1. Basic economic concepts

The aim of the lecture is to review basic economic concepts as they are applied to analysing and solving environmental problems. This introductory class can be summarised as follows. Every bullet will be explained in detail.

- Economics is about how people make choices in routine voluntary transactions
- Environmental and resource economics is about how people make choices regarding environmental quality and natural resource use
- People's decisions are based on the demand and supply of what they are interested in
- Willingness To Pay (WTP) and Willingness To Accept (WTA) are key concepts that help understand how people make choices

Many people think of economics as something which gives answers to questions: "how to become rich?", "why is the price a given stock sky-rocketing?", "what to do in order to eradicate poverty?" and so on. These are important questions, but they do not define economics. For at least 100 years, most economists have understood their discipline as the science of making choices. Of particular interest are choices when people cannot have everything they would like to have, so that when they wish to have something, they need to give up something else.

It is important to emphasise that economics is about choices made by ordinary people – not necessarily by the clever and the virtuous ones. We often say: "she must be an idiot if she spends so much on something" or "if he were ethical, he would not have done this". As citizens, neighbours, teachers, parents we can say so, but economists are not expected to judge choices made by other people. They are supposed to analyse people's decisions, but they should not call them right or wrong.

Transactions establish typical proportions certain goods are exchanged for each other. For instance, two apples go for one orange. We say that one orange is worth two apples, or one apple is worth a half of an orange. Often one good becomes so common that people refer to it as a *numeraire*. They say: one orange is worth 1 €. The good referred to as a *numeraire* is called money. We are used to consider some currencies – like euro – as money. Yet many other goods served in this capacity for a long time: cattle, gold, vodka, man-hour, or what have you.

It is also important to emphasise that economics is about routine transactions. Economists have little to say about transactions that are unique. If a criminal kills somebody "in exchange" for 100 € it does not mean that the victim's life was worth 100 €. Or if somebody spends 1 million € to cure herself it does not mean that the sickness was worth 1 million €. Only routine transactions are analysed in economics. Unique – i.e. non-routine – choices can be studied in psychology (or in other disciplines), but not in economics.

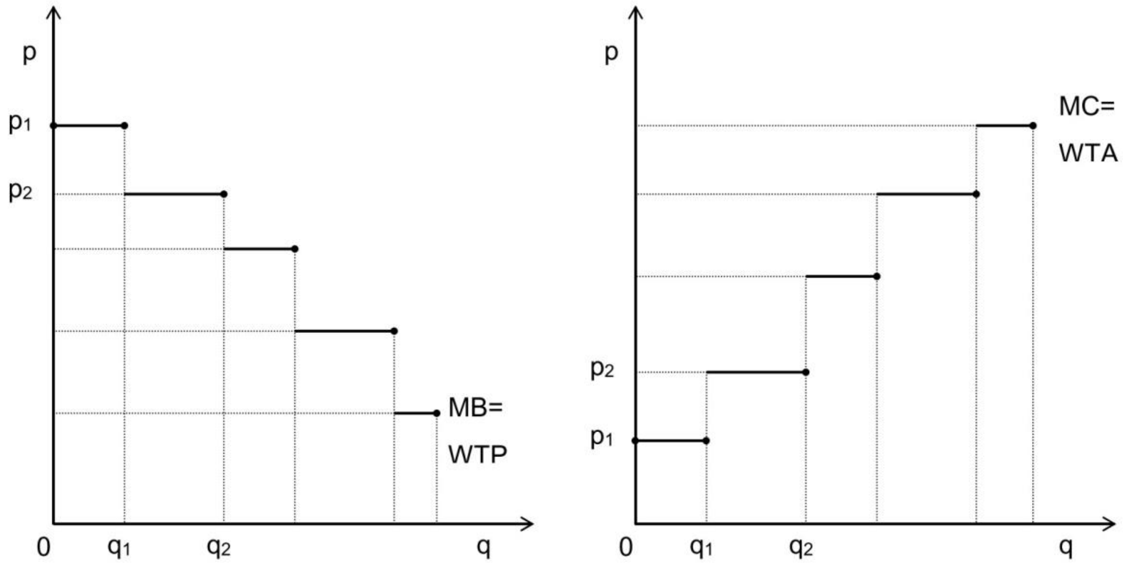
The last adjective in the first bullet point reads "voluntary". Economics does not analyse transactions which are involuntary (that are forced). For instance, the behaviour of slaves in ancient Egypt, or prices charged in Soviet stores cannot be explained by economic methods. Voluntariness makes an important assumption, since many economic analyses are based on it.

For example, economists argue that a purchase makes the buyer better off. Unless it is voluntary, the argument cannot be applied.

Thus economics is about how people make choices. They make choices every day, whenever they take decisions about what to do, what to buy, what to sell, whom to support, etc. In particular, they take decisions to what extent protect the air, or how much water to use. Sometimes we have an impression that there is nothing to be chosen. Indeed, often the choice is obvious. If I asked you, what do you prefer? Is it better to be young, clever, healthy and rich, or to be old, stupid, sick and poor? Most of us – as confirmed by the students in my class – would prefer the first option. Many choices we have to make are not that obvious. Let us assume that we are asked: do you prefer to breathe with clean air or faulty air? The answer is that we prefer clean air. Yet clean air does not come for free. A large part of air pollution is caused by cars. Are we ready to reduce the car traffic in order to enjoy cleaner air? This is a choice we need to do, and economics helps us to understand which option is justified given our preferences revealed by our activities.

Key concepts used in economic analyses are so-called Willingness-To-Pay (WTP), and Willingness-To-Accept (WTA). They are crucial in order to understand the demand and supply. These terms sound familiar whenever we talk about market goods such as apples and oranges, but they are indispensable in environmental and resource economics as well. It is fairly easy to imagine the demand for environmental protection – this is what we are ready to sacrifice in order to enjoy a better environmental quality. But what do we understand as "the supply of environmental protection"? Briefly speaking, the supply of environmental protection is the availability of options that let us enjoy cleaner air, cleaner water, nicer landscapes, more resilient ecosystems, etc.

The following graph explains how prices and quantities can be combined in order to construct demand and supply schedules.



Vertical axes measure prices (p) and other economic variables, and horizontal axes measure physical variables, such as quantities (q). Symbols WTP and WTA were explained already, so the only unexplained acronyms are: MB and MC. Those of you who studied economics know that MB stands for Marginal Benefit, i.e. benefits brought by an additional unit of a good, and MC stands for Marginal Cost, i.e. costs required by an additional unit of a good.

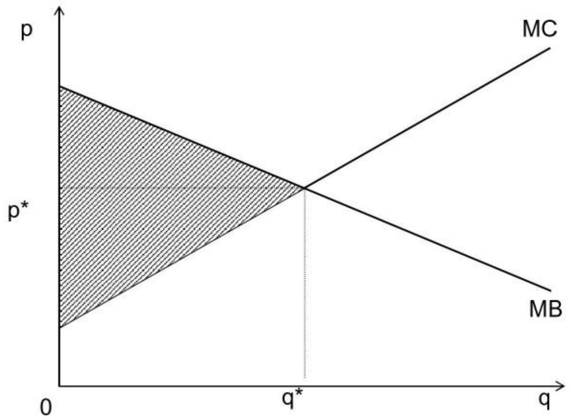
We will assume that both left and right pictures refer to what happens in the market for apples. Let us start to analyse the left picture. If the price of apples is very high, say, 50 €/kg, then nobody is willing to pay such an amount and the demand for apples is 0. If the price considered is lower, say, p_1 = 4 €/kg, then there may be some people ready to pay such a price, and the demand will be q_1 (see the highest step in the graph). If the price considered is still lower, say, p_2 = 3 €/kg, then there may be some more people ready to pay such a price. Those who paid the higher price p_1 – their demand was q_1 – will belong to this group (if they were ready to pay a higher price, they will certainly pay a lower one), but there may be some other people who were not ready to pay p_1, but they are willing to pay p_2. Their additional demand for apples is q_2-q_1, so the total demand from both groups is q_2 (see the second highest step in the graph). The story goes on. The lower the price, the higher the demand. That is how the left picture – the demand schedule – is constructed.

The analysis of the right picture starts with the lowest possible price, i.e. 0. The supply of apples is 0 at this price (no seller is willing to accept such a price). If the price considered is higher, say, p_1=0.2 €/kg, there are some suppliers who are willing to sell q_1 at this price (see the lowest step in the graph). If the price considered is still higher, say, p_2=0.5 €/kg, then there may be more suppliers to accept such a price. Those who accepted the lower price p_1 – their supply was q_1 – will belong to this group (if they were ready to accept a lower price, they will certainly accept a higher one), but there may be some other people who were not ready to accept p_1, but they are willing to accept p_2. Their additional supply of apples is q_2-q_1, so the total supply from both groups is q_2 (see the second lowest step in the graph). The story goes on. The higher the price, the higher the supply. That is how the right picture – the supply schedule – is constructed.

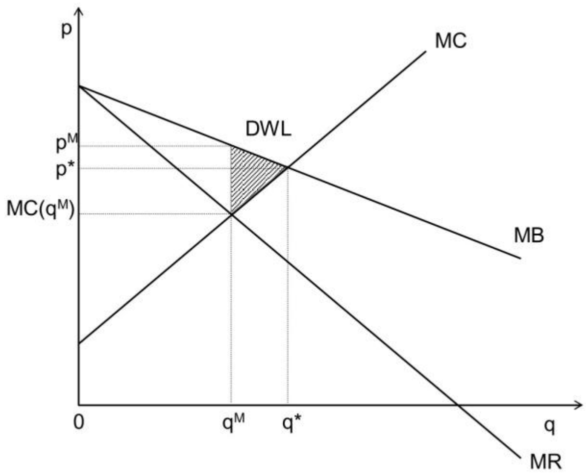
In your high school mathematics, you usually followed a convention that independent variables are measured along horizontal axes, and dependent variables – along vertical ones. Here it seems as if the quantities were independent and the prices were dependent. Sometimes we feel that it is the other way around: the quantity depends on the price. Economists do not claim that one variable is independent, and the other one – dependent. They simply indicate that prices and quantities are interrelated.

One last note in this exposition is about the continuity. The demand and supply curves are usually printed as continuous lines. Please look at the picture above. It shows the same line, except that what you see on the left can be zoomed-in 8 times (in the middle) and 64 times (on the right). In other words, what you get depends on how detailed a view you are interested in. In the case of apples it is whether you are interested in kilogrammes, tonnes or thousands of tonnes. If you wish to see a broad relationship, you are likely to draft a line like the one on the left. If you are interested in kilogrammes rather than tonnes, you should expect a picture like the one on the right (resembling steps from the previous page).

The graph below summarises the relationships studied previously. The MC line is a typical supply relationship, while MB represents a typical demand relationship. The shaded triangle is called economic surplus. It consists of two parts: the upper one (above the p^* level) is called consumer surplus, and the lower one (below the p^* level) – producer surplus. Let us combine the demand and supply schedules from page 2, and let us refer to the apple example. If the price is very high the demand is small, and the supply is big; their numbers are not consistent (more apples are available than requested). If the price is very low, it is the other way around: the demand is big, but the supply is small; their numbers are not consistent either (less apples are available than requested). There is an intermediate price (economists call it equilibrium price p^*), such that the demand and the supply are equal to each other (economists call it equilibrium quantity q^*). The combination (q^*,p^*) is called market equilibrium.

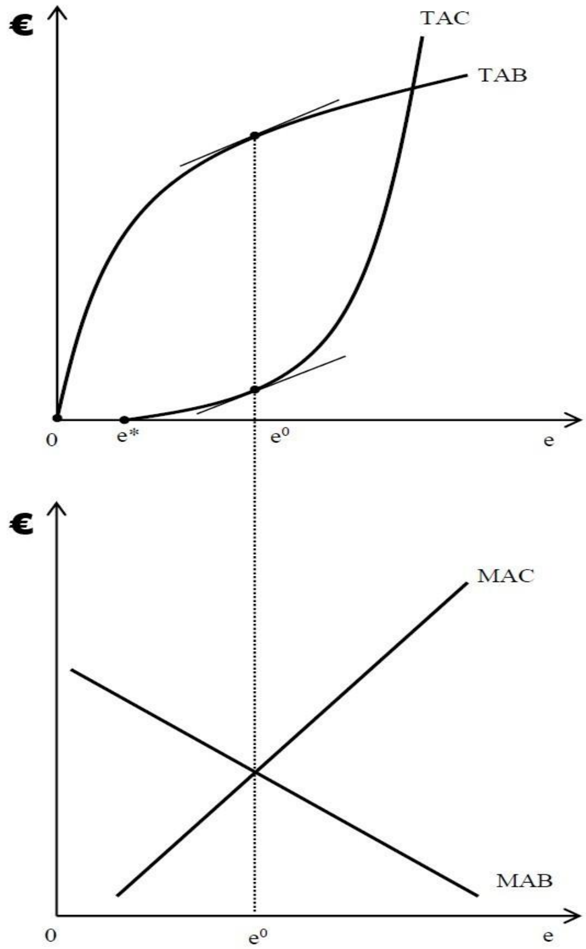


Market equilibrium makes the supply and demand equal to each other. It has another important property: if buyers and sellers are price-takers (i.e. they take prices as given; they cannot manipulate prices) it maximises economic surplus. If the quantity is lower than q^*, then – as a trapezium – the surplus is visibly smaller. If the quantity is higher than q^*, then the surplus consists of two triangles, but the right one should be subtracted from the left one, since it illustrates a deficit rather than a surplus (the additional units of a good in question require a higher price to produce than what the consumers are ready to pay: MC>MB; hence it is not a "surplus" – it is a deficit).



Economists analyse picture above in order to demonstrate what happens if buyers or sellers are not price-takers (they can manipulate prices). For instance, a monopolist can elevate the price from p^* to p^M by lowering the supply from q^* to q^M, since the monopolist knows that the marginal revenue is lower than the marginal benefit; the latter (i.e. the marginal benefit) is equal to the price obtained, but the former (i.e. the marginal revenue) is lower, since the price goes down if the supply go up. By manipulating quantities and prices, the monopolist reduces the economic surplus by so-called *Deadweight Welfare Loss* (DWL). Students who are not familiar with microeconomic concepts do not have to analyse the case in a detailed way. They simply should be aware of the fact that market equilibrium is not always a guarantee for the economic surplus maximisation. More examples of such "market failures" will be provided later on.

For the time being, we will study the case where nothing strange happens, and the level of environmental protection can be determined in a way consistent with the first picture from page 4. Please note that the lower picture below is almost identical with the earlier one. The difference is MAC and MAB instead of MC and MB, as well as e^0 instead of q^*. The earlier picture identified the quantity q^* which equated the demand and supply of a good (such as e.g. apples). The picture below identifies the abated (reduced) emission e^0 which equates the "demand" and "supply" of environmental protection ("A" stands for "abatement").



The horizontal axis measures abated emission (i.e. emission after abatement), while the vertical one – economic variables such as costs and benefits. TAB, and TAC stand for Total Abatement Benefit, and Total Abatement Cost, respectively (by the way, there may be some abatement possible without any cost: e^*). The difference TAB-TAC is the net benefit from abatement. It is maximised if abated emission is equal to e_0. This level of environmental protection can be interpreted as a socially justified one (a social optimum). If the required level of abatement is moved to the left or right from e_0, then the net benefit will shrink.

The same conclusion is derived when one looks at derivatives rather than total values. MAB and MAC stand for Marginal Abatement Cost and Marginal Abatement Benefit, respectively. High school graduates easily recognise that what economists call 'marginal values', mathematicians call 'derivatives'. The lower picture demonstrates that the net benefit is maximised when the MAB and MAC are equal to each other. In other words, a socially justified level of environmental protection is achieved when **the Marginal Abatement Cost is equal to the Marginal Abatement Benefit**. If the marginal abatement cost starts to be higher than the marginal abatement benefit, then further abatement is no longer justified economically. And *vice versa*. If the marginal abatement cost is lower than the marginal abatement benefit, then less abatement does not make economic sense (more abatement is justified).

MAC=MAB is thus a fundamental guiding rule of socially optimal environmental protection. What does happen if the rule is violated? There are two cases:

- MAC>MAB – "overshooting" caused by neglecting costs of protection
- MAC<MAB – inadequate protection caused by underestimation of damages

Let us analyse them in more detail. MAC>MAB means that abatement undertaken is more expensive than benefits it leads to. This can be the case, for instance, if regulations require extremely strict standards for drinking water quality. The European legislation sets the 200 μg/l limit for iron. Of course, it would be better to have even less iron in our tap water, say, 150 μg/l. There will be some benefits (MAB) from such an improvement, but probably the cost (MAC) of enforcing such a standard would be very high. If this cost was to be neglected then people could demand the standard to be stricter, even though this would not be socially optimal.

It can be the other way around. MAC<MAB means that abatement undertaken is in fact more beneficial than assumed by regulators. For instance, the European legislation set the standard of 25 $\mu g/m^3$ for fine particulates suspended in the air (so-called PM2.5 – particulates with diameter less than 2.5 μm). Setting a stricter standard, say, 20 $\mu g/m^3$, would provide benefits MAB which – according to some analysts – would be higher than the costs MAC required. If they are right, then the existing standard is inadequate. An economically justified level is lower than the standard. The society will be better off if a stricter standard is adopted.

Questions and answers to lecture 1

1.1 Prices of CDs with pop music are usually higher than prices of CDs with classical music. Is it caused by the fact that pop music is more expensive to record?

Not necessarily. Prices of CDs with various categories of music (e.g. pop *versus* classical) are determined by the supply and demand. Apparently the demand for pop CDs is stronger than for classical CDs. The supply of pop CDs and classical CDs responds to market stimuli. If prices of pop CDs are considered attractive then high-cost suppliers can enter the market. Recording technology is probably the same for classical and pop music. However pop artists can ask for higher honoraria than the classical ones.

1.2 Many students criticise the WTP concept. They argue that rich people can afford to pay more than the poor ones. Is this criticism justified?

Irrespective of whether we deem this fair or not, rich buyers can always afford more. A rich person is willing to pay more for a dress than a poor person. A rich person is willing to pay more for a travel than a poor person. The same applies to willingness to pay for clean air. Many of us will be appalled by this, as clean air is something we would like to have available in the same quantity and quality to everybody irrespective of their WTP. However, if a poor person is given an amount of money sufficient to "buy" a better quality environment, he or she may choose to buy something else. In order to prevent such a behaviour, we may be inclined to force the person to do what we think he or she should do. This is how centrally planned economies functioned, but they were ultimately rejected by societies that did not enjoy non-market allocation. There is no better organisation of an economy but letting people indicate what they are willing to pay for various goods. There are some limits to this freedom. In every market economy there are extensive legal systems narrowing the freedom of choice irrespective of whether a person has or does not have money to pay for it. Environmental protection is an example of an area where we do not want to leave everything to individual choices. Nevertheless, when taking decisions to provide certain goods (like clean air or water) obligatorily, public decision makers should be aware of the fact that people have limited WTP for these goods.

1.3 Sometimes people say "I do not know anybody to pay such a price", or "Everybody is willing to pay this amount". Are these statements justified in economics?

No, they are not. Economics is about choices made by people – not necessarily those whom we know, or who are well educated, honest, or otherwise distinct. If we do not know anybody to pay such a price (very low, or very high), it does not mean that it is unacceptable for some other people. Likewise, let us assume that everybody we know is willing to pay certain amount. This does not imply that the amount corresponds to an average person's WTP.

1.4 The "supply" of environmental protection was defined in the class as the availability of environmental protection measures. Can it be assumed to be an upward sloping curve?

Yes, we can assume this in typical circumstances. In economics we usually assume that people are rational. Therefore if something can be done in a less expensive or in a more expensive way, we choose a cheap option first. If we want to cut emission of something by, say, 40 tonnes, we may have options which deliver this for 2000 € or 3000 €. It would be ridiculous to choose the more expensive option, so we abate 40 tonnes for 2000 €. Let us assume that we would like to abate additional 40 tonnes. The cheap option was already exhausted, so we have to use the more expensive one. Thus the supply of environmental protection requires the price of 50 €/tonne if we contemplate abatement up to 40 tonnes, but from 41 to 80 tonnes it requires the price of 75 €/tonne; and so on. The availability of environmental protection is likely to resemble the right picture from the overhead ERE-1-2 (or page 2 above).

An important exception to this argument has to be mentioned. The measures available initially read "40 tonnes for 2000 € or 40 tonnes for 3000 €". It can be the case, that – because of the technological progress – when the additional 40 tonnes abatement is contemplated, the available option is much cheaper, say, 1500 € instead of 3000 €. It looks as if the supply curve was downward sloping. This paradox is caused by the fact that the first "step" (2000 €) was carried out earlier, while the second one – later. In order to argue that the supply curve is upward sloping, it has to be emphasised that options available refer to the same "state of the world".

1.5. The demand for environmental protection was assumed to be downward sloping. Please comment on this.

In economics we assume that the demand schedule is downward sloping. The lower the price, the more we demand – which is quite convincing. Yet there are some problems with empirical confirmation of this rule. Let us try to imagine the demand for waste segregation. Households are requested to segregate waste. If you do not segregate at all, you spend 0 minutes per day. If you segregate into, say, three containers, you spend 10 minutes per day. If you segregate into, say, four containers, you spend 20 minutes per day. Very sophisticated segregation systems require ten containers, and absorb you for 60 minutes every day. There is a discussion in economics how to translate time into money. In order to keep things simple, let us assume that the value of time is what you can earn. For some of us it is 12 € per hour. Hence it can be calculated that no segregation (throwing away everything without any deliberation) costs you 0, segregation into three containers – 2 €, segregation into four containers – 4 €, and segregation into ten containers – 12 €. You should sacrifice your time if you estimate your benefit from segregation to be higher than the value of the time spent.

How does this translate into environmental protection? The more you segregate, the better for the environment. Thus, if you do not segregate at all, the environment is not protected. If you segregate into three containers, the environment is protected somewhat. The more sophisticated the segregation system is, the better protection of the environment is implemented. For the time being, everything looks consistent with the theory: the demand for environmental protection is downward sloping: the better environment you want to have (you are closer to perfectly clean environment), the more you have to pay for it.

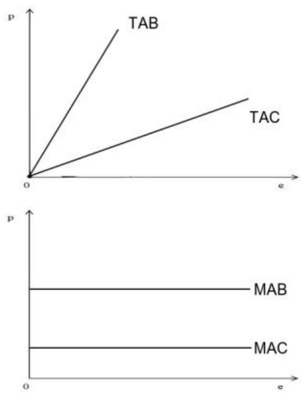
Many of us see waste segregation as something we do not derive any pleasure from, and we do this because we do not want to incur losses from environmental disruption. There is, however, some research demonstrating that our preferences may evolve. Initially, we looked at segregation as something which requires sacrificing our time. After a while, we may look at segregation as something which gives us also satisfaction (some people start deriving satisfaction from waste segregation). In other words, achieving better environmental quality may provide us with higher rather than lower marginal benefit. Hence the demand is not necessarily downward sloping, if preference evolution is taken into account. Let us refer to the numbers quoted. Initially there was a person who segregated into three containers, because he or she estimated his or her satisfaction from the better state of the environment between 2 € and 4 €. Once he or she started to derive some pleasure from segregation, the number of containers was increased to four since the benefit from better environmental protection (combined with pleasure from the segregation) was estimated at more than 4 €.

In order to argue that the demand for environmental protection is downward sloping, we assume that preferences do not change (please compare this to the last sentence of the 1.4 answer above).

1.6 As defined in the class, economic surplus can be interpreted as a shaded area below the MB curve and above the MC curve in the upper graph on page ERE-1-4 (or page 4 above). Can you use steps from picture printed on page ERE-1-2 (page 2 above) to illustrate the surplus concept?

Yes, the surplus can be interpreted by referring to step-wise functions from page ERE-1-2. Let us assume a small quantity of the good analysed (like q_1 from ERE-1-2). There are people who are willing to pay quite a lot to enjoy this quantity, and there are suppliers who are ready to accept quite a little for selling this quantity. Irrespective of the price agreed, (somewhere between WTA and WTP), both sides will gain something from the transaction. The area of the bar constrained by WTP from above and by WTA from below indicates the surplus obtained. Now let us consider the next portion of the quantity (like q_2-q_1 from page ERE-1-2). There are people who are willing to pay somewhat less to enjoy this quantity, and there are suppliers who are ready to accept somewhat more for selling this quantity. As before, irrespective of the price agreed, (somewhere between WTA and WTP), both sides will gain something from the transaction. The area of the somewhat shorter bar constrained by WTP from above and by WTA from below indicates the surplus obtained from q_2-q_1. The argument can go on until an equilibrium quantity (q^* from page ERE-1-4) is obtained. The total surplus can be interpreted as the sum of areas of all the bars between 0 and q^*. Students who are not afraid of mathematics may observe that the shaded area from page ERE-1-4 can be approximated by the area of all the bars concerned.

1.7 Is it possible to have TAB and TAC without an intersection point other than e_0=0 (see picture on page ERE-1-6 or page 5 above)?



It is possible, but rather unlikely. Picture above corresponds to such a strange situation. The TAB curve must be very steep always, while the TAC curve must be very flat always. This corresponds to the lack of intersection point of MAB and MAC curves (the equation MAB=MAC has no solution since MAB>MAC always for e>0; there is no economically justified level of environmental protection, because no matter what was done, if you do more, you increase net benefits (TAB-TAC)).

1.8 Please state the mathematical theorem the graphs on page 5 (or ERE-1-6) are based on.

Joseph Louis Lagrange was one of the greatest mathematicians, and he proved a number of fundamental results. Therefore there are dozens of "Lagrange theorems". The one the graph on page ERE-1-6 is based on states that the difference between two functions (here TAB and TAC) is the largest when the straight lines tangent to their graphs are parallel to each other. You see this in the upper picture. Tangent lines correspond to derivatives (here MAB=TAB', and MAC=TAC'). You see this in the lower picture: parallel tangent lines (i.e. the equality of their slope coefficients) correspond to the equality of the relevant derivatives. Do you remember the Lagrange theorem from your high school (or from your calculus class in the university)?

1.9 Is the MAB=MAC rule a popular one among environmental policy makers?

No. It is not. Environmental policy makers try to ignore the MAC=MAB rule officially. They prefer to declare that the standard they established for water, or air is determined by "objective" criteria, recommended by doctors. This is not true, because any pollution affects human health adversely. Thus the only "objective" standard is no pollution at all. However, establishing a ban on everything is not practical. Thus if a non-zero pollution standard is adopted, it is based on some kind of economic argument: "we do not recommend tightening the standard, because the additional benefits achieved by this would be too cumbersome for the economy (or society)". This is equivalent to saying that MAC>MAB. Policy makers hesitate to refer to the rule, because they are afraid of being accused of putting a price tag on human health or something like that.

2. Externalities

Externalities were introduced to economics by the end of the 19th century. However, it was only in the 1931, when Jacob Viner made a distinction between a pecuniary (financial) and a non-pecuniary (non-financial) externality. The latter takes place whenever a firm's profit depends on other agent's actions or a consumer's utility depends on other agent's actions and the impact does not confine to the price mechanism. The former happens if the impact confines to the price mechanism.

The following examples should demonstrate the difference. Let us assume that somebody buys ice-cream every day, and pays the price of 2 €. All of a sudden, eating ice-cream becomes fashionable, and the demand increases significantly. As a result, the price of ice-cream increased from 2 € to 3 €. The person who used to buy one ice-cream daily is "penalised" by what others did. Yet the impact confined to the price mechanism, so the externality is called a pecuniary one.

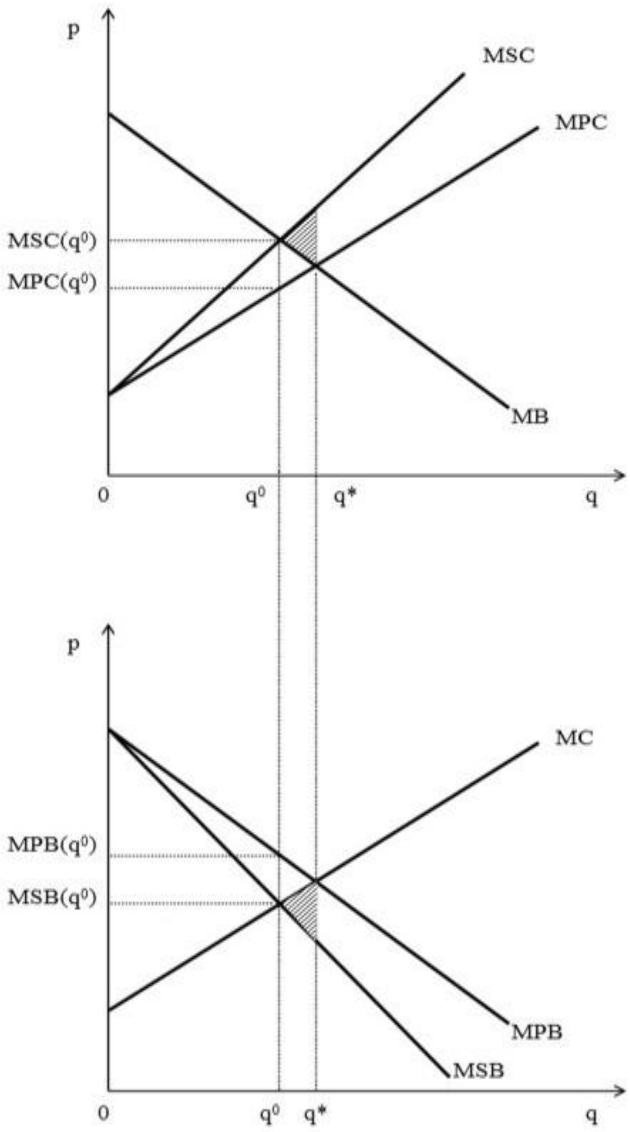
Now let us assume that a plant discharges wastewater and pollutes a lake. The lake contamination makes fishermen suffer from lower profits (the quantity or quality of fish goes down). This is an externality, like in the previous example, but its impact does not confine to the price mechanism. The polluter imposes some losses not because the price of fish goes down or the price of raw materials goes up, but imposes them directly (not through the market mechanism). We call this a non-pecuniary externality.

The distinction between the two is not only a theoretical one. It has important practical consequences. While pecuniary externalities do not imply market failures (market equilibrium can still maximise economic surplus like on page 4 (ERE-1-4), non-pecuniary externalities imply that equilibrium quantities do not have to maximise economic surplus, as pictures on pages 11-12 below demonstrate.

From now on we will confine to non-pecuniary externalities. There can be positive or negative externalities. A positive externality (external effect) increases a firm's profit or a consumer's utility, while a negative externality (external cost) decreases them. If the externality is caused by a consumer, we call it a consumption externality. If it is caused by a producer, we call it a production externality. Let me quote four examples of positive *versus* negative and consumption *versus* production externalities.

The lake pollution (referred to above) is a negative production externality. A positive production externality is imposed by a bee keeper whose bee hives are located near an orchard. The orchard owner benefits from the fact that bees pollinate trees thus making the orchard more profitable.

