.2 Active banking operations

Active operations of banks are those providing loans to customers directly or indirectly. In case of a direct loan the bank contracts its debtor and provides the contracted funds directly on the debtors account. In case of indirect loans the bank contracts with a third party, i.e. a mediatory who contracts with the ultimate user of the provided funds.

Credits are classified into three groups based upon their economic content.

➡ We talk about ***money credit*** *if the client obtains the amount of credit based on a credit application according to the classical rules of lending*.

The bank maintains a credit limit at the other party's disposal for a commission Its essence is that the debtor can take out the loan in whole or in parts, automatically, at any time at his own discretion within a specific period. We talk about a loan if the client takes out a specified amount of money from his bank based on the loan contract, and he assumes an obligation to pay it back together with interests.

➡ We talk about ***credit-like lending***, *if the bank grants credit not on the basis of credit application but due to any other legal relationship*, for example joins in discounting bills of exchange, factoring happens, etc.

A typical type of credit like loan is the commodity credit. The debtor selects the suitable product, enters into the loan agreement (nowadays, in most of the cases it can be done at the place of purchase), and he can leave with the goods. The duration of the credit can be short-, medium- and long-term.

➡ In case of ***commitment credit*** *the credit institution does not grant the amount immediately but only assumes an obligation that in the future it will fulfil its borrowing obligation under certain conditions*. Such transactions are for example guarantee, letter of credit or credit availability.

Credit types according to the character of the coverage:

- *collateral loans,*
- *non collateral loans (blank credit).*

We talk about ***collateral loans*** *if the ownership of any of the receiving party's assets* (land, real estate, equipments and even amounts on the debtor's bank account) *is transferred to the bank, in case the receiving party cannot repay the loan as agreed*. Such lending has a relatively low risk, as the bank secures itself against non-payment. The bank has to value the collateral properly The credit is 100% covered if the value of the asset offered as coverage is equal to the amount of the debt.

In case a default a problem still can occur, if the bank cannot turn the asset into cash or at a lower price than the market value. Thus the bank usually applies a cut on the value of the collateral, and provides the loan accordingly. The ***Loan To Value*** (LTV) measure s*hows the ratio of the provided loan to the value of the collateral*.

It is also conceivable, however, that the credit is only partially collateralised, i.e. the object pledged as collateral is worth less than the amount to be repaid. The bank is in the worst situation when the credit is an uncovered credit, namely no asset is transferred to the ownership of the bank if the debtor does not pay. Banks are trying to avoid the provision of these loans.

Banks have to face a wide range of risk with credits: the risk of changing interest rates or of economic conditions, non-payment, fraud, the debtor's bankruptcy prospects, etc. They try to reduce these risks somehow. As we have seen, one way to guard against the interest rates and economic conditions is that not fixed but variable interest rates are applied.

The really serious problem is if the debtor cannot pay for some reason. One of the solutions, as we have seen, the emission of only covered loans, when in case of non-payment the asset pledged as collateral is transferred to the ownership of the bank.

The types of securities are as follows:

- *security deposit,*
- *lien,*
- *suretyship,*
- *guarantee.*

The ***security deposit*** *is a form of financial collateral*s. If security deposit is offered for the insurance of an obligation, the holder can satisfy his claim directly from the amount of security deposit in case of non-compliance or non-contractual compliance of the contract. The holder may not use the object of the security deposit, it has to be separated from his assets and it may be used only for the purpose of satisfaction. If the contract signed for the underlying collateral are fully satisfied, the security deposit is refunded to the obligor. The security can be marketable and less marketable collateral.

Absolute marketable: cash, government, banking securities, fixed-term foreign exchange or deposits. Less marketable securities are: listed securities, bills of a first-class debtor.

The ***lien*** is a *physical collateral*. Its key objective is that the debtor should provide coverage to satisfy the deferred claims becoming due later, with property tied up in advance. If the debtor is unable to repay his debt, the holder can secure his claim with the realisation of the pledge being used to ensure the claims. Satisfying claims using the pledged item is usually based on a court decision, by way of compulsory execution. After completing the contractual obligations of the debtor, the lien on the pledged item is automatically terminated.

The *most common type of lien* is the ***mortgage***. To create a mortgage, the pledge agreement is required to be out down in writing and the mortgage is required to be received in a public register (of mortgages). The subject of mortgage may be real estate, vehicles, and tangible assets. In case of real estates, records are done by the Land Registry. The pledged property remains in the possession of the owner, (s)he is intended to use all the while, but he has to ensure the preservation. We know pledge, when the transfer of the pledge is also required in addition to the mortgage contract. The transfer of the pledge can happen to a third person (chattel mortgage holder)

In case of ***suretyship***, *the surety is a person, who obliges to pay if the original debtor cannot.* If the principal obligation is not enforceable judicially, the surety cannot be enforced, either.

In case of ***guarantee*** *the bank pays a fee to the guarantor, who secures repayment of the loan if the original debtor defaults*. The difference between the surety and the guarantor is that the former is brought into the business by the the debtor. The guarantee is an independent obligation, which means that it is a payment promise independent from the principal obligation. So it is indifferent why the obliged does not perform, the guarantee is obliged to accomplish if the obligor of the underlying transaction is innocent in the breach of contract.

The price of loans

- ➡ *The* ***interest*** *on loans is charged according to the specified annual interest rate. The types of interest rate have already been discussed in connection with passive banking operations.*
- ➡ ***Potential nominal provision fee*** *should be paid based on the unused amount of the credit limit. Its level is specified by the bank, usually on an annual basis up to 0-2%.*
- ➡ *Handling charge is a single one-time fee charged for credit assessment and binding of contract. Its extent compared to the full amount of credit is 0-2%.*

- *Credit institutions may charge a **credit assessment fee** regardless of whether the decision was negative or positive. It is generally a specific amount.*
- *The **loan disbursement fee** may be charged after signing the loan agreement, at the time of disbursing the loan. It is usually 0-2% of the loan amount.*
- *Banks charge **an amendment fee** in all cases when the client's file has to be handled.*
- *In case of **prepayment fee,** the bank charges a commission for the client's unexpected cash flow due to liquidity and interest risk.*
- *In case of guarantees and letters of credit, credit institutions charge annual fees rather than interest rates. The extent is between 0-3%.*
- *The payment of **risk taking commission** occurs in case of purchasing the receivables (factoring), which reduces the amount of money paid to the customer.*
- *The bank may oblige the client to the payment of **default interest** if he does not fulfil the current obligation in time.*

The *indicator that includes all costs to be paid on capital* is the ***Annual Percentage Rate (APR)***. The APR is an internal rate which shows the real expense of the loan, in terms of present value calculation: it is an internal rate, in addition to which the repayable capital and lending fee is equal to the sum of credit reduced by expenses paid by the client to the credit institution at the time of the liquidation.

So the ***total charge*** is *the amount to be paid by the borrower based on the credit agreement in addition to the principal amount*. Determining the interest rate and contributions of the credit is called credit pricing. First, the base rate is determined, its parts are the cost of liabilities and interest margin. So the minimum credit interest to be paid by the client is the average (of the) interest rates paid for the bank's liabilities (deposits, refinancing, etc.) and the interest margin. The bank increases the basic interest rate with the interest rate risk premium (due to the position of the client and the transaction).

CHAPTER 4 *BANKING RISKS AND REGULATION*

4.1 Financial intermediation

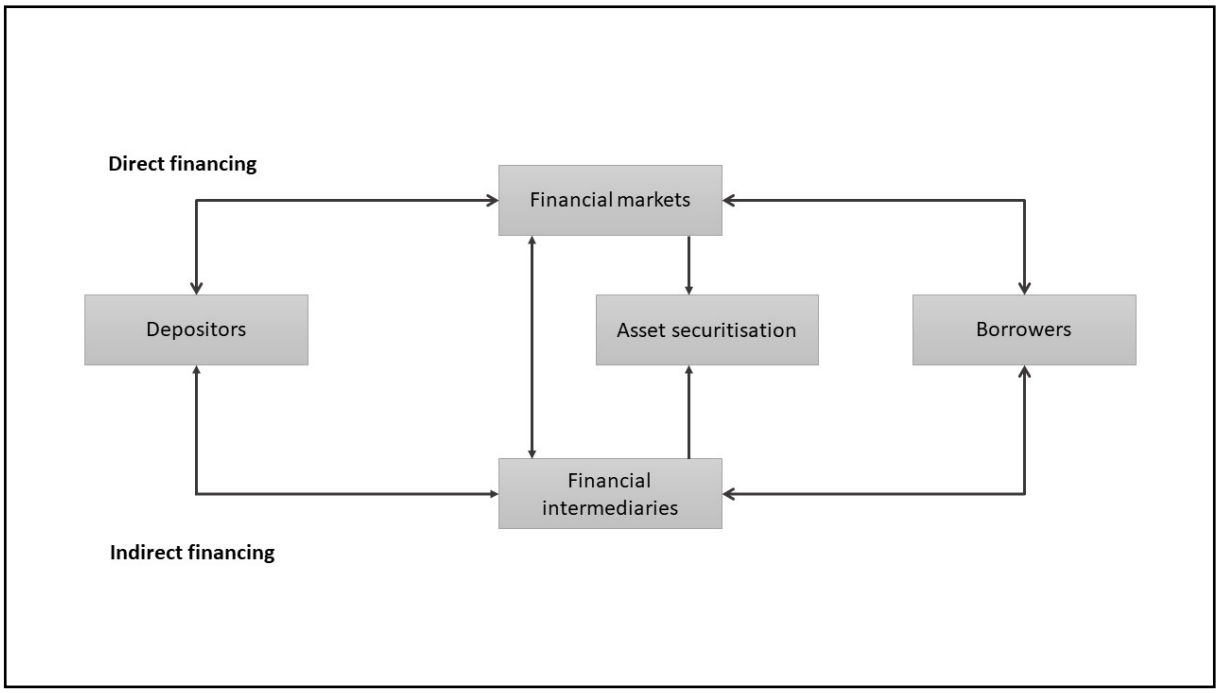
In order to understand how banks work, it is essential to understand financial intermediaries and their role in the economy. The process will be presented in a simplified way: based on the assumption that money already exists in an economy, the following economic units can be distinguished:

- *the state,*
- *companies and*
- *households.*

Among these economic units several transactions can take place on the goods and money market. Certain entities may spend more than their revenue (they are in deficit) or may have more revenue than their planned expenditure (they have a surplus).

In the absence of financial intermediaries, at first sight you would think that the units that have deficit (borrowers) would pair with units that have surplus (lenders). However, some barriers can be identified: the needs of the borrowers and lenders may be incompatible and even if their needs can be matched, it may still be difficult and expensive. These difficulties between the borrowers and lenders can be bridged by financial intermediaries.

FIGURE 1. MODERN FINANCIAL INTERMEDIATION



Source: Casu et al (2015)

Financial intermediation obviously entails costs, but the costs are lower than those for individual economic entities to directly access each other. Financial intermediaries cover their costs and profits from interest rate spreads between loans and deposits.

4.2 The role of banks and the different types of banking

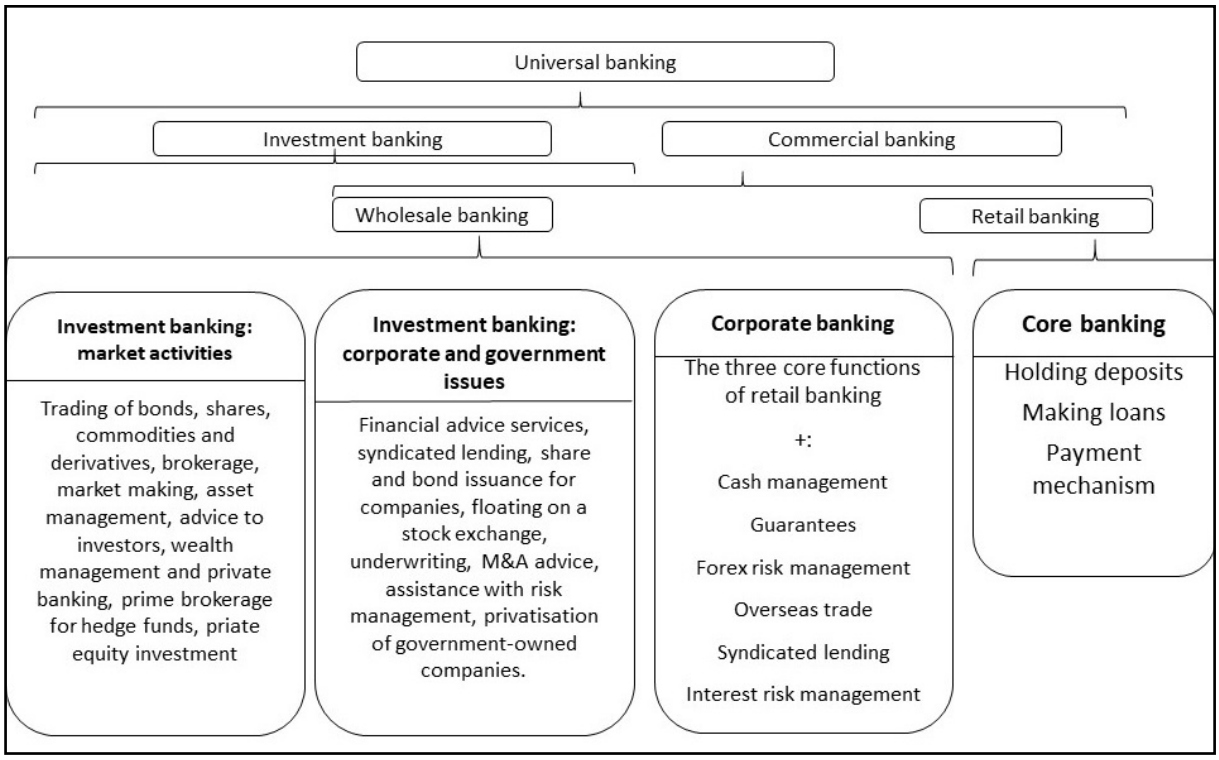
The main function of banks is to collect money (deposits) from units in surplus and lend money (loans) to units in deficit. Banks perform a ***transformation*** function, with regards to

- *size (generally, the amounts needed by the borrowers exceed the amount the lenders are willing to lend),*
- *maturity (generally, funds are lent for a short period of time into medium- or long-term loans),*

- *and risk (borrowers carry a risk that they might not be able to repay the amount of money that was lent to them).*

Initially, banks dealt with core banking operations, taking deposits, making loans and providing a payment mechanism. However, nowadays they conduct a much wider range of activities. Figure 2 provides one possible grouping to understand the different types of banking.

FIGURE 2. DIFFERENT TYPES OF BANKING.



Source: Arnold (2014)

The main characteristics of retail banks and wholesale banks are summarised in Table 1. The fundamental differences can be observed in the clients, the size and the number of the transactions and in the branch network. ***Universal banks*** *offer a wide range of commercial and investment banking activities*.

TABLE 1: MAIN CHARACTERISTICS OF RETAIL BANKING AND WHOLESALE BANKING.

	retail	**wholesale**
clients	households, small firms	large firms
number and size of transactions	numerous small	fewer large
branch network	extensive	one or few

4.3 Risks faced by banks

Risk measures the degree of uncertainty of an expected outcome. There are various types of risks and they originate from different situations. Taking into consideration that banks basically deal with others' money (mainly deposits) and not with their own (equity / bank capital), it is of crucial importance to pay more attention to the banking activities. As a rough breakdown, Table 2 shows the typical liabilities of banks.