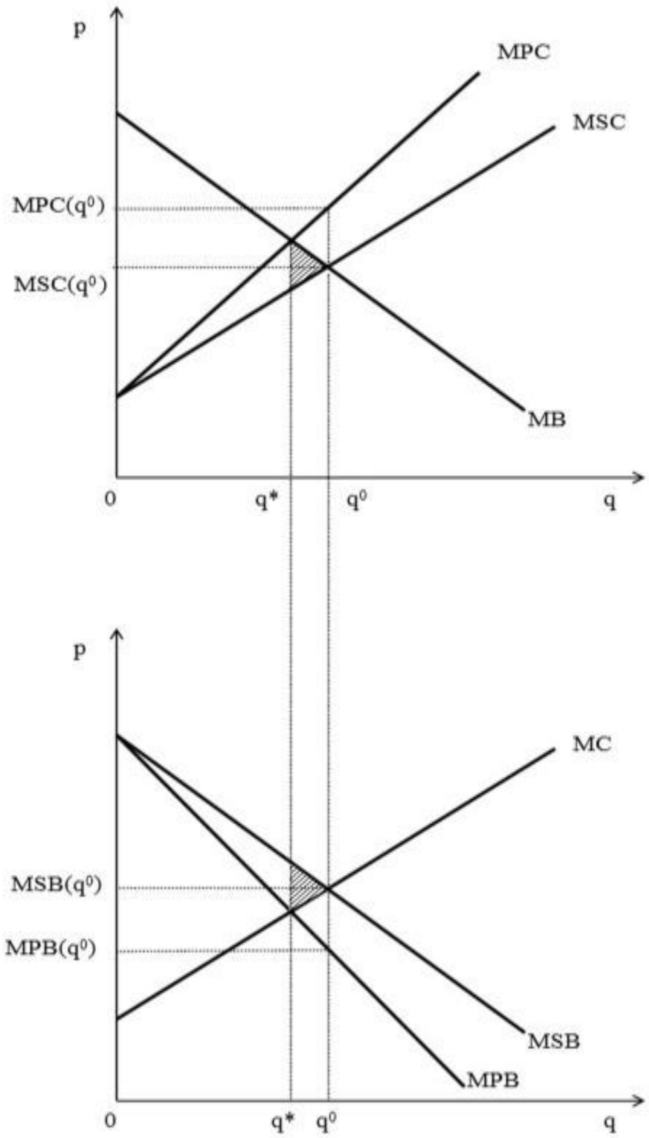
Noise makes a textbook example of a negative consumption externality. If somebody listens to the music loudly, his or her neighbours suffer. In contrast, nice fragrance of somebody's perfume makes others happier (think of entering elevator which was used a little while ago by a person who applied the fragrance you like).



Environmental improvements are often considered positive externalities. Those who decide about such improvements (e.g. pollution abatement) create externalities for others. The following analyses explain why the market fails to determine the socially justified level of environmental protection in the presence of externalities.

In the presence of a negative externality (see picture above), market establishes q^* as the equilibrium quantity, whereas q^0 is the socially desirable quantity: $q^0 < q^*$.

In the presence of a positive externality (see picture below), market establishes q^* as the equilibrium quantity, whereas q^0 is the socially desirable quantity: $q^0 > q^*$.



We need to explain acronyms used in these pictures. MSB and MSC stand for marginal social benefit and marginal social cost. MPB and MPC stand for marginal private benefit and marginal private cost. MB and MC – as before – stand for marginal benefit and marginal cost. Quantity q^0 is the economically justified level, while q^* is the market equilibrium.

The difference between MSC and MPC, that is the difference between marginal social cost and marginal private cost, is the marginal external cost (denoted as MEC). The difference between MSB and MPB is the marginal external benefit (denoted as MEB). These definitions allow to state to following equalities: MSC=MPC+MEC, and MSB=MPB+MEB. A positive externality can be considered a negative external cost (MEB=-MEC), and *vice versa*.

This is a standard economic terminology. In some textbooks, the term "social cost" is used to denote "external cost". This not appropriate since private cost is also a component of the social cost. For instance if the production of electricity requires a power plant to spend, say, 0.10 €/kWh, and the pollution implied by this production (unpaid by the plant) is, say, 0.15 €/kWh, then the social cost (i.e. what the society has to pay in order to enjoy electricity) is 0.25 €/kWh, not just the external cost.

Now let us interpret all the pictures. Page 11 illustrates a negative externality (e.g. lake pollution), and page 12 illustrates a positive externality (e.g. pollination by bees). Let us analyse the upper picture on page 11 (referring to the example of the power plant that pollutes the lake). The power plant looks at its MPC (the cost incurred to produce electricity) and MB (benefit from selling electricity). Having compared the two, the plant finds it profitable to produce q^* electricity. At the same time, the society looks at its MSC (the private cost incurred to produce electricity <u>and</u> the external cost imposed on fishermen, who suffer from the lake pollution) and benefits from electricity production MB. Having compared the two, the society determines q^0 as the economically justified level of electricity production. We see that $q^0<q^*$.

The same conclusion can be derived by analysing the lower picture. The power plant looks at its MC (the cost incurred to produce electricity) and MPB (private benefit from selling electricity). Having compared the two, the plant finds it profitable to produce q^* electricity. At the same time, the society looks at MC and its MSB (the benefits from electricity <u>net</u> of the damages suffered by the fishermen). Having compared the two, the society determines q^0 as the economically justified level of electricity production. Again we see that $q^0<q^*$.

The difference between the upper and the lower picture is just whether we attach externalities to the cost side or to the benefit side. If we attach it to the cost side, MSC=MPC+MEC. If we attach it to the benefit side, then MSB=MPB-MEC. We have to subtract it rather than add, since MEB=-MEC. It does not matter whether we look at the upper picture or at the lower one. They lead to the same conclusion. One has to remember though, that the externality has to be taken into account only once (either at the cost side or at the benefit side).

Let us move to pictures from page 12 (referring to the example of a honey producer, whose bees pollinate trees in a nearby orchard), and let us analyse the upper picture first. The honey producer looks at his or her MPC (the cost incurred to produce honey) and MB (benefit from selling honey). Having compared the two, the producer finds it profitable to produce q^* honey. At the same time, the society looks at its MSC (the cost incurred to produce honey <u>decreased</u> by the benefit enjoyed by the orchard owner, whose trees are pollinated by the bees) and MB. Having compared the two, the society determines q^0 as the economically justified level of honey production. We see that $q^0>q^*$.

The same conclusion can be derived by analysing the lower picture. The honey producer looks at his or her MC (the cost incurred to produce honey) and MPB (private benefit from selling honey). Having compared the two, the producer finds it profitable to produce q^* honey. At the same time, the society looks at MC and its MSB (the benefits from honey <u>and</u> orchard pollination). Having compared the two, the society determines q^0 as the economically justified level of honey production. Again we see that $q^0>q^*$.

The difference between the upper and the lower picture is just whether we attach externalities to the cost side or to the benefit side. If we attach it to the cost side, MSC=MPC-MEB. If we attach it to the benefit side, then MSB=MPB+MEB. As before, we observe that – by definition – MEB=-MEC. It does not matter whether we look at the upper picture or at the lower one. They lead to the same conclusion. One has to remember, that the externality has to be taken into account only once (either at the cost side or at the benefit side).

We can summarise these analyses by observing that in the presence of externalities the market fails (economists call it "market failure"). The failure is caused by the fact that $q^0 \neq q^*$ (a socially justified quantity is different from what the market determines). It can be either $q^0 < q^*$ (if the externality is a negative one) or $q^0 > q^*$ (otherwise). Shaded triangles on pages 11-12 illustrate welfare losses from adopting q^* rather than q^0. On page 11 the loss is caused by the fact that there is too much of the externality, while on page 12 it is caused by the fact that there is too little of the externality. In the lake example, electricity production is excessive if fishermen's losses are ignored. In the bee/pollination example, the number of bees is too small if their pollination "services" are ignored.

Economists observe that an externality arises when there is no market for the factor responsible for the externality (for instance, when property rights are ill-defined). In our lake pollution example the missing market means that the fishermen are not asked if the plant can discharge its sewage (the power plant does not have to buy any pollution permit). In the bee keeping business the missing market means that the orchard owner does not have to pay for the pollination (the bee owner is not offered any remuneration for the pollination service). "Transactions" (e.g. polluting the lake or pollinating trees) are not voluntary. They simply take place without waiting for anybody's consent.

Most students are perhaps familiar with the definition of a Pareto optimum. Let me recall it briefly just in case somebody forgot it. A Pareto optimum is such an outcome that nobody can be made better off unless somebody is made worse off. Assuming that there are 2 apples to be allocated to 2 people (both of them like apples), both can be given 1 apple. This is a Pareto optimum, since neither can be given 2 apples unless the other person is deprived of an apple (but this would make him or her worse off). The allocation can be different. One privileged person can be given 2 apples while the other one is left without any. Surprisingly, this is a Pareto optimum too, since in order to make the deprived person better off, at least 1 apple should be taken away from the privileged person making him or her worse off.

Many different allocations – some of them considered unfair – can satisfy the definition of a Pareto optimum. To see an allocation which violates it, let us consider 2 apples distributed in the following way. 1 apple goes to one person, the other person gets nothing, and 1 apple is left aside. This is not a Pareto optimum, because one person can be made better off by getting the apple left aside, and nobody will be made worse off. Intuitively, a Pareto optimum can be understood as an allocation which uses up everything what is available (not necessarily in a fair way).

Please note that there was no money referred to in this definition. If money is introduced things may look somewhat differently. Let us assume that both persons have identical preferences with respect to apples. They would be willing to pay 0.4 € for the first apple and 0.2 € for the second one (experiments confirm such preferences often). The allocation which gives one person both apples and leaves the other one without any is a Pareto optimum. But this allocation can be improved. Let us assume that the deprived person offers the privileged one a deal: "you give me one apple, and I pay you 0.3 € for it". Both of them will be better off. The privileged person will sell for 0.3 € the apple which was worth to him or her only 0.2 €. The deprived person for 0.3 € will get an apple which was worth to him or her 0.4 €. Both of them will be made better off by 0.1 € as a result of this transaction. Economists say that the total welfare will increase by 0.2 €. By doing this we achieve a so-called Generalised Pareto Optimum. Its formal definition states that nobody can be made better off unless somebody else is made worse off even after monetary compensations.

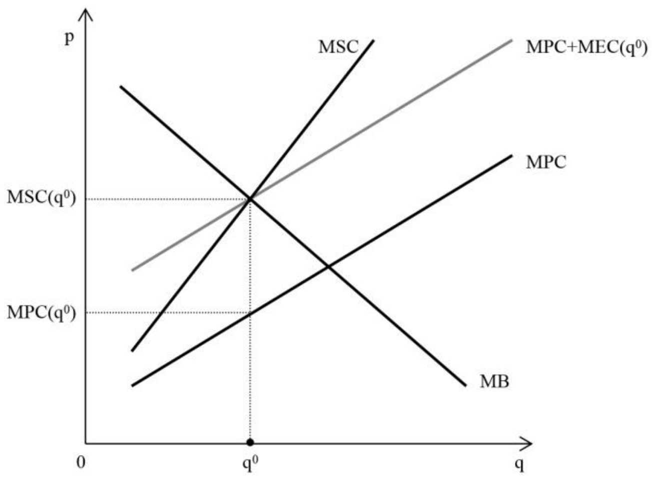
In examples above allocations (1,1) and (2,0) were Pareto optima, but only (1,1) was a generalised Pareto optimum. In the definition of a Pareto optimum, only physical quantities matter, and no compensations are possible. In a generalised Pareto optimum, not only physical quantities are looked at, but possible monetary compensations are contemplated too. By referring to the Generalised Pareto Optimum concept, an economically justified level of environmental protection can be characterised as an outcome which maximises economic surplus, i.e. TSB-TSC (total social benefits net of total social costs).

Economists look at so-called first order conditions which have to be satisfied in order to achieve an optimum subject to some mathematical properties. In particular, if TSB and TSC are differentiable, TSC is convex, TSB is concave, and $q>0$, then TSB-TSC is maximised if

MSB=MSC.

As before, MSB=TSB' and MSC=TSC'; MEC = MSC−MPC. The equation above is called the first order condition for achieving an economically justified level of an activity (for instance, environmental protection).

Now let us look at how a Generalised Pareto Optimum can be achieved.



The upper picture from page 11 can be redrawn as above. The straight line MPC+MEC(q^0) was added. This is the MPC line shifted by the amount MSC(q^0)-MPC(q^0)=MEC(q^0). If there is an externality generating agent (like the power plant from the lake example), then the quantity q^* determined by the market is where the lines MPC and MB intersect. However, if the agent is supposed to add a fixed amount of $MEC(q^0)$ to the marginal cost, then the decision will be q^0 rather than q^*.

This is the idea of the so-called Pigouvian tax (named after Arthur Pigou, who suggested it in 1920 as a solution to the market failure caused by externalities). The tax (PT) is calculated according to the formula:

$$PT(q)=MEC(q^0)(q-q_{threshold}).$$

If the agent maximises the profit then the following problem needs to be solved:

$$Max_q\ (B(q)-PC(q)-PT(q)).$$

Under typical mathematical assumptions this is equivalent to differentiating the function

$$B(q)-PC(q)-PT(q),$$

and checking where its derivative vanishes. In other words, the following equation has to be solved:

$$(B(q)-PC(q)-PT(q))' = 0, \text{ i.e. } MB(q)-MPC-MEC(q^0) = 0, \text{ or } MB(q) = MPC+MEC(q^0).$$

The solution to the last equation is envisaged on the page 15 picture.

Arthur Pigou worded his idea in the following way: "market failure caused by an external cost will disappear if the externality generating agent has to pay marginal damages implied by the externality". For 40 years a Pigouvian tax was considered the solution. In 1960 Ronald Coase questioned this view (this will be the topic of the next lecture). Before Coase, every textbook considered the tax an obvious solution to problems of environmental protection: polluters will stop polluting the environment excessively, if they are forced to pay Pigouvian taxes. Over the last one hundred years, in many countries environmental pollution declined significantly. Yet improvements were forced by instruments other than Pigouvian taxes. One reason that the tax invented by Arthur Pigou was almost never implemented was a fear that a high tax would kill the domestic industry.

For instance damages caused by sulphur dioxide emissions can be estimated at 2000 €/tonne. The sulphur dioxide emission in Poland is around 0.5 million tonnes. If you multiply the two numbers, you will get an impressive estimate of the tax obligation corresponding to 1 billion €. Most of this amount would have to be paid by the power plants which are responsible for a large part of the Polish sulphur dioxide emission. Nobody would dare to impose such a burden on power plants. For similar reasons, in every other country, politicians have hesitated to implement Pigouvian taxes.

An important reason for avoiding Pigouvian taxes in environmental policies was a poor understanding of the formula $PT(q)=MEC(q^0)(q-q_{threshold})$; the polluter pays for every unit of pollution above a threshold $q_{threshold}$. If the threshold $q_{threshold}=0$, the polluter pays for all the units, and arguments about "killing industry" can be real. But the threshold does not have to be 0. If $q_{threshold}=q^0$ (the threshold is set at the economically justified level), the tax payment is 0 if the polluter emits exactly what is socially desirable. If the threshold is set at a sufficiently high level (so that $q<q_{threshold}$), the polluter gets a subsidy. Only in 1989 it was demonstrated

that in fact the threshold does not have to be zero (as tacitly assumed by Pigou), but it can be any number without losing the incentive property of the tax.

Questions and answers to lecture 2

2.1 The energy bills may go up as a result of the necessity of paying for externalities. Is it caused by pecuniary or non-pecuniary externalities?

It can be either way. One reason for our energy bills to go up is the fact that other households use more electricity. If they increase their demand for electricity, more expensive suppliers find opportunity to sell their products, and the average price increases. We are – so to speak – penalised for what other customers do. The outcome is transferred through the market mechanism, so the externality is a pecuniary one. Another reason for our energy bills to go up is the fact that power plants are under increasing pressure to constrain their adverse impact on the environment. For instance, they need to reduce particulate matter emission. In order to do so, they install scrubbers (from 1990 to 2017 the emission of particulate matter from stationary sources fell by 83% in Poland). They also need to reduce sulphur dioxide emission (from 1990 to 2017 the emission of sulphur dioxide from stationary sources fell by 82% in Poland). In addition, they need to buy permits for carbon dioxide emission (although many power plants receive them for free; you will learn more on that in lecture 14). All these additional requirements translate into making production more expensive. If our energy bills include these costs, we participate in paying for non-pecuniary externalities.

2.2. Bee/pollination externalities have been recognised in many circumstances. Sometimes orchard owners are asked to pay to bee owners, and sometimes it is the other way around; farmers who cultivate crops that are attractive for bees expect bee owners to pay them something for the fact that increased revenues from selling honey originate in their fields. Who should pay whom?

It is important that parties to this externality have incentives to keep their activities at the level which is economically justified. In other words, bee owners should take into account the fact that their bees provide benefits for farmers, and the farmers should take into account the fact that honey production increases when they cultivate certain crops. Ideally bee owners should pay farmers for the pollen their bees collect from flowers (which is then used to produce honey), and farmers should pay bee owners for the pollination (which results in higher crops). But it would be very difficult to quantify these effects and to calculate relevant payments. Therefore these externalities are often ignored, but the resulting market failure is perhaps rather low.

2.3 Painting the façade of your house brings advantages for you and your neighbours. How much are you willing to spend on this?

If real estate owners are 100% egoistic they choose technology and materials according to the equation MC=MPB. If they are 100% altruistic they choose technology and materials according to the equation MC=MSB. In the former case they are motivated by their own benefits only (pleasure from having a nicer façade). In the latter case they are ready to spend more on this activity, since they take into account not only their own benefits, but also benefits enjoyed by the neighbours (pleasure from looking at someone else's façade). Sometimes there are government regulations on the quality of the façades, so the amount spent on renovations is higher than what real estate owners would be willing to pay for this purpose. Perhaps these regulations try to estimate external benefits from having a nicer façade.

2.4 In pictures on pages 11-12 (ERE-2-2 and ERE-2-3), an externality is referred to either at the cost side or at the benefit side. What would have happened if the externality was referred to both at the benefit and cost side?

A double counting would take place, and q^0 would be miscalculated. Pictures on page ERE-2-2 and ERE-2-3 indicate that the intersection of MSC and MSB curves (set together incorrectly) would be to the left of true q^0 in the case of a negative externality, and to the right of true q^0 in the case of a positive externality. Including the externality both at the cost and at the benefit side will result in underestimation of q^0 in the case of a negative externality, and in overestimation of q^0 in the case of a positive externality.

2.5 Can a positive externality be analysed as an external cost?

Yes, it can. There is an overall identity EB=-EC. Thus a positive externality is a negative external cost. If someone's activity lets enjoy the external benefit of, say, 10 €, it is equivalent to an external cost of -10 €. The negative external cost means that a "victim" (in fact a beneficiary) has to add the sum of -10 € to its private cost; in fact the private cost is decreased by 10 € (which is equivalent to increasing the benefit by 10 €). In some economic analyses all positive external benefits are calculated as negative external costs or *vice versa* (all external costs are calculated as negative external benefits).

2.6 Examples referred to in the class demonstrate that several different allocations can satisfy the definition of a Pareto optimum. Can you envisage methods to identify "better" and "worse" ones?

All Pareto optima are "efficient" in a sense that the total consumption of the analysed good cannot be increased. The only change implied by their reallocation is fairness, i.e. a feeling that a good in question is distributed in a way which maximises the well-being of the society. For instance, if the well-being is compromised by an unequal distribution of the good, then a Pareto optimum which leaves everybody with a similar allocation is "better". Please note that this method does not rely on monetary valuations (like in 2.7 below).

2.7 In the definition of a generalised Pareto optimum it was assumed that goods can be exchanged for money. Can you think of a different approach to comparing alternative Pareto optima?

The class definition of a generalised Pareto optimum referred to the possibility of calculating WTP for separate units of the analysed good. To simplify things, let us consider allocations of two apples between two people. WTP for the first apple is 0.4 €, while for the second one – only 0.2 €. This allows us to compare allocations (1,1) and (2,0) to conclude that the first corresponds to the total value of 0.8 €, and the second one – to 0.6 € only. The outcome will not change if the WTP was measured in British Pounds rather than euro. Neither will it change if any other monetary measure is introduced (gold, man-hour or what have you). In economic analyses, we use yet another concept in order to compare various outcomes, namely the utility people derive from the consumption. Utilities help to compare alternative Pareto optima in a very much similar way to the method based on WTP.

Comparing outcomes that are Pareto optima in order to identify a generalised Pareto optimum makes use of so called Kaldor-Hicks improvements. Nicholas Kaldor (1908-1986) and John Hicks (1904-1989; Nobel in 1972) developed the following criterion. When changing allocations, some will gain and some will lose. If those who gain can pay compensations to those who lose (the compensations are lower than the gains of the former, but higher than the losses of the latter), the new allocation is better than the old one. This is called a Kaldor-Hicks improvement. If no further Kaldor-Hicks improvements are possible, a generalised Pareto optimum is achieved. Please note that in the example above allocation (1,1) was identified as a generalised Pareto optimum. Indeed, no Kaldor-Hicks improvements are possible. Both persons would prefer to have 2 apples rather than 1. However, in order to have 2 apples rather than 1, the potential gainer is willing to pay no more than 0.2 €. At the same time, the potential loser is willing to accept not less than 0.4 €. There is no possibility to find a price of an apple that would be acceptable for both.

2.8 The formula for a Pigouvian tax suggests that the threshold level ($q_{threshold}$) is irrelevant for the tax to work. Please discuss advantages and disadvantages of establishing a non-zero threshold.

The tax rate MPT=MEC(q^0) is independent of $q_{threshold}$. A zero threshold imposes a tax on every unit q, and can be considered a heavy burden. Threshold $q_{threshold}=q^0$ makes PT(q^0)=0. A generous threshold $q_{threshold}>q^0$ turns a Pigouvian tax into a subsidy (PT(q^0)<0). Any threshold provides incentives to abate (if q goes down the Pigouvian tax goes down as well; or a Pigouvian subsidy goes up). Low thresholds make polluting activities more expensive. Thus less polluting activities are likely to survive in the market in the long run. Offering non-zero thresholds (especially generous ones) is controversial from the point of view of environmental protection. In addition, calculating non-zero thresholds is politically difficult, since the government can be accused of offering unequal treatment of various polluters.

Zero thresholds seem to be easier and politically safer, but governments are afraid of being accused of "killing industry". Hence there are no good solutions. Consequently Pigouvian taxation has been popular among economists, but virtually unused by practitioners.

3. Coase theorem

There are few names that trigger such a tremendous aggression as in the case of Ronald Coase (1910-2013; Nobel in 1991). If you type "Coase curse" into Google you will find a number of results starting with "Coase's Curse: How the Coase Theorem has Biased Ecological Policy Debate". Ronald Coase and his famous theorem are accused for almost everything bad we can see in the world: poverty, environmental disruption, drug dealing, arms race, and more. Coase theorem was published in 1960, and all the tragedies have existed for centuries. Nevertheless many people address their aggression or at least frustration towards economic theory embodied in the Coase theorem.

Before we proceed, let me summarise the previous lecture. In the presence of externalities markets fail. Unless the demand is perfectly inelastic (i.e. unless MB=const), and if MPC<MSC (i.e. MEC>0), then $q^0<q^*$, where q^0 is a generalized Pareto optimum, and q^* is a market equilibrium allocation (private optimum). In other words, a market equilibrium will not be a generalised Pareto optimum (it does not have to be a Pareto optimum at all). From 1920 to 1960 a Pigouvian tax was considered the obvious solution to the problem of

externalities. Ronald Coase questioned this view by demonstrating that sometimes a generalised Pareto optimum can be achieved without government intervention. But before stating the theorem a definition of a transaction cost (introduced by Coase in 1937) has to be stated. It is namely the cost of preparing and carrying out a transaction (including the cost of contract enforcement). Before Coase transaction costs were ignored in many economic analyses.

The precise wording of the Coase theorem reads:

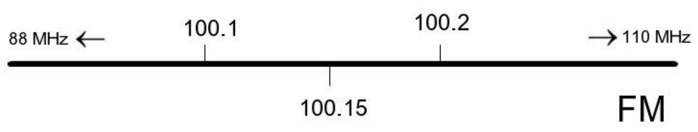
In the absence of transaction costs, if two rationally behaving agents can negotiate about the amount of an externality imposed by one of them on the other, if property rights are well defined and if distribution of welfare does not affect marginal values, then

> ***(1) the final allocation of resources will be Pareto optimal (thus there will be no "market failure"); and***
> ***(2) the final allocation will not depend on the allocation of property rights (initial allocation).***

The (1) holds even if the distribution of welfare does affect marginal values.

Before we explain what this wording means, let us look at how the theorem emerged. In many countries – including Poland – governments have problems with allocating FM frequencies. There exist National Councils for Radio and TV that are charged with allocating frequencies. The FM systems use the spectrum 88-110 MHz, and if a newcomer asks the Council about a frequency, the most likely answer is "Sorry, all the frequencies are taken already". The newcomer is likely to respond: "Nonsense – 100.1 and 100.2 are taken, but 100.15 is unallocated, and hence I can be given this frequency". The Council says: "Sorry, 100.15 is too close to 100.1 and 100.2". The newcomer responds: "We demand 100.15! If you think it is too close to the others, then deprive the others of their frequencies, and allocate them to us".

If the Council reallocates a frequency to the newcomer, then the deprived station accuses the government of stupidity, incompetence, and corruption (since the frequency reallocation required probably a bribe that the newcomer paid). If the Council refuses to reallocate the frequency, then the newcomer accuses the government of stupidity, incompetence, and corruption (since the frequency reallocation required probably a bribe that the newcomer did not pay). In other words, no matter what the Council does, it will be accused of stupidity, incompetence, and corruption. A picture below illustrates this problem.



The American government was accused of stupidity, incompetence, and corruption in 1959 when all TV frequencies were allocated, and newcomers could not be satisfied. The government approached Coase who at that time has already gained a reputation of a good economist. Asked about what they should do, Ronald Coase wrote the report with the conclusion "Do nothing". It took some time, before administrators understood this conclusion. According to the report, "doing nothing" meant leaving the allocation business to the market. Frequencies should be owned privately. If a newcomer wants to get a frequency, he or she has to buy one, or to rent one. The government should not give them to anybody. Neither should it expropriate anybody.

The Coasian solution is shocking at the first glance, but this is how addresses are allocated. If an individual wants to start a business, he or she needs to register the address of the firm. Nobody expects the government to allocate an address. The address should be bought or rented (real estate should be bought or rented). The idea with frequency allocation replicates this logic.

Let us analyse what may happen if the government yields to the newcomer's demand to allocate a frequency, say, 100.15 which is close to 100.2. The newcomer starts broadcasting at 100.15, and listeners of 100.2 complain about interference with some different station. The 100.2 decides to invest in a new antenna – say, 150 kW instead of 120 kW. Then listeners of 100.15 start to complain, and they motivate the station to invest in an even more powerful antenna – say, 200 kW. The other station responds with an even stronger device, and so on. The two stations fight a war, and listeners have to pay for it (in terms of fees or in terms of suffering from additional advertisements).

In the frequency allocation problem Ronald Coase recognised externalities. By broadcasting at neighbouring frequencies, the stations impose external costs on each other. A textbook solution to the problem would be to invite the government with some sort of Pigouvian taxation. The Coasian solution is to let the market solve the problem without any government intervention.

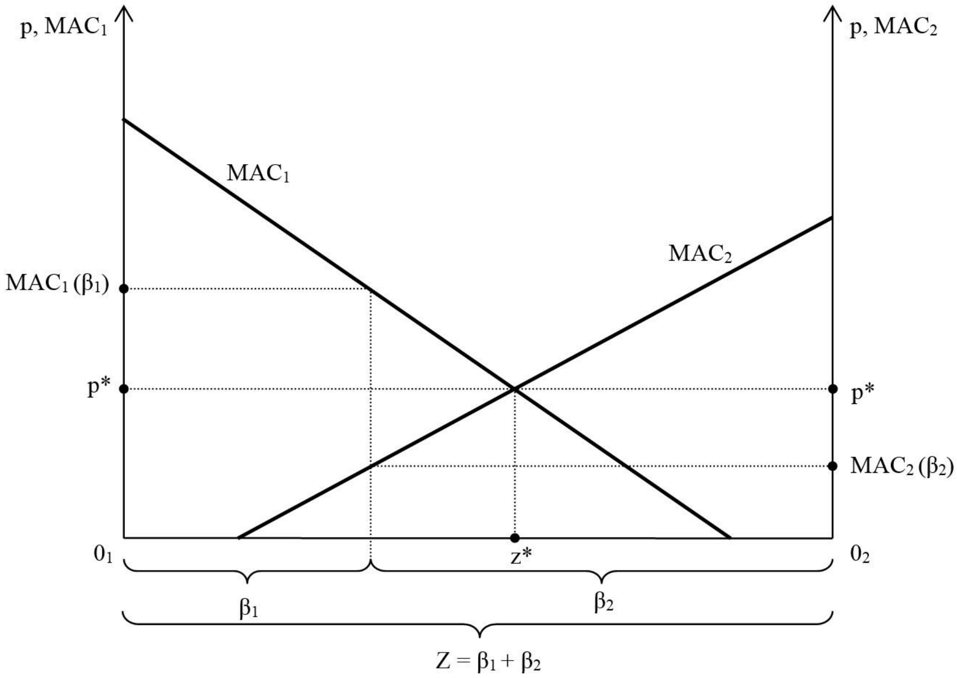
The solution is based on the idea of Kaldor-Hicks improvements (see the previous lecture). Let us assume that a frequency is owned by a station which makes a profit of, say, 1000 $. The newcomer estimates its profit at, say, 1200 $, and hence it is willing to pay for the frequency less than 1200 $. If he (or she) offers 1100 $, then the price should be accepted. The old owner will be better off by getting 1100 $ for something which gave the profit 1000 $. Also the newcomer will be better off, enjoying the profit of 1200 $ having paid 1100 $ (thus enjoying 100 $ in net terms). The frequency will go to the new owner who is able to use it more efficiently.

Now let us assume that the newcomer estimates its profit at, say, 800 $. He (or she) is willing to pay less than that. The old owner will not accept such an offer, and the frequency will stay with him (or her), i.e. where it can be used more efficiently. Kaldor-Hicks improvements are not possible, and the *status quo* turns out to be a generalised Pareto optimum.

The Coase theorem requires that the two externality generating agents are rational and there are no transaction costs. What happens if these assumptions are violated? If the agents are not rational, they do not have to take advantage of mutually beneficial transactions. In the first example above, the old station's profit is 1000 $, the newcomer's profit is 1200 $, so the frequency should go from the former to the latter. If the former is not rational, then he or she may choose not to sell the frequency, i.e. to lose an opportunity to make an improvement.

Likewise, transaction costs may prevent the improvement from happening. If the transaction requires a legal service of, say, 250 $, then the potential efficiency gain of 200 $ will be too small to trigger the transaction. That is why in the Coase theorem, it is assumed that there are no transaction costs.

An additional assumption reads "property rights are well defined". Nobody would like to make business with a thief. Unless property rights are well defined, a buyer risks making a deal with a thief. A prospective buyer would like to make sure that the seller has the property right to what is sold. In the case of uncertainty, even an attractive transaction does not have to be carried out.



Ronald Coase was a great economist, but he was not a skilled mathematician. If phrased in mathematical terms, the proof of his theorem is quite simple. In his original article (published in 1960 and based on his "TV frequency" report of 1959), the proof took many pages, and contained a mistake that was found a couple of years later. Coase provided an argument illustrated in my picture above, and based on a bargaining process between two agents. He thought that the argument can be generalised to three or more agents, but it cannot. Thus in my wording of the Coase theorem above you see "two agents". Some textbooks overlook this constraint, and the impression is left that it works for more agents. It does not.

Picture above illustrates how the proof looks like. We refer to the following problem. There are two plants (1 and 2) that emit pollution. Their abatement costs are MAC_1 and MAC_2, respectively. Their abatement cost curves – MAC_1 and MAC_2 – are downward sloping: the more they emit, the less expensive abatement is. The no emission variant is the most expensive one. If they are allowed to emit as much as they want, they will choose emission such that $MAC_1=0$ and $MAC_2=0$. The vertical axes measure economic variables (costs and prices). The horizontal axis measures emissions.

The left axis refers to the first plant and it has a standard orientation. The second plant is represented by a mirror image. Thus its vertical axis is on the right, and the direction of the horizontal axis is from right to left.