Table 2: Typical liabilities of banks.

	Proportion of assets (%)
Sight deposits (current accounts)	10 - 40
Time deposits (saving accounts)	10 - 50
Money market borrowing (repos, interbank)	10- 40
Bank capital (equity)	5 - 15

Source: Arnold (2014)

In the event of a bank failure, not only the bank owners, but also the depositors sustain a loss. That is why the depositors have to be protected. Since banking and the banking system are based on trust, a bank failure might trigger a chain-reaction that – in extreme cases – might also jeopardise the stability of the banking system. There are several risks associated with banking activity. Table 3 provides an overview.

Table 3. Overview of the main types of risk.

Risk type	**Definition / Main characteristic**
credit risk	the risk that a borrower will fail to meet its obligations in accordance with agreed terms
operational risk	the risk of loss resulting from inadaquate or failed internal processes, people and systems or from other external events
market risk	the risk that the value of an investment will decrease due to changes on the financial market / in market factors
liquidity risk	the risk that over a specific horizon the bank will be unable to settle obligations when due
systemic risk	the risk that an event will trigger a loss of economic value or confidence in a substantial portion of the financial system

Interest rate risk or interest risk is defined as *a negative effect on the cash flow of the bank from changing interest rates*. Interest rate risk derives from the case, that the interest rate sensitivity of the bank's assets and liabilities are different. An asset (e.g.: loan) or a liability (e.g.: term deposit) is interest rate sensitive, if the pricing can change during a specified period. At the expiration of a specific asset or liability the environment might change and consequently the conditions of the new assets or liabilities also change.

For example, if the bank is financing a certain fixed rate asset (e.g.: 5-year maturity, fixed rate mortgage loan) from a floating rate liability (e.g.: 1 year term deposit), after the expiration of the liability, an increase in interest rates has an unfavourable outcome for the bank's income. Interest rate risk can be managed by synchronising the expiration time of the certain assets and liabilities. This means that the interest rate sensitive assets are financed from interest rate sensitive liabilities.

Interest rate risk can be measured by several methods. One of the simplest risk assessment method is the interest gap analysis. The formula:

$$GAP_t = RSA_t - RSL_t \quad (1)$$

where: GAP: interest rate GAP during the given period
RSA: interest rate sensitive assets
RSL: interest rate sensitive liabilities

Example: Calculating interest rate risk
The GAP Bank's balance sheet consists of fixed and floating rate EUR loans and on the liabilities side fixed and floating rate EUR term deposits. Interest rate income is calculated by the difference of the interest revenue received from the loans and the interest expenses paid to the depositors.

TABLE 4: BALANCE SHEET OF GAP BANK.

GAP Bank			
Assets (million EUR)		***Liabilities (million EUR)***	
Loan (fixed rate)	200	Deposit (fixed rate)	100
Loan (floating rate)	300	Deposit (floating rate)	400

Assuming, that initially the loan's interest rate is 10% and the deposit interest rate is 5%, the bank's income is evaluated by the following method:

Interest rate income	**Amount**	**Interest rate**	**Interest**
Loan (fixed)	200	10%	20
Loan (floating)	300	10%	30
Sum			**50**
Interest rate expenses			
Deposit (fixed)	100	5%	5
Deposit (floating)	400	5%	20
Sum			**25**
Net income			**25**

As a result of the calculation the interest income of the bank is 25 million EUR.
Using the GAP formula:

$$RSA_t = 400,\ RSL_t = 300 \ \rightarrow \ GAP_t = 300 - 400 = -100$$

Assuming that the interest rate increases 100 basis points, the banks interest revenue and expense from fixed rate loan and deposit remains unchanged, however the revenue and expenses from the floating rate deposit and loan changes, as the table shows below:

Interest rate income	**Amount**	**Interest rate**	**Interest**
Loan (fixed)	200	10%	20
Loan (floating)	300	11%	33
Sum			**53**
Interest rate expenses			
Deposit (fixed)	100	5%	5
Deposit (floating)	400	6%	24
Sum			**29**
Net income			**24**

In this case the bank's interest income decreased by 1 million EUR. As a consequence of an interest rate raise the bank's profit will be lower than initially when the base rate was 5%. The fact that the bank holds more interest rate sensitive assets than liabilities results a decreasing interest income, after raising the base interest rate.

The change of income triggered by the change of the interest rate can be calculated as follows:

$$\Delta y = \Delta i * GAP \qquad (2)$$

where Δy is the change of income, and
Δi is the change of the interest rate

In the above example:
$\Delta i = +1\,\%\,,\ GAP = -100 \ \rightarrow\ \Delta y = 1\% * (-100) = -1$

Table 5 sums up the relation between the GAP and the interest rate risk.

TABLE 5: THE EFFECT OF GAP ON THE INTEREST RATE INCOME.

	positive GAP	negative GAP
increasing rates	**+**	**-**
decreasing rates	**-**	**+**

4.4 Banking regulation

Comparing to other companies in the real economy, banks operate with a much higher leverage. Moreover, information asymmetry is also typical for the banks and their customers. Information asymmetry means that the banks and the clients do not have the perfect information about each other. Because of the risks and risky activities the probability of bank failures is high. The avoidance of the bank failure is one of the main purposes of the regulatory authorities. That is why banks have to comply with regulations that essentially relate to capital, liquidity and risk-taking.

Banking regulation is complex since it is determined by legal, institutional and cultural factors. Despite of different national and regional characteristics, an international cooperation can be observed in banking regulation.

The ***Basel Committee on Banking Supervision*** (BCBS) is *the primary global standard setter for the prudential regulation of banks and provides a forum for regular cooperation on banking supervisory matters*. Its mandate is to strengthen the regulation, supervision and practices of banks worldwide with the purpose of enhancing financial stability. As far as its legal status is concerned, BCBS does not possess any formal supranational authority and its decisions do not have legal force. The BCBS relies on its members' commitments. However, the banking regulation practices of the major economies in the world are generally in line with regulatory packages of the BCBS. These regulatory packages are known as Basel I, Basel II and Basel III. Table 6 illustrates the development of risk measurement in the different regulatory packages of BCBS.

TABLE 6: THE EVOLUTION OF RISK MANAGEMENT.

Types of risk	Basel I (1988)	Basel II (2004)	Basel III (2010)
Credit risk	✔	✔	✔
Operational risk	✖	✔	✔
Market risk	✖	✔	✔
Liquidity risk	✖	✖	✔
Systemic risk	✖	✖	✔

The ***Basel Capital Accord*** of 1988 or simply Basel 1 *established an international standard around capital ratio of 8 per cent on credit risk*. There are four risk classes in the risk-weighted system related to credit risk exposure. Table 7 summarises the framework for capital adequacy risk weighted assets.

TABLE 7: RISK-ASSET RATIO APPROACH IN BASEL I.

Risk class	Weight	Example
No risk	0%	cash, bonds issued by OECD governments
Low risk	20 %	short term claims
Moderate risk	50 %	mortgages
Standard risk	100 %	commercial loans

According to this approach, the banks have to hold a minimum capital that is more than 8% of the total risk-weighted assets (RWA).

This method ignores other types of risk that banks face.

Example: RWA in Basel I
Let us take a bank with the following assets in its balance sheet:

- *cash:* 100 million EUR
- *governments bond*: 100 million EUR
- *interbank loans* (loans to other banks): 500 million EUR
- *mortgages:* 600 million EUR
- *commercial loans:* 1200 million EUR

In this case, the minimum capital ratio should be calculated as follows:

Assets	Risk class	Weight	Quantity	risk-weighted value
cash	No risk	0%	100	0
goverments bond	No risk	0 %	100	0
interbank loans	Low risk	20 %	500	100
mortgages	Moderate risk	50 %	600	300
commercial loans	Standard risk	100 %	1200	1200

As the value of risk weighted assets is 1600 million EUR the minimum capital the bank must hold is 8% x 1600 million EUR = 128 million EUR.

Basel II uses a *„three pillars" concept*:

minimum capital requirements
It deals with maintenance of regulatory capital calculated for credit risk, operational risk and market risk

supervisory review
Supervisors are supposed to control the banks' risk frameworks

market discipline
It encourages effective disclosure about risk exposure, capital adequacy and risk assessment processes

Basel II uses a more risk-sensitive approach and the biggest changes relate to the calculation of minimum capital requirements for credit risk. Since market risk and operational risk are taken into consideration in Basel 2, the minimum capital requirements should be calculated as follows (BCBS, 2014):

$$\frac{\textbf{Total capital}}{\textbf{Credit risk + Market risk + Operational risk}} = \textbf{The bank's capital ratio (8\%)}$$

The third instalment of the Basel Accords or ***Basel III*** was developed in response to the problems in financial regulation revealed by the 2008 financial crisis. Basel III is *intended to strengthen bank capital requirements by increasing bank liquidity and decreasing bank leverage*.

Basel III build upon the Basel II three pillar approach and all the pillars are strengthened, especially the first pillar due to the enhanced minimum capital and liquidity requirements. Basel III introduced two liquidity ratios:

- *Liquidity Coverage Ration (LCR) and*
- *Net Stable Funding Ratio (NSFR).*
-

FIGURE 4. LIQUIDITY AND LEVERAGE RATIOS IN BASEL III.

$$\textbf{LCR} = \frac{\textbf{Stock of high quality liquid assets}}{\textbf{Total net outflows over the next 30 calendar days}} \geqq 100\,\%$$

$$\text{NFSR} = \frac{\text{Available amount of stable funding}}{\text{Required amount of stable funding}} \geqq 100\,\%$$

$$\text{Leverage ratio} = \frac{\text{Capital measure}}{\text{Exposure measure}}$$

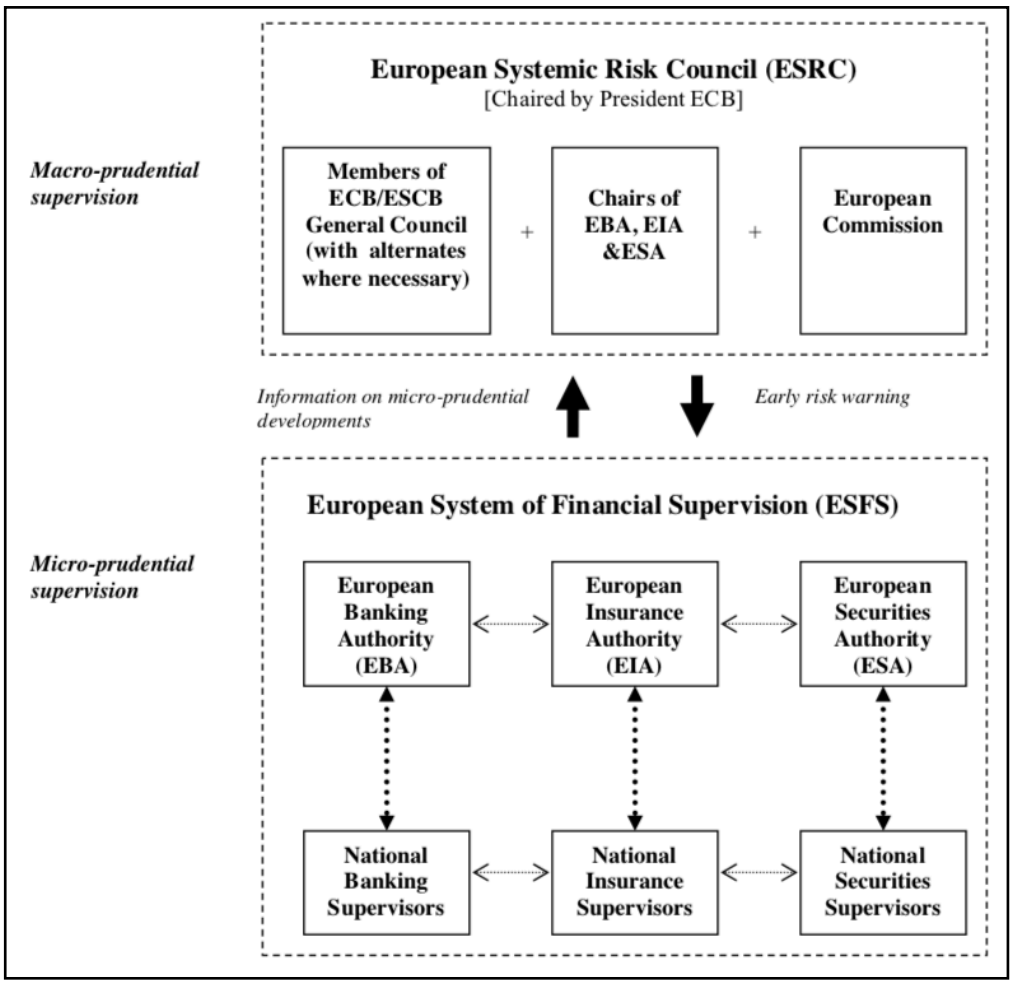
Source: BCBS, 2014.

There were significant changes in banking regulation after the financial crisis of 2008. A new regulatory system was established in the European Union after the crisis. The European System of Financial Supervision (ESFS) was introduced in 2010. It consists of: the European Systemic Risk Board (ESRB) and three European supervisory authorities (ESAs), namely:

- *the European Banking Authority (EBA)*
- *the European Securities and Markets Authority (ESMA)*
- *the European Insurance and Occupational Pensions Authority (EIOPA).*

Whilst the national supervisory authorities remain in charge of supervising individual financial institutions, the objective of the European supervisory authorities is to improve the functioning of the internal market by ensuring appropriate, efficient and harmonised European regulation and supervision. Figure 5 provides an overview of the new regulatory system.

FIGURE 5: THE NEW EUROPEAN FRAMEWORK FOR FINANCIAL STABILITY.



Source: European Economic and Social Committee (2009)

The financial crisis of 2008 made it clear that the harmonisation of supervisory practices and better coordination among the regulatory agencies is needed. The transformation of the new European framework serves this purpose.

As far as banking regulation is concerned, the European Banking Authority has an important role. The main task of the EBA is to contribute to the creation of the European Single Rulebook in banking. Its objective is to provide a single set of harmonised prudential rules for financial institutions throughout the European Union. EBA also plays an important role in promoting convergence of supervisory practices and is mandated to assess risks and vulnerabilities in the EU banking sector.

The European Central Bank has also a central role in the new supervisory system. After the crisis, the regulatory approach has changed with regard to central banks. In addition to taking into account price stability considerations, central banks pay more attention to financial stability and growth considerations. While the pre-crisis period separated the regulation of individual market players from the market as a whole, macroprudential supervision is aimed at reducing market-level risk for the market as a whole.

Bibliography

Basel Committee on Banking Supervision (2014): "*Basel III leverage ratio framework and disclosure requirements.*" Available at: https://www.bis.org/publ/bcbs270.pdf

Casu, B. – Molyneux, P. – Girardone, C. (2015): "*Introduction to Banking*", Pearson, UK.

European Economic and Social Committee (2009): "*ECO/259 de Larosière Report.*" Available at: http://webapi.eesc.europa.eu/documentsanonymous/ces1476-2009_ac_en.doc

Glen, A. (2014): "*FT Guide to Banking.*" Pearson, UK.

CHAPTER 5 *SECURITIES MARKETS*

5.1. Basic terms

Financial assets or securities (unlike cars and houses) *are intangible assets in the sense that their value does not depend on their physical characteristics*. Moreover, in most of the cases, they even do not have any physical attributes, since they are electronic, dematerialised. (Some financial assets, called certified securities, are issued in printed form, but the value of these instruments has still no connection with the quality of the paper or the ink.)

> Financial assets and securities are not precisely the same; the former term usually refers to a broader group. However, the legal definitions may vary by countries, and discussing these differences is beyond the scope of this book.

The most important common feature of financial assets is that they are *claims and, at the same time, obligations on future cash flows*. (Here we refer to Chapter 2 where we recognised that financial goods always appear in two balance sheets simultaneously: a future cash flow can be a claim for somebody only if it is an obligation for somebody else.) Hence, when determining the value of a financial asset, we have to calculate the present value of its future cash flows. This is why present value calculation is so important in finance. We remark that these valuations might be very complicated, for two reasons. First, the nominal sums of the future cash flows are not necessarily determined in advance (e.g., dividends of a stock). Second, the appropriate rate for discounting might be hard to find.

It is worth to observe that not only financial assets can be the subject of investment. We can invest our money in oil, paintings, real estate, and so on. A derivative contract (see Chapter 5.5) on oil is a financial asset, but one barrel of Brent oil is not.

As for the basic features of the claim embodied, we can talk about debt and equity instruments. Bonds and loans are typical examples of ***debt instruments***; t*heir holders require fix payments* (i.e., payments not depending on the economic performance of the borrower) *in return of lending funds*. ***Equity instruments*** - common stocks, for example - *promise payments dependent on the earnings of the issuer and paid after all debt instrument holders are paid*. Because of this, equity instruments are also called residual claims.

Financial markets are markets *where financial assets are issued and traded*. Different aspects might group these markets. According to the products traded, we can talk about debt, equity, derivatives, etc. markets. As for the tenor of the investments, we distinguish between the ***money market and the capital market***: t*he former is the market for short-term* (i.e. shorter than 1 year) *financial assets, the latter is for long-term assets*. The ***primary market*** i*s the expression for new issuances*, while the ***secondary market*** i*s where already issued assets are traded over*. Finally, it is important to mention that the "place" of the trading might be an exchange or the so-called over-the-counter (OTC) market. Several securities are traded in both; moreover, some stocks are listed in more exchanges at the same time.

5.2. Bond markets

Bonds are *debt instruments* similar to loans: the issuer is the debtor, the investor or the buyer of the bond is the lender. The issuer *promises to pay the interest and the nominal or face value of the security in the future*. However, there are at least two significant differences between loans and bonds. First, bonds are traded securities, which means that the 'lender' does not have to keep the asset until maturity, he can sell it on the secondary market. Second, the interest rate that the bond pays - also called coupon rate - is not necessarily equal to the return expected by the investors. Because of this, the loan amount that the issuer receives initially and that he pays back at maturity might be different.

The future cash flows of a bond can be determined according to the following five parameters: face value, coupon rate, interest payment period, principal payment schedule, time to maturity. The ***face value*** (or nominal value) is *the amount the issuer has to pay back* to the lender in the form of principal payments. The ***coupon rate*** of the bond is always expressed as *an annual percentage of the outstanding face value*. If the interest payment period is shorter than one year, we calculate the time-proportional interest linearly. The interests might be paid annually, semi-annually, quarterly, and so forth. In the case of government bonds, annual and semi-annual interest payments are the most frequent. The simplest and most usual principal payment schedule is when the entire face value is paid in one sum, at maturity. In this case, we call the instrument bullet bond, while in the case of more, partial principal payments, we call it amortising bond.

Example: cash flows of 4-year bonds with different cash flow schemes
Consider four bonds: A, B, C and D. For all of them, the face value is 1,000 EUR, the coupon rate is 4% and the time to maturity is 4 years. A, B and C pay interest annually, D pays semi-annually. A and D are bullet bonds, B and C are amortising bonds. Bond B repays the face value in two parts at the last two cash flow dates, in 20-80% proportion. Bond C pays equal principal payments during the 4 years.

TABLE 1: CASH FLOW SCHEMES OF 4-YEAR BONDS

Year*	A	B	C
1	40	40	290
2	40	40	280
3	40	240	270
4	1,040	832	260

Year*	D
0.5	20
1	20
1.5	20
2	20
2.5	20
3	20
3.5	20
4	1,020

**Year = the period that has to pass for the CF to be paid, expressed in years*

It is worth to mention that the coupon rate of a bond can be fixed or floater. In the first case, the coupon is constant over time and determined at issuance – this was the case in the previous example. Floater bonds pay interest that varies according to a reference rate. This reference rate is usually an interbank rate like Libor or Euribor. In the case of floater bonds, it is only the next cash flow that we know in advance.

Example: cash flows of a 5-year floater bond
A floater bond pays interest annually, the coupon is the 12-month USD Libor + 30 bps. At issuance, the Libor is 3.0%, so the first interest payment of the bond is fixed at 3.0% +0.3%=3.3%. The second payment will be fixed at the end of the first year, and so on (see Table 2). At maturity, the last coupon payment and the face value is paid as well. You can observe that the payments follow the market movements, which means that floater bonds bear less interest rate risk than fixed bonds. On the other hand, they carry cash flow risk which is excluded in case of the fixed securities. Hence, there is a trade-off between the two types of risk.

TABLE 2: CASH FLOWS OF A 5-YEAR FLOATER BOND

Year	CF*		Libor
0			3.0%
1	3.3		2.5%
2	2.8		2.9%
3	3.2		2.5%
4	2.8		2.0%
5	102.3		

**CF=cash flow that the bond pays at the end of the given year (expressed in percentage of the face value.)*

Finally, we have to mention some special types of bonds, those with embedded options. Convertible bonds give the right to the holder to exchange his bond to the stocks of the issuer company. These are instruments half-way between debt and equity. Callable bonds grant the right to the issuer to "call back", that is to repay the bond's face value before maturity. In the case of putable bonds, the holder has the right to "give back" the instrument to the issuer, that is to enforce the redemption before maturity. In all of these cases, the proper conditions of exercising the options must be determined in advance.

5.3. Credit rating

The most significant issuers of bonds are central governments. Others are local governments, financial institutions, and companies. Corporate bonds are more common in the US than in Europe.

Central governments issue bonds to finance their budget deficit. In the US, the issuances are organised by the Department of the Treasury. The most important types of securities in this market are T-bills, T-notes, and T-bonds. (T stands for Treasury.) T-bills are money market instruments, with 12-month maturity at most. These securities do not pay any interest, so they have only one cash flow: the face value at maturity. In case of positive discount rates, their price is under their face value and vice versa. T-notes and T-bonds are coupon-bearing instruments, with semi-annually interest payment period. T-notes are medium-term assets, T-bonds are issued with at least 10 years of maturity. We may observe that in the case of the US Treasury securities, only the longest instruments are called bonds. However, the economic and financial sense of bills and notes are the same, and they belong to the same broader class of financial assets. It is quite general that the money market securities of the central governments are zero-coupon instruments, and that they have a distinct name.

Government bonds are often called risk-free instruments, since governments (countries) cannot go bankrupt. This is a very simplistic statement, for two reasons. First, we have to

clarify that in this statement risk refers exclusively to credit risk, that is to the risk that the issuer is not able or not willing to pay its obligations. There are several other types of risk that a bondholder may take, the most important being the interest rate risk. Second, sovereign defaults occurred more frequently and in more countries than one might think - not only in South America, not only after military coups, not only in foreign currency, not only with very high debt/GDP ratio. However, we admit that government bonds issued in local currency are usually the safest investments within a country, concerning credit risk.

Table 3: The Credit Rating categories of fitch

	Category	Short description
Investment grades	AAA	Highest credit quality
	AA	Very high credit quality
	A	High credit quality
	BBB	Good credit quality
Speculative grades	BB	Speculative
	B	Highly speculative
	CCC	Substantial credit risk
	CC	Very high levels of credit risk
	C	Near default
	RD	Restricted default
	D	Default

Source: Fitch (2018)

Credit risk is a risk-type that is hard to measure. Specialised knowledge and expert is needed, and the data required for the calculations might not be available for every single investor. That is why credit rating agencies specialised to this task. The most well-known agencies are Moody's Investors Service, Standard and Poor's (S&P) and Fitch Ratings. They use alphabetical categories to assess the credit risk of a given issuer. The categories are similar, but not exactly the same at the different agencies, see for example the grades of Fitch in Table 3.

We may observe that the four best categories are called investment grades, while the riskier ones are speculative – they are also called junk bonds. It is important to mention that the agencies are not prophets. Their ratings rely on sophisticated models, but these are still only models subject to mistakes.

5.4 Stock exchanges

The below description of stock exchanges is based on how the Xetra trading system works. Xetra is an all-electronic trading system, which was originally developed in the Frankfurt Stock Exchange, but has expanded to be used by various stock exchanges throughout Europe, including Hungary. The trading system combines "continuous trading" periods with "auction" periods. In the Budapest Stock Exchange there are "auctions" from 8:30 to 9:00 a.m. and from 17:00-17:05 p.m. to collect the orders and after that in two minutes an algorithm knowing later determines the opening and the closing price of the shares. But there could be midday auctions as well, which enables liquidity to be concentrated in the middle of the day.

To trade on the stock exchange the investor must open a securities account at a bank or at a brokerage firm. Both securities and cash can be placed on this account. The balance of the securities account is the value of the investor's portfolio. Investors do not trade directly with each other but through an intermediary who can be a dealer or a broker.

Dealers have their own securities accounts, from where they sell securities at the ask price, and buy securities at the bid price. Thus their profit is the difference between the two prices, the so called bid-ask spread. The bid-ask spread reflects the liquidity of the security: the more liquid a security is the narrower is the bid-ask spread.

Brokers do not have securities accounts, thus they do not trade on the stock market, instead they match the trade orders of investors. The most typical form of a trade order is the limit order, when the investor sets a price limit at which the given security can be bought or sold.

A buy (or bid) order of price 80, size 120 can be interpreted as follows: the investor wants to buy 120 shares for 80 or cheaper. A sell (or ask) order of price 85, size 50 is an order, where the investor wants to sell 50 shares for at least 85 per share.

Limit orders that have not been executed yet are collected in the order book. The trading can be continuous, in this case matching orders are executed immediately, thus the order book contains orders that cannot be matched, i.e. the best ask price is somewhat higher than the best bid price. ("Best" from the point of view of the other side.)

The order book below (Table 1A) contains the prices at which orders were given in a descending order. There is a one dollar difference between the best ask price (45) and the best bid price (44). These are the best prices because the cheapest price a share could be bought for is 45 at the moment, the highest price a share could be sold for is 44.

TABLE 1A: ORDER BOOK.

BID / ASK	PRICE	SIZE
ASK	47	110
ASK	46	45
ASK	45	34
BID	44	78
BID	43	23
BID	42	81

The order book can be rearranged to give a more transparent view of the demand and supply sides of the market (Table 1B).

TABLE 1B: THE ORDER BOOK REARRANGED.

BID (BUY) ORDERS		ASK (SELL) ORDERS	
SIZE	PRICE	PRICE	SIZE
78	44	45	34
23	43	46	45
81	42	47	110

To avoid price manipulation only a limited number of the rows is public, i.e. displayed by the trading system. Every new order (except stop orders) placed in the central limit order book is immediately examined to see whether it can be executed against orders on the other side of the order book. Depending on the price limit of the incoming orders and the orders in the order book, execution can be performed at different prices. Execution is performed in accordance with the time/price priority, i.e. the order with a higher buy limit and lower sell limit are executed first. Following execution of the orders, all transactions are published immediately. For each transaction, the price, volume are displayed.

The trading system combines continuous trading (described above) periods with auction periods.

The auction begins with the outcry phase, followed by price determination. In the meantime buy and sell orders are collected in the order book according to price/time priority without executing them. The price will be determined and the orders will be executed at a particular point in time. The outcry phase ends randomly to avoid price manipulation. At the end of auctions all executable orders are executed so, as to prevent a “crossed” order book, i.e. there are no purchase and sale orders that overlap in terms of price, enabling continuous trading to continue. All orders that were not executed remain in the order book. The auction price is determined in accordance with the principle of highest volume transacted. This means that the auction price is the price at which the highest executable order volume is evident.

Example 1: continuous trading

Table 2 shows a part of the order book at a moment of the continuous trading period. The [86;1398] on the bid side means that there are orders to buy 86 stocks at price of 1398 or less. The [1405;110] on the ask side means that there are selling orders of 110 stock at a minimum price of 1405.

TABLE 2A: EXAMPLE 1 - ORDER BOOK

BID		**ASK**	
Size	**Price**	**Price**	**Size**
120	1400	1405	110
86	1398	1407	140
210	1395	1410	320
134	1392	1411	80
340	1389	1414	75

What are the impacts of the following orders on the book? What transactions can be executed if any can be done at all?

➡ a limit selling order: price 1403, size 80
This is better for the buyers than the best sell offer so far, but still is not good enough. It enters the top row of the ask side of the order book and the selling offers so far slide down.

TABLE 2B: EXAMPLE 1 - CONTINUED

BID		ASK	
Size	**Price**	**Price**	**Size**
120	1400	1403	80
86	1398	1405	110
210	1395	1407	140
134	1392	1410	320
340	1389	1411	80
		1414	75

➡ a limit buying order: price 1403, size 100
At this time, a transaction is possible: since the top of the ask side is 1403, it can be matched with the new bid offer. Thus 80 stocks will be sold immediately, and the remainder will be placed at the top of the bid side.

TABLE 2C: EXAMPLE 1 - CONTINUED

BID		ASK	
Size	**Price**	**Price**	**Size**
20	1403	1405	110
120	1400	1407	140
86	1398	1410	320
210	1395	1411	80
134	1392	1414	75
340	1389		

➡ a limit buying order: price 1406, size 90
The 90 stocks will be sold. The price will be the previous one (1405) which was already in the order book according to the price/time priority.

TABLE 2D: EXAMPLE 1 - CONTINUED

BID		ASK	
Size	**Price**	**Price**	**Size**
20	1403	1405	20
120	1400	1407	140
86	1398	1410	320
210	1395	1411	80
134	1392	1414	75
340	1389		

In case of market orders traders give the quantity to be bought or sold but they do not specify the price, which means that the order must be executed at the best price(s) available. A market order to sell (to buy) is executed buy going down on the bid (ask) size of the order book.

Example 1 - continued

➡ *market order to sell 70 stocks*

This order will be executed immediately at the best prices which are on the top of the bid side: 20 shares will be bought / sold for 1403, 50 share for 1400.

TABLE 2E: EXAMPLE 1 - CONTINUED

BID		ASK	
Size	**Price**	**Price**	**Size**
70	1400	1405	20
86	1398	1407	140
210	1395	1410	320
134	1392	1411	80
340	1389	1414	75

Stop limit orders are conditional orders that can be executed after the market price have reached a predetermined value. In case of stop limit orders to buy, the trader sets an activation price (stop price) higher than the prevailing market price, because (s)he thinks that reaching the activation price gives the share an upward momentum. The investor's reasoning is based on some market analysis that will be proved or disproved by reality later. The stop limit order to buy gives a chance to the investor to trade according to his/her believes about the future conditional movements of the price.