If the two systems are to be included in the same picture and the horizontal axis is the same for both, one needs to determine how far from each other the vertical axes should be. It was assumed here that the distance is equal to the total emission from the two plants that we can tolerate, i.e. Z. An externality referred to in this proof of the Coase theorem is not the externality imposed on the environment. It is the externality that one plant imposes on the other one by emitting something (x) and thereby forcing it to abate more (i.e. emitting Z-x).

Let us assume that the two plants are given emission permits β_1 and β_2, respectively. The first plant can emit β_1, and the second one: β_2. The sum of their emissions is $\beta_1+\beta_2=Z$, i.e. what can be tolerated. The plants have to abate in order to emit what they were permitted. Their marginal abatement costs read $MAC_1(\beta_1)$, and $MAC_2(\beta_2)$, respectively. In my picture $MAC_1(\beta_1) > MAC_2(\beta_2)$. Therefore the first plant has an incentive to approach the second one and to ask that the second plant abates more (i.e. emits less) and sells the unused part of its emission permit. If the first plant buys this part of emission permit unused by the second plant, it can emit more and thus avoid some of its abatement cost. The first plant is willing to pay less than $MAC_1(\beta_1)$, and the second plant should be willing to accept anything more than $MAC_2(\beta_2)$. Since $MAC_1(\beta_1)>MAC_2(\beta_2)$, they should agree on a mutually attractive price of the unused part of the emission permit. By looking at the picture, you see that they have incentives to trade until they hit z^*, i.e. an allocation such that their marginal abatement costs are identical. The price of the emission permit is p^* then. This ends the proof of the Coase theorem.

What did we prove? We proved that if two plants take rational decisions, property rights are well defined (both plants know what they are permitted to emit), there are no transaction costs (they can buy and sell permits without any obstacles), then they will achieve the least cost allocation (z^*) irrespective of what was the initial allocation of their property rights (permits to emit β_1 and β_2, respectively). This is the essence of the Coase theorem.

Let us analyse its implications for the example of the noise externality problem mentioned in the previous lecture. Neighbour 1 listens to the music loudly late in the night and does not allow neighbour 2 to sleep. Let us assume that the pleasure derived by neighbour 1 is 10 €, and the loss incurred by neighbour 2 is 12 €. With the help of the Kaldor-Hicks improvement concept, we see that in the generalised Pareto optimum, neighbour 1 should stop making the noise, and neighbour 2 should get an opportunity to rest.

We will see that this outcome can be achieved irrespective of how property rights are assigned. First, let us assume that the law does not permit to make noise in the night. Neighbour 2 can force neighbour 1 to stop the noise. Neighbour 1 can try to persuade neighbour 2 to permit the noise in exchange for a compensation. Neighbour 2 is willing to accept at least 12 €, and neighbour 1 is willing to pay at most 10 €. No price can meet these requirements. Second, let us assume that the law permits to make noise. Neighbour 1 can make the noise, but neighbour 2 can try to persuade him (or her) not to do it in exchange for a payment. Neighbour 2 is willing to pay less than 12 €, and neighbour 1 is willing to accept more than 10 €. If they agree to, say, 11 €, a deal will be struck. Thus irrespective of what the law says (in Coase's language irrespective of how property rights are assigned), if neighbours can negotiate a deal, an optimum solution will be achieved. In this case, neighbour 1 will not listen to the music loudly (and neighbour 2 will sleep).

Now let us assume that the pleasure derived by neighbour 1 is 12 €, and the loss incurred by neighbour 2 is 10 €. In the generalised Pareto optimum neighbour 1 listens to the music loudly, and neighbour 2 suffers the loss. As before, we look at two cases. First, the law does not permit the noise. Neighbour 1 cannot listen to the music loudly, unless he (or she) bribes neighbour 2 to let him (or her) do so. Neighbour 1 is willing to pay less than 12 € to obtain a permission from neighbour 2. Neighbour 2 is willing to accept anything more than 10 € before he or she gives the permission. If they agree to the payment of, say, 11 €, a deal will be struck. Second, if the law permits the noise, neighbour 1 will listen to the music loudly, unless neighbour 2 offers an adequate compensation. Neighbour 1 is willing to accept at least 12 €, and neighbour 2 is willing to pay at most 10 €. No price can meet these requirements. Thus irrespective of what the law says (in Coase's language irrespective of how property rights are assigned), if neighbours can negotiate a deal, an optimum solution will be achieved. In this case, neighbour 1 will listen to the music loudly (and neighbour 2 will suffer the noise).

For many people Coase theorem sounds odd. The pollutee is to bribe the polluter to stop the pollution?! The government is to stay away?! Indeed, this is contrary to what many people presume, since there is a widespread expectation that the government should take the side of a victim. The following example demonstrates that one should be careful when establishing such principles for policies.

Let us assume that there has been a recording studio located near a large city. An international airport was built in the neighbourhood. The studio is a small one. Its annual profit is around 10,000 €, and it employs 5 people. The airport is a big one. Its annual profit is around 10 million €, and its employment is more than 1000 people. Because of noise, the recording studio cannot co-exist with the airport. No matter what measures it undertakes, the noise precludes any recording. According to Coase, if they are free to negotiate, a socially desirable outcome (a generalised Pareto optimum) will be achieved. There is no doubt that in this case the socially desirable outcome is moving the studio elsewhere, and letting the airport continue. Let us analyse two cases. First, the law attaches the priority to the recording studio. The airport should disappear unless it "buys" a permission from the recording studio to operate. The airport is willing to pay for such an outcome less than 10 million €. The studio is willing to accept more than 10,000 €. They are likely to find a mutually attractive price to strike a deal. Second, the law attaches priority to the airport. The studio should disappear unless it bribes the airport to stop its operations. The studio is willing to pay the airport less than 10,000 €, but the airport is willing to accept at least 10 million € if it were to discontinue its operations. They will not strike any deal.

The example shows two things. First, the government should not take the side of a victim by default. Second, Coase theorem suggests a solution that does not require any government intervention.

Finally, let us come back to the note which stated that the first conclusion of the Coase theorem (the elimination of the "market failure") holds irrespective of a distribution of property rights. The difference between the first and the second conclusion is slight. The first says that – as a result of negotiations – a socially desirable outcome will be achieved. The second states that – as a result of negotiations – the same (socially desirable) outcome will be achieved.

Let us look at the picture on page 22. It demonstrates that an optimum outcome (z^*) is achieved where the two marginal abatement cost curves intersect. One of the assumptions of the Coase theorem reads "if distribution of welfare does not affect marginal values". Hence in particular, it says that irrespective of who is given what, marginal abatement cost curves are the same. If the curves are the same, then their intersection point is the same, and the derivation of z^* is unaffected. Does welfare distribution affect marginal values? Distribution of permits (abatement tasks) has certain welfare consequences. If a plant is given a generous permit, it does not have to abate much, and it is wealthy. On the contrary, if a plant is given a challenging permit, it has to abate much, and it is not that wealthy. A wealthy plant may apply a different abatement technology than a less wealthy one. Thus – as a result – its abatement cost curve can be different. In order to prove the second conclusion of the Coase theorem an additional assumption is needed that irrespective of the initial allocation (β_1 *versus* β_2) marginal abatement cost curves intersect over the same point (z^*).

The Coase theorem sheds light on what environmental policy should take into account, and under what circumstances a government intervention is inevitable. In many cases the theorem does not apply. Then eliminating a market failure requires some sort of intervention, such as:

- quantity regulation, i.e. imposing a constraint $q \leq q^0$ (in the case of a negative externality) or $q \geq q^0$ (in the case of a positive externality); or
- imposing a Pigouvian tax, i.e. $PT(q)=MEC(q^0)(q-q_{threshold})$, where $q_{threshold}$ is an arbitrary threshold; or
- merging agents that create and suffer externalities ('*institutional internalisation*')

The first two types of intervention are based on the graph from the first lecture (ERE-1-6) or the second lecture (ERE-2-6). An economically justified outcome is where the MAB and MAC curves intersect (i.e. over e^0 or q^0 – depending on whether the terminology emphasises emission or abatement). The quantity regulation means that the polluter is forced to constrain the externality to the economically justified level. Imposing a Pigouvian tax leads to the same outcome except that the polluter is given an incentive to do so voluntarily. The last type of intervention obliges the polluter to merge with its victim in order to make the external cost an internal one. By doing so, the pollution is expected to be brought down to a socially desirable level as a result of Kaldor-Hicks improvements. Such an "institutional internalisation" looks odd (think of creating an enterprise consisting of a power plant and fishermen), but it can be conceived theoretically.

Questions and answers to lecture 3

3.1 The summary of the previous lecture contains words "unless the demand is perfectly inelastic". What happens if the demand is perfectly inelastic?

Perfect inelasticity takes place when the demand does not respond to price changes, i.e. when the demand is a vertical line (we assume usually that it is a downward sloping curve). Please recall the upper picture in my overheads ERE-2-2 (or picture in ERE-2-6) but with a vertical MB line (see picture on page 26 below). Then $q^0=q^*$, because irrespective of whether you take the intersection of MB with MSC or MPC, it corresponds to the same coordinate on the horizontal axis. Hence there is no market failure. Perfect inelasticity is rather unlikely. What can be observed in practice – is a very steep demand curve. The market failure exists in such a case, but it is small, because the difference between q^0 and q^* is small.

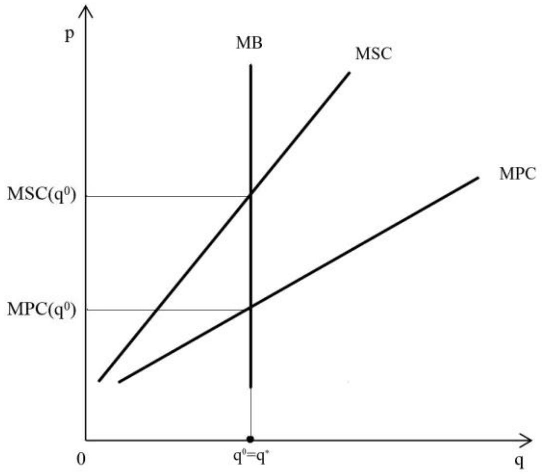


Illustration to question 3.1

3.2 Assumptions of the Coase theorem require that there are no transaction costs. Sometimes a slightly different Coase theorem is proved with the assumption that transaction costs are "small enough". Can you comment on that?

Please refer to the following example mentioned in my lecture. The old broadcasting station enjoyed the profit of 1000 $, while the newcomer expected the profit of 1200 $. Transferring the frequency from the old station to the new one would imply a welfare gain of 200 $. But the transaction would cost 250 $. Hence irrespective of who was to pay for it (the old, the new, or both, say, 125 $ each) it will not be carried out. The transaction cost of 250 $ turned out to be prohibitive. If the transaction cost was lower, say, 150 $, it could have been afforded. Assuming that the stations agreed to the price of 1100 $, each could get 100 $. If they split the transaction cost into halves, then each would have to contribute 75 $. The transaction can take place, because each station will gain 25 $. In this case the transaction cost was not zero, but small enough not to prohibit the transaction. This example helps to see what economists understand by "small enough" in the context of the Coase theorem.

3.3. Please provide an example of what can go wrong if there are three externality generating agents.

The argument you heard in the class was based on the fact that there are two negotiating agents. If there are three of them, then any two can make a coalition against the third one, and offer a deal which is better than no deal (so it will be accepted), but it will not be a generalised Pareto optimum.

3.4 Why do we tolerate real estate market (buying addresses), but we hesitate to accept the Coase solution to FM frequency allocation (buying frequencies)?

It is difficult to explain. Some people see the reason caused by the fact that frequencies are invisible and hence we do not want the market to decide on their allocation; real estate is visible, while FM frequency is invisible, and people may doubt if something invisible can be sold. This may explain a part of our attitude. An additional likely reason why we tolerate real estate market is that it has existed for centuries, and we got used to it. In contrast, FM frequency question emerged in the 20th century. We feel bad if we see that somebody acquired an ownership title all of a sudden (instead of inheriting it from somebody else earlier). On top of that, an overall distrust in market mechanism is often caused by a fear that somebody can utilise his or her market power to crash somebody else who is less powerful. The next class will address problems of this sort in greater detail.

3.5 Please refer to the "noisy neighbour" example. What difference will be made if we assume that there are several households suffering when the neighbour listens to the music loudly?

In the next class the problem of so-called "free riding" will be analysed in greater detail. Here it is sufficient to observe that in the case when there are at least two victims of the same externality, each of them may hesitate to negotiate with the externality generating agent, expecting that the other one will do this.

3.6 Once again, please refer to the "noisy neighbour" example. Instead of relying on the Coase theorem, many municipalities simply enforce obligatory "night silence" regulations. Why?

An important reason why this problem cannot be practically solved by the Coase theorem is that the number of agents is higher than two (please see 3.5 above). Transaction costs may be prohibitive as well. For instance, if the noise is suffered by 10 households, then an agreement between them requires a massive logistical effort. Moreover the agents may not be rational; if they react emotionally, they will not pay attention to whether something is more or less beneficial for them.

3.7 In the graphical proof of the Coase theorem on page 22 (or ERE-3-5), there are two kinds of externalities present. One is between the two polluters and the environment, and the other is between the polluters themselves. In environmental economics we are often interested in how the government decides about Z, i.e. the sum of emissions from all the polluters. Can the Coase theorem be applied to this problem as well?

If there is one polluter (perhaps several emitters treated as "industry") and one victim (the government acting on behalf of all the pollutees), then it is possible to conceive a Coasian bargain. However, if there are two polluters and one victim (three agents), the Coase theorem does not apply. What form a Coasian bargain can achieve, if there is one polluter (say, "industry"), and one victim (say, the government)? The government looks at the demand for environmental protection reflecting WTP for abatement. "Industry" looks at the supply of environmental protection reflecting WTA for abatement. If they behave rationally, then the government is ready to pay in order to "buy" abatement to let "industry" abate. In equilibrium, WTP=WTA (MAB=MAC) which means that abatement measures undertaken (and financed) are justified by damages avoided thanks to the abatement. This is how Z is determined. The graph serving as the proof of the Coase theorem is not about this. The level Z is taken as given, and the argument is about how the two polluters negotiate between themselves how to determine z^*, that is who abates how much in order to achieve Z jointly. By emitting z_1 the first plant forces the second plant to emit $z_2=Z-z_1$, and *vice versa*. This is the externality imposed on each other (not the fact that their pollution is environmentally harmful).

3.8 Coase theorem is sometimes understood as a suggestion that government should stay away from business. Please comment on this.

The theorem suggests that sometimes externality generating agents and their victims can achieve an optimum solution without a government intervention. Nevertheless specific assumptions of the Coase theorem oblige the government to do a number of things. Government decisions influence transaction costs. Decreasing legal fees and speeding up administrative procedures cannot drive transaction costs to zero, but they can lower them. Also clear definitions of property rights (another assumption in the Coase theorem) is a domain of government activities.

3.9 Should the government take side of the victim always?

No. If the government took the victim's side always, there would be no incentive for the pollutees to defend themselves, whereas optimum solutions require both sides to act. For instance, tap water has to meet certain criteria, but it does not have to be completely aseptic. It would be ridiculous to force water companies to pump water free from any chemical substances and any microorganisms. It is expected that water consumers will take appropriate measures if they want to use the water for some medical or technical purposes which require the absence of anything.

3.10 Should the government act against those economic agents whose history is shorter?

No, as the class example demonstrats. Again, let us assume that there has been a recording studio located near a large city. An international airport was built in the neighbourhood later. The studio is a small one. Its annual profit is around 10,000 €, and its employment is 5 people. The airport is a big one. Its annual profit is around 10 million €, and its employment is more than 1000 people. Because of noise, the recording studio cannot co-exist with the airport. No matter what measures it undertakes, the noise precludes any recording, so they can negotiate a deal to stay or to go away. According to Coase, if they are free to negotiate, a socially desirable outcome (a generalised Pareto optimum) will be achieved. There is no doubt that in this case the socially desirable outcome is moving the studio elsewhere, and letting the airport continue.

It was irrelevant who was the first in the area. Even though the recording studio was located there before the airport, it would be inappropriate for the government to refer to this argument. Only if agents involved in an externality problem are equally important from the social point of view, then their history can provide an argument.

3.11 Can you indicate environmental policy implementations of the Coase theorem?

So-called marketable pollution permits (see lecture 14, ERE-14) are considered the best known implementation of the Coase theorem. They are based on the picture on page 22 (ERE-3-5) serving as the graphical proof of the theorem. Their overall idea is that a group of polluters have emission permits. They are allowed to emit what their permits imply. If a plant emits less, it can sell the unused part of its permit to another plant which then can emit more than its original permit allowed. Thomas Crocker is considered the "father" of this instrument, but his most important contribution to environmental economics (in 1966) is the definition of the "supply of environmental protection". In the first class (ERE-1), it was described as the availability of environmental protection measures. This is precisely what potential sellers of abatement outcomes should do: they need to estimate the cost of abatement, and compare it with what potential buyers would be willing to pay for it.

Apart from the marketable pollution permits, the Coase theorem serves as a reference for a number of other environmental policy instruments. The one which is very fashionable now is called "Payments for Ecosystem Services". Those who steward valuable ecosystems do whatever they do, motivated by their benefits. At the same time, the stewardship creates external effects which benefit others as well. The idea of "Payments for Ecosystem Services" is to provide incentives for stewards to do even more – to do what brings benefits enjoyed by the others too. The idea is consistent with the Coase theorem: creating a market for what was an externality.

4. Public goods

Many goods analysed in economics – such as apples, cars, or pencils – are used individually, and everybody is responsible for their consumption. People reveal their willingness to pay for them, and unless the price is higher than what they revealed, they buy such goods. But there are also goods that have different characteristics. Two principles need to be referred to in this context. The non-exclusion principle reads: if a unit of a good was provided, then nobody can be (easily) excluded from using it. The non-rivalry principle reads: the same unit of a good can be simultaneously used by more than one user. Goods that comply with these two principles are called public goods.

In contrast, a private good does not comply with non-exclusion, and does not comply with non-rivalry principles. For instance, an apple is a private good, because once somebody got it, he or she can prevent (easily) others from using it; hence non-exclusion principle does not hold. The non-rivalry principle does not hold either; if somebody ate the apple, it cannot be eaten by anybody else.

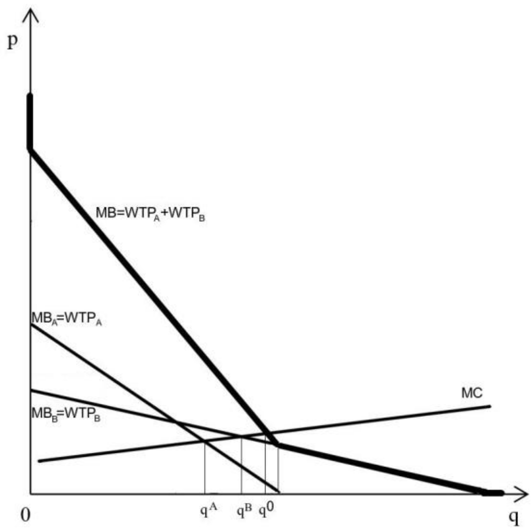
Lighthouse has been the first textbook example of a public good. If a lighthouse operates, then it does not matter how many vessels look at it; hence the non-rivalry principle holds. The non-exclusion principle holds as well, since everybody – also those who have not contributed to establishing the lighthouse – can benefit from it.

Air defence is another example of a public good. A modern air defence system consists of radars, fighter planes and anti-missiles. If an alien object is detected by a radar, fighter planes take off, and an anti-missile is fired. The same air defence system is required irrespective of whether it protects just a royal palace, or a ten-million city. Hence the non-rivalry principle holds. If a missile is approaching, it would be ridiculous to guess whether it will hit a pacifist's house or a military barrack. The anti-missile has to be fired immediately. Therefore the non-exclusion principle holds too.

There are numerous examples of public goods analysed in environmental economics. Clean air, biodiversity, and climate protection satisfy both principles. It does not matter whether climate is protected for 5 billion people or for 10 billion people. The damages caused by its destruction will be suffered by everybody – also by those who tried to protect it.

Whether a good is considered a public or private one may depend on the level of analysis. For instance, so-called adaptation to climate change is a public good from the point of view of a country, and is a private one from the point of view international community. Adaptation consists of measures – such as planting trees in cities – to let people survive heat waves resulting from the global warming. A shaded area in a city can be used and appreciated by many people, so it is a public good. However, from a global point of view it is considered a private good in the sense that decisions whether to take such measures are autonomous, and they do not require international cooperation. They do not comply with the non-rivalry and non-exclusion principles (projects undertaken in one country do not benefit other countries). In contrast to climate protection which is a public good for sure, adaptation is private or public depending on the level of analysis.

Picture below illustrates how a demand schedule for a public good is constructed. It will refer to the following story. Two neighbours, Mr. Ant and Mr. Bat (let us call them A and B, respectively) contemplate lighting in a dark hallway they share. Both of them have certain preferences with respect to the lighting, reflected in their marginal benefits from the lighting, and equal to what they are willing to pay for it ($MB_A=WTP_A$, and $MB_B=WTP_B$). Their joint benefits are equal to their joint willingness to pay ($MB=WTP_A+WTP_B$).



The decision about the lighting is not a dichotomous one (to have or not to have). Once the lighting is installed, a decision can be taken on the power of the bulbs (for instance, 20 W or 60 W), their quality, reliability, and so on. Thus the horizontal axis measures the quantity of lighting as a continuous (not a dichotomous) variable.

The lighting is a public good, since non-rivalry and non-exclusion principles work. The former works, because it does not matter how many neighbours benefit from the light. The latter works, because once it has been provided, then nobody can be excluded from its benefits (even the neighbour who did not contribute to it).

Constructing a demand schedule starts from the right end of the horizontal axis. Please recall from the first lecture (left picture on page 2) that drafting the demand schedule for a private good – unlike now – started from the upper end of the vertical axis. The procedure applied now is different. Nobody (neither A nor B) is interested in the very high quantity of lighting (the horizontal – flat – part of the $MB=WTP_A+WTP_B$ curve that you can see on the right). If the quantity considered is somewhat lower (a little bit to the left), then the demand comes from B (and only from B); the demand curve is identical with $MB_B=WTP_B$ (not horizontal but still fairly flat). When the quantity decreases further, the demand comes from B and A together. This is the steeper segment of the demand curve. When the quantity hits zero, the demand curve starts to be vertical (nobody is willing to pay for lighting a very high amount).

If the neighbours reveal their preferences, the equilibrium quantity of lighting is q^0 (where the demand meets the supply represented by the MC curve). This is an economically justified amount of the lighting (public good).

But the story can look differently. Let us assume that the initiative originates from A. He comes to B, explains that the hallway requires lighting, both of them will benefit, and both of them should contribute to the project. But B may say that he is not interested; if A wants to have lighting, then he can provide it himself. It can be the other way around. The initiative originates from B, and A declares that he is not interested. If A was to decide on the lighting himself, he would choose q^A. If B was to decide on the lighting himself, he would choose q^B. In either case, the quantity chosen would be lower than the economically justified (socially optimal) one: $q^A<q^0$, and $q^B<q^0$. The behaviour revealed in these cases is called free-riding. People are said to "free-ride" on what others do when they declare that they are not interested in financing a public good, knowing that – by the non-exclusion principle – they will be able to use it once somebody else has paid for it.

The story may have a totally different conclusion. Irrespective of whether A or B initiated the discussion, if the second neighbour declared that he was uninterested, the first one could decide not to invest at all, and the quantity of lighting is zero (q=0). This is what may happen when there are too many potential users who decide to free-ride.

When I ask the students how would they feel in the role of A or B, they say (or write) that they would be angry to face refusals. Some say that they would agree to participate if asked by a neighbour; there are people who do not free ride even when they can. Unfortunately, some of us absolve themselves from free riding. The justification of free riding can be cynical: "I do not want to pay, because I do not have to." Sometimes it is less cynical: "I do not want to pay, because others paid already." On top of that, people who – in principle – are willing to pay, may change mind having learnt that others free ride: "I do not want to pay, if free-riders are to benefit from my money". No matter what the justification is, the fact that some people hesitate to pay for public goods leads to a market failure.

The overall conclusion from these analyses is that the non-exclusion principle implies a problem; users avoid purchases of the good, anticipating that they will use the good purchased by somebody else. As a result, in an unregulated market, the supply of a public good is lower than socially optimal.

Environmental improvements often boil down to reducing an externality. If such an improvement complies with non-rivalry and non-exclusion principles, it can be considered a public good. Please recall the lake pollution example from the second lecture (pages 10-19). If the plant abates, then fishermen gain and may be motivated to persuade the plant to abate. If there are several fishermen, then nobody can be excluded from the benefits. What happens if some of them decide to free-ride? The following example will shed light on such problems.

We will look at prospects for spontaneous cooperation. The story will be about a typical externality problem. There is a steel mill (the polluter) and a laundry (the victim). We use the language of game theory, and call them 'players'. The steel mill makes the air dirty, and the laundry requires clean air for some of its operations (e.g. drying). If the air is dirty, the laundry's profit goes down. If the steel mill abates, then its own profit goes down, as the table below ('payoff matrix' in game theory language) demonstrates.

(1) One-polluter, one-victim case

Profits without protection (-) and with protection (+)

Steel mill's profits	Laundry's profits	
	-	+
-	20; 10	20; 13 (NE)
+	17; 15	17; 17 (PO)

Signs (-) and (+) refer to whether some protection measures are taken or not. If nobody takes any protective measures, then the steel mill enjoys the profit of 20, and the laundry – the profit of 10. If both of them take protection measures, then the profits are 17 and 17. If the steel mill abates, and the laundry does not, their profits are 17 and 15. If the laundry takes some protection measures, and the steel mill does not abate, the profits are 20 and 13.

There are acronyms NE and PO. All of the students of this course guess that PO stands for Pareto optimum, but not all of them know that NE stands for Nash equilibrium – the most important concept in game theory. The simplest definition of Nash equilibrium is the outcome of the game such that none of the players has an incentive to unilaterally change the decision.

Numbers in the table above can be interpreted as the results of a 'game' between the steel mill and the laundry ('players' in the game). Each of the players has to take a decision to protect (+) or not to protect (-). The results ('payoffs' in game theory language) depend on their decisions. It is easy to see that (17,17) makes a generalised Pareto optimum: the joint profit – irrespective of how they share it – is 34 and it cannot be made larger. For (-,-) it is 30, for (+,-) it is 32, and for (-,+) it is 33. Let us check that (20,13) is a Nash equilibrium. If the steel mill switches from - to + (assuming that the laundry sticks to +), its profit will decline from 20 to 17. If the laundry switches from + to - (assuming that the steel mill sticks to -), its profit will decline from 13 to 10. Hence 20,13 is a Nash equilibrium, an outcome likely to be observed since nobody has a motivation to change unilaterally.

Let us analyse whether it is possible to move from NE to PO. As checked above, the players do not have a motivation to do it unilaterally. Things look differently, if the players can negotiate with each other. Coase theorem suggests that the laundry may approach the steel mill and offer a bribe of 3.5 to switch from - to +. By accepting the bribe, the steel mill will be better off, since 17+3.5=20.5>20. The laundry will be better off either, because 17-3.5=13.5 >13. The bribe of 3.5 is not the only one to solve the problem. The laundry's WTP is lower than 4, and the steel mill WTA is higher than 3. Thus numbers such as 3.1 or 3.9 can do the job as well.

The story can be changed somewhat by assuming that the laundry consists of ten identical independent units (small laundries) characterised by the following table (see page 33). The numbers are replicated from the previous table, except that laundries' profits are equal to the old numbers divided into 10 (reflecting the fact that each small laundry is 1/10 of the original one). Numbers for the steel mill do not change.

In the previous case, a bribe of 3.5 was adequate to motivate the steel mill to move from NE to PO. If each of the laundries contributes 0.35, then they collect 3.5 – sufficient to motivate the steel mill to abate. However, abatement is a public good, and some laundries may choose to free-ride, because they know that they cannot be excluded from benefits, once the steel mill abates. Let us assume that one laundry decided to free-ride, and only nine of them collect the money. 9x0.35=3.15>3, so the amount is sufficient to bribe the steel mill. But what if two laundries choose to free-ride? 8x0.35=2.8<3 which is too little for the steel mill to be persuaded to abate. Nevertheless, they can collect 0.38 rather than 0.35. 1.7-0.38=1.32>1.3, so they are better off even having paid 0.38 each. The collected sum will be 8x0.38=3.04>3 which will make the steel mill better off either.

(2) One-polluter, multi-victim case

Profits without protection (-) and with protection (+)

Steel mill's profits	Profits of a laundry (1 of 10)	
	-	+
-	20; 1.0	20; 1.3 (NE)
+	17; 1.5	17; 1.7 (PO)

However, it can be the case that three laundries choose to free-ride, and only seven laundries are ready to collect the money. Even if they decide to pay 0.4 – the maximum of what they can pay without making them worse off – they will collect 2.8 which is too little to compensate the steel mill. If the number of free-riders is four or more (greater than three), the laundries will not be able to move from NE to PO either.

Let us complicate the story even further by assuming that the steel mill consists of ten independent units. Their profits are given in the table below (case 3). If every small laundry is polluted by one neighbouring small steel mill, then we have simply 1/10 of the first case (1). In order to move from NE to PO, every laundry has to bribe "its" steel mill with 0.35. The geographical pattern of pollution dispersion can be more complex though, and laundries may depend not only on what they decide, but also on what other laundries do. The public good aspect of the problem may motivate some laundries to free-ride, making the prospects of moving from NE to PO uncertain.

(3) Multi-polluter, multi-victim case

Profits without protection (-) and with protection (+)

Profits of a steel mill (1 of 10)	Profits of a laundry (1 of 10)	
	-	+
-	2.0; 1.0	2.0; 1.3 (NE)
+	1.7; 1.5	1.7; 1.7 (PO)

In cases (2) and (3) above problems were caused by the fact that economic agents free-ride, because they do not reveal their preferences truthfully. In plain language, they simply cheat. Therefore economists who analyse public good questions, have tried to find mechanisms to motivate agents to reveal their preferences truthfully (i.e. not to cheat). It is virtuous to tell the truth, and if whatever people say is true, the problem disappears. Yet the truth-telling assumption is unrealistic, and it cannot be assumed in economic analyses.

Economists identified a mechanism – called a Groves-Clarke tax (GCT) – to motivate agents to truthfully reveal their preferences with respect to a public good. We will define GCT by referring to a decision to go for a specific project. Let us assume that a public good whose cost is c is to be financed by k people (numbered 1,…,k). Their contributions are to be

$c_1,\dots,c_k$ ($c_1+...+c_k=c$) sufficient to pay the total cost. The value of the project to potential users is $v_1,...,v_k$. The project will be carried out, if its total value exceeds its cost ($v_1+...+v_k \geq c$). In other words, net values n_i defined as $n_i=v_i-c_i$ for i=1,...,k, must sum up to a non-negative number ($n_1+...+n_k \geq 0$) for the project to proceed. But net values n_i (like the gross values v_i) are known to users themselves only. Actual negotiations are based on *declared* net values $s_1,...,s_k$ that are not necessarily equal to the true net values $n_1,...,n_k$.

The GCT is defined as follows:

- $GCT_i = \sum_{j\neq i} s_j$ if $\sum_{j\neq i} s_j \geq 0$ and $\sum_j s_j < 0$, (1)
- $GCT_i = -\sum_{j\neq i} s_j$ if $\sum_{j\neq i} s_j < 0$ and $\sum_j s_j \geq 0$, (2)
- $GCT_i = 0$ otherwise. (3)

If (1) or (2) holds for a given *i*, the agent *i* is called *pivotal*.

The formulae in (1)-(2) look complicated. In fact, they are not. But rather than analysing them theoretically, let us calculate GCT for a specific case which should clarify the idea. The case is about a typical externality problem (two polluters and two victims). The abatement (to be undertaken by polluters) creates a public good for victims. Individual negotiations *à la* Coase do not make sense, because the number of agents is higher than two. The government is expected to intervene and to make abatement obligatory. But its decision whether to intervene or not is based on declarations of the four agents. Unfortunately they can cheat: polluters may exaggerate their benefits from "no intervention" (avoided cost of abatement), and victims may exaggerate their benefits form "intervention" (avoided losses from pollution).

A referendum is planned to compare pros and cons of the intervention. If the total (declared) benefits of intervention exceed the total (declared) benefits of its lack, the government will come in. Otherwise the government will stay away. In order to motivate agents to truthfully reveal their preferences the following GCT is imposed on them. The tax is calculated according to the formulae in (1)-(3) above.

Benefits	Government intervention	No government intervention	GCT
Firm A (polluter)		20	0
Firm B (polluter)		10	0
Firm C (victim)	28		8 = -(22-20-10)
Firm D (victim)	22		2 = -(28-20-10)
Total	50	30	10

The outcome of the referendum (50:30) is to intervene, as obviously seen by all who compare the two numbers. Now it will be checked who turned out to be a pivotal agent. If the firm A was absent, the outcome of the referendum (50:10) would be for the intervention anyway. The presence of A does not change it, and thus it is not a pivotal agent. If the firm B was absent, the outcome of the referendum (50:20) would be for the intervention too. The presence or absence of B does not change it either, and thus B is not a pivotal agent. If the firm C was absent, the outcome of the referendum (22:30) would be against the intervention. The C changes it, so it is a pivotal agent. Finally, if the firm D was absent, the outcome of the referendum (28:30) would be against the intervention. The D changes it, so it is a pivotal agent. Non-pivotal agents (A and B) do not pay GCT. Pivotal agents (C and D) do pay GCT taxes. Once again please note that calculations made in the last column apply the formulae from (1)-(3).

GCT can be interpreted as a *sui generis* Pigouvian tax. By their votes, pivotal agents impose an externality on the rest of the economy, since – in the case of their absence – the outcome of the referendum would be different. The externality is simply the change in the sum of benefits, and they are supposed to pay this amount as the GCT.

The Groves-Clarke tax is rather small. Nobody likes to pay taxes, but by cheating firms risk to lose even more. This is the idea of GCT: you may turn out to be a pivotal agent and you may have to pay a tax, but when you cheat you risk losing even more; so, do not cheat!

The GCT mechanism does not allow the tax revenues to be paid back to taxpayers. A recirculation scheme would undermine the incentive mechanism: taxpayers would not be afraid of paying taxes, if they knew that what they pay would be paid back to them (perhaps indirectly). Hence it is crucial that GCT revenues are withdrawn from the economy. This questions the rationality of the tax. It leads to the optimal supply of the public good, but decreases the supply of the private good (money) below its optimum, since it deprives the (pivotal) agents of some of their private goods: they will have less money to buy them. In other words, it improves the allocation of a public good, but lowers the allocation of a private one.

This is an important theoretical reason why GCT has not been applied. But there are also practical reasons why the mechanism exists in textbooks only. Think of the cost of organising a referendum where people are asked to reveal their preferences with respect to a project to be financed from their taxes. Policy makers prefer to take relevant decisions based on what they think people's preferences are. If they are wrong – that is if they miscalculate people's preferences, and undertake intervention decisions contrary to what people expect – they will not be re-elected. This is not a perfect mechanism of assessing the demand for public goods, but perhaps an alternative would be even less satisfactory.

The first lecture (pages 1-6) outlined the general idea of environmental economics based on the fundamental principle MAB=MAC; protection activities are justified as long as they imply more benefits than costs; if they become too expensive, they should be stopped. The next lectures (pages 10-16 and 19-25) explained that some benefits and some costs may be recorded by different economic agents which leads to so-called externalities. Coase theorem suggests that under certain assumptions market failures caused by externalities can be avoided without a government intervention. The last lecture (pages 29-35) explained that government intervention (not necessarily leading to budgetary involvement) is called for in order to overcome free-riding.

Questions and answers to lecture 4

4.1 Can traffic lights be analysed as a public good?

Yes, they can. Traffic lights comply with both principles. If there are traffic lights, then it does not matter whether they operate for 1 person, or for 100. Thus non-rivalry works. In order to cross the street safely you wait for the green light, because you know that when there is the red light for cars, they will have to stop. The traffic lights are not free. The cost of installing them is quite high; some 100,000 € was probably spent from the municipal budget to install them. Nevertheless, while crossing the street you are not asked whether you paid appropriate taxes, because the non-exclusion principle works as well.

4.2 So-called "open access" takes place if the non-exclusion principle works, but non-rivalry does not. Please provide examples of environmental resource abuse under the open access regime.

The best known example is perhaps whale extinction. Whales are large animals that require abundant space to survive. They live in the ocean. Everybody can navigate in the ocean, and everybody can hunt whales. Of course there are international conventions to protect whales, but they are not very effective. Consequently non-exclusion principle works: nobody can be excluded from hunting whales. At the same time non-rivalry principle does not work: a whale killed and used by somebody, cannot be hunted by anybody else (consequently this is not a public good). As a result of the open access, whales become extinct.

The mechanism has been studied as "tragedy of the commons". Let us assume that there is a common grazing land for cows. If you feed your cow on this pasture, you benefit from it, but the fertility of the land decreases by, say, 1/10 (assuming that the pasture can accommodate up to 10 cows). If the total number of cows is less than 10, then the net benefit is positive. If it is 10, the cows can survive, but farmers make no profit. If it is 11 or more, then no cow can survive. The tragedy is that everybody has the same incentives to feed their cows even though the pasture cannot accommodate an unlimited number of cows.

This is a mechanism responsible for the collapse of many ecosystems. At the same time, Elinor Ostrom (a Nobel laureate in 2009) collected evidence from several communities where open access has not resulted in natural resource abuse. Apparently if a community establishes effective cultural constraints on using the common resource, the "tragedy of the commons" is not inevitable. In the case of whale hunting such constraints did not exist, and there are very weak prospects for establishing them in the future.

4.3 If for some good the non-rivalry principle works, but non-exclusion can be overcome, we call it "a club good". Is coded TV a club good?

Yes, it is. The coded TV signal can be decoded by an unlimited number of users. Thus the non-rivalry principle works. Free-riding can be overcome by the necessity to apply a de-coder, a device capable of translating the coded signal into something that can be used by a TV screen. The de-coder is a private good. It has to be purchased individually: a de-coder installed in one household cannot be used by another one. Moreover, firms which sell coded TV programmes, require purchasing their de-coding software every month. As a result, the non-exclusion principle (otherwise making a conventional TV signal a public good) is overcome.

4.4 Should free-riders be penalised somehow by those who pay?

From the moral point of view, probably yes. But it would be difficult to design a fair mechanism. A free-rider (like everybody else) can indicate zero WTP for a public good. Nobody can be forced to consume anything. At the same time, by the non-exclusion principle, he or she can use the good (that he or she did not pay for), but this is not his or her decision: public goods (like clean air or beautiful landscape) are used by everybody. Thus it would be difficult to penalise such people.

In a sense free-riders (like all of us) are penalised. Public goods are often financed by the budget. Thus – indirectly – all tax payers finance them. If financed by users voluntarily, public goods are to be undersupplied. Therefore, if an economically justified level of their supply is to be provided, they are paid by the budget on behalf of the society. Potential free-riders pay too.

4.5 How can an adequate supply of a public good be provided?

Sometimes a public good can be privatised. Police is an example of a public good. If people felt that it was undersupplied, then they tried to privatise it. One of the fastest growing businesses in former centrally planned economies after 1989 was private security. Shopping malls, other organisations, and apartment houses hired employees whose job was similar to what the police was supposed to do, that is to protect their assets and activities. This trend was not against people's welfare, since the police (who are underpaid) could focus their activities in other areas – benefitting those who cannot afford private protection.

Sometimes a public good can be transformed into a club good. Swimming pools serve as examples. The non-rivalry principle is obviously satisfied. They can be open to the public, but it is possible to fence them and to exclude free-riders. One needs to pay in order to enter such a swimming pool. TV – which was once a public good – has developed a coding system; it can be sold, and managed as a club good now (see 4.3 above). By transformation into a club good, the problem of free-riding is solved.

A typical method of providing an adequate supply of a public good is budgetary financing. The government decides how many units of a good is needed, and it pays for it.

A combination of public decision regarding the supply, and private decisions to "produce" it is often applied too. For instance, the government may decide that polluters should abate by a certain amount so that their emission is what the environment can cope with. Environmental improvement caused by this decision is a public good. Nevertheless, the government does not pay for the abatement. The polluters are expected to finance the abatement themselves. If they fail to abate they shut down their operations. In this case the government may be under the social pressure to loosen some of its environmental requirements, and to modify its decision about the environmental improvement (public good). In particular, it can subsidise the polluters in order to continue their operations. Separating decisions regarding the supply of a public good and financing is never quite strict.

4.6 In the lake pollution example (lecture 2, ERE-2), the emission from the power plant imposes an external cost on fishermen. If the plant abates, it provides a "public good". How could the supply of this good be financed?

According to the Coase theorem, the power plant and the fishermen can achieve an economically justified level of lake pollution irrespective of what environmental regulations state. Let us assume that the law says that the lake has to enjoy clean water. If the water quality deteriorates, fishermen's profits (depending on the abatement in the plant, MAB) go down. The plant would like to discharge its waste of, say, 30 tonnes into the lake, but it has to have fishermen's consent. Assuming that its abatement cost is MAC, the criterion MAB=MAC is satisfied for the emission level of, say, e^*=10 tonnes. The plant has to abate 20 tonnes, and compensate the fishermen for the remaining 10 tonnes. Now let us assume that there is no such law; if the fishermen wish to enjoy clean water, they need to "bribe" the plant to reduce its emission. They will do so as long as MAB>MAC. They will stop when the criterion MAB=MAC is hit. As before, the plant will emit 10 tonnes, and 20 tonnes will be abated. However, fishermen will bear the cost of abatement under this arrangement.

The supply of the public good is financed either by the polluter or by the victim, assuming that there are only two agents bound by the externality. If there are several independent fishermen, and the law does not force the plant to abate, then achieving an economically justified level of environmental protection is difficult, because of free-riding. Please recall the second and the third case of the steel mill / laundry example (pages 32-33 above; or in my overheads: ERE-4-5, and ERE-4-6).

4.7 If there are more than two agents bound by an externality, can the Coase theorem be applied?

The Coase theorem applies to two agents only, so it does not apply directly to this case. Nevertheless it can inspire an approach based on negotiations. Let us suppose that 1 million pedestrians and 1 million drivers negotiate how to organise the car traffic in their city. There are 2 million agents, but one can imagine negotiations between a representation of pedestrians and a representation of drivers; there are only two agents now who – according to the Coase theorem – can achieve a generalised Pareto optimum irrespective of whether the law favours pedestrians or drivers. As a result of negotiations an agreement can be reached regarding traffic lights, speed limits, pedestrian bridges, tunnels, and so on.

Such an agreement solves the problem if the representations of both sides can be treated as disciplined entities. In other words, if the representation of pedestrians or car drivers agreed to pay for something, it can deliver; it has an effective mechanism to collect the money. Otherwise free-riding will make the agreement unenforceable.

4.8 Why is it essential that GCT revenues are permanently withdrawn from an economy?

If a GCT-payer expects to get some money back (in the form of a direct subsidy, or in the form of lowering other taxes), GCT will not provide incentives that the theory of rational behaviour predicts. Attitudes of people who are exposed to the tax, can be similar to the attitudes of people who take part in lotteries. Assuming that they base their decisions on expected values, if the average prize is, say, 2 €, they should not pay for the ticket more than 2 €. However, if proceeds from the lottery finance something that people care for, they may not be afraid of paying for the ticket more than 2 €.

4.9 In the absence of a Groves-Clarke Tax, is there any chance that people will truthfully reveal their preferences?

Sometimes we do not cheat. If the public good we are supposed to reveal our preferences for is a part of our culture, we do not free-ride. For instance, if we are asked about budgets for certain social care projects, we do not expect that the money will be squeezed from others. On the contrary, economists study so-called "warm glow" effect. It refers to the empirical observation that people enjoy some emotional reward when they give to others. Accordingly they are inclined to declare even more than they are actually willing to pay to support the project.

Over the last several decades smart surveying techniques were developed to check if the declared WTP reflects people's true preferences. If it does not, then economists at least try to estimate the error.

5. Uncertainty

Uncertainty is something that adds to the complexity of environmental protection issues. No matter what we do, we can never be certain about the result. Nature can surprise us by responding not in a way we expected. We may protect coastal areas against floods, and the flood damages go up. We may spend quite a lot on protecting a species, and the species becomes extinct anyway. We may protect human health against certain injury, and in fact we expose people to an even greater damage.

We often refer to "uncertainty" and "risk" as synonyms, but in 1921, Frank Knight made an important distinction between the two. The former is a broader concept, and its meaning is consistent with what we attach to the word in our everyday language. In contrast, the latter is a special case of uncertainty, when we do not know what will happen, but we can estimate probabilities of alternative outcomes. Unlike uncertainty in general, risk can be insured. For instance, if someone's real estate is located by the river, the risk of flooding can be calculated by hydrologists. If it is estimated at, say, 2%, and the owner would like to get a compensation of 10,000 € if flood occurs, then he (or she) can buy insurance by paying 10,000 x 2/100 €, that is 200 €. Nobody knows when the flood comes, but if it comes the owner of the real estate can be paid what his (or her) insurance policy states.

Not all uncertainty is insurable. For instance, an entrepreneur who invests in a project cannot be sure whether the investment will pay back. Nevertheless it is impossible to buy insurance against business failure, because its probability cannot be calculated in a meaningful way. There are a number of uncertain situations that do not allow to buy effective insurance. Some economists claim that this is the essence of business activity, and the source of entrepreneurial profits.

Despite these terminological details, people usually understand by environmental risk unknown consequences of environmental degradation, irrespective of whether probabilities can or cannot be attached. 'Environmental uncertainty' would be a more adequate term, but 'risk' has become more popular in this context.

In 1980, an article "Witches, floods and wonder drugs" was published in a book on *Societal Risk Assessment*. One may wonder why these words were listed in a row, but in fact this is how people have tried to deal with uncertainty.

Everybody heard about medieval witch hunting. If a woman was accused of being a witch, a procedure was initiated to check whether indeed she was one. A typical test the woman was exposed to was to immerse her in water. If she drowned, then it turned out that perhaps she was not a witch (but it was too late to rescue her). If she did not drown, her witchcraft was proved (since some non-natural process must have been manifested). In other words, if she was suspected to be a witch, there was no effective procedure to prove that she was not (unless she was drowned).

It turns out that many contemporary procedures to eliminate risk resemble those used for witch hunting. For instance, how a chemical can be checked whether it is a carcinogen? So-called bioassays are used in order to check it. The bioassay requires that a box with animals, say, 50 mice is observed for the time of several months, or a year. This is called the control group. Simultaneously there is another box – called the experimental group – with 50 mice fed with the chemical to be tested. When the experiment is over, the number of cancer cases in both groups is compared. If the number of cancer cases in the experimental group is higher than in the control group, the carcinogenicity has been proved. But if the number of cancer cases in both groups is the same, then does this mean that the chemical is safe? No. Another bioassay needs to be performed, with a higher concentration of the chemical dosed in the experimental group.

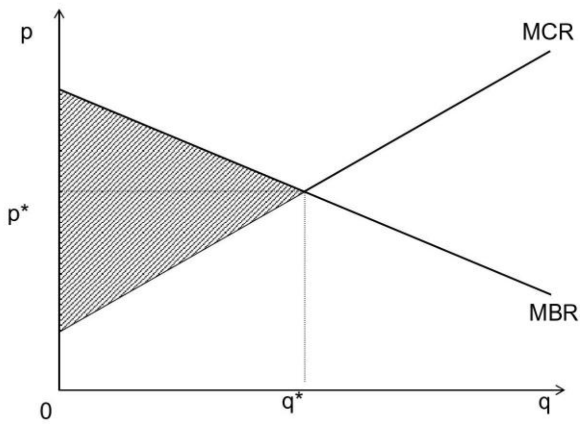
However, even the additional bioassay cannot prove the safety of the substance. If the outcome does not prove the carcinogenicity, its concentration in the food has to be increased even more. If the food becomes uneatable for the mice, they can be fed by force. If experiments with mice do not prove the carcinogenicity of the substance, scientists can switch to rats or larger animals.

This procedure resembles witch hunting. If a substance is suspected to be a carcinogen, then there is no procedure to "absolve" it. Either it is confirmed to be a carcinogen, or additional experiments are necessary.

This is inevitable whenever societies try to eliminate risk. It is impossible to eliminate risk completely. Societies need to learn how to live with it. Attempts to eliminate the risk of flooding proved to be ineffective. Even if large investments are carried out, floods occur from time to time. Their complete elimination is not possible. Instead, societies need to adapt.

Students from Germany (Germans are aware of the fact that floods affect the city of Köln from time to time) know that it would be ridiculous to try to make the river Rhine fully risk-free. Even the German economy (not to speak of lower income countries) should not spend too much money on anti-flood measures. It is better to acknowledge that some (not excessive) risk is inevitable.

Also drugs are risky. Attempts to eliminate any risk of medical treatment would result in sentencing to premature death millions of patients who are cured by drugs that – unfortunately – happen to give unwanted side effects sometimes.



Instead of trying to eliminate risk, societies have to accept some level of it. Economists define the acceptable level of risk by referring to the fundamental MB=MC criterion (see lecture 1, page 4). If you assume that both MB and MC depend on the level of risk (denoted as q in picture above), policy makers need to check whether its changes provide certain benefits and require certain costs. We can replicate the earlier graph with the letter 'R' added. Instead of MB we have MBR (meaning marginal benefit from additional risk reduction), and MCR (meaning marginal cost of additional risk reduction) instead of MC. The acceptable risk is where the two curves intersect (see picture above).

In other words, the more one tries to eliminate the risk, the lower marginal benefits from its elimination are, and the higher the costs of this elimination are. A rational solution would be to accept certain level of risk (q^*). Tolerating a low level of risk elimination (lower than q^*) helps to avoid the cost MCR, but loses higher benefits MBR. Trying to eliminate risk (more than q^*) provides benefits MBR that are lower than costs MCR.

Despite what economists advocate for as an acceptable level of risk, practitioners follow a much simpler approach. If the carcinogenicity is to be referred to, let us assume that experiments with mice demonstrate that concentration x_0 of certain substance in food is not toxic (which does not mean that it is completely safe). An average mouse lives two years. An average man lives 40 times longer. Thus one can expect that the concentration of $x_0/40$ should not be toxic for men. Yet, among the humans, there are individuals that are 5 times more sensitive to poisons than average. Thus the concentration of $(x_0/40)/5$ should not be toxic even for such people. Finally, an additional factor of 3 can be applied, "just in case". Consequently the concentration of $((x_0/40)/5)/3$, i.e. $x_0/600$, is accepted for consumption or medication purposes.

This procedure does not have any scientific base, but it is practised in medicine and engineering. This is how engineers arrive at various standards for bridges and architectural elements. They call the coefficients of 40, 5, or 3 interpreted as "safety factors". As in the case of toxicity, there is no scientific justification for adopting such a solution, but the method works. There are occasional medical scandals, and bridges collapse sometimes, but we follow doctors' prescriptions usually, and we are not afraid of crossing a bridge.

So-called *Precautionary Principle* (it will be discussed in more detail in lecture 12) states that in the case of uncertainty, decisions should take into account the worst possible outcome. Applied to managing environmental risk, it implies its complete elimination which is a very bad answer. Economists recommend using the MB=MC criterion, and practitioners stick to much less sophisticated pragmatic solutions.

Questions and answers to lecture 5

5.1 Is there a possibility to buy an insurance against failing on an exam?

No. Some professors pass all the students, and some have a "failure rate" of 30% or so. These numbers can be fairly stable, so some students may claim that there is e.g. 10% probability that they fail and therefore they may think of buying an insurance against failure. For example, if they value their "failure damage" at 500 €, they could expect to buy an appropriate insurance for 50 €. I am afraid, however, that no insurance company would be willing to offer such a policy. The reason is that – unlike flood – failing on the exam is not a random process. Even though some probabilities can be attached, it is not random, because –

to a large extent – students' behaviour may influence the result. If they study hard, the failure is less likely. If they do not, they are more likely to fail. An insurance company will not offer a policy if the mechanism of a damage is not random.

By the way, this is the reason why businessmen cannot buy insurance against bankruptcy. Even though statistics demonstrate that every year, say, 5% of registered companies turn out to be insolvent, the process is not random. Likewise, nobody can buy insurance against pregnancy, even though doctors say that the probability of success in the human fertilisation process is much less than 100% and they can estimate it. Students may identify dozens of processes where probabilities exist, but there is no randomness.

5.2 Can measures aimed at risk reduction result in greater losses?

Yes, they can. In the 1930s and 1940s the American government spent a lot of money invested in flood protection infrastructure (walls, retention reservoirs, etc.). At the same time much higher flood damages were recorded. Some analysts were surprised, but the mechanism was fairly simple and predictable. Motivated by improved protection, investors were more eager to build homes, stores, and manufacturing plants close to the rivers. Irrespective of protection measures, floods occur sometimes anyway. Once they come, they imply higher damages if there are more assets to be hurt.

When undertaking anti-risk measures, one needs to expect that these may induce more risky behaviour. For example, some road safety measures do not lead to lower accident rates if they encourage people to drive less carefully.

5.3 Why do procedures to assess carcinogenicity of chemical substances resemble that of witch hunting?

Because there is no possibility to give a negative answer. The answer must be a positive one (the chemical is proved to be a carcinogen), or ambiguous (not proved to be a carcinogen). Like in the witch hunting procedures: if a woman was accused of being a witch, there was no way to demonstrate that she was not a witch. The only statement possible was that she was not proved (for the time being) to be a witch.

5.4 Why are the mice used in bioassays most frequently?

Because they are cheap. Rats are bigger, so experiments with rats require more expenditures. Larger mammals are even more expensive to fence and feed. Besides, people are less embarrassed when they kill a mouse rather than a cat or a monkey.

5.5 Flyers attached to medical drugs list unwanted effects and their probabilities. One of the unwanted effects is the death of a patient. When we read that a death can be caused by using the drug, shall we reject the prescription?

The probability that a drug used for medical purposes can cause a patient's death is rather small. Nevertheless, it can be positive. Drugs known to cause patients' death (with very low probability) are administered usually when the alternative (i.e. no drug) leads to a higher probability of death. If I was prescribed a drug leading to unpleasant unwanted effects, I would check if my chances of survival without the drug – according to what doctors say – are high enough. The prescription of the drug should not be rejected if unwanted effects show up with sufficiently low probabilities.

5.6 In 1981 the US *Food and Drug Administration* adopted a standard of 2 PPM (*Parts Per Million*) for PCB (*Polichlorinated Biphenyls*) concentration in fish meat. The standard was motivated by preventing cancer cases. If PCB was proved to be carcinogenic, why did not they adopt the zero standard?

Adopting a zero standard will bring the number of PCB-related cancer cases to zero which is a benefit. However, at the same time, it will preclude selling fish caught in aquifers that are contaminated even slightly. Environmental policy makers must judge whether the cost of a standard (or any other regulation) is justified by its benefit for the society. They are reluctant to compare health related benefits with pecuniary costs. In some cases, however, it is possible to evaluate benefits and costs in the same physical units (rather than money). For instance, alternative anti-COVID-19 measures were assessed in terms of fatalities both at the benefit and the cost side. Certain social distancing regulations lead to a lower number of COVID-19 deaths which is a benefit. Simultaneously they lead to a higher number of deaths caused by other factors, which is a cost. Striking a balance is a challenge policy makers cannot escape.

5.7 Do "safety factors" have a scientific justification?

No, they do not have any scientific justification. The only reason why doctors or engineers adopt the number 5 rather than 10 (or *vice versa*) is experience and common sense: this is what has been practised, and it works.

5.8 The Precautionary Principle suggests that people should not fly airplanes. Why do people ignore it?

The worst possible outcome of a flight is a crash leading to the passenger's death. Everybody knows that airplanes crash sometimes. Nevertheless, people ignore the precautionary principle, and they choose air travel from time to time. Perhaps there are some people who calculate probabilities, and discover that per kilometre travelled, planes are safer than cars and bicycles. However, most of us choose the airplane, because we appreciate the benefit of visiting a distant place without spending too much time travelling. Once again, this emphasises the adequacy of what economists consider "an acceptable level of risk". This level equates MBR with MCR. It looks at whether the marginal cost of reducing the risk (i.e. reducing the risk of dying in a plane crash) is larger than the marginal benefit (i.e. depriving ourselves of the benefit from moving to the destination quickly). Apparently whenever we decide to embark on a plane, we judge that MBR≥MCR.

5.9 Scientific analyses of climate change suggest that the global warming may trigger irreversible adverse modifications of weather patterns. What environmental policies do such analyses imply?

Modelling climate change demonstrates that global warming results from excessive anthropogenic emission of carbon dioxide. Carbon dioxide abatement requires some costs. At the same time, global warming will deprive the world of important benefits. According to many assessments, the MBR=MCR criterion implies a drastic reduction of carbon dioxide emission. This means that the <u>global</u> carbon dioxide emission should be reduced.

Unfortunately it has been growing despite abatement undertaken unilaterally by some regions over the last three decades. Quite paradoxically, unilateral abatement action may cause an increase rather than a decrease in the global emission (unilateral abatement may encourage the growth of emission elsewhere). This is one of the topics of my class of International Environmental Cooperation (in Spring semester). Climate protection is a public good, and unless appropriate mechanisms are deployed, its supply will be inadequate (see lecture 4). In particular, unilateral (national or regional) policies to abate carbon dioxide are not effective.

6. Discounting

If I asked you to give me 100 euro today, and promised to give you 100 euro back one year from now, most of you would say "no". But if I ask "why did you disagree", many students explain that they are afraid of inflation: 100 euro a year from now – because of inflation – is worth less. They are polite, so they do not raise another reason often, but there is also a risk that I will die within the next year, and consequently they will never get their money back. There is also yet another reason of uncertainty. What if I survive, but I turn out to be a thief who denies any promises to give the money back?

So let us assume that there is no risk. If you give me the money, you will get it back for sure. Also inflation can be excluded. If there is inflation, say, 2%, I will give you 102 euro so that the money you get has exactly the same purchasing power as when you gave it to me. Despite that, most of us would not agree to such a deal.

If I changed my question: "you give me 100 euro today, and I give you 130 euro next year (no inflation, no risk of losing the money); do you agree?" Some of you will probably say "yes". Let us bargain then. I ask: "you give me 100 euro today, and I give you 101 euro next year". You answer "no". I say "you give me 100 euro today, and I give you 125 euro next year". You answer "yes". I say "you give me 100 euro today, and I give you 102 euro next year". You answer "no". And the negotiation process can go on. If there is an amount, say, 105 euro to be paid back after a year, and you answer "I am indifferent – I can agree to such a deal, but I do not insist; I can do without such a deal either", then we say that 5% is the discount rate applied by the person who agreed.

The discount rates we apply can be very different. As a rule, they are positive (nobody has a zero or a negative discount rate), but they depend on many circumstances. If somebody has very attractive investment opportunities, then his (or her) discount rate is high. If somebody does not have such opportunities, then the discount rate is lower. Also, if the sacrifice one has to envisage in order to decrease the consumption today is large, then the discount rate is high. If this sacrifice is not large (for instance, if you come from a wealthy household) then the discount rate is low.

Discount rates are very often identified with bank interest rates. This is true, but it would be an oversimplification to equate the two. One reason is that bank interest rate is typically combined with a number of financial services (such as e.g. credit cards), and therefore it is difficult to claim that you agree to, say, 3% because this is your discount rate; perhaps your discount rate is higher, but you agree to 3% because of the other benefits the bank offers you. Or it can be the other way around. You take advantage of the 7% rate of interest not because your discount rate is that high. The bank is threatened by bankruptcy, you are aware of the risk, and you require that the bank offers you some extra premium for the risk.

Interest rate on Swiss government bonds is considered a useful proxy for a discount rate (at least for those who buy such bonds). They are considered safe (the Swiss government is unlikely to go bankrupt), and they do not offer any benefits (like credit cards or so). The interest rate on Swiss government bonds is around 3%, and this is sometimes considered a useful reference for what our discount rate can be.

The formula for finding equivalents of money amounts in different time moments can be iterated. In general, it is (t stands for some moment in the future):

$$X_t = X_0(1+r)^t, \text{ or } X_0 = X_t/(1+r)^t,$$

where r is a discount rate. Please note that in my examples the time span between "now" and the "future" is one year. For one year the formula then reads:

$$X_1 = X_0(1+r)^1, \text{ or } X_0 = X_1/(1+r)^1,$$

For two years it will be

$$X_2 = X_0(1+r)^1(1+r)^1=X_0(1+r)^2,\text{or } X_0 = X_2/(1+r)^2,$$

and so on. The formula linking X_t with X_0 is called the "Present Value" (PV) of X_t in moment t=0. If it is applied to X_1, X_2, X_3, and so on, the formula of PV reads:

$$NPV = X_0/(1+r)^0 + X_1/(1+r)^1 + X_2/(1+r)^2 + ... + X_T/(1+r)^T,$$

where T is the last year that the decision (project) is expected to imply a cost or a benefit (costs are entered with "minus" signs, and benefits are entered with "plus" signs). The letter "N" stands for "net", and consequently the acronym NPV reads "Net Present Value". The word "net" refers to the basic philosophy of Cost-Benefit Analysis: we are interested in benefits net of costs (i.e. benefits minus costs). If benefits and costs belong to different time periods, then we need discounting to make them comparable. Namely, we recalculate everything in present value terms, as if everything was to happen at once, in the beginning (that is at the time of the analysis), in t=0.

If X_t=X=const – then the formula for the present value simplifies:

$$NPV = X/(1+r)^0+X/(1+r)^1+X/(1+r)^2+...+X/(1+r)^T = X(1+1/(1+r)+...+1/(1+r)^T)$$

It is good to know, that there is such a simplified version (like the one above), but we will not use this formula extensively (I used it when I calculated the tables you will see in a moment).

Formulae of how to calculate NPV are used to define the so-called Internal Rate of Return (IRR). Namely, IRR is defined as a discount rate which makes the NPV=0. In other words, IRR is a discount rate r such that:

$$X_0/(1+r)^0 + X_1/(1+r)^1 + X_2/(1+r)^2 + ... + X_T/(1+r)^T = 0$$

If $X_0, X_1, \ldots, X_{\tau-1} < 0$, and $X_\tau, X_{\tau+1}, \ldots, X_T > 0$, then IRR is the only r which solves the equation above (IRR is unique). For typical projects this condition is satisfied (i.e. one needs to bear some investment cost in the beginning, reflected by negative X_S, and then one benefits from it, reflected by positive X_S afterwards).

IRR is a useful indication of profitability (it answers whether the project makes economic sense):

- If IRR is higher than the interest rate available for investor (to borrow the money in order to finance it), the project is efficient (and it is worth financing)
- If IRR is lower than the interest rate available for investor (to borrow the money in order to finance it), the project is inefficient (and it should be abandoned)
- If IRR is negative then the project does not make sense irrespective of the terms of availability of the money (unless you do not have to pay for the credit; the bank pays you to borrow the money).

Projects with low IRR need to be scrutinized carefully. Even if you do not have to borrow money (if you can finance the project yourself), low IRR indicates that the advantage of benefits over costs is small. If there are compelling non-economic reasons to carry out the project, it may be worth promoting, but please be careful. Perhaps there are better ways to spend the money.

Now let us go back to discounting. It is fairly easy to convince people about the need of discounting. If there is no discounting (or if the discount rate is zero) then you should accept the deal of giving me 100 euro now, and I will give it back to you a year from now. Everybody understands that this would be ridiculous, and therefore some discounting is inevitable. Problems arise when we try to envisage long-term consequences of discounting.