Stop limit orders to sell are the reverse case: the investor sets an activation (stop) price lower than the prevailing market price, because (s)he does not want to sell until the market price is above this value. By this order (s)he can maximise the potential loss on the investment or can avoid a downward momentum triggered by reaching the stop price. (The existence of this sudden price drop is, of course undecided…)

Example 1 - continued

➡ stop limit order to buy: stop price 1450, price 1480, size 100

➡ stop limit order to sell: stop price 1398, price 1389, size 80

Neither of the above orders the order book until the market price reaches the given stop prices.

Example 2: Price determination at auction

After collecting the buy and sell orders the overlapping part of the order book looks like the below (Table 3A):

TABLE 3: THE ORDER BOOK AFTER THE AUCTION PERIOD.

Bid	Price	Ask
40	1380	120
70	1385	40
30	1390	90
100	1400	80
30	1405	70

First, quantities are accumulated at each side, based on the fact that a limit order to buy (to sell) can be executed at a lower (higher) price. Thus, the accumulation is top-down at the ask side, and bottom-up at the bid side. Now, at each price the possible tradable amount can be determined by taking the minimum of the appropriate cumulated bid and ask offers. The price at the end of the auction will be the one that maximises the traded quantity.

In this example two prices (1385 and 1390) apply for this maximum principle. As the trading system has to come up with only one price, a secondary optimisation is introduced over the prices that maximise the traded quantity. The "winner" is the price where the offered quantity that cannot be traded is minimal.

The prices where the highest transacted amount (160 stocks) can be executed are 1385 and 1390. The better one of them is 1385 because at this price there is only 70 stocks which could be transacted (there are orders to buy them at this price) but there are no selling orders to pair with them. At the price of 1390 there would be orders for 90 stocks which cannot be realised.

TABLE 3B: THE ALGORITHM DETERMINING THE AUCTION PRICE.

cum Bid	Bid	Price	Ask	cum Ask	Tradable *min[cum Bid; cum Ask]*	Cannot be realised *max[cum Bid; cum Ask] - - tradable*
270	40	1380	120	120	120	
230	**70**	**1385**	**40**	**160**	**160**	**70**
160	30	1390	90	250	160	90
130	100	1400	80	330	130	
30	30	1405	70	400	30	

The secondary optimisation could happen in other ways as well. For example, we could strive to achieve the higher price or the lower price movement. But the method of the Xetra trading system follows the method above so the auction price will be 1385. All cumulated asks below and at this price, and all cumulated asked over this price can be paired. From the 230 shares accumulated at the determined price 70 remains according to time priority. The order book after the execution of all orders (that could be paired) is shown by Table 3C and 3D.

TABLE 3C: THE ORDER BOOK AFTER THE AUCTION.

Bid	Price	Ask
40	1380	
70	**1385**	
	1390	90
	1400	80
	1405	70

TABLE 3D: THE ORDER BOOK AFTER THE AUCTION REARRANGED.

Bid		Ask	
Size	Price	Price	Size
70	1385	1390	90
40	1380	1400	80
		1405	70

5.5 Derivatives

A ***derivative*** is *a security whose value depends on the value of the underlying (asset)*. The underlying (asset) can be a share, a bond, some kind of commodity, foreign exchange, interest rate, a stock index, etc. Derivatives can be used for hedging and for speculation.

There are three groups of derivatives:

- *forwards,*
- *options,*
- *swaps.*

The ***forward contract*** *is a mutual commitment of two parties to buy/sell the underlying at a specified price (F_T) on some future date (T)*. The ***future contracts*** are *special type of forward contracts traded on organised exchanges* known as futures market. The value of the forward at maturity depends on the value of the underlying (S_T). The buyer of the forward is in long position, that is (s)he gains profit if $S_T > F_T$: (s)he pays for the underlying the forward price, and as (s)he can sell it immediately for the prevailing market price, the difference between the two prices is his/her profit or loss. The seller is in short position, (s)he gets the forward price and pays the market price. (See Figure 1 for graphical details.) Thus the joint profit of the two parties is always zero. This gives rise to pure cash settlement on the futures markets: the parties do not need to deliver the underlying asset physically, it is enough if the "looser" pays the appropriate amount (i.e. the absolute value of the difference between the futures price and the market price at maturity) to the "winner".

> In the case of having a LF(K=500) and a SF(K=800) position with the same maturity (T) the investor will have income of 300 at the maturity of the contracts. If the transaction costs of the contracts are lower than the present value of the 300, then the investor can get risk-free profit which means that there are discrepancies in the market creating the possibility of an arbitrage.

Futures contracts and markets have further specialties. Considering that one of the goal of the financial markets is to concentrate demand and supply, the future contracts are

standardised, which implicates higher trading volume and lower trading cost. On the futures markets counterparties do not make the deal with each other. The ***clearing house*** *is an intermediary between buyers and sellers ensuring that the transaction happens as planned*: the clearing house buys the underlying and pays the seller even if the buyer refuses to pay, and vice versa. The payable fee for this service is low because of the risk mitigation by the margin accounts. Each investor agreeing the contract has a margin account to receive the profit and pay the loss on a daily basis. The profit (or loss) comes from the one-day appreciation (or depreciation) of the future price. Because of this the clearing house is on the hook for only the risk of one-day price movement. Moreover, investors are required to have a deposit as well on that account which likely covers the one-day loss. If there is not enough fund on the account, the clearing house require the investor to deposit additional funds (margin call). Finally, only an extremely volatile market can result in the loss for the clearing house. See Table 4 for a detailed comparison of futures and forward markets.

TABLE 4: COMPARISON OF FUTURES AND FORWARD MARKETS.

Futures	**Forward**
traded on organised exchanges	traded on OTC (Over The Counter) markets
standardized	non-standardized
low counterparty risk	high counterparty risk
regulated market	less regulated market
daily financial settlement	settlement at maturity
easy to close position	difficult to close position
cheap	expensive
highly liquid	rather illiquid
margin is required	margin is not required

Options *give their buyers the right* – but not the obligation – t*o buy/sell the underlying asset at a specified price in the future, while the option seller is obliged to accept the buyer's decision*.

The investor having the right is in long position, the other party having the obligation is in short position. In option markets, the procedure for shorting an option is commonly referred to as writing an option. The ***call option*** gives *the right to its owner to buy the underlying;* in the case of a ***put option*** *the investor in long position* (i.e. the buy of the option) *has the right to sell the underlying*. Therefore there are four possible positions depending on whether it is a call or put option and the investor is in long or short position

CALL option

- ***long call (LC)****: is the position of the option's buyer having the right to buy the underlying asset*
- ***short call (SC):*** *is the position of the option's seller (writer) having the obligation to sell the underlying asset, if the other contractor wants to buy it*

PUT option

- ***long put (LP)****: is the position of the option's buyer having the right to sell the underlying asset*
- ***short put (SP)****: is the position of the option's seller (writer) having the obligation to buy the underlying asset, if the other contractor wants to sell it*

At the time or in the period when the option expires, the value of the underlying determines whether the right to buy or sell will be exercised. Figures 5A and 5B show the values of the four positions.

FIGURE 5A: THE CALL OPTION.

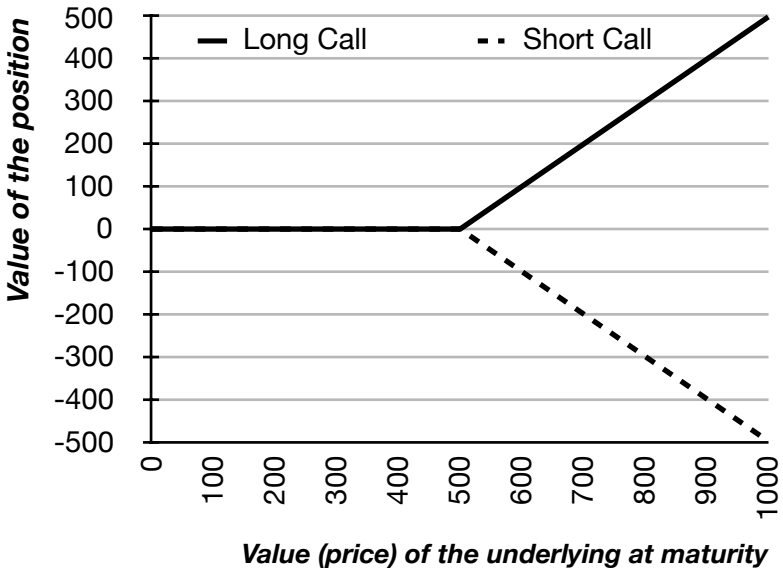
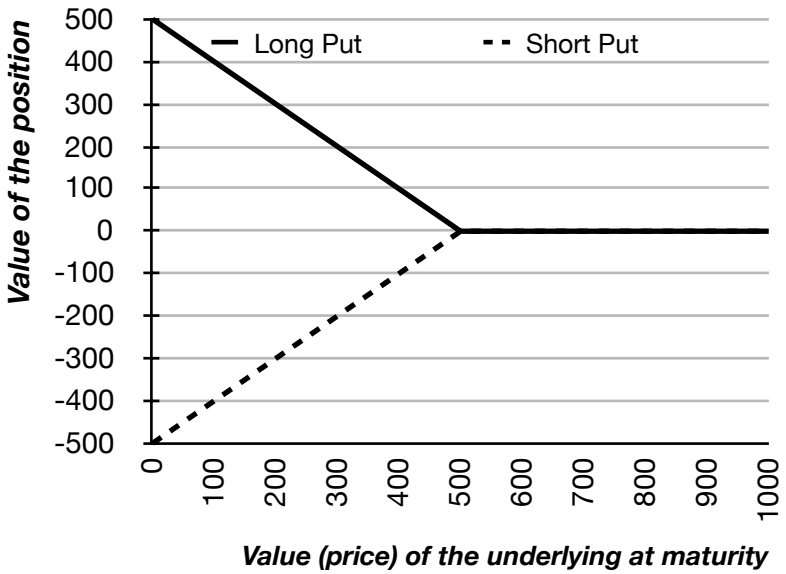


FIGURE 5B: THE PUT OPTION.



Based on the above, the investor in the long position seemingly can only win with the option (otherwise (s)he would not exercise his/her right), while the investor in short position can loose or make a profit of maximum zero. This is obviously known by both investors, therefore the counterparty in the short position will only enter into such contract if the other party pays for it (option fee).

There are many types of options, the most typical ones are the following. The American option can be exercised any time until its expiration date. The European option can be exercised only on the date of expiry. The Bermuda option can be exercised on the expiration date and on certain days before that. The value of an Asian option is based on the average price of the underlying asset, not on the prompt price of it at expiry. The barrier option can only be exercised if the price of the underlying asset is met or did not meet a predefined value. These options are preferred by the buyers of the options because these are cheaper than normal options due to the lower probability of payment. The binary option is practically a bet where the option pays a certain amount only if the condition in the contract is met (e.g. if there is a hurricane it pays 1000 dollars; if there is no hurricane it does not pay anything). Exotic options are complex structured products, where the value of the option is derived by many complicated factors. (Sebestyén 2016)

The third type of derivatives are the ***swaps***. They are about *exchanging future cash flows, out of which at least one is subject to uncertain factors* (such as interest rate, exchange rate, stock price or the price of goods). These deals are personalized, OTC agreements. The two most common forms are interest rate and exchange rate swaps. In case of the former one party pays fixed while the other pays floating interest rate on the same nominal value. To avoid excessive transaction costs and risks, the payments are netted, only the difference is paid.

> For example, the A firm pays 4% fixed interest to the B bank, while it pays EURIBOR +3,8% interest to the A firm under a swap contract with a face value of 15 million euros. The payments are fulfilled annually. If the EURIBOR for the actual payment is -0,18%, then the bank must pay the firm 543.000 euro. Since this amount is higher than the fixed interest (600.000 euro), after netting the payments the firm pays the bank 57.000 euro.

The reason behind an interest rate swap is usually that the buyer of the swap can have cheaper fixed / floating rate loan, while businesswise it would be more beneficial for it to have a floating / fixed rate. With an interest rate swap however, it can achieve both goals.

> For example, if the movements of interest rates are indifferent for the costs and incomes of a firm then the firm prefers to get fixed interest loan otherwise the increase of the interest rate increases its costs as well diminishing the profit while the incomes and the operating costs are the same. Getting a fixed rate loan, the movements of the interest does not affect its profit. In this case the firm gets a floating interest rate loan (if it's cheaper) of 50 million euros and looks for a bank who is willing to contract an interest rate swap for the face value of 50 million euro. By this position the firm pays the bank the fix interest, and the bank pays the floating interest the firm, which is equal to the interest of the loan.

In case of an exchange rate swaps, the interest rate, the duration and the face value are fixed, only the currencies are different. Thus, the parties pay each other the interest in different currencies in defined timing. Netting is not evident in this case as the currencies are different. The reason behind an exchange rate swap is usually that the buyer of the

swap can have cheaper loan in one of the currencies, while business wise it would be more beneficial for it in the other one. With an exchange rate swap however, it can achieve both goals.

> For example, a Hungarian firm supposes its bond would be preferred by German investors, which means–they tend to give loan at lowest interest rate, however the firm's costs and incomes are in Forint. In this case it should issue euro bonds and contract an exchange rate swap to convert the euro cashflow to Forint, so that it can get the lowest debt interest and keep off currency risk. The bank as the counterparty of the swap pays the firm in euro the same amount as the firm has to pay the bond-owners, while the firm pays the bank the payment of the swap in forint (like the interest of the debt). Finally, not only the costs and incomes of the firm are in forint but the interest to be payed for the debt.

Bibliography

Fitch (2018): Rating Definitions, Fitch Ratings. Available: https://www.fitchratings.com/site/definitions

Xetra (2018): Trading. Available: http://www.xetra.com/xetra-en/trading

CHAPTER 6 *THE BALANCE OF PAYMENTS*

In this chapter, we are getting prepared to understand international economics by studying the ***Balance of Payments*** (hereinafter "BoP") which is both a certain framework for thinking and a practical statistical "tool" to analyse international economic relations. The BoP has been internationally developed to provide policy decision-makers with consistent and comparable information on economies' external relations. The International Monetary Fund (IMF) first created a manual for reporting balance of payments statistics in 1948. Since then, the IMF has been regularly updating the *Balance of Payments Manual* to respond to economic and financial developments and the accumulation of experience by compilers (IMF, 200)). Nowadays, the notions of the BoP and the data reported accordingly are widely used in economic decision-making while also discussed and even challenged in the literature.

This chapter is structured as follows. In Section 1, the purpose of the BoP is described while taking a first glance at its overall structure. Section 2 introduces basic definitions and principles necessary to compile the BoP of an economy. This is indeed what the Reader is going to do in an example (Section 4), after having studied the content of the rows and columns step by step (Section 3). Finally, the chapter concludes with some remarks on the interpretation and application of the BoP in economic analysis.

6.1 Purpose of the BoP

The BoP *summarises all economic transactions between residents (i.e. the domestic economy) and non-residents (i.e. foreign economies) for a period of time.* It does so in a structured way, with standard rules for the content and entry of data. Figure 1 shows the main structure of the BoP. Section 3 is going to describe each component in detail. Here, it is enough to note that the table is split into two main parts (horizontally): there is a real economic (I-II) and a financial block (III-IV).

This splitting is important because it reveals a more profound purpose of the BoP. By collecting all foreign transactions, it actually aims to determine the domestic economy's net external position, arising from real economic links. As economies trade in goods and services and provide means of production (labour, capital) to each other (with incomes sent and received), deficits and surpluses are likely to arise. This means either foreign borrowing (i.e. getting financed by other economies, in case of a deficit) or foreign lending (in case of a surplus) in net terms in a given period. Thanks to the financial block, the BoP also shows the financial transactions underlying the net foreign borrowing or lending. The net external position is further discussed in Section 3. At this point, the Reader should note that the BoP includes important information on an economy's international financing patterns.

Figure 1. The Balance of Payments.

	Credit	Debit	Net
I. Current account (CA)			
I.1. Goods			
I.2. Services			
I.3. Incomes			
I.4. Current transfers			
II. Capital account (KA)			
II.1. Non-produced, non-financial assets			
II.2. Capital transfers			
III. Financial account (FA)			
III.1. Direct investment			
III.2. Portfolio investment			
III.3. Other investment			
IV. Change of reserves (dRES)			
V. Net errors and omissions (NEO)	**- - -**	**- - -**	
Balance			**0**

6.2 Basic Definitions and Principles

The BoP is compiled from the viewpoint of an economy regarded as domestic. That is, one can look at the BoP of the U.S. or China or Russia etc. Transactions are recorded vis-á-vis the rest of the world. It follows from the foregoing that notions with bilateral relevance, e.g. exports, imports, claims and liabilities refer to those of the domestic economy.

Residents are the *actors (households, firms, non-governmental organisations [NGOs] and the government) of the domestic economy.* This is an exhaustive definition only at the first glance, though, as the international mobility of people and entities raises questions about what the attribute "domestic" really means. For example, is a French citizen living in the USA an American resident? Is a foreign-owned company resident in the economy where it operates or where its owners reside? To solve this problem, there is a clear rule to classify actors as domestic, i.e. resident. *An actor is resident if its economic interest is predominantly attached to the territory of the domestic economy.* In more practical terms, this means that

- *people are regarded as residents of a certain economy if they stay there for at least one year;*
- *firms and NGOs are regarded as residents of the economy where they operate, i.e. perform the core activities of their business.*

FIGURE 2. THE BALANCE OF PAYMENTS EXPLAINED.

Real economic components of transactions (I-II)

CA + KA = net external position

Financial components of transactions (III-IV)

	Credit	Debit	Net
I. Current account (CA)	**Revenue, income or transfer received** by the domestic economy	**Expenditure, income or transfer sent** by the domestic economy	**= CA**
I.1. Goods			
I.2. Services			
I.3. Incomes			
I.4. Current transfers			
II. Capital account (KA)			**= KA**
II.1. Non-produced, non-financial assets			
II.2. Capital transfers			
III. Financial account (FA)			**= FA**
III.1. Direct investment			
III.2. Portfolio investment			
III.3. Other investment			
IV. Change of reserves (dRES)			**= dRES**
V. Net errors and omissions (NEO)	- - -	- - -	**= NEO**
Balance			**≡ 0**

Here, it should be noted that the "economy" is to some extent different from the geographical territory of the country in question. An "economy" also includes units geographically located in another country, but being under direct domestic control, such as embassies, consulates, military and scientific bases. By the same token, it excludes such units geographically located within the domestic country. The actors of such units are the residents of the controlling economy.

The Reader should find some examples helpful here. Let's decide if the next actors are residents of the U.S.:

a) an American shopkeeper living in Mexico since 1991;
b) John Deere's subsidiary in Russia;
c) Deutsche Bank subsidiary in the U.S.

In case of a), the answer is no because this person has been staying outside the U.S. for more than one year. Regarding b), the answer is no again as this firm operates in

Russia. The case of c) is just the reverse, this firm operates in the U.S. and therefore, it is resident there.

Transactions between residents and non-residents are recorded in the BoP on the basis of the following principles:

- ***Transactions at market value:*** *Transactions are accounted for at market prices even if partners agree otherwise.*
- ***Accrual principle:*** *Transactions are included if and when agreed on by partners, not when payed.*
- ***Double-entry accounting of the transaction value****: Each transaction involves a credit entry and a debit entry (see columns of the BoP in Figure 1). In most cases, this seems logical as transactions are usually exchanges, i.e. two-sided. There may be a real and a financial "side" (e.g. selling a product in exchange for money), or there are purely financial exchanges (e.g. taking out a bank loan, buying a share). However, in case of transfers, the double-entry accounting seems unjustified as transfers are unilateral transactions by definition (e.g. providing food aid does not result in any obligations on the part of the recipient). Nonetheless, this problem is solved by having extra rows for transfers. Thus, the double-entry principle is applied throughout the BoP without exception. This implies the final zero balance (further explained in the following section).*

6.3 Constructing the BoP Step by Step

This section explains the structure of the BoP along with the exact method of entering transactions data into it. Having discussed the double-entry accounting principle, the question naturally arises what "credit" and "debit" mean. In general, they can only be defined rather broadly. The reason is that these terms are interpreted slightly differently in the different parts of the BoP. So generally, it can be stated that *values in credit increase the balances of the parts of the BoP while values in debit decrease them.* This is so because *net value* of each row is calculated as *credit – debit*. By summing up all net values of a part, we get the corresponding aggregate balance (CA, FA, KA, dRes). (The same result is obtained by summing up all credits and debits of a part separately and subtracting the latter.)

Still in general, credit could be loosely labelled as "revenue/inflow" and debit as "expenditure/outflow". Figure 2 points out the accurate meanings which become clear once having studied the contents of the horizontal parts.

As mentioned earlier, the BoP consists of two main horizontal blocks (separated by the dashed line in Figure 2). Into the ***first block (I-II),*** *the real economic components of transactions* are entered. Throughout this block, values are accounted for in the ***credit*** column ***if the domestic economy obtains revenue or income in the transaction*** in question. ***Debit*** includes the reverse, i.e. *expenditure or the outflow of income*. This horizontal block is further split into the current account and the capital account.

The ***current account (CA)*** *shows flows of goods and services (trade) and flow of incomes and current transfers*. Particularly, the following rows and contents appear here:

Goods: *This row covers the international trade of products. Export of goods brings about revenues for the domestic economy, so it is entered as credit. Conversely, imports are accounted for as debits.*

Services: *This row covers the international trade of services. Credits and debits in this row are interpreted in the same way as in "Goods".*

Incomes: *In this row, the international flow of incomes of means of production (labour, capital) is accounted for. These incomes are obtained either by workers or by holders of financial or other non-produced assets who provide/rent these for foreign actors. Practically, this row contains salaries and wages, rents, interest incomes and dividends sent and received. If received by the domestic economy, these are entered into the credit column. Debit includes incomes sent abroad.*

> *Compensation of workers appears in this row only if the employer and the employee are resident in different countries. That is the case of workers who work in a foreign economy for less than one year as they are still residents of their economy of origin while getting compensation from a non-resident (a foreign employer). Staying at least one year abroad, these workers themselves become non-resident, though. Consequently, their compensation appears in the BoP only if they decide to send it home. If so, this amount is not regarded as an income any more, but as a (current) transfer because it is unilaterally provided to the home economy.*

Current transfers: *The purpose of this row is to account for a certain type of unilateral transactions. Transfers are considered as "current" if they directly affect the level of disposable income (and hence, consumption possibilities). (By contrast, capital transfers included in the capital account affect the level of wealth.) Typical examples of current transfers are aides, taxes on income and wealth and social contributions and social benefits (e.g. pensions). Moreover, membership fees paid to international organisations are also entered as current transfers. If received by residents, transfers are regarded as credit. Conversely, transfers provided by residents to non-residents are debits.*

The aggregate balance of these rows is known as the current account balance (also denoted by CA).

The **capital account (KA)** includes the following:

Non-produced, non-financial assets: *The term "non-produced, non-financial assets" refers to natural resources and rights to utilise them (e.g. land, mines/mining rights, radio spectra); permissions to undertake specific activities; and patents, copyrights, trademarks, brand names and franchising. The BoP includes transactions of these assets. Entries follow the same logic as in the first two rows of the current account. If residents sell non-produced, non-financial assets to non-residents, the transaction value is entered as credit (revenue to the domestic economy). If such assets are, on the contrary, bought by residents (expenditure), debits appear in the BoP.*

Capital transfers: *Here, the description of "Current transfers" applies except that capital transfers affect wealth.*

The aggregate balance of these rows is known as the capital account balance (also denoted by KA).

The *sum of the balances of the current and capital accounts (CA + KA)* defines ***the net external position of the domestic economy***, i.e. whether it needs foreign financing or is able to finance foreign economies *as an outcome of its provision and use of real economic resources in a given period*. To put it differently, the domestic economy may run an *external deficit* (net borrowing) *or* an *external surplus* (net lending). As all values in the BoP, the net external position is a flow indicator, capturing changes in a stock. An external deficit (in itself) increases the liabilities of the domestic economy vis-á-vis the

rest of the world (or decreases its claims). Conversely, an external surplus (in itself) results in an increase of its claims (or the decrease in its liabilities).

> The stock in question is most precisely the ***net international investment position (NIIP)*** which is defined as *the difference between the foreign liabilities and claims of the domestic economy.* At the start of a period, the domestic economy can have either net liabilities or net claims which is changed by the net external position recorded in the BoP for that period. It must be noted, however, that the change of the NIIP is not exclusively driven by the net external position. Revaluations, i.e. changes in prices of assets and exchange rates also affect its value. Consequently, an economy running an external deficit may indeed end up with (higher) net claims if its assets appreciate in the meantime. While revaluations can have particularly important effects in the short run, Obstfeld (2012) nevertheless concludes that regarding long-term patterns, the NIIP is driven primarily by the net external position. (See

The net external position is conceptually equal to the balance of the financial account and the change of reserves (*FA + dRES*). In other words, the ***second block of the BoP (III-IV)*** *shows how the net borrowing from or the net lending to non-residents is financed* as it reveals the various foreign financial transactions in a given period. The **financial account (FA)** records the following:

> ***Direct investment:*** *Also referred to as foreign direct investment (FDI), this type of cross-border investment arises when an investor acquires significant degree of influence on the management of an enterprise resident in another economy. According to the IMF's BoP standards, there is a significant degree of influence if the investor obtains equity giving it voting power of at least 10 percent in the enterprise. If a non-resident invests in a resident firm, this is entered in the BoP as credit because the liabilities of the domestic economy increase. If residents invest abroad, that is precisely the reverse (increase of claims) and consequently, debit.*
>
> > *Owning equity in a firm gives rise to claims on that firm. (The best-known signs of that are dividends which, however, trace back to the very fact that the owner has invested part of its wealth into that company. This shareholding appears as an asset in the owner's balance sheet.)*
>
> ***Portfolio investment:*** *Portfolio investment involves trading financial assets without acquiring significant degree of influence on an enterprise (if this were possible at all). Trade of shares, bonds and other securities is accounted for in this row. Credits and debits are interpreted in the same way as in "Direct investment". Investors involved in portfolio investment can be mostly regarded as financial investors striving for exchange rate gains. By contrast, direct investors are more likely to be strategic investors having long-term interests in the specific business (industry).*
>
> ***Other investment:*** *Other investments include, among others, bank loans and trade credits. Credits and debits are interpreted analogously as in the previous two items, with residents taking out loans as credit (increase of liabilities) and granting loans as debit (increase of claims).*

Taking a simple approach, ***reserves*** cover the currency stock of the domestic economy. More precisely, reserves refer to *monetary gold and foreign currency (FX) assets held by monetary authorities (central banks) to meet balance of payments financing needs* (besides other purposes). Here, for simplicity, we assume that all international (FX) payments involve central banks (or less restrictively, that currency is deposited within the domestic two-tier banking system, "channelling" to the central bank). Hence, we can

interpret reserves as “the domestic currency stock” used to pay for imports and including revenues on exports.

The BoP records changes in this stock (***dRES***). If the domestic economy obtains currency (e.g. export revenue), the value is entered into the debit column. Currency outflow (e.g. payment for imports) is accounted for as credit. This is consistent with the interpretations of credit and debit in Figure 2 as a currency is a claim of the domestic economy. It is easy to see that residents with currency “in their pockets” can anytime decide to purchase goods, services etc. from abroad. More sophisticatedly, a currency is a domestic claim because at its very origin, it is a foreign bank’s liability. (Remember Chapter 2 on money creation.)

Having discussed the content of the BoP row by row, the section concludes with drawing attention to the final 0 balance. As already mentioned, $\sum \text{credit} - \sum \text{debit} = 0$ as both columns contain exactly the same values. By calculating the balances of the parts (CA, KA, FA, dRES) first and adding them up, only the order of addition and subtraction is changed. Hence, the following identity holds:

$$\sum \text{credit} - \sum \text{debit} = CA + KA + FA + dRES = 0$$

Note that in practice, the BoP is compiled by statistical authorities on the basis of a huge number of transactions reported by many economic agents. Therefore, the final 0 balance is usually not obtained due to statistical errors and omissions. *To achieve the conceptual zero, the value of the actual deviation is entered appropriately into* the technical row ***“Net errors and omissions” (NEO).***

In this **example**, let Hungary be the domestic economy. Transactions are then accounted for in Hungarian forints (HUF) (though overwhelmingly FX denominated in reality). After having recorded all transaction values, the Reader shall be able to determine the net external position of Hungary in this hypothetical example.
Transactions taking place over a period are the following:

(1) A Hungarian company buys goods worth 10 million HUF from a foreign company, and it pays in euros.

(2) A Hungarian company sells services worth 6 million HUF to a foreign partner who promises to pay later.

(3) A domestic, American-owned company sends dividends in USD worth 20 million HUF into the U.S.

(4) A domestic foundation grants food aid worth 4 million HUF to an African country hit by drought.

(5) The state budget repays debt denominated in EUR (worth 100 million HUF) to a foreign bank (principal payment).

(6) The foreign buyer settles half of its debt arising from transaction 2).

(7) A domestic company decides to set up a subsidiary in Italy. Aiming to develop its production there, it makes the investment (50 million HUF) in kind: it hands over production lines and raw materials to the subsidiary.

(8) A foreign media company successfully applies to broadcast on a domestic radio frequency. The media authority charges 200 million HUF for the use which the company promises to pay.

(9) A domestic company buys industrial robots from its foreign parent. The robots are worth 50 million HUF, but the company pays only the equivalent of 30 million HUF for them.

(10) A foreign investor buys shares worth 20 million HUF at the Budapest Stock Exchange, acquiring 1% of the shares of a domestic public limited company. (Payment in EUR.)

Solutions:

(1) This is import of goods, i.e. expenditure of Hungary, settled by immediate payment.

Credit: Reserves, 10 million (outflow)
Debit: Goods, 10 million (import)

(2) This is export, so almost the reverse of (1), except that services are traded and payment is expected to occur at a later date.