The present value of the future amount of X_T=1,000,000

	T=1	*T*=5	*T*=10	*T*=20	*T*=50	*T*=100	*T*=200
r=0%	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
r=1%	990,099	951,466	905,287	819,544	608,039	369,711	136,686
r=4%	961,538	821,927	675,564	456,387	140,713	19,800	392
r=8%	925,926	680,583	463,193	214,548	21,321	455	0.21
r=12%	892,857	567,427	321,973	103,667	3,460	12	<0.01

Please look at the table above. It gives you the present value of a high amount (1,000,000) which is going to be offered to you next year, 5 years from now, 10, 20, 50, 100, or 200 years from now. The present value depends on the discount rate applied. If the discount rate is *r*=0% (no discounting), then the present value is always the same and it is equal to the nominal value of 1,000,000. If a modest discount rate is introduced (say, 4%), the present value is less than a half of the original amount if offered 20 years from now. If the discount rate is *r*=12%, then the present value almost disappears in 100 years. This is something that most people are not ready to acknowledge. Yet it is an unavoidable consequence of applying discount rates.

The next table reveals another uneasy result of discounting. Let us assume that we are to receive 100 annually forever. The present value of 100 obtained now is obviously 100. However, we are going to get this amount over and over again. If no discounting is applied (or if *r*=0%), then the value of this payment for 10 years is 1,000, 5,000 in 50 years, 10,000 in 100 years, and infinity if there is no finite time horizon. The situation changes when some discounting is introduced. The present value of such a constant amount increases with number of years we look forward to, but it is finite even if $T=\infty$. For example, if the discount rate is *r*=4%, and we are interested in the present value of a gift of 100 per annum we are going to receive for 10 years, it is not 1,000; it is only 811 (mere 81% of the product 100*10).

The present value of the flow of a constant amount of X=100

	T=10	*T*=50	*T*=100	*T*=∞
r=0%	1,000	5,000	10,000	∞
r=1%	947	3,920	6,303	10,000
r=4%	811	2,148	2,451	2,500
r=8%	671	1,223	1,249	1,250
r=12%	565	830	833	833

The calculation of all columns can be carried out in a spreadsheet. Only the last column can be calculated in an easier way (e.g. on your calculator). To see this, please recall the formula from page 45:

$$NPV = X(1/(1+r)+...+1/(1+r)^T)$$

This is a geometric series with the quotient $q = 1/(1+r)$. Please note that the first payment is to be received by the end of the first period; thus its PV is $X/(1+r)$. The formula for the sum of such a geometric series (I assume that you remember your high school mathematics) is

$$a_1(1-q^T)/(1-q),$$

where a_1 is the first term of the geometric series, and q is its quotient. If $r>0$ then the quotient q is less than 1. If you go with $T\rightarrow\infty$, then q^T goes to zero (hence the numerator goes to a_1), and the denominator is constant (it is equal to $1-q$). This lets us calculate the last column as $(X/(1+r))/(1-q)$, that is $(X(1+r))/(r(1+r))$ or simply X/r.

The following case study will help us to see how these concepts can be applied in practice. We will calculate the IRR of a wind-mill project.

Let us assume that the project is characterized by the following parameters. The wind mill of 1 MW capacity

- works 2000 hours per annum
- costs 2 million euro
- sells electricity at 50 euro/MWh
- requires no maintenance costs for 30 years.

The assumptions are not far from reality. A typical wind-mill can work for 30 years or so. It is very expensive to build it and to connect to the network, but once this is done, it operates almost for free. Annual maintenance costs are 2% of the investment cost (40,000 euro in this example), and they can be simply neglected; wind-mills do not require any repairs usually. The wind mill works 2000 hours per annum (on average less than 6 hours per day). Of course it can work longer, but with the capacity lower than 1 MW. Thus, every year, it produces around 2,000 MWh. It sells electricity at the price of 50 euro/MWh (5 eurocent/kWh). The price you pay for electricity is much higher, but wind-mill sells at wholesale (not retail) prices; what households pay includes the cost of distribution, taxes etc. Hence the annual revenue of the wind-mill is 100,000 euro (if you do not like the maintenance cost of 40,000 euro to be considered negligible, you may assume that the wholesale price of electricity is higher by this amount, and then you can subtract the maintenance cost; this will not change the calculations).

In other words, the project costs 2,000,000 euro to be paid on January 1 of the first year, and then it brings 100,000 euro every December 31 for 30 years. Its present value depends on the discount rate applied. If we want to calculate IRR, we shall answer the question what discount rate makes the NPV=0. The question can be rephrased: what is the highest discount rate which applied to this project does not imply losing money?

$$\text{NPV} = -2{,}000\text{k}/(1+r)^0 + 100\text{k}/(1+r)^1 + 100\text{k}/(1+r)^2 + \ldots + 100\text{k}/(1+r)^{30} =$$
$$= -2.000\text{k} + 100\text{k}(1/(1+r) + 1/(1+r)^2 + \ldots +1/(1+r)^{30}) =$$
$$-2{,}000\text{k} + 100\text{k}((1-q^{30})/(1-q)), \text{ where } q=1/(1+r), \text{ if } q\neq 1, \text{ i.e. } r\neq 0$$

(like in the language of computer scientists, here k is understood as 'kilo', that is one thousand). In order to solve the equation NPV=0, one needs to solve an algebraic equation of the 30th degree. We know how to solve 1st degree (linear) equations and 2nd degree (quadratic) equations. Some people know how to solve the 3rd or the 4th degree equations, but there are no formulae to solve 5th degree and higher degree equations. They can only be solved numerically. By trial and error, we can check when the NPV is positive, and when it is negative. Then we can expect that somewhere "in between" it is equal to zero.

If you have access to spreadsheets, then finding an IRR (finding *r* or *q* such that NPV=0) can be made automatically, not by trial and error. I did it, and I found in this example that NPV=0 if and only if q=0.97, i.e. r=0.03. Please note that *r* is not simply 1-0.97. *r* needs to be calculated from the formula: q=1/(1+r). Hence r=1/q-1. Incidentally, it is 0.03.

IRR=0.03 for this project can be interpreted in a straightforward way. If the discount rate is higher than 0.03, then the investment will never pay back. If the discount rate is lower than 0.03 then the investment makes sense. It can also be phrased in the following way. If the money is available to the investor at the rate lower than 3%, then investment makes sense. If the money is available to him (or her) at the rate higher than 3%, then investment should be abandoned, since it will never pay back. Finally, please note that if discounting is abandoned (or *r*=0), then the investment will pay back after 20 years (because 2,000,000 = 20*100,000). Please look at the first row in the table on page 47: if you have a steady flow of revenues, and if the project works for a sufficiently long time period, then it will repay its investment cost sooner or later. Discounting demonstrates that this argument is wrong.

Discounting is inevitable. If you do not discount, then it is not possible to make any reasonable economic analysis. However, several decades of my discussions with environmentalists demonstrate that they do not accept discounting (or – what is logically equivalent – they are of the opinion that the discount rate should be zero). There are two case studies which illustrate this.

The first is a wind-mill story, or any analysis involving renewable resources: you make a one-time investment, and then you get benefits "for free", year by year forever. If you receive benefits forever, then it is just a matter of a sufficiently long time horizon to see that their sum can be made arbitrarily high; in particular it can be made higher than the initial investment cost. Therefore the argument is: build wind-mills at any cost – they will pay back after some time. But all of a sudden there are economists who say (as in the example above), that NPV depends on a discount rate. Wind-mills make economic sense only under specific assumptions. Some environmentalists are not convinced.

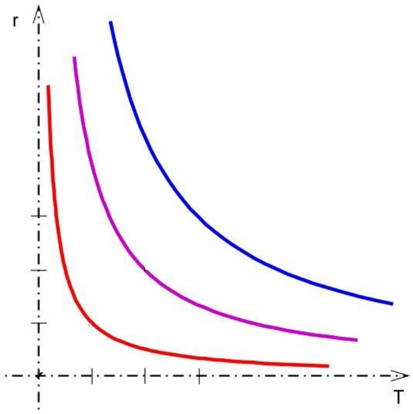
The second case is a nuclear controversy. Many people know that a nuclear power plant is excellent, and – as a rule – environmentally safe (Chernobyl and Fukushima are exceptions which do not contradict this rule). There is one weak point in this analysis. Namely they require a huge decommissioning cost. There are no ideas what to do with scrapped reactors and radioactive waste. The cost of this can be quite high, perhaps the same as the cost of investment. Nevertheless it is to be borne after 50 years (since a typical nuclear power plant is likely to work for 60 years – even more than 50). The table on page 46 shows you, that if you apply a modest discount rate, say, 4%, the PV of this postponed cost is less than 14% of the nominal value. If you apply 8%, then the PV is less than 2%, i.e. almost nothing. The environmentalists say: nuclear power plants should not be built. If they are built, then this is because of discounting: under these circumstances economists can ignore their huge and perhaps prohibitive decommissioning costs.

There is a pressure either to ignore discounting or to apply a much lower rate than implied by empirical research. The latter suggests that many of us apply a discount rate close to 4%. This is based on calculations, experiments, and observations on how we behave if we choose between now and the future. But the future we can capture is typically of one or several years ahead. It would be impossible to run an experiment extending over, say, a 50-year period. Therefore we do not know what are our preferences with respect to very long-term decisions. We simply extrapolate what we observe with respect to short-term behaviour to our hypothetical behaviour in the long run.

This is justified by so-called time consistency principle. This principle says that we can calculate NPV in two stages. Let us assume that we would like to calculate NPV of 1,000 to come 10 years from now, if the discount rate is 4%. This is simply $1{,}000/1.04^{10} = 1{,}000/1.48 = 676$. In other words, the NPV of 1,000 in 2034 is 676 in 2024. We can arrive at this number by discounting it from 2034 to 2028 (six years), and then from 2028 to 2024 (four years). The "intermediate" value (NPV discounted from 2034 to 2028) will be $1{,}000/1.04^6 = 1{,}000/1.26 = 790$. When this amount is discounted from 2028 to 2024 we get $790/1.04^4=790/1.17=675$ (it is not 676 because of the rounding errors; if all calculations were carried out with higher accuracy, then the second number would be equal to the first one exactly).

The principle is based on a rule that everybody remembers from school: $x^{a+b}=x^a x^b$. If you want to calculate 2^5 (which is 32), you can calculate $2^3=8$, and $2^2=4$, and then multiply the two numbers. Until recently economists were convinced that the principle of time consistency is unquestionable. Hence, if you calculate NPV over a ten-year period you have to use the same discount rate when you discount over, say, 4 years, or 6 years. The first empirical doubts were voiced in the 1980s. Yet it was much later that economists provided a convincing explanation that – indeed – if you discount over a long period of time, you may apply a lower discount rate than is appropriate for a shorter one. This idea is often called "hyperbolic discounting", as the following pictures explains.

The graph – displaying three examples of hyperboles r=A/T (with three positive constants A) – illustrates the tendency that the longer the time horizon T, the lower the discount rate r. For instance, if we look at short time periods (say, one or three years) then the rate we discount with can be 4% or so. But if we look at a long time period (say, forty years) an appropriate discount rate can be much lower, perhaps 2% or so.



Discounting with variable discount rates violates the time consistency principle. For instance, if a long-term investment is planned (60 years in the case of a nuclear power plant) then a low discount rate may be appropriate. If a shorter-term investment is planned (30 years in the case of a wind-mill) then a higher discount rate may be appropriate. Calculations of NPV cannot be carried out in two stages like in the previous example (if the long period is divided into two shorter ones then discount rates applied do not have to be the same, and the formulae we used on page 49 do not hold).

Questions and answers to lecture 6

6.1 Most of us think that if inflation rate is 3%, then the nominal discount rate of 8% corresponds to the real rate of 5%. This is not quite correct. Why?

Discount rate makes the two amounts – X_0 this year and $X_1=X_0/(1+r_{nominal})$ next year – equivalent. But because of inflation the future amount is worth only 1/(1+CPI) of its nominal value (where CPI – *Consumer Price Index* – is the inflation rate). We can thus write $X_1/(1+CPI)=X_0/(1+r_{nominal})$ or – equivalently – $X_1=X_0(1+CPI)/(1+r_{nominal})$, and we would like to compute r_{real} such that $X_1=X_0/(1+r_{real})$. By combining both equations we get $X_0(1+CPI)/(1+r_{nominal})=X_0/(1+r_{real})$, that is $(1+CPI)/(1+r_{nominal})=1/(1+r_{real})$, or $(1+r_{nominal})=(1+CPI)(1+r_{real})$. Hence $r_{nominal}=1+CPI+r_{real}+CPI*r_{real}$. The identity $r_{nominal}=1+CPI+r_{real}$ neglects the term $CPI*r_{real}$. This term is small indeed if both CPI and r_{real} are few percent. In our example CPI=3%, and r_{real}=5%, so their product is equal to 0.0015, and it can be neglected. Nevertheless, the general rule that in order to arrive at real values you simply subtract inflation rate from nominal values is not precise.

6.2 So-called Ramsey formula reads $r=\rho+\eta*g$, where r is the discount rate, ρ is the pure time preference, η is the elasticity of the marginal utility of money, and g is the growth rate Analysed in the class discount rate r is higher than "pure time preference" ρ. Why?

Educated at the University of Cambridge, Frank Ramsey (1903-1930) made important contributions to mathematics and economics. What we call "the Ramsey formula" ($r=\rho+\eta*g$) is perhaps the best known of these. The formula was developed in order to solve an optimisation problem of how to split production between consumption (which pleases us immediately) and investment (which will satisfy our needs in the future). In trying to determine r in the class discussion, we probably had in mind all three components, i.e. ρ, η, and g. Let us explain how these parameters may influence our choices.

The pure time preference ρ is perhaps the most obvious motive of our desire to have now rather than in the future. Yet we also expect to be wealthier in the future. If the growth rate is, say, 4%, then we expect that instead of having, say, 10,000 € today we will have 10,400 € next year. Will this additional money make us better off? Surely it will: we will have 400 € more. But how much better off will we feel? Economists say, that the marginal utility of money (whatever it means) increases somewhat slower than the money. This means that $\eta<1$. If $\eta=½$, then these additional 4% correspond to only 2% as far as our well-being is concerned. Thus $r=\rho+2$. Hence if we discovered that our discount rate was, say, 5%, this number consisted of the pure time preference of, say, 3%; and 2% corresponding to $\eta*g$ (if we expected that our wealth would grow by 4%, and the marginal utility of money was estimated at ½). If we expect our income to grow, and additional money makes us happier, then $\eta*g>0$. Consequently $r>\rho$.

6.3 IRR of the project #1 is lower than IRR of the project #2. Does this mean that #1 is worse than #2?

If the projects have the same duration, it does. But if the second project has a longer duration than the first one, then it does not have to be worse. In order to compare projects with different time spans economists developed the concept of so-called levelised costs. This term is especially popular among electricity specialists.

The *Levelised Cost of Electricity* (LCOE) tries to capture not only the distinction between investment and maintenance cost, but also tries to find a comparability between projects characterised by different time scales. The most relevant dilemma to be addressed in this context is to compare windmills and nuclear power plants. The former are typically built for 25-30 years. The latter can work 50-60 years. As expected, discounting proves to have a crucial role in determining the adequacy of either technology, but – unfortunately – there are no convincing arguments to advocate for a discount rate of, say, 4% or 6%. Arguments for using variable rates (perhaps hyperbolic discounting for very long-term projects) seem to be appropriate, but this obscures the analysis even further.

The definition of LCOE is fairly straightforward:

$LCOE = (I_o + \Sigma_i M_i/(1+r)^i)/(\Sigma_i E_i/(1+r)^i)$, the summation is for $i=1,\ldots,n$,
where

- n is the lifetime of a project,
- I_o is the investment cost,

- M_i is the maintenance cost in year i,
- E_i is the electricity production in year i,
- r is a discount rate

LCOE (e.g. measured in € per kWh) takes into account discounting (like IRR), but also the fact that various projects may have different time spans. No matter what their respective IRRs are, electricity generating projects can be compared by looking at their LCOE numbers.

6.4 As the winner of a lottery you have a choice of getting your prize, i.e. $ 1,000,000, immediately. Alternatively you can wait one year to receive $ 1,030,000. What will you choose?

The answer is simple. You check whether 1,030,000/(1+r) is greater than 1,000,000 or not. If your discount rate is lower than 3% you will wait. If your discount rate is higher than 3% you will claim your prize immediately. But please see also 6.5.

6.5 As the winner of a lottery you can get your $ 1,000,000 immediately, or in two instalments: $ 500,000 now, and $ 500,000 a year from now. What will you choose?

Unless your discount rate is 0%, you should claim your entire prize immediately, because the present value of the second instalment is lower than its nominal value. Nevertheless, shrewd investors calculate their tax obligations too. It is possible that your income of $ 1,000,000 is subject to a higher tax rate than $ 500,000. If this happens then in order to choose between the two variants you should compare the hypothetical loss of decreased present value with the hypothetical gain from lower taxation.

6.6 Please refer to the table on page 46. The present value of 1,000,000 to be received 5 years from now is 821,927, and to be received 10 years from now (the time doubles) is 67,564 when the 4% discount rate is applied. If the waiting time doubles from 50 years to 100 years, the present value shrinks from 140,713 to19,800. Why are these proportions different?

The proportions should be the same if the numbers change linearly, but they do not. $X_{10}/X_5=(1+r)^5/(1+r)^{10}=(1+r)^{-5}$ and $X_{100}/X_{50}=(1+r)^{50}/(1+r)^{100}=(1+r)^{-50}$ The second number is smaller than the first one (unless r=0). Even though the numbers do not change linearly, their logarithms do: log67564/log821927/≈0.8, and log19800/log140713≈0.8.

6.7 Please refer to the table on page 47. If the discount rate is 12%, then the present value of the annual payment of 100 over 50 years is 830 (not 5,000) It almost does not grow afterwards. Why?

The present value of 100 to be received in the 50th year is 0.35, that is a negligible amount. Later on it is even smaller (for instance, the present value of what you receive in the 51st year is 0.31). If you keep adding these smaller and smaller numbers the sum increases only slightly. The present value of 100 to be received in the 100th year is 0.0012, that is virtually zero.

6.8 The prince of Milan, Ludovico Sforza, gave a vineyard near the village of Fiesole, instead of paying Leonardo da Vinci what they agreed upon. How should this transaction be judged?

Leonardo as a painter was a genius, but he had problems with arithmetic. He probably agreed to this transaction without deep analyses. The prince of Milan would not be able to carry out necessary calculations either, but his officials were probably sufficiently competent to prepare the transaction. Its adequacy should be judged by looking at the table on page 47. If the vineyard gives an annual revenue of 100, the NPV of such a permanent flow depends on the time horizon, and on the discount rate. It can be assumed that a vineyard can live 50 years, and the discount rate is 1% (this is what it could have been in the 15th century Lombardy). Under these assumptions, the market value of such a vineyard was 3920. If this is what the prince owed to Leonardo then the transaction was adequate. If he owed more, then he cheated. If he owed less, then he turned out to be generous.

6.9 IRR=3% for a wind mill means that if money can be borrowed at 4%, then the investment is not economically justified. The availability of money is often even more difficult (than 4%), and yet you see hundreds of wind mills being constructed around. How is it possible?

Please look at the assumptions of the class example. The wind mill is supposed to work 2000 hours per year, and the retail price of electricity is 50 €/MWh. These numbers can be changed. An average year has 8766 hours. It was assumed that a wind mill operates only 2000 hours, that is less than 25% of the time. There are maps of wind velocity. They show that the average wind velocity varies quite a lot. If the wind mill is located in a windy place, then it can operate more than 2000 hours per year. It was also assumed that the retail price of electricity is 50 €/MWh. But the government has instruments to make it higher. For instance, electricity from renewable sources does not need carbon dioxide emission permits. If other power plants have to buy such permits, the retail price of electricity may go up. It is a matter of environmental policy to regulate the market for carbon dioxide emission permits. If the number of operation hours, and if the retail price of electricity are higher than assumed in our calculations, then the IRR of a wind mill may exceed the rate the money is available at.

6.10 Why did economists consider the time consistency principle as an unquestionable thing?

We are so used to the rule "$x^{a+b}=x^a x^b$" that the time consistency principle seems obvious. It allows for easy calculations of NPV. It was unimaginable that an analysis can be performed for variable discount rates; using variable discount rates would not allow to combine results carried out in shorter time intervals. Yet empirical observations – suggesting that people use higher discount rates for shorter periods – cast some doubts in the 1980s. Nevertheless, the principle was rejected only fairly recently, when a reasonable explanation was developed whether a project can be divided into smaller segments. By the way, whether projects are "divisible" or not depends on analyses of how a newly acquired information (as a result of technological progress) can be used. For instance, a nuclear power plant will use an old technology for 60 years even if a new one became available after 30 years of its operations. But this is just a curio for students who are interested in more sophisticated principles of economic analysis.

7. Economics of exhaustible resources

There are two, or perhaps three, types of natural resources. In many textbooks resources are considered either renewable, or exhaustible. Iron ore and coal are examples of the latter; their supply is limited, and if exploited, it will be exhausted in the future. Fish in a lake and trees in a forest are examples of the former; their supply is renewable, and they can last forever if we use them reasonably (what "reasonably" means will be explained in the next lectures).

Some textbooks define a third category of resources, called non-depletable, whose the only well-studied example is natrium chloride (table salt). In principle, these resources are constrained (since they do not breed), but in practice they cannot be exhausted. The annual consumption of salt is few kilogrammes per capita. Let it be, say, 10 kg. There are about 8 billion people in the world. Thus our annual consumption is around 80 billion kg, that is 80 million tonnes. If you add salting the streets in countries where it is practiced in winter, the total consumption does not exceed 150 million tonnes. At the same time 1 km^3 of the ocean contains 27 million tonnes of salt. Given the fact the ocean's volume is gigantic, the global consumption of salt is negligible. In some countries people extract geological resources (originating from ancient seas). In Poland their supply is estimated at 90 billion tonnes. In some countries people use the sea salt directly, just evaporating the water in artificial shallow lagoons. The world resources of table salt are so huge that their current consumption is meaningless. Hence they are not analysed in economic theory (by the way, the salt consumed will reach the sea sooner or later, and it will become available again). If they are of interest to economics, it is not because any intertemporal trade-off our consumption of salt implies.

We will start our analysis of natural resources with exhaustible resources. The Hotelling rule (developed by Harold Hotelling in 1931) provides a reference for this category. An extraction rule has to be found to maximise the present value of the flow of profits from sales of an exhaustible resource in a competitive market. Thus we look for a rule to be followed by a (hypothetical) rational owner of the resource. If postponing extraction turns out attractive, he (or she) will do so. Otherwise he (or she) will continue the extraction, perhaps its speed will be even increased. In order to judge the attractiveness of alternative scenarios, an assumption is made that the size of the resource is known and it cannot be increased (i.e. no future discoveries). This assumption is questionable, because everybody knows that unexpected discoveries may happen. Nevertheless, it would be impossible to make any predictions, if the future supply was undetermined. The Hotelling rule reads:

$$\mathbf{(d(p-MC)/dt)/(p-MC) = r},$$

where:

- p – price of the resource,
- MC – marginal cost of extraction,
- r – discount rate;
- p-MC – rent from the extraction

If a separate symbol R for the rent is introduced (R=p-MC), then the formula reads:

$$\mathbf{(dR/dt)/R = r.}$$

It can be stated in the following way: the rate of expected rent growth is equal to the discount rate. The mathematical derivation of the Hotelling rule requires using a fairly sophisticated theory of optimal control. Here we will simply analyse what happens if the equation $(dR/dt)/R=r$ does not hold.

If the equation $(dR/dt)/R=r$ does not hold, then there are two possibilities: either $(dR/dt)/R > r$, or $(dR/dt)/R < r$. In the first case, the owner expects that the rent will increase more than the discount rate. The decision he (or she) is likely to make is to postpone the extraction. The extracted (and sold) amount of x will subtract from the deposit and it will not give the increased rent $x(dR/dt)/R$. Instead the owner will be given xr. So he (or she) will be given less than if the resource was sold later, when the rent would be higher. In the second case, the owner expects that the rent will increase less than the discount rate. The decision he (or she) is likely to make is to extract as much as possible, sell it and invest the money in an alternative project. The extracted (and sold) amount of x will subtract from the deposit and it will not give the expected rent $x(dR/dt)/R$. Instead the owner will be given xr. So he (or she) will be given more than from keeping the resource and selling it later.

The only case that the resource owner does not have any incentive to postpone or to speed up extraction is when $(dR/dt)/R=r$, that is when the Hotelling rule is satisfied.

Although the Hotelling rule is the most important economic finding about exhaustible resources, it applies to a number of other cases too. For example, it explains what motivation art collectors may have when they buy or sell specific paintings. Derivation of the Hotelling rule is based on the assumption that the supply is limited, and it can only be lower (as a result of "extraction"). Rembrandt produced a large number of paintings (maybe 400 or so). From time to time we hear that a new picture was found, but this is a rare event. A more likely event is that the supply will disappear from the market (e.g. as a result of fire). That is why the Hotelling rule is used to analyse works of art prices. There are investors who buy Rembrandts not because they are crazy about the 17th century Dutch painting, but because they consider it a good investment.

The Hotelling rule can be used to analyse any market where the supply cannot grow. If a product is not manufactured any more, the supply of its spare parts comes from wholesalers. The Hotelling model explains how quickly the wholesalers are likely to sell what they have.

The Hotelling rule can be used to draw some surprising conclusions regarding market structure and environmental protection. For obvious reasons we do not like monopolies, and we prefer to have competitive markets, since monopolies try to elevate prices by constraining the supply. On the other hand, environmentalists voice their concerns about the growing scarcity of exhaustible resources. We think that it is better if something is to be exhausted in 200 years rather than 100 years. If the Hotelling rule logic is applied to a resource extracted in a monopolistic market, it turns out that competitiveness increases the supply, and thus speeds up its exhaustion.

There are insurmountable problems with empirical verification of the Hotelling rule. It was analysed in the context of raw material statistics many times, but it proved impossible to make any reliable price projections. Owners of exhaustible resources do not know how to predict the rents they are going to enjoy. For example, oil prices depend on the political situation in the Middle East more than on economic circumstances.

Economists analyse exhaustible resources with the help of the Hotelling rule. It turns out that discount rate has to be used in order to explain what motivation people have to postpone or speed up their extraction. In the case of renewable resources, their exhaustion is not inevitable. Nevertheless time dimension is important, since what is done in the present may have some impact on the future. We will see that discounting is unavoidable in this case as well.

Questions and answers to lecture 7

7.1 Are there any examples of non-depletable resources other than table salt?

To be considered a non-depletable one, a resource should be abundant so that its use by the economy does not decrease its future supply. Table salt is the only example referred to in environmental economic textbooks. Nevertheless, some other factors important for our well-being have similar characteristics too. Oxygen is perhaps another raw material we all use (and need in order to survive) whose supply is so abundant that policy makers do not have to worry about. Its concentration in the atmosphere is around 21%, and its quantity is huge. It is "produced" by plants, and used by all living organisms in the process of breathing. Occasionally there are concerns voiced that we may lose it if we destroy forests or something like that. However, it is not the problem of losing oxygen but rather the problem of carbon dioxide emission. A calculation was done what would have happened if the entire biosphere (including mankind) was burnt. The amount of oxygen would be decreased somewhat. Its atmospheric content would be lowered by the amount corresponding to what we are exposed to when we increase our elevation by 300 m. We would hardly sense any difference in its availability. The concentration of carbon dioxide would increase drastically, but this is another story. Oxygen is inevitable for our life, but we do not have to be concerned by the fact that its supply is threatened.

7.2 The South African firm De Beer is considered a monopolist in the diamond market, since it owns most of the deposits of diamonds in the world. What does the Hotelling rule say about the diamond extraction?

From the very beginning De Beer owners knew that in order to keep the price of diamonds high, they need to constrain their supply, and they did it. If they expected that $(dR/dt)/R < r$ (corresponding to the Hotelling rule recommendation to extract as much as possible), they did not have to speed up the extraction. They knew that in order to maximise their profit they did not have to adjust to the market price; they knew that by manipulating the supply they could change the price (and this is what they did). When the demand curve is sufficiently steep, if the quantity goes down by, say, 10%, the price may go up by much more, so that the revenue goes up. De Beer coined the phrase "A Diamond is Forever" (which is considered the best advertising slogan of the 20th century). Enjoying its monopolistic position, it drove the price of diamonds (and the rent from their extraction) to a higher value, than determined by (hypothetical) competitive circumstances.

7.3 If diamond extraction is so profitable, why are there only few firms active in this market?

Because unlike cars and computers diamond supply is constrained by natural factors (their production cannot be increased in response to market signals).

7.4 Some people buy and store wines, expecting that their price may increase over time. What is the right moment to sell the wine?

The quality of some wines improves with age. Consequently their prices go up. The cost of storing them is not zero. Even if you do not have to do anything, just the storing space costs money. But storing good wines requires temperature and light control. If you take into account all the costs, the marginal net benefits from storing the wine – while positive – are likely to go down. Once their annual rate equates with the level of your discount rate there is no sense in keeping your inventory; you should sell the wine. This is what the Hotelling rule logic suggests you to do.

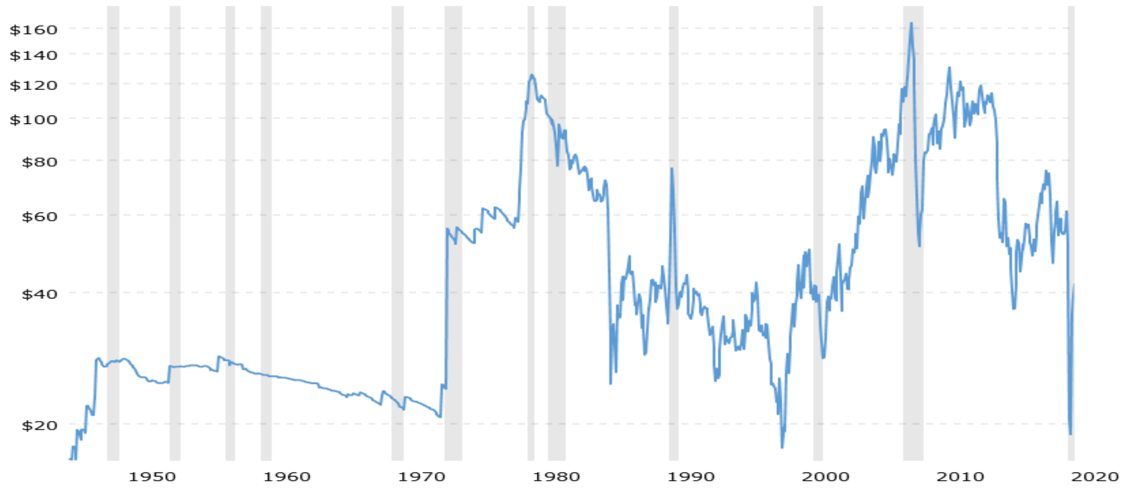
7.5 Coal is an example of an exhaustible resource. What are important factors which influence the expected rent from its extraction?

The price of coal depends on a number of its characteristics. It is related to the price of oil (if the oil becomes more expensive, the price of coal goes up too since they are close substitutes). It also depends on the coal contamination. The coal extracted in many places (in Poland too) is contaminated with sulphur. In some coal deposits the content of sulphur can be as large as 1% or more. When burnt, it pollutes the atmosphere with sulphur dioxide. Its abatement – mandatory for environmental protection reasons – makes the power plants' operations more expensive. Coal combustion leads to higher carbon dioxide emission (per unit of electricity generated) than oil and gas. In many countries carbon dioxide emission is regulated which makes fossil fuels – and especially coal – less attractive for power plants. In order to sustain the demand, fossil fuel producers sell their products at prices that take into account abatement cost. In particular coal prices reflect contamination with sulphur compounds. They also reflect climate protection measures. Besides, renewable energy sources become substitutes for fossil fuels. As a result, the rent from coal extraction is constrained severely.