Credit: Services, 6 million (export)
Debit: Other investment, 6 million (trade credit extended by Hungary

(3) Note that this company is resident in Hungary, so its transactions with its non-resident U.S. owners are rightly to be entered in the BoP. In this case, dividends are income of the U.S. owners earned by providing capital to Hungary.

Credit: Reserves, 20 million (outflow of currency)
Debit: Incomes, 20 million (income sent abroad)

(4) This is the "classic" case of current transfers, with food increasing the consumption possibilities of recipients. The use of the row "Current transfers" is evident from the above while the transaction itself materialises as an outflow of goods. In fact, export is registered which, however, does not result in any claims on the part of the aid provider. (Note that credit and debit cancel out each other, with no balance appearing in the CA.)

Credit: Goods, 4 million (export)
Debit: Current transfers, 4 million (transfer sent)

(5) This is a decrease of liabilities of Hungary, involving payment in EUR.

Credit: Reserves, 100 million (outflow of currency)
Debit: Other investment, 100 million (debt repayment)

(6) There remains 3 million EUR still outstanding as trade credit. The other half is obtained in cash. This item involves a change in the form of a domestic claim only.

Credit: Other investment, 3 million (decrease of claims)
Debit: Reserves, 3 million (inflow of currency)

(7) Here, a direct investment is made by the domestic company as it sets up a fully owned enterprise abroad. The capital is granted in kind (flow of goods).

Credit: Goods, 50 million (export)
Debit: Direct investment, 50 million (investment made abroad)

(8) The right to use a radio frequency is a non-produced, non-financial asset. In this case, this right is bought by a non-resident, providing revenue to Hungary. The promise to pay later is regarded as a trade credit already seen in (2).

Credit: Non-produced, non-financial assets, 200 million (revenue)
Debit: Other investment, 200 million (claim on the foreign company)

(9) The industrial robots are imported from abroad. This is import as assets are sold and purchased by the parties. (So this is not an investment that would bring about the increase of the subscribed capital of the subsidiary, regarding the full ownership here). Still, there is an element which reflects the relationship between the two companies. Note that the industrial robots are worth 50 million EUR at market prices. However, the subsidiary is

offered a more favourable price, and no further liability arises in this transaction itself. That is, a transfer is provided affecting the stock of assets in Hungary.

Credit: Capital transfer, 20 million (transfer received)
Credit: Reserves, 30 million (outflow of currency)
Debit: Goods, 50 million (import)

(10) As less than 10% of the shares is taken over, this is a portfolio investment.

Credit: Portfolio investment, 20 million (investment in a Hungarian entity)
Debit: Reserves, 20 million (inflow of currency)

As a result of all the transactions, Hungary ends up with an **external surplus of 196 million HUF** in this period. This, *ceteris paribus*, means that its NIIP "improves", with net liabilities decreased or net claims increased at the end of the period.

6.4 Concluding remarks

In this chapter, the international balance of payments has been presented. The purpose of the BoP is to summarise all economic transactions between residents and non-residents for a period of time. A key notion of the BoP is the net external position which is the balance of the real economic transactions recorded. Although the net external position is an important marker of changes to external indebtedness, the net approach in itself is not sufficient to uncover all relevant financing patterns**.** It is apparent from the BoP compiled in the example in Section 4 that gross flows are much bigger than the net external position. This is of particular relevance in the case of the financial block of the BoP. In a world of huge global capital flows and complex financial links, stability concerns may well arise from financial transactions mostly appearing only in the financial block.

Note that purely financial transactions appear only in the financial account and the row „change of reserves", but not in their final (net) balance as they cancel out when calculating it. Examples are transactions 5), 6) and 10) in Section 4.

Attention to gross values is also vindicated by the related fact that claims (resulting from surpluses) of some domestic actors are not directly available to settle the liabilities of others. So even if the domestic economy runs large external surpluses and hence, accumulates huge foreign claims in net terms, there are possibly actors who are (externally) indebted. If these debts happen to be unsustainable for whatever reason, a debt crisis may erupt due to spill-over effects.

However, the importance of the gross approach does not annul the relevance of the net external position in economic analysis. Sustained net patterns may signal underlying disturbances, "imbalances" in an economy or a group of economies. Recurrent external deficits and countervailing surpluses may thus provide early warning of crises, e.g. international debt crises. Empirical evidence shows that external positions had become increasingly polarised in the case of the U.S. and China and within the euro area before the 2008 crisis.

In case of emerging economies, catching up usually involves external deficits for several consecutive periods as advanced technologies can possibly be accessed only from abroad, or to put it differently, there is no sufficient internal capacity to invest. If external financing is really used to achieve higher competitiveness (boosting export revenues to repay debts), there is no threat of unsustainability. It is hard, however, to come clear on this in advance.

While it is difficult to tell *ex ante* if sustained external positions are truly problematic in a certain case, this chapter concludes with arguing that they are definitely worth of attention and further analysis.

Bibliography

IMF (2009): *Balance of Payments and International Investment Position Manual*. International Monetary Fund, Sixth Edition (BPM6), Accessible at https://www.imf.org/external/pubs/ft/bop/2007/pdf/bpm6.pdf.

Obstfeld, M. (2012): *Does the Current Account Still Matter?* American Economic Review, Vol. 102(3), pp.: 1-23

CHAPTER 7 *FOREIGN EXCHANGE MARKETS*

7.1 Introduction to FX-markets

The price of one nation's currency in terms of another is the so called FX-rate or exchange rate; the place where different currencies are exchanged is the foreign exchange market. This is the largest financial market: with a global trading of 7/24, the total amount of FX traded daily is about 5 trillion USD. The global trading of currencies is the fundamental of international economic relations, and therefore a key variable for the economic policy as well.

This is a meeting point between many participants of the economy. Firms engaged to international trade are exposed to the movements of exchange rates. Importers of goods are having expenses in foreign currencies and revenues in the domestic currency; exporters are having revenues in foreign and expenses in the local currency. However, usually they do not change their exposure directly, but seek the adequate liquidity on the global market, usually via banks and other institutions. FX-risk management provides hedge against volatile exchange rate movements. Some participants take this risk on their balance sheets for profit seeking reasons thus providing hedge to those who want to get relieved of their exposure.

"Key currencies" are the currencies of the major economies. The US Dollar (USD) is called the 'world currency', the euro (EUR), the Swiss franc (CHF), the Japanese yen (JPY) are the most traded currencies besides the US dollar. Other currency pairs, such as Hungarian Forint (HUF) / South African rand (ZAR) cannot be traded directly, but usually via key currencies.

A ***foreign exchange quotation*** (or quote) i*s a statement of willingness to buy or sell at an announced rate*. The CUR1/CUR2 quotation may seem a bit confusing or non-intuitive. The currency to the left of the slash is called the base currency or the unit currency. The currency to the right of the slash is called the price currency or quote currency. The quotation always indicates the number of units of the price currency, CUR2, required in exchange for receiving one unit of the base currency, CUR1.

> For example, the most commonly quoted currency exchange is that between the US dollar and the euro. The quotation **EUR / USD 1.2174** designates the euro (EUR) as the base currency, the dollar (USD) as the price currency, and the exchange rate is USD 1.2174 = EUR 1.00.

A ***direct quote*** is *the price of a foreign currency in domestic currency units*. An ***indirect quote*** is *the price of the domestic currency in foreign currency units*. In retail exchange in many countries (such as currency exchanged in hotels or airports), it is a common practice to quote the home currency as the price and the foreign currency as the unit.

> A woman walking down the Avenue des Champs-Elysées in Paris might see the following quote: **EUR 0.8214** / **USD 1.00** Since in France the home currency is the euro (the price) and the foreign currency is the dollar (the unit), in Paris this quotation is a direct quote on the dollar or a price quote on the dollar. Verbally, she might say to herself, "0.8214 euros per dollar," or "it will cost me 0.8214 euros to get one dollar." These are European terms.
>
> At the same time a man walking down Broadway in New York City may see the following quote in a bank window: **USD 1.2174 / EUR 1.00** Since in the US the home currency is the dollar (the price) and the foreign currency is the euro (the unit), in New York this would be a direct quote on the euro (the home currency price of one unit of foreign currency) and an indirect quote on the dollar (the foreign currency price of one unit of home currency). The man might say to himself, "I will pay $1.2174 dollars per euro." These are the American terms.

The two quotes are obviously equivalent (at least to four decimal places), one being the reciprocal of the other: 1 **EUR 0.8214/USD** = **USD1.2174/EUR**

Although a newspaper or magazine article will state an exchange rate as a single value, the market for buying and selling currencies, be retail or wholesale, uses two different rates, one for buying and one for selling. The ***bid*** *is the price in one currency at which a dealer will buy another currency*. The ***ask*** *is the price at which a dealer will sell the other currency*. Dealers bid (buy) at one price and ask (sell) at a slightly higher price, making their profit from the spread between the prices. The bid-ask spread may be quite large for currencies that are traded infrequently, in small volumes, or both. Bid and ask quotations in the foreign exchange markets are superficially complicated by the fact that the bid for one currency is also the offer for the opposite currency. A trader seeking to buy dollars with euros is simultaneously offering to sell euros for dollars. For simplicity we regard to FX-rates as mid rates (average of bid and ask) from now on.

7.2 FX markets: demand and supply

FX rates can be modelled in a simple supply-demand framework. To avoid ambiguity the market is observed from a domestic point of view. Demand and supply of foreign currencies can be connected to the international transactions of the home economy. Imports of goods and services and exports of capital generate FX demand, while exports of goods and services and imports of capital generate supply.

If a Hungarian company buys foreign goods - for example Swedish furniture - it has to buy foreign exchange first: ***import of goods → FX demand***

If an investor who has HUF-denominated assets in Hungary - for example a bank deposit - wants to reallocate his portfolio by investing in American treasuries, has to buy US dollar first: ***export of capital → FX demand***

If a Hungarian manufacturer sells her product to a French merchandise retailer and is paid in euros, the she has to sell the euros for Forints, because she pays salaries, taxes, her domestic suppliers in Forint: ***export of goods → FX supply***

If German investors want to buy shares on the Budapest Stock Exchange, they have to sell euros for forint: ***import of capital → FX supply***

The market is in equilibrium when the quantity of foreign currency demanded equals the quantity supplied (Figure 1). If export of goods or import of capital increases, the supply curve shifts to the right, and the equilibrium exchange rate decreases (Figure 2, left hand panel). This is called the appreciation of the home currency (or the depreciation of the foreign currency). An increase of import of goods or of export of capital shifts the demand curve right, and the home currency depreciates (Figure 2, right hand panel).

FIGURE 1: THE FX MARKET.

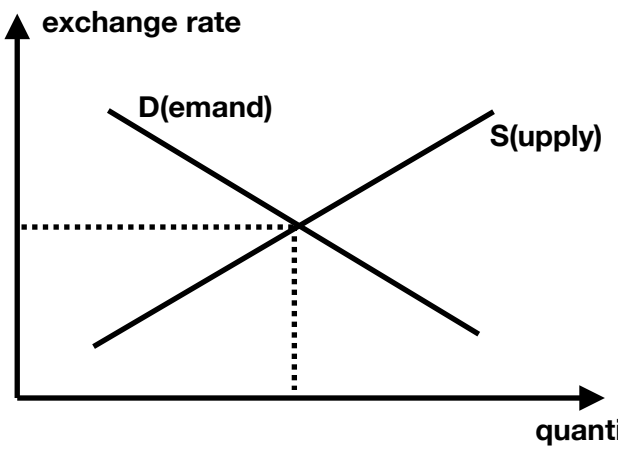
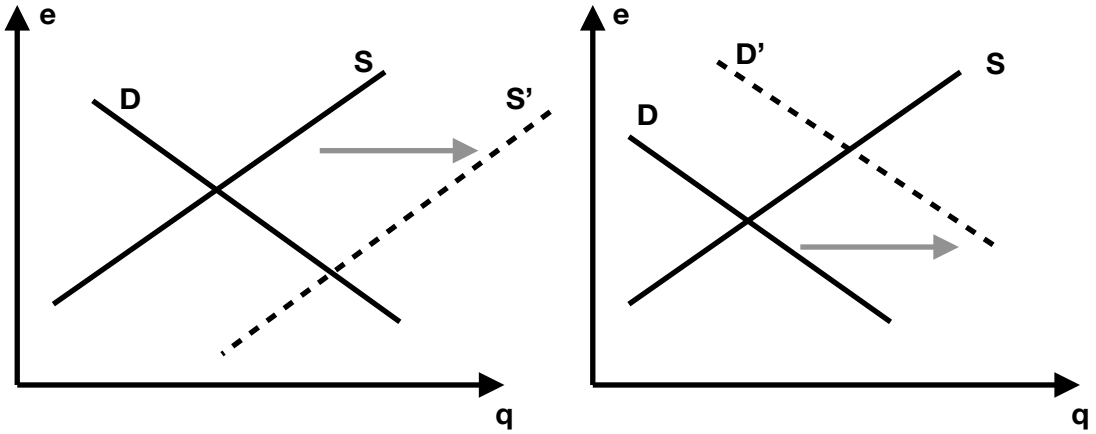


FIGURE 2: APPRECIATION (LEFT) AND DEPRECIATION (RIGHT) OF THE HOME CURRENCY.



7.3 Exchange-rate theories

A practical point for understanding how exchange rates are determined is a simple idea called the law of one price. If two countries produce an identical good, and transportation costs and trade barriers are very low, or non-existing the price of the good should be the same throughout the world no matter which country produces it. One of the most prominent theories of how exchange rates are determined is the theory of ***purchasing power parity*** (PPP). It states that *exchange rate between any two currencies will adjust to reflect changes in the price levels of the two countries*. The theory of PPP is simply an application of the law of one price to national price levels.

> Suppose that German steel costs 100 euro per ton and Japanese steel with the same quality costs 10,000 yen per ton. For the law of one price to hold, the exchange rate between the yen and the euro must be 100 yen per euro (€0.01 per yen), so that one ton of German steel sells for 10,000 yen in Japan (the price of Japanese steel) and one ton of Japanese steel sells for €100 in Germany. If the exchange rate were 200 yen to the euro, Japanese steel would sell for €50 per ton in Germany or half the price of German steel, and German steel would sell for 20,000 yen per ton in Japan, twice the price of Japanese steel. This situation generates excess demand for the Japanese steel, excess supply of the German steel and a demand for the yen (as German companies buy yen for euro). As a result, the price of Japanese steel increases, the price of German steel decreases, and the EUR/YPN exchange rate decreases (the yen appreciates) until the exchange rate equals the ratio of the two prices.

Applying the law of one price to the price levels in the two countries produces the theory of purchasing power parity, which claims that if the Japanese price level rises 10% relative to the Eurozone price level, the euro will appreciate by 10%. As our example illustrates, the theory of PPP suggests that if a country's price level rises relative to another's, its currency should depreciate (the other currency should appreciate). Another way of thinking about purchasing power parity is through a concept called the ***real exchange rate***, *the rate at which domestic goods can be exchanged for foreign goods*. In effect, it is the price of domestic goods relative to the price of foreign goods denominated in the domestic currency.

> If a basket of goods in New York costs $50, while the cost of the same basket of goods in Tokyo costs $75 because it costs 7,500 yen while the exchange rate is at 100 yen per dollar, then the real exchange rate is 0.66 (= $50/$75). The real exchange rate is below 1.0, indicating that it is cheaper to buy the basket of goods in the United States than in Japan.