7.6 What would be the Hotelling rule for a resource owned by a monopolist?

The Hotelling rule is the same, i.e. (dRent/dt)/Rent=r, except that the monopolist is not a price-taker. Consequently, a monopolistic firm does not have to adjust the extraction to the market price. It can make this price higher or lower in order to maximise its profit.

7.7 OPEC is the best known example of a raw material cartel. The chart below lists crude oil prices (2020 USD per barrel) from 1946 to 2020 (grey bars indicate recession periods). Can the Hotelling rule explain the dynamics?



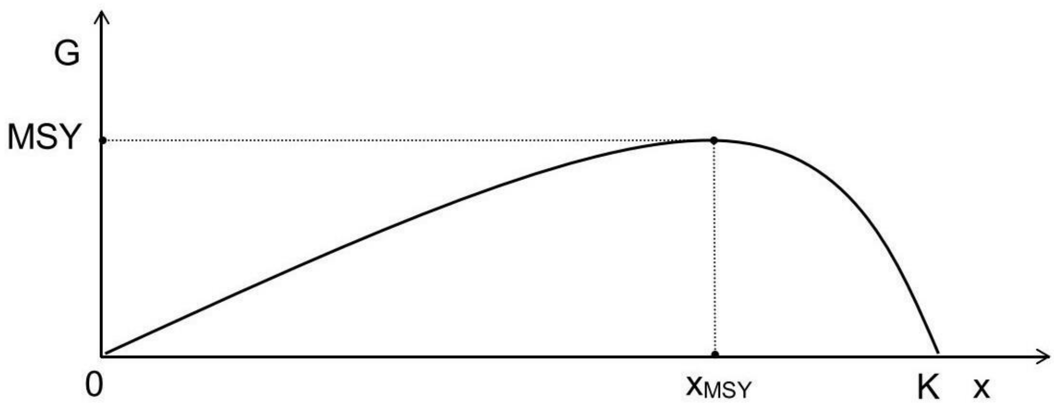
No. There are much more factors affecting oil prices than the Hotelling model takes into account.

OPEC (*Organisation of the Petroleum Exporting Countries*) is a cartel established in 1960. It had five founding members: Iran, Iraq, Kuwait, Saudi Arabia and Venezuela. Now there are 13 member countries. They control 70% of oil deposits, and 40% of oil extraction. They meet periodically in Vienna in order to negotiate and allocate extraction quota to maximise their profits. OPEC enjoys a monopolistic position, but despite its homogeneity (almost all members belong to the Islamic tradition), members are not quite loyal to each other. Having allocated extraction quota, they are notorious for exceeding them. As a result, the global supply of oil is usually higher than they planned, and the price cannot be sustained.

If you look at the chart you see the price volatility which is correlated with recession periods (something that can be consistent with the Hotelling model). Nevertheless, the most drastic price hikes accompany political events such like the Arab-Israeli wars in the 1970s. Hence the Hotelling rule cannot explain the oil price dynamics satisfactorily.

8. Economics of renewable resources (immobile resources)

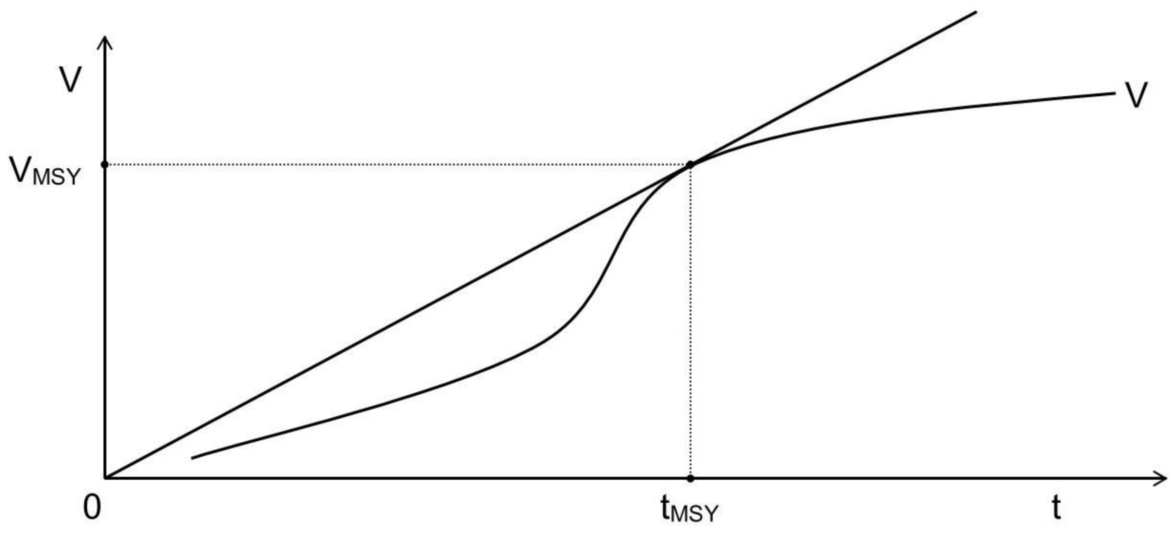
Maximum Sustainable Yield (MSY) is the most obvious concept that comes to our minds when we analyse hypothetical management regimes of renewable resources. Its idea is illustrated in the picture below. We refer to a story about a lake where fish can breed. The vertical axis measures the annual net increase of the fish population (the difference between the number of births and deaths). The horizontal axis measures the quantity of fish in the lake. This quantity is between 0 and K (it is customary to use the letter K, since *Kapazität* means the volume in German, that is the quantity of fish that the lake can accommodate; German economists started to study the problem earlier than others). Zero (0) corresponds to the lack of fish, and K – to their maximum quantity that the lake can accommodate. In both extreme cases the annual net increment is zero. In the first case it is zero, because there are no fish. In the second case it is zero because the number of births is the same as the number of deaths; if the number of births were higher than the number of deaths the population would grow, but – by definition – K is its maximum quantity, so the population cannot grow above K.



The graph above illustrates the relationship between the quantity of a resource (e.g. the number of fish in a lake), and its annual increment. If one wishes to harvest sustainably (i.e. only the annual increment, so that the stock is not depleted), and if one wishes to enjoy the highest possible yield, the stock should be kept at the level of x_{MSY}. This level is somewhere between 0 and K. Empirical research suggests that sometimes it can be found in the middle of this segment (i.e. x_{MSY}=K/2). But it can be located closer to 0 or closer to K (as in the picture). The graph resembles an inverted "U", perhaps not quite a symmetric one. The quantity x_{MSY} serves as a reference for various concepts of renewable resource management, but – as you will see – it is not an optimum solution in general.

Renewable resources can be either mobile or immobile. Principles of their management are different. As the name suggests, mobile resources can move, while immobile resources cannot. This has fundamental consequences for their users. If a user of a renewable resource does not harvest it now in order to save it for a future use, the resource will not "escape" if it is an immobile one. In contrast, if it is a "mobile" one, it can "escape" and it can be used by somebody else. Trees in a forest provide an example of an immobile resource, while fish in a lake – of a mobile one. This lecture will be on immobile resources.

It may not be clear how the picture on previous page 58 relates to forestry (the story was about fish in a lake). But the maximum sustainable yield concept is easy to grasp in this case too. In a typical Polish forest, the annual harvest is around 8 m^3 per hectare. It is close to the MSY quantity. If the harvest was lower, than the volume of timber stored in the forest would grow quickly. If the harvest was much higher, then the volume of timber would shrink. Interpreting the maximum carrying capacity (K) is more difficult. Everybody can envisage a hectare without any trees; this corresponds to x=0; but what is the maximum volume of timber that a hectare of the forest can accommodate? It depends on natural conditions, but it can be envisaged as what happens if no trees are removed. Old trees will die and fall, and new ones will germinate. The volume is likely to be high, but not growing; this is how K can be interpreted. Foresters never let the forest achieve this state; they cut the trees before these become very old and die naturally.

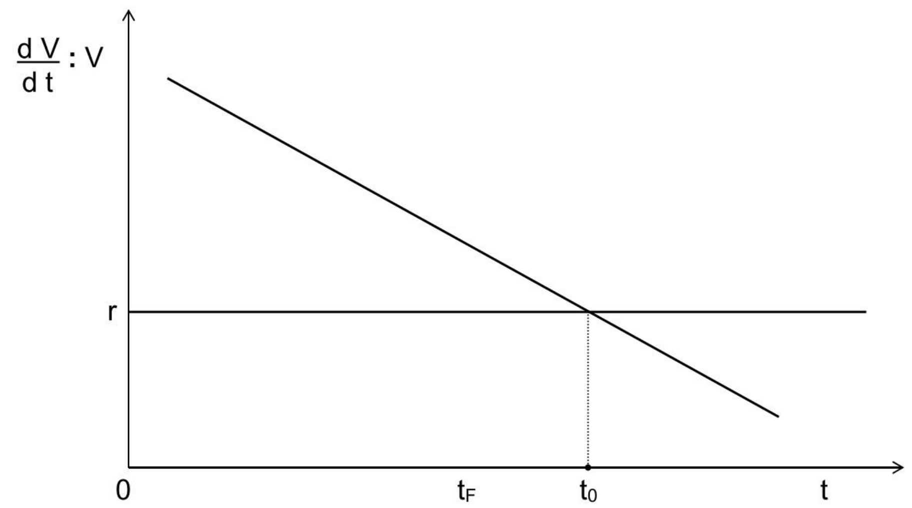


Let us look at the tree growth process. Initially its volume V (and the mass) is close to zero. After a while, it grows, and its annual increments increase. Yet once the tree gets larger, its volume does not grow as quickly as before. After some time it may not grow at all, or its growth rate becomes very low. The graph above illustrates this process.

The vertical axis measures the volume of a tree, and the horizontal axis – the time. The volume of a tree resembles a stretched letter "S". If you try to find the steepest straight line passing through the origin and touching the graph of V, it will be a tangent line like in the picture above. Let us denote this tangency point (t_{MSY},V_{MSY}). Symbols suggest that it has something to do with the maximum sustainable yield. We will analyse this in greater detail.

The moment t_{MSY} is when the average annual increment of the volume achieves its maximum value (the tangent line emerging from the origin (0,0) is the steepest); please check that any ratio V:t cannot be made higher than V_{MSY}:t_{MSY}. This means that at the moment t_{MSY} a tree has maximised its average annual growth. If the forester wishes to identify the moment such that a tree has maximised its average annual growth (from the time of planting the tree), t_{MSY} is what he (or she) looks for. The maximisation results from the arithmetic. Sustainability means that the forester plants the tree, waits until t_{MSY}, cuts the tree, and plants a new one. If he (or she) follows this schedule, then the management regime is sustainable. It is not a robber economy.

The argument above can be questioned. In reality, the forester is an investor. He (or she) bears the cost of planting a tree, and spends money on caring for it, but expects to get the money back once the tree has reached an appropriate value. When he (or she) decides to cut the tree, a comparison needs to be made – like in the Hotelling rule – what is more attractive: to sell the timber as it is, or to wait an additional year when the volume of the tree will be higher. The decision should result from comparing the expected growth rate of the value of a forest, and a discount rate. If the former is higher than the latter, then the forest should grow. However, if the former is lower than the latter, waiting is not attractive, and the trees should be cut immediately, right now. Martin Faustmann, a German economist, made this argument in 1848. Picture below illustrates it.



The downward sloping curve illustrates the rate of a tree volume growth, i.e. (dV/dt)/V. Initially it is higher than the discount rate r. In t_0 it is equal to r, and afterwards it becomes lower than r. Thus t_0 is the moment when the forester should cut the tree, since the expected increase in tree volume (dV/dt)/V becomes lower than the discount rate r (one needs to cut the tree and sell timber, because otherwise expected benefits will grow at the pace lower than r).

Martin Faustmann observed correctly that even when all the trees are cut, the barren land (without any trees) has some value. Thus if the forester wishes to maximise the profit, he (or she) should cut the trees somewhat earlier, namely in t_F rather than t_0.

Foresters call t_F "an optimum rotation period". It is understood as the time between planting a tree and its felling immediately followed by planting a new tree. Rotation periods depend on a number of natural and economic factors. For many conifers growing in Poland it is around 50 years. For broad-leafed trees it is somewhat longer. In both cases it is fairly short, if one takes into account the fact that older forests are more valuable from the point of view of their ecological and recreational characteristics. What incentives may foresters have to let the trees live longer than implied by their "optimum rotation periods"?

One can simply order the foresters not to cut the trees. If they comply, then they will not maximise their profits, but they will provide the economy with other benefits the society expects. They will prove to be altruistic by letting somebody else benefit while sacrificing their own benefits. Is it possible to achieve the same effect of allowing the old trees to live without referring to the altruism of foresters?

The answer is: yes. Without a government intervention foresters enjoy revenues by caring for the forest and felling the trees that achieved the rotation period. Yet the forest provides other additional benefits that the foresters cannot sell in the market. Perhaps the government – acting on behalf of the society – could "buy" from the foresters the unsaleable benefits (they are usually called *non-timber benefits*)? Actually this is what goes on in many countries. If a forester signs a contract to maintain the forest to satisfy certain criteria, the government pays a subsidy to compensate for losses caused by not cutting the trees that reached their rotation periods. Often foresters do not sign any contracts; forest owners simply have to comply with certain environmental protection regulations, and they are required to provide certain *non-timber benefits*. If the expected subsidies are paid, then there is a chance for forest protection. If the government fails to compensate foresters for old trees, it may happen that the foresters burn their forest in order to reclaim the control over the land which is owned by them. Experts say that most of the forest fires that plague some places in the world are deliberate arsons commissioned by frustrated forest owners.

Efficient management of renewable resources is difficult, but Faustmann solved the problem theoretically. His "optimum rotation period" is a reference for any form of forestry organisation in any country.

Nowadays forestry is often looked at in the context of climate protection. A popular wisdom says "trees produce oxygen"; forests are considered gigantic lungs letting people survive in the planet. Indeed the forest makes a wonderful ecosystem and deserves to be protected carefully. Nevertheless stories about "oxygen production" are not quite correct.

The forest stores carbon, and burning a hectare leads to an emission of carbon dioxide. An average hectare of a Polish forest stores roughly 260 m^3 of timber. Following chemistry experts, one can estimate that every cubic meter of timber stores 1,000 kg of carbon dioxide (not in the form of the gaseous CO_2, but in the form of C included in plant tissues). If the trees were burnt, then the carbon dioxide emission from an average hectare would have been 260 tons roughly. Forest fires happen, but they do not pose a catastrophic problem in Poland. A large part of timber – after felling – is sent to manufacturing, especially to furniture and paper industries. Even though carbon dioxide stored in the trees is not released into the air immediately, it will end up there anyway. Timber products will become a waste sometime in the future, and they will emit carbon dioxide. The only difference with the forest fire is that emission will be postponed.

Even in the absence of the human population forests could not transform carbon dioxide into oxygen without any constraints. Trees do not live forever. Thanks to photosynthesis they increase their mass for several decades (by taking atmospheric carbon dioxide and releasing oxygen). After some time this growth process weakens and levels off completely (see picture on page 59). Decay processes start to dominate and the mass of trees may even decrease. When tissues disintegrate they return carbon dioxide to the atmosphere. The dead wood moulders and becomes the source of nutrients for other organisms, but at the same time it emits carbon dioxide into the atmosphere.

Only a young forest absorbs carbon dioxide and produces oxygen. A mature forest – ecologists call it a climax ecosystem – is neutral for the carbon balance. Deforestation negatively contributes to this balance, while planting trees where there was no forest contributes positively.

It should be emphasised that there are two climate-related issues that forests contribute to. One is carbon dioxide balance and the second is carbon sequestration. As old trees may emit more carbon dioxide than they fix in photosynthesis, the long-run contribution of trees to carbon balance is neutral. Nevertheless their contribution to carbon sequestration is valuable. There is less carbon dioxide in the atmosphere, if carbon is stored (sequestrated) in the forest. The role of forests is even more important, because some carbon is stored in the soil. Therefore it will not be released immediately even once the forest is cut (or burnt).

From the point of view of environmental protection a young forest is not so valuable as an old one. And conversely: what we appreciate in old tree stands, i.e. hollows, rotten wood, landscape amenities etc., is not useful from the point of view of carbon dioxide balance. Hence addressing the climate change by afforestation cannot be reconciled with environmental protection easily. In the long time perspective what was captured by trees can be released into the atmosphere anyway. In the short run – when the forest is young and it contributes to the atmospheric carbon dioxide balance in a good way – it will not start to perform ecological functions that are essential for nature. As explained by Faustmann, it has to be cut down and substituted with newly planted trees before vital ecological processes start to develop.

Interest in climate change triggered speculations about a potential contribution of forestry. Protecting the global climate calls for a reduction in carbon dioxide emission. The emission comes from deforestation, among other things, but climate protection cannot succeed if the burning of fossil fuels continues.

The so-called Paris Agreement was arrived at in 2015 by signatories of the United Nations *Framework Convention on Climate Change*. The agreement suggests that carbon sequestration by forests is an important factor of climate protection. It is indeed. However, in developed countries – such as Poland – which adopted binding emission reduction commitments a long time ago, carbon sequestration by forests cannot substitute for abatement. Such countries will be expected to abate emission not by documenting land use changes, but by promoting low carbon technologies in manufacturing and transport.

Yet some people in Poland hope to point at planting trees as an excuse for not reducing carbon dioxide emissions from manufacturing and transport. Paris Agreement alludes to such a possibility, but its careful reading leaves no doubt that this may be relevant for countries that have not taken any binding commitments. It would be useful to quote the part of Article 5, where this possibility is mentioned.

> *1. Parties should take action to conserve and enhance, as appropriate, sinks and reservoirs of greenhouse gases as referred to in Article 4, paragraph 1 (d), of the Convention, including forests.*
>
> *2. Parties are encouraged to take action to implement and support, including through results-based payments, the existing framework as set out in related guidance and decisions already agreed under the Convention for: policy approaches and positive incentives for activities relating to reducing emissions from deforestation and forest degradation, and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries; and alternative policy approaches, such as joint mitigation and adaptation approaches for the integral and sustainable management of forests, while reaffirming the importance of incentivizing, as appropriate, non-carbon benefits associated with such approaches.*

In other words, the agreement encourages to assist developing countries to build infrastructure adequate for multifunction forestry management. It does not include any provisions for developed countries to substitute their commitments to abate carbon dioxide with any domestic activities aimed at improving forest management.

The Faustmann model serves as a reference for forestry management, but its logic is applicable to other activities – such as agriculture – where human efforts are combined with what nature gives us. The productivity of natural systems varies, and as a rule, it is highest in their initial development stages. Natural systems are subject to ecological succession leading to so-called climax ecosystems. These climax ecosystems do not grow; they live and they reveal a number of marvellous characteristics, but their mass and volume are approximately constant. The "optimum rotation period" aims at keeping ecosystems in early succession stages and preventing them from reaching the climax.

This is also what farmers do, except that in the agriculture the "optimum rotation period" is much shorter – usually one season. Unlike trees, grass and grain grow just a couple of months. It would be unprofitable to harvest them too early. Likewise, it would be unprofitable to wait with harvest until the next season (neither grass nor grain will increase their mass; they will evolve towards more natural ecosystems). If one was to replicate picture from page 60, the $(dV/dt)/V$ line would have been much steeper (crossing the r level by the end of the first season).

Questions and answers to lecture 8

8.1 If the Arctic Ocean can accommodate no more than 100,000 individuals of a given whale species, what would be their annual net increment if the quantity is exactly equal to this amount?

If the stock has exactly the quantity of the "maximum capacity" (K), the net annual increment is equal to zero (the number of births is exactly the same as the number of deaths).

8.2 The S-shaped curve on page 59 is tangent in (t_{MSY},V_{MSY}) to the straight line passing through the origin. In a small neighbourhood of t_{MSY}, does the volume of a tree grow faster or slower year after year?

If you look at this picture, you see that the curve corresponding to V (the volume of a tree) is concave in the neighbourhood of the tangency point. This means that the volume grows, but its pace goes down. The same conclusion can be arrived at without studying the picture. In the tangency point the annual increment is the highest. This could not happen when annual increments grow over time (if they grow, it would be advantageous to wait in order to make this average annual increment even larger).

8.3 The fact that the barren land (lacking trees) has some value implies that in the Faustmann model t_F is lower than t_0. Why?

The owner of the forest waits for the moment that the profit from selling timber is as high as possible. If the entire wealth consisted of timber only, it would be t_0. But the wealth is implied not only by the trees, but also by the land. Taking into account that the owner can sell the land too (even if the land does not have trees anymore), timber growth is not that attractive; one can cut trees even earlier – in t_F. If you prefer mathematics rather than intuition, you can apply the following argument (outlined in my overhead ERE-8-4). The owner of the forest (and the land) has two revenue sources: selling timber and selling land. The revenue is thus Z(t) = X+V(t)p, where X – the price of land (constant over time), and p – the price of timber (also constant over time). He (or she) should cut the trees and plant new ones when (dZ(t)/dt)/Z=r. The moment t_0 was calculated as the solution to a somewhat different equation: (dV(t)/dt)/V=r. We need to demonstrate that t_F (solving the equation with Z rather than V) is lower than t_0. We start by observing that (dV(t)/dt)/V-r=0 for t_0. Let us find a solution to (dZ(t)/dt)/Z-r=0. By using the definition of Z, that is X+V(t)p, and multiplying both sides of the latter equation by Z, we get dX/dt+pdV/dt-rZ=0. Please note that the first term is equal to 0 as the derivative of a constant with respect to t. The equation thus simplifies to pdV/dt-rZ=0. Dividing both sides into Vp, we get (dV/dt)/V-rZ/(Vp)=0. This is like the equation (dV/dt)/V-r=0 (solved by t_0), with rZ/(Vp) instead of r (the product of r and Z/(Vp) instead of r). Z/(Vp) is higher than 1, because Z>Vp. Hence the second equation is solved by t lower than t_0 (the solution of the first equation). If we denote this lower number by t_F the proof is complete.

8.4 Which forestry products can be commercially sold?

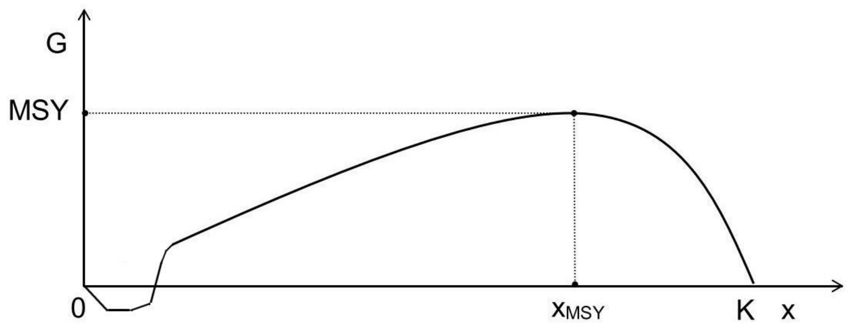
Private goods can be commercially sold, while public ones – cannot. The most obvious private good produced by foresters is timber. Foresters can sell it commercially, and they do it all over the world. Recreation is also a private good, and foresters can sell it commercially to some extent. In many places there are camping sites that charge admission fees. However, in some countries (especially in Northern Europe) there is so-called "Everyman's Right", or "right of public access to the wilderness", or "the right to roam". It means that forest owners cannot charge anything if people seek recreation. In some places it is regulated by a custom such that up to one or two days, anybody can enter the forest, pick berries, and even pitch a tent. More disruptive exploitation, such as hunting, making fires, and driving off-road vehicles is not permitted usually. Foresters are constrained in selling some private goods linked to recreation.

When it comes to public goods, such as stabilising water level, stabilising temperatures, sustaining biodiversity etc., commercial transactions are precluded because of free-riding. The most obvious way such goods can be financed is through budgetary subsidies. In some countries experiments are carried out with "payments for ecosystem services". Such payments are offered by environmental NGOs. They provide some opportunities for "buying" forestry products that cannot be commercially sold.

8.5 Can a renewable resource become an exhaustible one?

Yes. Any renewable resource can be exhausted if its harvest is higher than the net increment (the difference between what naturally adds to the stock and what naturally subtracts from the stock). A "reasonable" harvest (a concept referred to when discussing a sustainable yield) means that what we remove from the stock is – in principle – the net increment only. It does not mean that the harvest must be equal to the net increment always. If the stock is very large (it corresponds to a quantity between x_{MSY} and K on the page 58 picture), and if the harvest exceeds the net increment, the population stock goes down, but net increments go up; the harvest does not have to lead to extinction. It is important though, that when the x_{MSY} level is reached, subsequent harvests do not exceed the net increments any more. If they do, they may lead to extinction.

8.6 Empirical research demonstrates that in the neighbourhood of the origin (x=0) net increment curve may look not like on page 58, but rather like in the graph below.



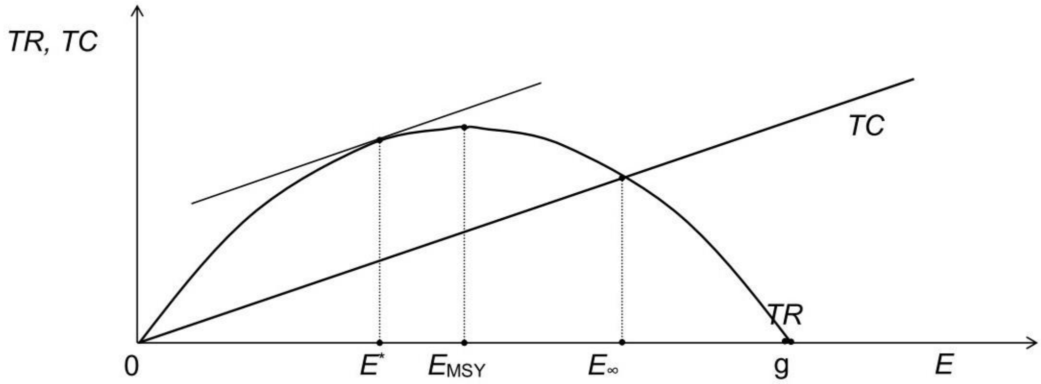
How can you interpret what happens when the stock is close to zero?

When the stock is close to zero, the net increment is negative which means that the number of deaths is higher than the number of births. In other words, the population is doomed to extinction, even though there are still individuals to breed. Some species are driven to extinction by killing the last living individual. This happened to *Aurochs* in Poland in the 17th century, or *Stellers Sea Cow* in the 18th century. They were killed by hunters. The last *Pyrenean Ibex* was killed by a falling tree in northern Spain in 2000. But a more typical pattern is when a species is driven to extinction gradually, once its population is so small that random events (like temporary food shortage, sickness, etc.) let the last surviving individuals die without an opportunity to have children.

9. Economics of renewable resources (mobile resources)

Mobile resource management is much more complex. In addition to the intertemporal dimension (present also in the management of immobile resources) one needs to take into account the fact, that unused resources can "escape" and thus they will not be available for the future use. Management solutions depend on assuming a specific harvest mechanism. The Gordon-Schaefer model will be used here since it does not require any sophisticated mathematical foundations. Its idea is explained in the picture below.

As before, economic variables – such as total costs (TC) and total revenues (TR) – are measured along the vertical axis. The horizontal axis contains physical quantities related to the harvest effort E. The latter can be measured by several methods. One way to capture the harvest effort is to quantify the fishing fleet (e.g. number of vessels, number of fishermen, capacity of the fishing gear, and so on). Another way would be to look at energy outlays such as fuel consumption. Yet another way would be to quantify salaries of fishermen. There are many ways to measure the harvest effort E.



In our picture it is assumed that harvests are sustainable, that is they are confined to catching the annual net increment only, and leaving the stock at the previous level. The effort is measured in terms of catching a part of this net annual increment (0%-100%). There are two extremes when, despite efforts, nothing can be caught sustainably. Zero effort (0) is when the stock is as large as possible (K); fishermen do not have to do anything in order to catch whatever is available, but nothing is available (the net increment x_K=0). The maximum effort (g) – i.e. trying to catch 100% of what is available – corresponds to the stock being extinct (0); fishermen try hard to catch whatever is available, but nothing is available, since the stock does not exist (x_0=0). There are some catches possible between these extremes. The TR curve resembles the one for MSY (page 58); like before, it is inverted "U"-shaped, except that the earlier picture had a physical quantity along the vertical axis (net annual increment), and the present one has the net increment multiplied by the price of fish. The shape is the mirror image of the curve for MSY (0 corresponds to K on page 58, and g corresponds to 0 on page 58), and it is measured in monetary units (harvest times unit price). The ray TC reflects the assumption that costs are proportional to the effort.

There are several levels of effort which correspond to specific outcomes of the Gordon-Schaefer model. They are characteristic for:

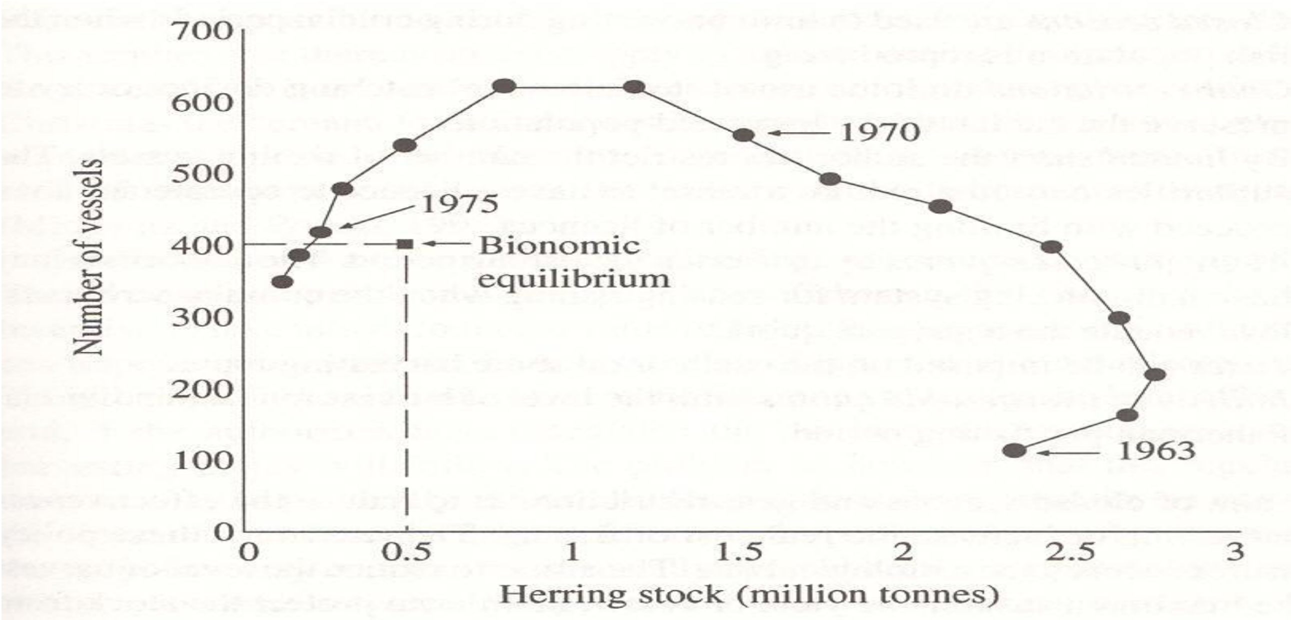
- E_∞ – bionomic equilibrium
- E_{MSY} – Maximum Sustainable Yield

- E^* – profit maximization

The name E_{MSY} sounds familiar. It corresponds to keeping the stock at the maximum sustainable yield (MSY) level, and harvesting the maximum net increment available. This does not maximise the profit though. The reason that it does not is that the cost is neglected. The difference between TR and TC is maximised, if the effort is E^* (not E_{MSY}). The result is obvious when looking at the graph, but a precise (mathematical) method of determining E^* is to find a point along TR where the tangent line is parallel to TC; please recall the argument from lecture 1 (overhead ERE-1-5, and also question 1.8).