The real exchange rate indicates whether a currency is relatively cheap or not. Another way of describing the theory of PPP is to say that it predicts that the real exchange rate is always equal

to 1.0, so that the purchasing power of the dollar is the same as that of other currencies such as the yen or the euro.

This prediction of the theory of PPP could work out in the long run. From 1973 to 2013, the British price level rose 102% relative to the U.S. price level, and as the theory of PPP predicts, the dollar appreciated against sterling, though by 60%, an amount smaller than the 102% increase predicted by PPP. Yet, as the same figure indicates, PPP theory often has little predictive power in the short run. From early 1985 to the end of 1987, for example, the British price level rose relative to that of the United States. Instead of appreciating, as PPP theory predicts, the U.S. dollar actually depreciated by 40% against the pound. So even though PPP theory provides some guidance to the long-run movement of exchange rates, it is not perfect and in the short run it is a particularly poor predictor. What explains PPP theory's failure to predict well? The PPP conclusion that exchange rates are determined solely by changes in relative price levels rests on the assumption that all goods are identical in both countries and that transportation costs and trade barriers are very low. When this assumption is true, the law of one price states that the relative prices of all of these goods (that is, the relative price level between the two countries) will determine the exchange rate. The assumption that goods are identical may not be too unreasonable for German and Japanese steel, but is it a reasonable assumption for American and Japanese cars? Is a Toyota the equivalent of a Chevrolet? Because Toyotas and Chevys are obviously not identical, their prices do not have to be equal. Toyotas can be more expensive relative to Chevys and both Americans and Japanese will still purchase Toyotas. Because the law of one price does not hold for all goods, a rise in the price of Toyotas relative to Chevys will not necessarily mean that the yen must depreciate by the amount of the relative price increase of Toyotas over Chevys.

Furthermore, PPP theory does not take into account that many goods and services (whose prices are included in a measure of a country's price level) are not traded across borders. Housing, land, and services such as restaurant meals, haircuts, and golf lessons are not traded goods. So even though the prices of these items might rise and lead to a higher price level relative to another country's, the exchange rate would experience little direct effect. As empirical studies find PPP can hardly explain the change of exchange rate, so it is not suitable for forecasting or operative decision making.

In the short-run exchange rates are determined by capital flows. With free capital mobility investors (re)-optimise their portfolios according to changes in expected returns and risks. Suppose that there are two portfolios, the home portfolio (H) and the foreign portfolio (F) characterised by the (r_H, σ_H) and (r_F, σ_F) expected return - risk pairs respectively. Assuming that investors are risk-averse, a change in either of the four elements leads to a shift in portfolio weights. If the return of the home portfolio increases, investors reallocate their investments towards the home country, thereby appreciating the home currency. A decrease of the risk of the home portfolio leads to the same result. These are the cases of capital import from a domestic point of view. Capital export and thereby depreciation can be triggered by either a decrease of the $(r_H - r_F)$ difference or by an increase of the $(\sigma_H - \sigma_F)$ difference.

7.4 Exchange-rate systems

Normally, the trend of the exchange rate is determined by the fundamentals of the economy. The real appreciation of the home currency in the long rung is the mirror image of a higher potential growth rate of the home economy. However, capital movements triggered by the changes of investors return-risk perceptions can divert the exchange rate from its long-run trend. Though they are temporary, these diversions can cost economic agents a lot.

Suppose that a Hungarian company imports raw material from Brazil, and sells its product in Germany. Both the import and the export are settled in euro. Suppose furthermore that the long-run equilibrium exchange rate is 320 EUR/HUF, and the company buys the euro to pay for the import at this rate. However, by the time it finishes manufacturing and sells the product, the Forint has appreciated by 5 percent, thus the rate at which the company can sell the euros received for the export is 304 EUR/HUF. If the competition is fierce in the industry, this change can turn the company's profit into a loss.

The higher the openness - measured by the total value of exports and imports relative to the GDP - of an economy is, the more sensitive its performance is to exchange rate movements. Especially in small open economies the exchange rate is the most important element of the price system. Therefore economic policies have to define their relation to the exchange rate by determining an exchange rate regime.

The exchange rate regime is called ***floating*** when *market demand and supply determines the FX rate* without any central bank intervention. The central bank can influence the exchange rate both as a market participant and both as an authority. In the first case it sell (buys) foreign currency if it wants to strengthen (weaken) the home currency. As an authority the central bank can restrict convertibility by imposing barriers to currency exchange. In formal terms a central bank intervention shifts either the demand or the supply curve to the right, restricting convertibility shifts these curves to the left.

In a ***fixed exchange rate system***, *exchange rates are either held constant or allowed to fluctuate only within an explicit band.* In general, the central bank must offset any imbalance between demand and supply conditions for its currency in order to prevent its value from changing.

In 2001 the Hungarian central bank introduced a fixed exchange rate regime with a $\pm 15\,\%$ band. The centrum of this band was 286 EUR/HUF the weak and the strong edges were 324 HUF/EUR and 240 HUF/EUR respectively. Market forces could freely determine the exchange between these two values. However whenever the exchange rate reached the weak (strong) edge the central bank had to sell (buy) euros to keep the exchange rate within the band.

Central banks can strengthen the home currency to the extent of their foreign reserves. Buying FX is not technically constrained, because the central bank creates the home currency. However, as money creation can threaten the central bank's inflation target (see Chapter 2), sooner or later intervention has to be suspended.

When the central bank bank bumps into one of these constraints, it can either reset or exit the exchange rate regime. Reseting means that the parameters of the band are changed. In case of a weakening home currency and insufficient FX reserves the central bank ***devalues*** the home currency by *setting the centrum of the band higher*. Devaluation is a legal act, an official admission that the home currency was overvalued before. In case of appreciation and fear of inflation the central bank ***revalues*** the home currency by *setting the centrum of the band lower*. A third case of reseting the fixed exchange rate regime is broadening the band.

In February 2008 a massive capital inflow appreciated the Hungarian Forint, and the exchange rate reached the 240 EUR/HUF limit. However, the Hungarian Central Bank did not intervene, instead it changed the exchange rate regime from a fixed to a floating one.

A fixed exchange rate can be beneficial to a country for several reasons. First, exporters and importers on the goods and services markets can engage in international trade without concern about exchange rate movements. Any firms that accept the foreign currency as payment would be insulated from the risk that the currency could depreciate over time. In addition, any firm that need to obtain that foreign currency in the future would be insulated from the risk of the currency appreciating over time. Firms can engage in direct foreign investment as well, as they hopefully can convert their foreign currency earnings into home currency without a major change in the exchange rate. Third, investors would be able to invest funds in foreign countries without concern that the value of their investments decreases because of an appreciation.

A disadvantage of a fixed exchange rate system is that there is still a risk of the central bank changing its currency's value. Furthermore, from a macroeconomic point of view, a fixed exchange rate system may render the country more vulnerable to economic conditions in other economies. Remember, PPP claims that the the home and foreign inflation rates are in a long-run equilibrium relationship with the exchange rate. If the exchange rate is fixed, the inflation rate of the other economy is imported. Thus with a fixed exchange rate a high foreign inflation can hurt the home economy. A floating exchange rate system can insulate the home economy from the inflation of other countries.

> If the United States experiences a high rate of inflation, then the resulting increased US demand for UK goods will place upward pressure on the value of the British pound. The appreciation will make UK goods more expensive for US consumers, even though UK producers have not raised their prices. The higher prices are caused by the pound's appreciation; that is, a greater number of dollars are now required to buy the same number of pounds. In the United Kingdom, the actual price of the goods (as measured in British pounds) may be unchanged. Even though US prices have increased, UK consumers will continue to purchase US goods because they can exchange their pounds for more US dollars.

Another advantage of a floating exchange rate is that a country is more insulated from unemployment problems in other countries.

> Under a floating rate system, the decline in US purchases of UK goods will lead to reduced US demand for British pounds. Such a demand shift could cause the pound to depreciate against the dollar. This depreciation will make British goods cheaper for US consumers than before, offsetting the reduced demand for these goods that may follow a reduction in US income. As was true with inflation, a sudden change in unemployment will have less effect on a foreign country under a floating rate system than under a fixed rate system.

These examples illustrate that in a freely floating exchange rate system, the problems experienced in one country will not necessarily be contagious. Exchange rate adjustments serve as a form of protection against "exporting" economic problems to other countries. An additional advantage of a freely floating exchange rate system is that a central bank is not required to constantly maintain exchange rates within specified boundaries. It is therefore never required, just for the sake of controlling exchange rates, to implement an intervention policy that could have an unfavourable effect on the economy. Furthermore, each government is free to implement policies irrespective of their effect on the exchange rate. Finally, if exchange rates were not allowed to float, then investors would invest funds in whatever country had the highest interest rate. The likely result would be governments of countries with low interest rates seeking to restrict the exit of investor funds from the country. Hence there would be more restrictions on capital flows, and financial market efficiency would be reduced.

In the previous example, the United Kingdom is somewhat insulated from the problems experienced in the United States because of the freely floating exchange rate system. Although

an advantage for the protected country (here, the United Kingdom), this insulation can be a disadvantage for the country that initially experienced the economic problems.

> If the United States experiences high inflation then the dollar may weaken, thereby insulating the United Kingdom from the inflation (as discussed previously). From the US perspective, however, a weaker US dollar causes import prices to be higher. This may increase the price of US materials and supplies, which in turn would increase US prices of finished goods. In addition, higher foreign prices (from the US perspective) can force US consumers to purchase domestic products. As US producers recognise that their foreign competition has been reduced by the weak dollar, they can more easily raise their prices without losing customers to foreign competition.

In a similar manner, a freely floating exchange rate system can adversely affect a country that has high unemployment.

> If the US unemployment rate is rising then US demand for imports will decrease, putting upward pressure on the dollar's value. A stronger dollar will then cause US consumers to purchase foreign rather than US products because the foreign products now can be purchased cheaper. However, that reaction can be detrimental to the United States during periods of high unemployment.

CHAPTER 8 *PUBLIC FINANCE AND TAXATION*

8.1 The economic functions of the government

The modern state has three economic functions according to Musgrave. These are:

- *allocation,*
- *redistribution,*
- *stabilization.*

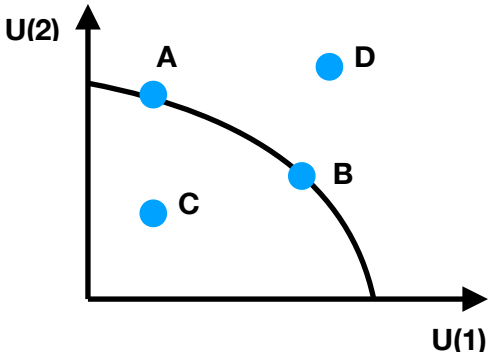
These functions represent the necessity of public sector in the economy. There is another so-called prefunction or zero function namely the creation of legal institutions. Examine the functions in order.

Allocation of resources

According to the ***First Fundamental Theorem of Welfare Economics*** the *perfect competition leads to Pareto-efficient allocation of resources.* In Pareto optimum no individual utility can be made better off cannot without worsening someone else's utility.

The Utility Possibilities Curve shows the maximal available utility of the second person (or social class) in function of the achieved utility of the first person (or social class) at a given level of resources of the examined economy. (Figure 1) Point "D" signs an unavailable output, point "C" is a suboptimal, while point "A" and "B" are Pareto-optimal.

FIGURE 1: UTILITY POSSIBILITIES CURVE



If the conditions of First Welfare Theorem are not met, the resource allocation will not be Pareto-optimal (point "C"). The market failures are the following:

- *public goods,*
- *externalities,*
- *asymmetric information,*
- *monopoly.*

Public goods have two characteristics: their *consumption is non-rivalrous and non-excludable*. A good or service is non-rivalrous when someone's consumption does not disturb others' consumption. Individual utility from consumption will not be less if other individuals consume the same good. A good or service is non-excludable if it is possible to prevent consumption in case of non-payment. By the two characteristics four sets of groups can be constructed (Table 1).

TABLE 1: THE SETS OF GOODS.

Consumption	Excludable	Non-excludable
Rivalrous	Private goods (*apple*)	Common goods (*air*)
Non-rivalrous	Club goods (*theatre*)	Public goods (*lighthouses*)

Why does the market fail in the case of public goods while in the case of the private goods it works efficiently? If no one can be excluded from the consumption of a good and at the same time the consumption does not disturb others then nobody will be interested in producing this good. ***Free riding*** is a rational habit/attitude: just *enjoying the produced good but not participating in the financing/funding*. If everyone thinks rationally no one will produce public goods, however, there is a massive demand for them. In the implementation of producing public goods the government can play a role. If the market cannot achieve the Pareto-optimal resource allocation (market failure), then the government has to intervene and solve the problem. In this case the public sector can improve the efficiency – we call it Pareto-improvement. Adam Smith also supported the state's intervention in the case of public goods, however, he devoted minimal role to the public sector.

The second type of market failures is ***externality***. When *an economic actor does not calculate with every consequence of his activities* we speak about external economic effect. Imagine a company that works near by a river and pollutes it.

Suppose, that the company does not care about the pollution while calculating the costs and returns. The marginal profit function is decreasing in line with the quantity of production as Figure 2 shows. This marginal profit becomes 0 when the production achieves the quantity Q'. Above this quantity the total profit will decrease because the marginal profit will be negative. At this point (Q') the profit is at its maximum so this represents the market equilibrium. Assume that there is a beach near by the river a few kilometres away from the factory. The pollution of the factory causes cleaning costs for this beach. This (external) costs are increasing in line with the quantity of production as Fig. 2 shows. The external costs and the marginal profit functions cross at quantity Q*. This is a special point. The whole marginal profit function for the society (hereafter the social marginal profit) comes from the distinction of individual marginal profit and the external cost functions (see Figure 3). At Q* the social marginal profit is zero. This point will be the Pareto-optimal quantity. If the factory produces less than Q*, its marginal profit would exceed the external cost of pollution so the company would be able to compensate the cleaning cost of the beach. If the factory compensates the beach indeed the company can increase its production until Q*. If the production exceeded Q*, the marginal profit wouldn't be able to compensate the external costs since it is less than the external cost. At point Q* nobody's utility can improve without damaging anyone else's utility.

FIGURE 2: EXTERNAL COST AND INDIVIDUAL MARGINAL PROFIT FUNCTIONS.