E_∞ has an interesting interpretation. It is called "bionomic equilibrium", and it corresponds to a situation such that TR=TC. In other words, the total revenues are used to cover the cost of the harvest, and no profit is left. The high effort is accompanied by low level of stock resulting in low annual net increments. As demonstrated below, this is not just a theoretical curio; it can be observed in real life circumstances.

Picture below illustrates herring harvests in the North Sea. It would be difficult to think of a more popular species than a herring. Its North Sea population does not seem to be endangered, because the sea is not semi-enclosed (like the Baltic Sea), and it has a safe link to abundant resources of the Atlantic Ocean.



In 1963 the stock of herring was almost 2.5 million tonnes, and the number of vessels was around 100. Fishery business was apparently attractive as the number of vessels increased. The stock of herring grew somewhat. Everything seemed to be OK until 1965. After this date the number of vessels kept growing (it achieved 300 in 1966), but the herring population began to shrink. The number of vessels achieved its peak of 600 in 1971-1972, but the herring stock fell below 1 million tonnes. In 1975 the numbers were close to the bionomic equilibrium (calculated theoretically as on page 66). The North Sea herring became almost extinct, and fishing revenues of the 350 vessels could not finance their costs. Please note that the number of vessels was approximately equal to its bionomic equilibrium while the population of herring was close to zero. The stock survived only because of the harvest moratorium ordered by the European Commission. Moratorium was lifted in 1983, and fishing was resumed.

One can ask the question why would the fishermen continue fishing if TC is higher than TR. When TC-TR is a positive number, then the gain from fishing is negative. A firm which makes losses should not survive in the market. And yet, European fishermen survived even though the cost of doing business was higher than revenues. This was possible thanks to subsidies offered to them by governments. Governments are likely to subsidise unprofitable activities in order to fight unemployment (or other social) problems. As evident from the graph above, fishermen survived in the late 1970s because of subsidies. The harvest effort was excessive, and the population to be harvested was close to extinction. Supporting employment policy was anti-environmental.

The history of the North Sea herring fishery demonstrates that managing mobile renewable resources is a challenge. It is more complicated than in the case of immobile resources. Both differ from a typical environmental protection problem because of intertemporal choice. If benefits and costs occur simultaneously, decisions can be taken by comparing them directly. In the case of natural resources, simultaneity does not exist. In forestry the distance between expenditures and benefits can last many years. In fisheries the problem is even more difficult, but the time dimension is significant too: caching less today in order to catch more in the future is an important motive.

The MSY concept is used for both mobile and immobile resources. Nevertheless its relevance for the mobile ones is somewhat weaker since resources saved for the future can be used by somebody else. Several decades ago environmentalists claimed that extinctions are caused by open access. If renewable resources – such as whales – had a single owner, the owner would manage them sustainably precluding their extinction. In 1973, Colin Clark (an excellent mathematician and biologist) developed a model demonstrating that even a single owner may find it profitable to drive a population to extinction.

The model is very simple. Let us assume that a species, say, a population of whales has a single owner who wants to maximise the profit from its use. Let us assume further that the population is kept at the level providing maximum sustainable yield. For whales the annual net increment cannot be higher than g_{MSY}=2%. If the population is X, then what can be harvested sustainably is gX. If the price of the resource is p, then the annual sustainable yield is gXp. On the other hand, if the owner decides to harvest the entire population X rather than gX, then the (one-time) revenue will be Xp. This revenue can be invested giving the annual income of Xpr, where r is the discount rate used by the resource owner. Which income is higher: gXp (that is sustainable) or Xpr (that is a robber economy)? The question boils down to solving the inequality $gXp < Xpr$. Its both sides can be divided into Xp, and we are left with $g < r$. If a discount rate is, say, 4%, and the net growth rate of the resource is 2%, then MSY is less attractive than extinction.

The model is simplified. It does not take into account harvest costs (which can rise sharply if the surviving population stock is low), neither does it take into account that the price is likely to go down when the market supply is X rather than gX. Nevertheless it demonstrates that even a single owner may have an incentive to harvest more than allowed by sustainability constraints.

The main problem of mobile resources is to keep their harvest at the sustainable level. Unconstrained access is likely to result in overexploitation. The most effective regime involves setting a total allowable catch limit. While calculating the limit is possible, its enforcement is difficult. So-called poaching is persecuted everywhere, but authorities do not have sufficient capacities usually. Systems of renewable resource management aim at making users require more effort to harvest. Several instruments were developed in order make harvests more costly than in their absence. They include:

- licensing;
- protection periods; and
- fishing gear regulations

Licensing is perhaps the easiest way to make harvests more expensive. A typical license requires paying some fee. By elevating its price authorities can limit the demand for harvests arbitrarily. Nevertheless they hesitate to do so not only for social or political reasons. A high price of a license creates an incentive to poach. The price of a license does not have to confine to paying certain monetary amount. For instance, in Poland hunters are supposed to be involved in some activities (to pay "in kind" by working) in forests where the hunting is to take place.

Another popular instrument is to create protection periods. Typically these protection periods are established when the population to be harvested has a breeding season, and individuals are more vulnerable at that time. Yet long protection periods may result in some adverse economic consequences. For instance, in one aquifer, a fish population was protected for 364 days per year. That is fishermen were allowed to catch one day only. But the day was not announced in advance. The fish was fairly attractive, so there were powerful fleets ready to take advantage of the harvest. Vessels were equipped with sophisticated expensive electronic instruments in order to catch as much as possible in just one day. From the social point of view this is not justified, as protection could have been carried out cheaper.

Fishing gear regulations are popular as well. One of the best known instruments is to control the mesh size of nets. If the mesh size is very small, then animals cannot escape when they are caught. If the mesh size is large enough, then all the small individuals (especially the very young ones) can escape, and the net is filled with mature ones only. This policy may lead to unwanted effects such as the gradual elimination of large individuals from the population, but it has been implemented widely. In terrestrial ecosystems there are hunting gear regulations. For instance, it is usually prohibited to kill animals from helicopters or using sub-machine guns. The idea is to let the hunters exert more effort before killing the game animal.

Questions and answers to lecture 9

9.1 What condition is necessary to make E^* identical with E_{MSY} in the Gordon-Schaefer model?

By definition of MSY, the revenue from selling the catch is maximum, if the harvest effort is E_{MSY}. The revenue neither grows nor declines, so the tangent line must be horizontal. At the same time, E^* takes place if the tangent line to the TR curve is parallel to TC. If this horizontal tangent line is parallel to TC, then TC curve must be horizontal too. If it is to go through the origin (0,0) – like on page 66 – it must be identical with the horizontal axis (the cost of harvest is zero irrespective of the effort). If the TC curve does not go through (0,0) – that is the cost of the zero effort is not necessarily zero which is conceivable (irrespective of what fishermen do, the cost does not change) – its "horizontality" means that changing the effort does not change the total cost. This is the necessary condition for E^* being equal to E_{MSY} in the Gordon-Schaefer model. Otherwise E^* is on the left from E_{MSY}.

9.2 Is it possible that the harvest effort E is higher than that observed in the bionomic equilibrium in the Gordon-Schaefer model (E_∞)?

Yes. Picture on page 67 illustrates this situation. This is what happened in the North Sea in the 1970s. After the 1973 the herring stock was lower than in the bionomic equilibrium while the number of vessels was higher than estimated for this equilibrium. Hence fishing fleet incurred losses (TC>TR). This is hardly conceivable in the long run, unless negative profits are "corrected" by subsidies. This is what happened in fisheries. Governments subsidised fishermen and encouraged them to catch herring even though the herring population was shrinking. Subsidies were motivated socially. If fishermen left their business, they would qualify for even higher unemployment benefits. But if environmental protection was taken into account, this would have been a better solution.

9.3 Is bionomic equilibrium possible in the absence of government subsidies?

Yes, it is. If the access is open then fishermen may have incentives to harvest, hoping that others would leave the market. They are ready to incur a temporary loss in order to enjoy profits in the future. In the long run, however, it would be difficult to survive without a subsidy. Therefore in the long run bionomic equilibrium (and – even more so – the effort greater than E_∞) is difficult to envisage without an external subsidy.

9.4 If the TC curve is not linear, but upward sloping (irrespective of whether it is convex or concave), can the conclusion from 9.1 (E^* is on the left from E_{MSY}) change?

No. The conclusion that E^* is on the left from E_{MSY} results from the fact that the TR curve has the maximum for E_{MSY}, i.e. it is upward sloping to the left of it, and downward sloping to the right. The Lagrange theorem (see question 1.8) states that E^* is where the tangent to TR is parallel to the tangent to TC. If the latter is linear, then its tangent is identical with it. If it is not linear, but upward sloping, the tangent is upward sloping as well. Thus it can be parallel to the tangent to TR only where TR is upward sloping, that is to the left of E_{MSY}.

9.5 The extinction of whales is often linked to the open access (everybody could hunt whales). Does the Colin Clark's argument support this view?

No. Clark's calculations imply that whales' extinction is caused by their low breeding rate (lower than discount rate). Hence even if the access was not open (even if a single owner decided what to do with their stock), then profit maximisation would encourage him (or her) to push towards extinction. Catching all individuals, selling them, and investing the revenue elsewhere would be more attractive than sustainable harvests. Open access does not help the whales to survive, but even without it, their renewability would not be attractive, since their natural regeneration rate is lower than typical discount rates that people use.

9.6 Can the Clark's argument (see 9.5 above) be reconciled with the Gordon-Schaefer model?

No. The Gordon-Schaefer model is based on the assumption that only sustainable harvests are analysed. In contrast, in Clark's argument the owner of the resource compares the sustainable harvest to an alternative which is not sustainable (and finds that the latter is more attractive than the former).

9.7 The average weight of cod caught in the Baltic Sea declined from more than 20 kg to less than 10 kg over the last three decades. Can this be linked to fishing gear regulations?

Yes. Scientists contemplate several explanations of this fact, and the "fishing gear hypothesis" belongs to the plausible ones. In cod populations (like in human populations) individuals differ is size (weight). The average size can be fairly stable, since – to a large extent – it is a hereditary characteristics. Applied in the Baltic region fishing gear regulations require that mesh size is not too small, so that only large individuals can be caught. Consequently large individuals have a lower chance to breed and to have their genetic material reproduced. On the contrary, small individuals have a better chance to survive and reproduce. As a result, the next generations of cod will have a higher proportion of small individuals, and hence a lower average weight.

10. Welfare indices and the environmental resource base

We are used to measuring our welfare by GDP (*Gross Domestic Product*). If GDP shrinks, politicians are concerned, and if GDP grows, many people are happy, they do not ask difficult questions, and they trust that everything can be solved thanks to the growing affluence. This is not quite correct. Environmentalists say that "GDP counts what does not count, and it does not count what counts". To see what they mean, let us assume that there is an oil spill in the sea. Of course, this is a bad news, and our welfare is affected adversely. What will be the reaction of the GDP? Despite what some may suspect, the GDP may actually go up, as a result of higher detergent production (to dissolve the oily "skin" in the water), and higher salaries of rescue workers. Or let us assume that the quality of a forest improved. People can pick berries and enjoy recreation. Of course, this is a good news, since our welfare is affected positively. What will be the reaction of the GDP? The GDP will probably go down, because picking berries for immediate consumption (rather than for sale) does not contribute to it, while selling timber does. If the improved quality of the forest was caused by saving old trees, then the reaction of the GDP will be a negative one.

GDP was defined in the 1930s for a very specific reason. Namely politicians were looking for a statistical device to help them control business cycles. In this role, GDP performed well, since it provided information on transactions carried out in an economy. Its success was so spectacular, that many people started to believe that GDP informs about everything important. It does not. The following social accounting matrix lists important statistical categories. It covers seven "accounts" (households, labour, capital, production, abatement, savings / investment, environment). Rows reflect where a given "account's" revenues originate, while columns reflect how a given "account" spends its revenues. Traditionally defined GDP looks at the first row only.

Social Accounting Matrix

	Households	Labour	Capital	Production	Abatement	Savings/ Investment	Environment
Households		Wages	Profits				Rent
Labour				Wages	Wages		
Capital				Profits	Profits		
Production	Consumption				Inputs	Gross investment	
Abatement				Abatement			
Savings / Investment	Households' savings			Depreciation	Depreciation		
Environment	Environmental services			Environmental damages	Environmental improvements	Value of the net change of resources	

Disappointment caused by the failure of a traditionally defined GDP to reflect "what really counts" led to defining a number of alternative indices. They capture "what really counts" (sometimes), but – at the same time – they are highly controversial because of their arbitrary assumptions. A more appropriate way to proceed is to "green GDP". The following definition offers an index which tries to overcome the shortcomings of the traditional GDP. It will be called a "greened" Net National Product (NNP). The definition reads:

NNP =

Consumption of marketed goods
+ Public expenditures on consumption
+ The value of the net change of real capital
– Flow of environmental damages
+ The value of the net change of human capital
+ The value of the net increase in the environmental resource base (– if the net change is negative)

The terminology needs to be explained. First of all, what is the difference between "domestic" and "national"? Statistics are collected either according to the nationality of economic agents or according to where a transaction takes place. For instance if a German firm produces something in France, then it can be attributed either to Germany or to France. In the first case it contributes to the German GNP, and in the second case – to the French GDP.

Now, what is the difference between gross and net numbers? The idea of GNP or GDP is to account for whatever new was produced in an economy. The new production can be consumed and then its role is not controversial: it contributes to GNP or GDP. But the new production does not have to be consumed immediately; it can be invested to produce something in the future. Yet there is a problem here. It can substitute for something that was already before (and now needs to be replaced) or it adds to the capital. In either case it is a part of the "gross investment". But in the first case – as a "replacement" – it cannot be considered a "net investment". Only in the second case it is a "net investment".

In principle the difference between "gross" and "net" is fairly obvious. Actually it is not. When you buy a new computer, you need to ask the question whether you substitute for an old one which does not have any value, or you add to what is still working and serving your needs. The answer is not obvious. In every country there are accounting rules which define so-called depreciation rates for any physical assets owned by an economic agent. In Poland, for instance, a company car depreciates in 5 years. This means that a car that was bought for 15,000 € after 2 years has the accounting value of 9,000 € only. If the company buys another car at that moment – and thus makes an investment of 15,000 €, this expenditure is considered "gross investment", but the "net investment" is only 9,000 € (6,000 € is considered to replace what was depreciated).

Accounting rules are arbitrary, and they do not necessarily reflect whether investment substitutes something that has gone, or adds to what is still valuable. Theoretically net numbers are more meaningful than the gross ones. Nevertheless economists do not trust net statistics, and they prefer to work with gross ones which seem to be less arbitrary. Consequently GDP or GNP are more popular than NDP or NNP statistics.

Finally the question of capital. In economics it is defined as anything that can provide us with a revenue (either monetary or in kind). It has three main components:

- man-made (real) capital;
- human capital; and
- natural capital

Man-made capital is what was produced and what can be used by us whether directly or indirectly. Cars, computers, and buildings are examples of this type of capital.

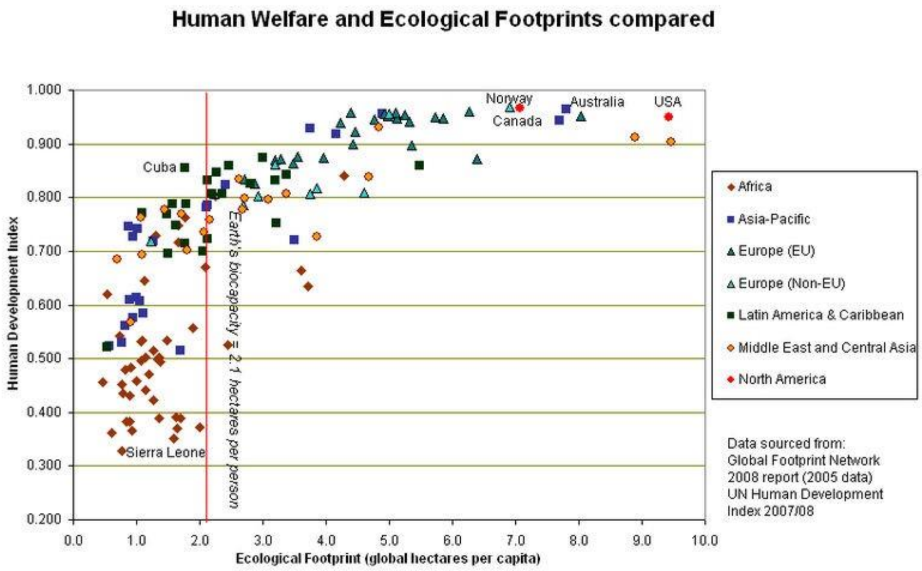
Human capital is what we have in our brains and our hearts. Engineering knowledge is an obvious example of it. But also trust and honesty belong to this category. Think of how our daily life would look if we did not trust each other. When we make shopping, the buyer gives the money to the counter clerk, and the seller gives the product we asked for. There is a moment such that the buyer got the product already and did not pay, or – conversely – the payment was done before getting the product. We do this, because we trust each other, and do not treat each other as thieves. If I suspect that the person I give the money to, or I get the money from, is a thief, then the transaction becomes much more cumbersome.

Natural capital is what we can use (directly or indirectly), but what was not necessarily produced by us. Trees in a forest or fish in a lake are examples of the natural capital.

Both investment and depreciation apply to all three types of capital. In the case of man-made capital they are obvious. Education is an example of investment in human capital. It can depreciate as well (what you learnt is less valuable now than it was when you acquired the knowledge). Also natural capital can increase (for instance, if we allow the fish to breed), or it can depreciate (for instance, if we deplete exhaustible resources).

The definition of a "greened" NNP includes our consumption (the first two items), subtracts flow of environmental damages (a sort of "negative consumption"), and adds net investment in all three types of capital (man-made, human, and natural). The traditionally defined (not "greened") NNP includes only the first three rows of the formula on page 72.

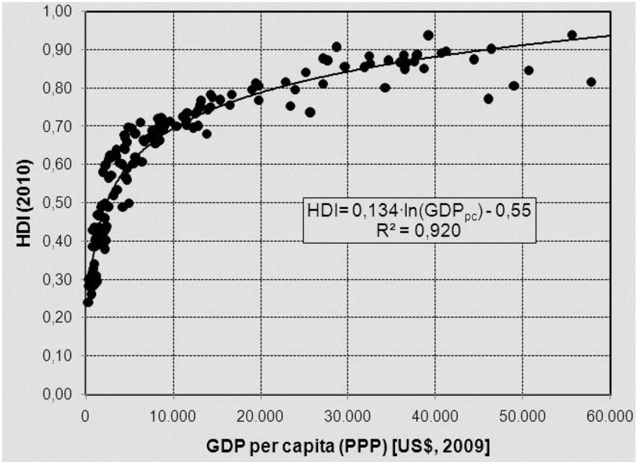
Human Development Index (HDI; see picture on page 74) is one of the most popular measures considered an alternative to GDP (or GNP). United Nations Organisation has used it in a number of documents. Its definition combines GDP per capita with a number of indicators aimed at reflecting what is not necessarily captured by GDP. For instance, it takes into account longevity, illiteracy rate, number of hospital beds, number of doctors, and so on. These additional indicators are arbitrary for two reasons. First, it cannot be determined objectively, what weight one should attach to, say, longevity: the same as to GDP or perhaps higher or lower? Second, the number of additional indicators cannot be determined objectively. For instance, the single number of doctors can be substituted with two numbers: the number of dentists and the rest separately. Or – depending on the age structure of the population – paediatricians or geriatrists can be assigned separate entries.



However, the most important objection to HDI is reflected in the picture on page 75. Two indicators are represented by each single dot in the graph: HDI and "ecological footprint" of a country. The latter reflects how many hectares one's consumption requires. In the case of potatoes consumed per year it is straightforward. In the case of clothes and cars it is more difficult to estimate, but it can be done at least in rough terms. If you divide the area of the land into the number of people in the world, you will get approximately 2 hectares per person. Citizens of the wealthiest countries like Norway and Unites States have "ecological footprints" several times higher. Citizens of poor countries, like Sierra Leone, have this indicator below 1 ha. As a rule, the higher the HDI, the higher the "ecological footprint". There are very few exceptions. Cuba uses less than 2 ha per head, but at the same time enjoys a very high HDI (almost like USA with the ecological footprint of almost 10 ha).

The most important (and embarrassing) conclusion from this picture is that – except for Cuba – high HDI numbers are correlated with excessive "ecological footprints". In other words, HDI numbers do not take into account environmental protection satisfactorily (and Cuba can be hardly considered a country many people would like to live in).

Picture below demonstrates an additional reason why HDI does not add much to what is captured by GDP. The two indicators are correlated rather strongly. The higher the GDP per capita, the higher the HDI is. This is not surprising, since if GDP per capita is high, the society can afford more doctors, lower the illiteracy rate, and let the people live longer. Additionally, the correlation is not surprising also because of the fact that GDP per capita is an element of HDI.



From the environmental point of view, GDP is not a satisfactory indicator. Nevertheless, it is probably better to "green" it rather than to use alternative numbers like HDI that are arbitrary. "Greening" has a very strong mathematical justification. In economic theory it is assumed that wealth is characterised by the so-called Bergson-Samuelson function of economic welfare $W(x_1,...,x_k,G)$, where the consumers are numbered from *1* to *k*, x_i is the consumption of the private good (perhaps different for every *i*), and G is the consumption of the public good (the same for every *i*). This hypothetical welfare function *W* can be approximated linearly (inspired by the first order Taylor series) as $p_1q_1+...+p_nq_n$, where $p_1,...,p_n$ stand for prices of the *n* goods and services produced in the economy, and $q_1,...,q_n$ denote their quantities. The "greened" NNP can be interpreted as an implementation of this formula. Please note that it does not introduce arbitrary numbers like longevity or something like that. It sticks to goods and services (q_j) that are provided to the people (not necessarily through the market). Please also note that prices (p_j) can be either market prices (especially for private goods) or other values (especially for public goods). The next lecture (overheads ERE-11) will be devoted to non-market good evaluation techniques.

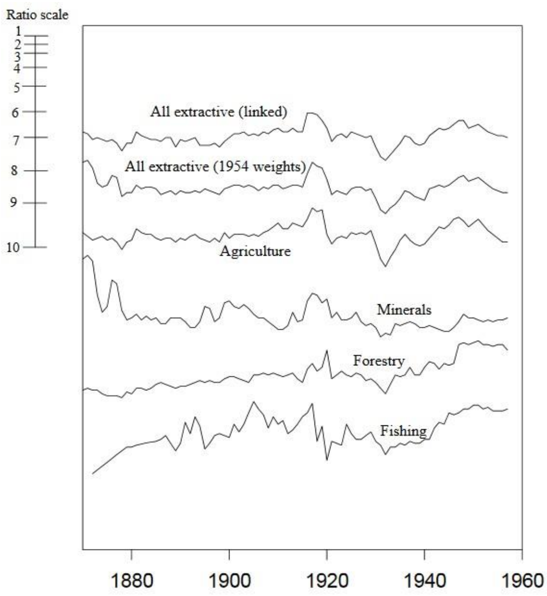
An appropriately defined NNP can be interpreted as a linear approximation of an unknown welfare function. The problem then boils down to what products to include and what prices to apply. One can further demonstrate that in the case of many goods pure market prices (net of subsidies and taxes) are sufficient. Only with respect to non-market goods – including environmental protection – alternative valuations need to be sought.

Thus an adequately "greened" NNP has a similar form to its traditional prototype except that certain goods are excluded from summation (if they do not contribute to welfare), certain goods are added (if they do contribute to welfare but they are neither sold nor bought in the market), and prices are not always market ones. The starting point is a "traditional" GNP which is the sum of consumption, savings and those environmental services that are consumed directly (not bought in the market). The "greened" NNP is then calculated by switching from gross to net values, i.e., instead of all savings one has to take into account only those which increased the value of capital, and from the value of environmental services one has to subtract environmental damages, i.e. the cost incurred in order to compensate for environmental amenities lost.

To sum up, the NNP "greened" in the way outlined above differs from the traditional one in three aspects. First, it includes direct consumption of environmental services. Second, it adds investment in natural resources or subtracts their depreciation. Third, it subtracts environmental damages. The revised indicator reflects changes in social welfare better than traditional GNP, but still this is just an approximation of the true value.

The "greened" NNP explains how environmental resources affect welfare. This is a special case of the problem what role environmental resources play in economic development. The question was asked (and answered) by economists already in the 18th century. They concluded that the environment was a crucial factor of economic development. At the same time, a question was asked whether the future economic development can rely on the availability of environmental resources forever. The answer was negative, but there were two different reasons identified. Thomas Malthus (1766-1834) emphasised the absolute barrier of environmental resource availability. In contrast, David Ricardo (1772-1823) emphasised the relative barrier of environmental resource availability.

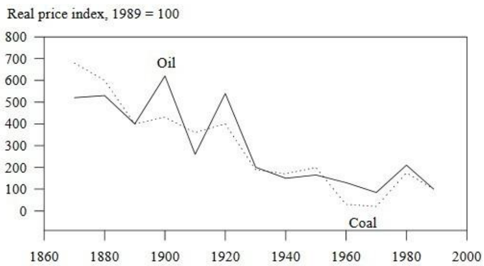
The Malthusian barrier is based on the concept of an absolute exhaustion of natural resources. Once this barrier is hit, further economic growth becomes impossible. The Ricardian barrier is based on the concept of gradual scarcity of natural resources manifested in their rising prices. According to Ricardo, even though the barrier will be approached slowly, ultimately it will make economic growth difficult (if possible at all). In the middle of the 19th century John Stuart Mill (1806-1873) suggested that both barriers can be alleviated by technological progress, and they do not have to stop economic development.



Until the middle of the 20th century, Mill's conclusion was accepted universally. In 1963 Harold Barnett and Chandler Morse checked empirically the hypothesis that natural resource scarcity was (or was not) observed in the market. They took time series from Britain and the United States, since they were the longest ones available. If Malthusian or Ricardian barriers existed, then they should be reflected in the rising costs of natural resource production. But no signs of rising costs were detected. Then Barnett and Morse tried to verify the hypothesis of rising relative costs. They suspected that there were no signs of cost increase, perhaps because – as a result of technological progress – all costs went down. If there was a growing scarcity of natural resources, then their costs would be declining more slowly than the overall cost index. But statistical tests did not detect any signs of such a trend either. Finally Barnett and

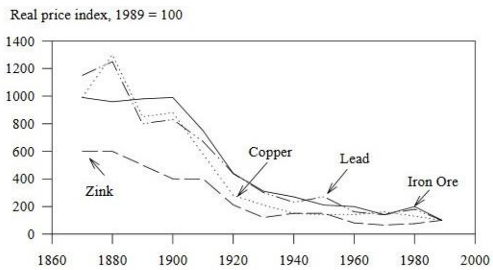
Morse assumed that even if the costs do not reveal any growing scarcity, perhaps market prices react to some scarcity symptoms sooner than costs do. Thus they looked at the relevant time series of relative prices (i.e. prices of natural resources deflated by the general price index). As before, no signs of scarcity were detected in geological resource markets (coal, iron ore, oil, and so on). However, they did detect some signs of scarcity in markets related to biological resources, such as fisheries and forestry in the United States (see picture on page 76).

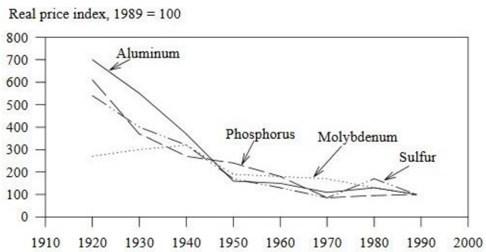
Barnett and Morse's research triggered lasting interest in their method of studying resource scarcity. There were numerous attempts to detect signs of scarcity by studying long-term time series. It turns out that if the time interval is modified, or quadratic trends rather than linear ones are looked at, resource scarcity is not an inexistent problem (as suggested by Mill).



For instance, long term fuel prices seem to go down. But if a quadratic trend is fitted, or if the entire period is divided into several parts, the last decades are different.

Similar pattern emerges when one looks at other raw materials. The second half of the 20th century does not resemble previous periods.





Environmental protection has an important impact on human welfare, and GDP does not reflect it satisfactorily. Rather than inventing new indicators to capture this impact, a better approach is to modify traditional measurements of economic performance. A "greening" strategy – as outlined in this lecture – defines a modified version of the relevant index. It accounts for environmental amenities enjoyed outside the market. As well, it subtracts environmental damages in order to reduce the impact of products purchased just to defend against these damages. In addition, it takes into account investment in (and depreciation of) three types of capital (not only the man-made capital).

Environmental resources drew economists' attention in the context of economic development. For several decades, John Stuart Mill's belief in the role of technological progress reduced economists' interest in the growing scarcity of environmental resources. The declining availability of living resources was observed first. Later on analysts detected signs of growing scarcity of other environmental resources.

Questions and answers to lecture 10

10.1 Why do critics say that GDP counts what does not count, and it does not count what counts?

This famous saying states that the definition of GDP is inadequate. It reflected (adequately) economic variables of interest to economists who, in the 1930s, wanted to detect and control business cycles. Environmental protection was not recognised as an important issue at that time. Thus GDP does not reflect things that are essential from this point of view. Moreover, it may react in a perverse way to what happens to the environment. If the environment is damaged (as in the case of an oil spill), it can grow as a result of the damage (contrary to what some people may suspect). On the contrary, environmental improvements can result in a decrease of GDP.

10.2 What elements that did not exist in the original definition of GDP show up in the Social Accounting Matrix?

Whatever is related to environmental resources. The matrix contains such entries as "abatement", "environmental damages", and "environmental improvements". Abatement was included in traditional GDP if it took the form of an investment expenditure (not a running production cost). Environmental damages were not accounted for at all. Even worse, they could result in GDP's growth. Environmental improvements were not accounted for either. Even worse, they could result in GDP's decline (if the improvement – like forest protection resulting in less trees being cut – was caused by stopping an activity that contributed to GDP).

10.3 Do all public expenditures contribute to GDP or NNP?

If the budget spends on investment it does contribute to GDP or NNP, and its contribution is already taken into account in investment numbers. Therefore, the definition on page 72 includes "Public expenditures on consumption" only. There is no risk of double counting with private consumption, since the latter includes only what is purchased by households in the market. If the public transport ticket costs 1 €, and additional 0.50 € is paid from the budget, then every ride is included in GDP; its value is 1€ (paid by the passenger) plus 0.50 € (paid by a public authority).

10.4 When you buy a desk to your room, do you make a gross or net investment?

It depends on whether the new desk adds to what you have or it substitutes for an old depreciated one. If it is a completely new item, then its total value of, say, 60 € should be considered a net and gross investment. If it is going to replace an identical worn out one, it is a gross investment (like before), but not a net one. If it is going to add to an old one whose value was depreciated by 40 €, then the net investment is only 20 €.

10.5 Human capital takes into account honesty (in addition to other characteristics). Can this be quantified at all?

My favourite example refers to protecting a bicycle against theft. Please think of appropriate protection in Warsaw, in Stockholm, and in Northern Norway. If you leave your bicycle on a Warsaw street, you need to protect it with a good padlock which is likely to cost 15 €. If you leave it in Stockholm, it has to be protected with a padlock, although not that strong; the padlock can be somewhat less sophisticated, perhaps worth some 8 € or so. If you travel in Northern Norway, you can leave your bicycle without any protection and you will find it untouched. Where do these differences come from? They can be explained by differences of human capital accumulated in an average inhabitant of Warsaw, Stockholm, and Northern Norway. Assuming that its level accumulated in Northern Norway is fairly high (the degree of honesty is high), then its deficit in Stockholm is reflected by the defence expenditure of 8 € (the cost of the padlock appropriate for Stockholm). Its deficit in Warsaw is reflected by the defence expenditure of 15 €. The padlock is an example of a man-made capital. In other words, human capital can be substituted by a man-made capital (whose valuation does not seem to be controversial). This can also be phrased in the following way. Accumulation of the human capital decreases the demand for man-made capital (economy can be used in order to satisfy other needs). Honesty can be quantified.