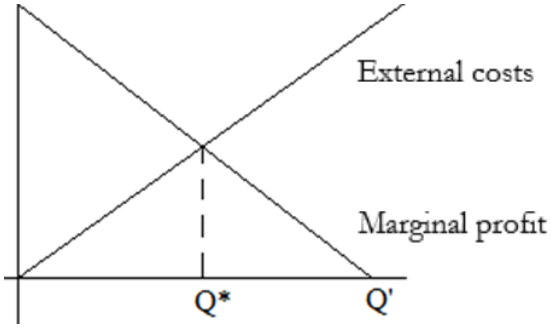
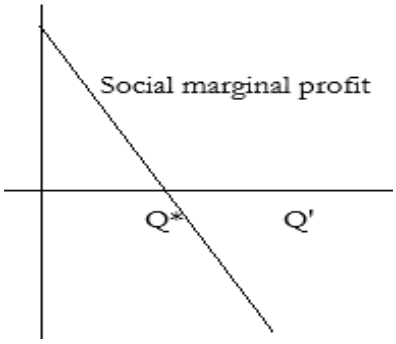


FIGURE 3: SOCIAL MARGINAL PROFIT FUNCTION



According to the ***Coase Theorem*** i*f the government makes the rights clear, the profit- and utility seeking behaviour of market participants lead to Pareto-efficient resource allocation.* From the aspect of efficiency the owner of the right does not matter. However from the aspect of equity the two cases are not identical since the owner of the right will get compensation. In this case the market solves the problem after the state has declared the rights. But usually there are a lot of victims of an externality far in time and in space. It is hard to define the compensation. In these cases the government uses *taxes for collecting the external costs from the causers*. Pigou suggested that the amount of tax should be equal to the amount of external costs. These taxes are called ***Pigouvian-taxes***.

The government can decide if the company has the right to pollute or the beach has the right for clear water. In the first case the company gets compensation from the beach for decreasing the production from Q'. While in the second case the beach gets compensation from the factory for increasing production from zero quantity. In both cases the equilibrium is at Q*.

The First Welfare Theorem is true in the case of perfect competition. According to conditions of perfect competition the ***information*** is symmetric however, it is usually ***asymmetric***. It can cause two types of problems: moral hazard and adverse selection. The Pareto-optimal resource allocation is not ensured automatically. The market can solve these problems (by using signals, filters or incentive mechanisms) in many times but not in every case.

Without governmental intervention none of the banks would supply students loan anywhere in the world. The governmental intervention has several modes, for example the warranty.

In the case of ordinary monopolist the pricing mechanism causes deadweight losses so the output is not Pareto-optimal. (In the case of price discriminating monopoly there is no deadweight loss but the problem of distributive justice is present - see the following chapter.) Usually the government prohibits the present of ***monopoly*** but in several cases it is impossible. When the average profit is increasing with the production (economies of scale) the present of monopoly is natural. It is usual if the costs of entry are high while the running expenses are much less relative to it. Consider for example drainage. The investment cost is high while the cost of the joining of an additional consumer is close to zero. Instead of prohibiting these natural monopolies the government controls them.

Redistribution of income

Redistribution of incomes is the second economic function of the state. At Figure 1 point "C" was not efficient and point "D" was unavailable. What about point "A" and "B"? Both of them are efficient but are they the same from other aspects? Point "A" signs a less symmetric distribution than point "B". What if the society prefers "B" to "A" because the first one is a more fair situation? According to the ***Second Fundamental Theorem of Welfare Economics*** *any of the Pareto-efficient resource allocations can be achieved by perfect competition in case of the adequate initial redistribution of resources*. It means that the government only has to intervene at the beginning of the process and thereafter it leaves the economic actors to decide alone. In this case the result will be Pareto-efficient and fair as well. This is the base of the ***mixed economy***: *the market and the government work together side by side*.

The main governmental instruments of the distribution function are (progressive) personal income taxation, social transfers and obligatory social security (health care and pension funds).

Macroeconomic stabilization

Finally, the third economic function is stabilisation. Besides efficiency and fairness there are four other important aspects which cannot be ignored by governments. The fiscal policy has to pay attention to the following four elements: price stability, employment, external balance and economic growth.

8.2 Revenues of the Government

The main instruments of the government to achieve its goals are taxation, transfers and expenditures. This chapter deals with the revenues.

Classes of taxes

The receipts of the government have 3 main forms: taxes, charges and borrowing. In the case of borrowing the government has to repay it and to pay some interest as well. Before examining the tax categories it is useful to determine the basic definitions of taxation.

Tax is a *statutory monetary payment liability imposed unilaterally by the state without any reward.* Statutory: determined in a law. Monetary payment: presently fulfillable only by money but in the past by product or services were usual. Unilaterally: the government decides without asking the citizens. Without any reward: the tax payer cannot claim anything against it.

Tax payer is a (legal or natural) person *who is imposed by the tax law*.

Tax base is *the quantity which is imposed by tax law*. It can be a nominal value or a natural measure of the taxable subject. The former is typical in the case of income tax while the latter occurs in the case of property taxes (for example a flat's area in square metres or the engine capacity of a car).

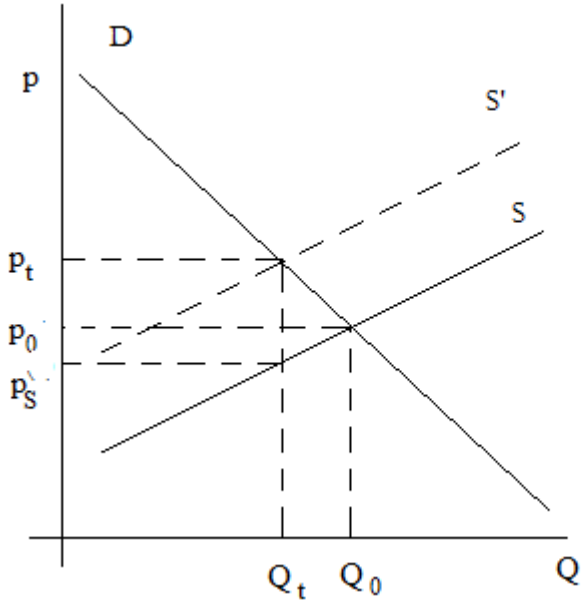
The design of ***personal taxes*** can r*eflect the economic characteristics* (mainly the ability to pay) o*f the tax payer*. The so-called ***In rem taxes*** *depend on the transaction or object which they are imposed on*, independent of the tax payer's circumstances. They can be called ***direct*** and ***indirect taxes***. These expressions are used to be approached from another aspect. In case of direct taxes the tax payer bears the burden of the tax while in the case of indirect taxes the tax payer can shift it. To make it clear the process of tax incidence has to be considered.

Figure 4 shows the market of a good. The supply is signed by *S* before levying any tax. Let's assume that the government decides to tax the selling of the good by an amount of *t*. After this decision the curve of supply will move forward by *t S',* parallel with the original. The original market price (p_0) increases, the new price is p_t. After the seller gets this amount he has to pay the tax (*t*) and keeps the difference (p_S) between the market price and the tax paid.

$$p_S = p_t - t \tag{1}$$

The burden of the tax borne by the buyer is: $p_t - p_0$ and by the seller is: $p_0 - p_S$. They shared the burden between themselves.

FIGURE 4: TAX INCIDENCE



Tax rate can be expressed as a *function which shows the amount of the tax payment at a given level of the tax base*.

$$T = t(Y) \tag{2}$$

Average tax rate function shows *the ratio of the tax payment to the tax base*. It is called effective tax rate as well.

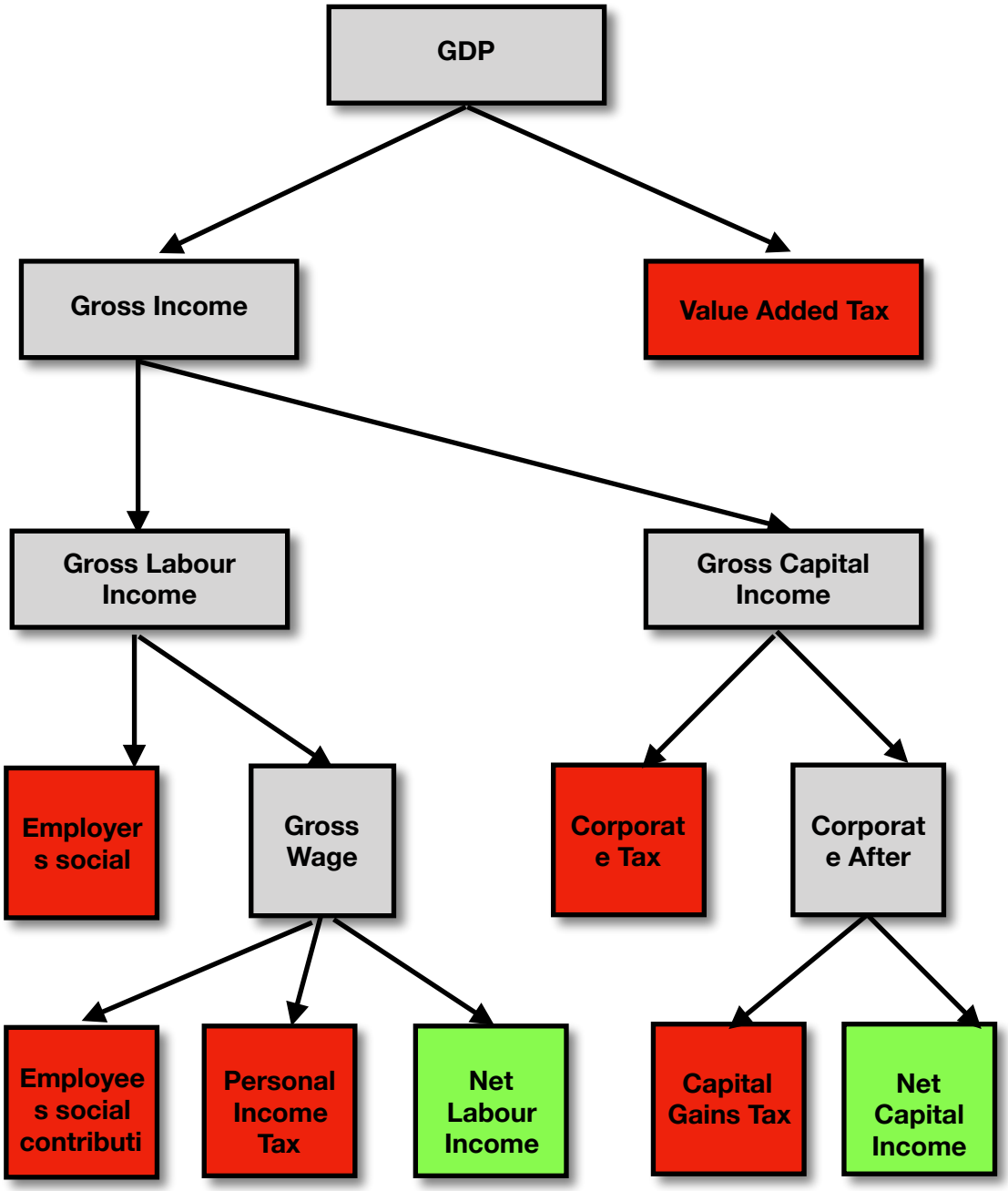
$$a(Y) = \frac{T}{Y} = \frac{t(Y)}{Y}. \tag{3}$$

Marginal tax rate function calculates *the tax burden of an additional unit of tax base*.

$$m(Y) = \frac{\delta t(Y)}{\delta Y}. \tag{4}$$

The tax rate is called ***progressive, proportional (flat rate) or regressive*** if its *average tax rate function is increasing, constant or decreasing in order*.

FIGURE 5: THE TAXTREE.



Value Added Tax (VAT) is an indirect tax on transactions. Theoretically the tax base is the value added of the product or service between phases of production. In practice the

taxpayer has to pay the rate of VAT based on the whole value but the previously paid VAT for the purchased products and services is deductible. The GDP (Gross Domestic Product) contains the VAT. If the rate of VAT is 20 %, the amount of it will be the 16.67 % of the GDP as the following equations give us. Equation (5) shows the decomposition of GDP.

GDP = VAT + taxbase of VAT = 0,2x taxbase of VAT + taxbase of VAT = 1,2 x taxbase of VAT (5)

From (5) the VAT in the function of GDP if the VAT rate is 20%.

$$VAT = 0.2 * \frac{GDP}{1,2} = 0.1667 * GDP$$

GDP can be divided according to the contribution of production factors, i.e. capital and labor. The gross labour income consists of the *employer's social contribution* and of the gross wage of the employees. The former can be calculated similar to VAT. If the rate of social security contribution of the employer is 20%, the contribution will be 16.67% of the gross income from labour. The employee's net wage is derived from the gross wage after deducting the *personal income tax* and the *employee's social contribution*. The gross wage is the tax base of both.

The gross income from capital is the companies' taxable profit charged by the ***corporate tax***. After paying the tax the company can reinvest and/or pay some dividend for the owners. They get the net capital income after deducting the ***capital gains tax***.

Besides the taxes on the taxtree there are several other types of taxes in the world. Duties, excise taxes on tobacco, alcohol and fuel, property or wealth taxes, ecotaxes or green taxes etc. The most important revenue is realised from the excise tax on tobacco, alcohol and fuel because the demand of these products is reasonably inelastic. The economic weight of property or wealth taxes are varying in the world between 0 to 10 %.

Principles of taxation

Many economists and social philosophers dealt with the requirements of a good tax structure. (They agree in many aspects but there are some differences obviously.) The most important principles are the following.

➡ *First of all the revenues have to cover the government's expenditures.*

The tax system has to take into consideration equity, that is taxes should be fair. Two principles deal with this topic.

➡ The ***benefit principle*** suggests that *everyone should pay contribution proportional to the rate as (s)he enjoys public services*. This concept does not calculate with the redistributional objectives. Usually those people require most public services (health care or education) who could not achieve private services.

➡ The ***principle of ability-to-pay*** deals with this question fundamentally. It says that *everyone should be charged according to their capacity to pay taxes*. Horizontal equity prevails if those with the same capacity are treated equally, whereas vertical equity claims that those with higher capacity must pay higher taxes. The so-called equal sacrifice principles give 3 different solutions for the tax determination problem.

The ***Principle of equal absolute sacrifice*** requires that *the absolute change in the utility has to be the same amount for every person* in the society, that is

$$U(Y_i) - U(Y_i - T_i) = \text{constan for every i} \quad (7)$$

The ***Principle of equal proportional sacrifice*** requires that *the percentage change in the utility has to be the same ratio for every person* in the society, that is

$$\frac{U(Y_i) - U(Y_i - T_i)}{U(Y_i)} = \text{constant for every i for every i} \quad (8)$$

The ***Principle of equal marginal sacrifice*** requires that *the marginal utility of the last tax unit has to be the same for every person* in the society, that is

$$\frac{dU(Y_i - T_i)}{d(Y_i - T_i)} = \text{constant for every i for every i} \quad (9)$$

➡ *The requirement of **efficiency** claims minimal deadweight loss should be caused by the taxation.*

In Figure 4 the so-called Harberger-triangle shows the amount of it. Only some part of the decrease in consumer's and producer's surpluses was realised by the state as taxes. The remaining part is lost for the society, this is the excess burden of the taxation. The less is the price elasticity the less is the deadweight loss. From the aspect of efficiency any good with inelastic demand is an adequate tax base. However these goods have usually low income elasticity as well, thus taxing them charges mainly the poor. As there is a trade-off between efficiency and equity it is hard to decide between them.

➡ *The tax system is **neutral** if it does not disturb the ranking of the tax payers or the economic projects.*

If person *X* has higher gross income or wealth than person Y, then their positions should not change with taxation. Similarly if the net present value of project *A* is higher than of project *B* their relation should not vary with taxation.

➡ *The costs of administration should be as low as possible.* These costs appear at the tax authority and at the taxpayers as well.

➡ ***Elasticity** of a tax system shows its ability to adapt to the varying circumstances.*

➡ *Other features of the good tax system are simplicity, transparency, and stability.*