10.6 Will HDI react to environmental catastrophes such as oil spills?

It will react to oil spills in the same way as GDP. Non GDP components of HDI – such as illiteracy or hospital beds – do not decline as a result of an oil spill. At the same time, GDP can react to an oil spill in a perverse way (as explained in the lecture). Consequently HDI is not a good indicator of environmental protection.

10.7 How "ecological footprint" should take into account the consumption of imported goods?

In principle, the problem can be solved. If you eat imported potatoes, your ecological footprint should include hectares used for potatoes in the country of origin. Likewise, if you buy an imported T-shirt, your ecological footprint should include hectares used for producing cotton used for your T-shirt in the country of origin. There is a statistical problem though. Import statistics use different statistical classification than production statistics. Consequently – unless one works with extremely disaggregated categories – it is impossible to estimate the "ecological footprint" of imports to a country.

10.8 What is the idea of approximating a non-linear function with its Taylor series?

We have to quote one of the most famous theorems of mathematical analysis. It reads: If a real-valued function $f(x)$ is differentiable at the point $x = a$, then it has a linear approximation near this point. This means that there exists a function $h(x)$ such that $f(x) = f(a) + f'(a)(x-a) + h(x)(x-a)$, where $h(x) \to 0$ if $x \to a$. Often we take $a=0$. Then the theorem reads: If a real-valued function $f(x)$ is differentiable at the point $x = 0$, then it has a linear approximation near this point. This means that there exists a function $h(x)$ such that $f(x) = f(0) + f'(0)x + h(x)x$, where

$h(x)\to 0$ if $x\to 0$. Thus an approximation $f(x) \approx f(0) + f'(0)\, x$ is justified. Both $f(0)$ and $f'(0)$ are constants. Let us call them A and B, respectively. Hence the formula can be rewritten as $f(x) \approx A + Bx$.

So far the real-valued function f had a scalar variable ($x\in R$). The theorem works for vector variables ($\mathbf{x}\in R^n$) as well. The derivatives need to be understood as vectors of partial derivatives: $f'(\mathbf{x})=(\partial f(\mathbf{x})/\partial x_1,\ldots,\partial f(\mathbf{x})/\partial x_n)$ and the product $f'(\mathbf{0})\,\mathbf{x}$ as a scalar product of two vectors. Likewise **B** has to be understood as an n-dimensional vector.

The idea of looking at the "greened" NNP as an indicator of welfare is based on the Taylor approximation of an unknown welfare function.

10.9 What ideas is Thomas Malthus best known of?

An Essay on the Principle of Population (1798) is probably the best known work by Thomas Malthus. In this book he analysed the relationship between demography and economics. In particular, he observed that – because of agricultural production constraints – the availability of food grows slower than the number of people. As a result, the availability of food plays the role of a "mitigation" factor of the population growth. His concerns can be ridiculed, since everybody knows that agricultural production grew tremendously over the last two centuries. But Malthus predicted also that conflicts over the use of natural resources may result in hunger, wars, epidemics, etc. When one realises that the world is plagued by hunger, wars, and epidemics indeed, Malthusian ideas cannot be rejected easily.

10.10 The upper graph on page 76 (summarising the results of Barnett-Morse's research) refers to "extractive" industries. What sectors are included there?

"Extractive" industries include agriculture, mining, forestry and fishing. By "extractive" industries Barnett and Morse understand extraction of natural resources (both exhaustible and renewable). Manufacturing does not belong to "extractive" industries, because it produces goods that cannot be considered natural resources. The distinction between "extractive" and "non-extractive" is a difficult one (especially in agriculture), since many products contain a large component that cannot be considered "natural".

10.11 Real prices of many raw materials declined from the middle of the 19th century to the middle of the 20th century. They do not seem to decline now. Why such a difference between the past and the present?

In the past technological progress (as predicted by John Stuart Mill) could mitigate the growing scarcity effectively. Now that the resources are largely exhausted, mitigation must be much more expensive. Besides, human economy encroached on more natural ecosystems than before. As a result, even renewable resources (like forestry and fisheries) are more difficult to "harness".

11. Valuation of the environment

If I were to deliver this lecture 60 years ago, I would probably say that economics could not evaluate the environment, like anything which stays outside the market. Environmental resources do not have economic values. Perhaps they may have scientific, religious, patriotic, or sentimental values, but we are not ready to calculate their economic value.

Today we are ready to do this. Over the last several decades economics developed quite rapidly. Please recall what economics is about (page 1 or ERE-1-1 in my overheads). It is about how people make choices. And what are values? Values are proportions certain goods are exchanged for each other. Customarily one good plays the role of a *numéraire*, i.e. the role of money. Economists apply now the concept of so-called total economic value which captures what seemed to be outside of economic analyses. Its definition reads:

$$TEV = UV+NUV = DUV+IUV+EV+BV,$$

where

- TEV – *Total Economic Value,*
- UV – *Use Value,*
- NUV – *Non-Use* ("Passive Use") *Value,*
- DUV – *Direct Use Value,*
- IUV – *Indirect Use Value,*
- EV – *Existence Value,*
- BV – *Bequest Value*

The total economic value consists of two elements called use value (UV) and non-use value (NUV); the latter is sometimes called "passive use" value. For instance, swimming in a lake or just watching the water implies enjoying its use value. If someone derives satisfaction from the mere fact that the lake exists (and perhaps confirms this by paying membership fees in an organisation active in its protection), then he (or she) enjoys its non-use value. Both elements can be decomposed into more specific components.

Use value can be split into direct use value (DUV) and indirect use value (IUV). Benefits from recreation at the lake coast provide an example of the former. An example of the latter is the benefit enjoyed by a neighbouring farmer who has water in his (or her) well thanks to the lake.

Analyses of non-use value are more complicated. According to the formula above, it can be decomposed into two categories: existence value (EV), and bequest value (BV). If you ask why somebody cares for something that he (or she) will never see and will never use, you may conclude that its value results from the fact that perhaps it can be enjoyed by his (or her) inheritors. This is bequest value. Nevertheless even when the inheritance motive does not apply (the person does not have children and does not feel attached to any potential "successor"), the value is caused by the mere existence of a good. This is existence value.

Some time ago economists tried to identify "option" and "vicarious" values. The former was expected in a situation when no benefits can be traced at the moment, but they can be anticipated in the future. The latter was expected in a situation when a good does not provide any benefits for a person who nevertheless enjoys the fact that somebody else may benefit from it. A significant risk of double counting caused these concepts to disappear. Components of the total economic value are usually confined to the four that our formula lists.

In the past economists quarrelled about the origin of values. There were only two serious candidates: labour and energy. Proponents of the first theory claimed that the value of a good reflects the amount of the labour used in its production. There is a problem with this approach, since the value of a fish caught in a lake would correspond to the effort of casting the fishing line, but this contradicts what we see. Proponents of the second theory claimed that the value of a good reflects the amount of energy necessary for its production. There is a problem with this approach too. A herbivorous fish should be several times less valuable than a predatory one (please note that with every trophic level the amount of solar energy we rely on increases many times; as a result a herbivorous fish requires several times less solar energy than a predatory one), but – again – it contradicts what we see. Since the late 19th century economists have shared the opinion that the value of a good originates from its use rather than production. The more people are willing to sacrifice for it, the higher its value is. The value of a good reflects what people are willing to sacrifice in order to have it.

Any classification – even the most consistent and convincing one – fails to be a practical one unless it allows measurements. Price informs about the value of market goods. We know the value of a car with certain characteristics, of a kilogramme of apples, of one square meter in an apartment, because we see the prices. These can be somewhat misleading – especially if the market is not a competitive one, or if it is distorted by taxes and subsidies – but at least they inform about values. Non-market goods by definition do not show up in markets, so they do not have prices. Hence tracing their value is more difficult, but it is doable. Quite paradoxically, these non-existent markets serve as a reference for their value. It can be so-called surrogate market or a hypothetical market which could exist theoretically, even though it does not. Valuation techniques depend on what market they refer to. The following methods are used by economists:

Indirect methods (based on surrogate markets)

- Travel Cost (TC),
- Hedonic Price (HP).

Direct methods (based on hypothetical markets)

- Contingent Valuation (CV),
- Choice Experiment (CE).

The first group of methods originates from the idea suggested by Harold Hotelling (the same person who developed the rule for managing exhaustible resources; see lecture 7). One of the American national parks was threatened by a dam project. An investor figured out that the canyon protected in the park is an excellent site for an artificial reservoir to produce hydroelectricity. He convinced the relevant politicians and constituencies that this would be the best way to use the local natural resource. The park director was distraught. He asked economists for a help to find arguments against the project. In 1946 Hotelling suggested him to check the license plates of cars by the entrance to the park. If a local license plate is found, then there is nothing to be excited about. However, if there is a license plate from a distant location, it means that somebody exerted quite an effort to visit the park. The value he (or she) attaches to the park must be at least that high (assuming his/her rationality – an overall assumption adopted in economics).

The approach proved effective, and the park was saved. Modern methods based on the "travel cost" are much more sophisticated, and they are aimed at squeezing more information from surveying visitors. By the way, the method does not have to look at a "real" travel. For instance, we may be interested in the value of silence. This is not a market good: you can neither sell nor buy silence. But there is a market for noise-proof windows. If a person lives in a silent area then traditional windows can be installed. However, if a person lives in a noisy area then he (or she) is willing to pay an extra, say, 100 € to install noise-proof windows. This additional payment informs about the value of silence. In this case the "travel" means simply a financial effort undertaken in order to enjoy a non-market good.

So-called hedonic price method is based on the following idea. Let us assume that there are two identical houses, except that the first one is located in a noisy area, and the second one – in a silent area. It can be expected that the second one will have a higher price than the first one. The price difference informs about the value people attach to silence. Of course, statistical inference based on just two observations does not make sense. Hedonic price studies are usually based on large data sets, where individual observations are characterised by a number of variables, including a parameter informing about the non-market good of interest (e.g. whether the location is silent or noisy). Statisticians know how to interpret the outcomes of calculations in order to estimate the value of the good.

TC and HP are linked to surrogate markets. There is no market for the good of interest. However there is one for a closely related good; we call it the surrogate market. By studying the latter we can say something about the former. In the first example above, the national park was a non-market good, and the travel was a related market one. In the second example silence was a non-market good, and noise-proof windows were market ones. In the third example silence was a non-market good (again), and real estate market prices were analysed to shed light on the value of silence.

Economists prefer indirect methods, since the values of non-market goods are based on people's actual behaviour revealed in real market transactions. Unfortunately there are situations such that an appropriate surrogate market cannot be identified. I will refer to my unsuccessful attempts to find such a market for the Baltic Sea clean-up (in professional terminology: "reduced eutrophication"; do not worry if you do not know the word; eutrophication means excessive inflow of nutrients resulting in serious damages to marine ecosystems). In the 1990s the University of Warsaw was involved in an international project to estimate benefits from reduced eutrophication of the Baltic Sea.

Our first attempt was to look at real estate prices near the Baltic Sea coast. We expected that they were related to eutrophication, with less eutrophicated locations revealing higher prices. Biologists explained that eutrophication is something to be observed hundreds kilometres from the shore, and therefore it cannot be reflected in real estate prices.

Our next attempt was to look at the tourist traffic from Stockholm to Helsinki and back. There are large ferries commuting between the two cities every day. The annual number of passengers is roughly 10 million. We expected that if the eutrophication was reduced, the demand for travel would go up thus indicating to what extent people appreciate the improved quality of water. We had to abandon this approach once we realised that most passengers take advantage of cheap alcohol served on board, and they absolutely do not care about whether the water in the sea is clean or dirty.

Having failed to use tourist attractiveness as a proxy for benefits we were interested in, we planned to interview ship owners in order to determine to what extent they can save on maintenance costs if the sea is cleaned-up. It turned out, however, that these costs do not depend on the water quality.

Our last attempt was to interview fishermen in order to learn how much they would gain by catching fish in a less eutrophicated sea. We realised that, in fact, harvests in a eutrophicated sea can be higher than in a less eutrophicated one, so we had to give up again.

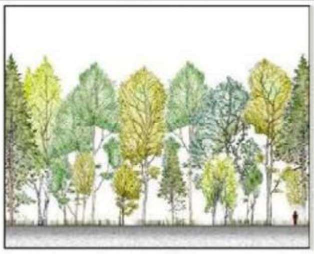
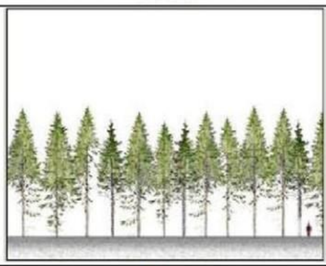
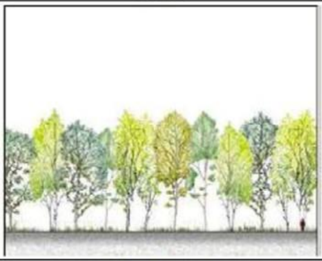
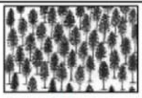
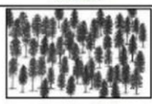
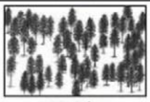
If no market can be easily related to the non-market good in question, economists rely on so-called "declared preference" methods. People are simply asked how much they would be willing to pay (WTP) to get something (e.g. reduced eutrophication of the Baltic Sea). If a simple question was asked "how much are you willing to pay for a less eutrophicated Baltic Sea", then some people would say \$2, some would say \$200, and most would not give any answer probably. The results of such a survey would be useless.

There has been a tremendous progress over the last fifty years in developing methods to ask WTP (*Willingness To Pay*) questions in a way which reduces the risk of getting random or misleading answers. In particular, economists developed methodologies to ask OE (*Open Ended*) or DC (*Dichotomous Choice*) questions. In the former case the question asked reads "how much are you willing to pay". In the latter: "you are asked to pay x dollars; do you agree? – please, say 'yes' or 'no'". A typical survey is prepared in a series of steps, starting with a 'Pilot' questionnaire, and ultimately followed by the 'Main' one. Typical surveys used to be face-to-face ones; now they rely on computers. Mail surveys were cheaper, but they were affected by the so-called "self-selection bias" (people who have particularly strong feelings about the topic of the survey are more likely to respond). An additional problem is linked with responses WTP=0. Some people are willing to pay 0 indeed; they are called legitimate "zero bidders". Others, however, state zero willingness to pay not because they mean it, but just because they want to protest against the survey (for whatever reason); they are called "protest bidders". There is a methodology of how to recognise who is a legitimate "zero bidder" (whose answer should be taken into account) and who is a "protest bidder" (whose answer should be omitted).

The name contingent valuation (CV) refers to the fact that before being asked about WTP for the non-market good in question, respondents are informed about a specific plan (scenario) of how this good can be provided, and why it is not free. Respondents state their WTP if they believe that the plan is credible. That is why the method is called "contingent valuation" (the valuation is "conditional"; namely the condition is that the plan is considered credible).

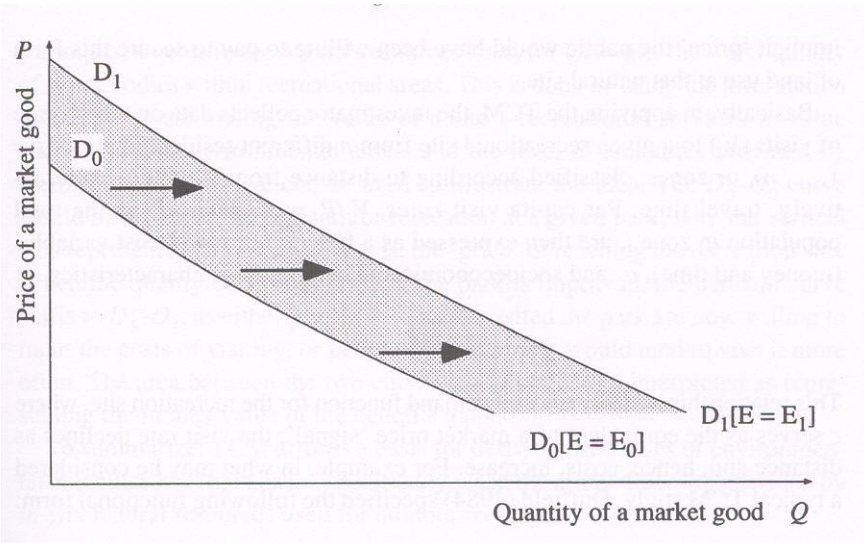
CV surveys were extremely popular in the 1980s and 1990s, but they were fairly expensive. The survey of 1,000 respondents (this was a standard sample to check WTP for a typical non-market good) required the expenditure of 10,000 € or more. The result was just one number, without any possibility to check how variations in the level of the provision of the good change WTP for it.

Marketing specialists have developed a method to survey potential clients about their WTP for certain changes in a series of choices where respondents are asked what their decisions would be if they were given certain choices. Environmental economists started to use this method widely some two decades ago. It has become the most popular method of soliciting people's WTP for environmental protection. The following picture illustrates how such a "choice experiment" (CE) looks. The aim of the survey (that this picture was taken from) was to determine people's preferences with respect to various forest characteristics.

	Forest 1	Forest 2	Forest 3	None
				
Forest type	Mixed	Coniferous	Broad-leaved	
Number of species	5	1	4	
Stand age	100 years	70 years	70 years	
Variation in tree age	Two-aged	Even-aged	Even-aged	
Ground vegetation	Medium	Absent	Medium	
Tree spacing	 Regular	 Quasi-regular	 Irregular	
Forest edges	 Straight edges & no ecotone zone	 Convoluted edges & ecotone zone	 Convoluted edges & no ecotone zone	
Distance	5 km	30 km	15 km	
Your choice	○	○	○	○

Respondents are asked to mark the bottom row ("your choice") in the column which is the most preferred one. If nothing seems preferable, he (or she) should mark the last cell ("none"). The table presents three types of the forest characterised by some words and icons. Under the icons there are 8 lines informing about the attributes of the forests. The first line refers to the forest type (in this table it can be coniferous, broad-leaved or mixed). The second line informs about the number of species; the first forest has 5 species, the second – 1, and the third – 4. The third line tells about the age of the forest (100, 70 and 70 years). The fourth line informs about variations in tree age (two-aged, even-aged, and even-aged). The fifth line tells about the ground vegetation (medium, absent, and medium). The sixth attribute refers to tree spacing; it can be regular, quasi-regular, or irregular. The seventh line characterises forest edges. These can be either straight or convoluted, and they can either contain ecotones or not (ecotone is a transitional ecosystem with certain characteristics taken from ecosystems it borders with). Finally, the last (eighth) attribute informs about the distance one needs to go in order to visit the forest.

If the respondent marked the column of the "Forest 2", it means that he (or she) would prefer to go 30 km in order to visit this forest, even though "Forest 1" was just 5 km away, and "Forest 3" was 15 km away. In a typical survey, respondents are shown several or even dozens such tables and asked to identify the preferred column. If a column displays all attributes being better than what can be found in the neighbouring ones, the respondent would mark this column obviously, and choice would be predictable (this is like choosing between being stupid, old, poor and ugly and between being clever, young, wealthy and pretty; of course, everybody prefers the second option). But if the combinations of various attributes are determined in a smart way (more attractive levels of some attributes are accompanied by less attractive levels of other attributes), then computer programmes can discover the respondent's preferences with respect to various characteristics of the analysed good (e.g. with respect to a forest characterised by several attributes) quite accurately. The second to the last row always contains the (monetary) cost of the option presented in a given column. In our survey we abandoned monetary valuation, since we assumed that a distance (measured in kilometres) is a better understood "cost" necessary to bear in order to enjoy a given forest.



The overall idea of valuation of non-market goods (especially by referring to surrogate markets) can be summarised in the picture above. There are two demand schedules for a market good (say, real estate, or travel to a national park): D_0 and D_1. The former reflects a worse state of the environment ($E=E_0$), while the latter – a better one ($E=E_1$). The area between the two curves can be interpreted as the welfare improvement caused by improving the state of the environment.

There have been thousands of studies trying to capture the value of the environment. The best known summary was carried out by Robert Costanza (and his students in the University of Berkeley), and published in 1997. It covers the entire world, and gives the average global value of annual ecosystem services. You will find its summary in the table below.

- Numbers in the body of the table are given in 1994 \$ ha^{-1} $year^{-1}$,
- Row and column totals are in 10^9 \$ ha^{-1} $year^{-1}$,
- Column totals are the sum of the products of the \$ per ha services and the area of each biome, not the sum of the per ha services themselves,
- Empty cells indicate services that do not occur, are known to be negligible, or no data are available for.

10^9 means the American billion, but a 'milliard' in some other countries. Thus the total value – the number from the lower right of the table – is 33 trillion USD (in USA) or 33 billion USD (in Europe) depending on who reads. The number was heavily criticised as a non-credible one, since it is almost what the world produces. This criticism does not seem to be correct, since the global GDP reflects what flows through global markets. Goods and services provided by the natural environment do not flow through the market often. Hence they do not have to be related to the global GDP in whatever way. A large part of the agricultural production (including column #13) flows through the market, but other "services" – like pollination (column #10) – do not.

The table is based on 17 categories of benefits provided by the world ecosystems, such as erosion control, raw material production, and recreation. Habitats providing these "services" were grouped into 16 categories called biomes. Some of these – like boreal forest and tropical forest – are summed up into broader categories – like forest in general – so that the total number of identified biomes is 22. The table could contain more than 300 numbers. Actually it contains much less, since some cells were left empty (either because appropriate publications were not found at the time of Robert Costanza's project, or because a given biome does not provide a given service).

Some analysts criticised Costanza for an alleged sacrilege: nature should not be sullied by monetary valuations. This is not a justified objection. Despite the fact that nature is indispensable, and sometimes we do not understand its logic, we do make choices which affect it. Therefore we should try to evaluate it in order to check if our choices are reasonable. The most interesting discovery would be to see that they are not.

Many analysts have observed that numbers from the table – as any numbers estimating the value of nature – are not reliable. Indeed they are not. But rather than concluding that nothing is certain in the world, it is better to estimate what emerges from people's decisions and declarations. In particular, if their consequences are bad or strange, then questions have to be asked, what policy instruments (something we are going to address in the next lectures – especially in ERE-13, and ERE-14) can be applied in order to solve problems.

The numbers from the table are neither perfect nor accurate, but they have been acknowledged as the first attempt to quantify economic values related to the global environmental protection. Costanza repeated his survey, and quantified the value of the same "services" almost two decades later. The results are interesting. First of all the total turned out to be much higher. What we receive from nature is worth more than what our economies produce. Secondly, the sum would have been even higher, but the area of the most valuable biomes shrank. For instance the area of tropical forest decreased from 1.9 billion ha to 1.3 billion ha. At the same time the area of deserts increased from 1.9 billion ha to 2.2 billion ha.

The data from the table need to be treated with caution. These are global average numbers. In a specific place they can be quite different. For instance the table quotes 50 $/ha as the average food production in a boreal forest (e.g. the forest that exists in Poland). Of course people hunt in Polish forests and the meat of the animals shot is eaten. As well mushrooms and berries are picked. Research carried out in the University of Warsaw shows that the number is much less than 50 $/ha. Perhaps elsewhere in the world people rely on food produced in the forest to a larger extent. On the other hand, the annual value of forest recreation Costanza estimates at 36 $/ha. The Warsaw University research (using the TC method, among other things) estimates its value in Poland much higher – perhaps even more than 2000 PLN/ha. International comparisons suggest that Poles appreciate forest visits much higher than in other countries. The table takes the average from societies characterised by diversified preferences.

Even though the numbers from the table are sometimes inconsistent with what has been estimated in some places, they play the role of an important reference. When other results are confronted with Costanza's indicators, researchers are expected to explain why their estimates

Biome	Area ($ha \times 10^6$)	Ecosystem serivces (1994 US\$ ha^{-1} yr^{-1}) 1 Gas regulation	2 Climate regulation	3 Disturbance regulation	4 Water regulation	5 Water supply	6 Erosion control	7 Soil formation	8 Nutrient cycling	9 Waste treatment	10 Pollination	11 Biological control	12 Habitat/ refugia	13 Food production	14 Raw materials	15 Genetic resources	16 Recreation	17 Cultural	Total value per ha (\$ ha^{-1} yr^{-1})	Total global flow value (\$ yr^{-1} $\times 10^9$)
Marine	36,302																		577	20,949
Open ocean	33,200	38							118			5		15	0			76	252	8,381
Coastal	3,102			88					3,677			38	8	93	4		82	62	4,052	12,568
Estuaries	180			567					21,100			78	131	521	25		381	29	22,832	4,110
Seagrass/ algae beds	200								19,002						2				19,004	3,801
Coral reefs	62			2,750						58		5	7	220	27		3,008	1	6,075	375
Shelf	2,660								1,431			39		68	2			70	1,610	4,283
Terrestrial	15,323																		804	12,319
Forest	4,855		141	2	2	3	96	10	361	87		2		43	138	16	66	2	969	4,706
Tropical	1,900		223	5	6	8	245	10	922	87				32	315	41	112	2	2,007	3,813
Temperate/boreal	2,955		88		0			10		87		4		50	25		36	2	302	894
Grass/rangelands	3,898	7	0		3		29	1		87	25	23		67		0	2		232	906
Wetlands	330	133		4,539	15	3,800				4,177			304	256	106		574	881	14,785	4,879
Tidal marsh/ mangroves	165			1,839						6,696			169	466	162		658		9,990	1,648
Swamps/ floodplains	165	265		7,240	30	7,600				1,659			439	47	49		491	1,761	19,580	3,231
Lakes/rivers	200				5,445	2,117				665				41			230		8,498	1,700
Desert	1,925																			
Tundra	743																			
Ice/rock	1,640																			
Cropland	1,400										14	24		54					92	128
Urban	332																			
Total	51,625	1,341	684	1,779	1,115	1,692	576	53	17,075	2,277	117	417	124	1,386	721	79	815	3,015		33,268

turn out to be higher or lower. Despite wide variations, some results turn out to be similar in most analyses. For instance the value of "services" provided by a hectare of the tropical forest is much larger than what can be found in the boreal forest. Likewise the value of "services" provided by a hectare of a meadow is a small fraction of what a hectare of a wetland provides.

Questions and answers to lecture 11

11.1 Does the value of nature confine to its economic value?

No. Nature – like everything else – can be assessed according to many criteria: in emotional terms, in religious terms, in patriotic terms, in ecological terms, and so on. According to some of them its value can be considered infinite. Yet economic values are basically finite; it does not happen that we are willing to sacrifice everything in order to secure something. Therefore nature too has a finite economic value. People will not sacrifice everything in order to protect nature. By actual decisions (choices) they reveal how much they are willing to sacrifice in order to save its part or to restore it. Economic values indicate how much people are willing to devote to it. This does not imply that other values (emotional, religious, patriotic, ecological etc.) are unimportant.

11.2 Why does economic theory acknowledge an economic value of natural resources even if their availability does not require making any effort?

Labour theory of value encountered insurmountable problems with many goods, and especially with natural resources. If their value depends on labour alone, then fruits do not have any value since they do not need anything to grow – just the solar energy and insects involved in the pollination process. Several generations of economists lean on conviction that the value does not originate in production; it originates in the course of using (consuming) the fruits. Hence irrespective of how much labour its production required, the value of a good depends on whether it is useful for somebody. If natural resources originate without any economic activity, they can be valuable anyway. In reality, they are useful not only because of natural processes, but also because of labour. Yet their value depends not on where they come from, but on what can be achieved by using them.

11.3 Why was the option value rejected as a result of concerns about double counting?

Let us assume that botanists identified a species that provides no benefits at the moment. Nevertheless it may provide some benefits in the future. Its value will be higher if it finds some attractive application. "Option" value suggests that – for the time being – the species is less valuable than in the future. Its assessment requires a reference to its NPV (see lecture 6 on discounting). Thus, if we expect that a species may provide some benefit in, say, 20 years, then these future streams have to be reflected in NPV. Consequently the value of this (so far) useless species can be captured in its TEV without the concept of the "option" value. An attempt to identify it separately may result in double counting.

11.4 How can a distinction be made between the direct use value (DUV) and the existence value (EV) of a specific forest for a given person?

I am afraid that this is a difficult task. If someone is willing to pay for a forest a specific amount, the payment includes both a DUV part (benefits from visiting the forest), and an EV part (benefits from the mere fact that this forest exists). There are some methods to split TEV into smaller parts (including DUV and EV), but they are not quite convincing. Often DUV is estimated by the travel cost method, and EV – by the contingent valuation method or by a choice experiment. If I were surveyed then – even though I am fairly competent in environmental economics and I appreciate forests – I would have a problem in separating components of NUV. In particular, I would have a problem to separate EV from BV. Components of TEV are defined theoretically on page 81 (ERE-11-1 in overheads), but practical measurement problems do not allow their decomposition in many cases.

11.5 In a hedonic price (HP) research of real estate prices carried out in the University of Warsaw, it turned out that proximity of green areas had a positive impact on apartment prices, but it had a negative impact on prices of detached houses. Can you interpret this?

This was not a statistical error, but a rational response of real estate agents. We like green areas because of many reasons. First of all, we like to look at them, and we like to visit them. Besides we prefer them over a dusty street. Yet their impact can be a negative one too. Proximity of playgrounds can be noisy. We are also afraid of the presence of unwanted visitors. Nevertheless positive aspects prevail usually. Proximity of green areas should have a positive impact on the prices of real estate, and our analysis of apartment houses confirmed this. The case of detached houses is more complex. The proximity of green areas is probably appreciated by inhabitants too, but there are also other circumstances. Firstly, inhabitants of a detached house have their own greenery. Every detached house is located in a plot of land covered with grass, flowers, shrubs and perhaps even trees. Hence the proximity of an external green area is not so attractive as for people who live in an apartment house. Secondly, inhabitants of apartment houses – unless they live in a base floor level – do not have to be afraid of bad people who may hide themselves somewhere in the shrubs. People living in detached houses are exposed to such unpleasant circumstances. That is why the proximity of green areas may result in a lower price of a house.

11.6 In a hedonic price (HP) research of real estate prices carried out in the University of Warsaw, it turned out that proximity of low emission sources (low emission – as opposed to high emission – is much more dangerous for human health) had a positive impact on prices. Do not people understand how dangerous is low emission?

This was the reaction of my students who participated in this research project. They kept saying that people are stupid. If they were more clever, they would be aware of the fact that a high stack is not so dangerous as the low stack of a local heating plant (not to speak of chimneys connected to small stoves). I told them that perhaps people are stupid and they do not know what those more clever do, but an economist should not make such an *a priori* assumption. He (or she) should assume that people are rational. Only when beyond any doubt this assumption has to be rejected, an economist may admit that – for whatever reason – people take irrational decisions. We solved the puzzle once we received more specific information from environmental administrators about low emission sources. A low emission source turns out to be a stack of a local heating plant (something we expected), but also an air conditioning system of a shopping mall (something we were not aware of). The air conditioning systems do not emit pollutants, but only the air we use for breathing. They are localised where we like to make our shopping. Thus their positive impact on housing prices does not result from people's stupidity and their willingness to pay more in order to live in a polluted environment, but from their rationality and their willingness to pay more in order to live where the shopping is easier.

11.7 In a choice experiment (CE) research of people's preferences with respect to forest recreation (using a different table than that shown on page 85, and in ERE-11-3) carried out in the University of Warsaw, it turned out that – *caeteris paribus* – people prefer the forest with a moderate amount of dead wood over the forest where the dead wood is removed immediately. Some foresters did not like this conclusion. Why?

The research led to the conclusion that people appreciate forest with a moderate amount of dead wood and do not like it when the amount of dead wood is very large (respondents were informed about what 'moderate' or 'large' means in this context). There is too little dead wood in the Polish forests. As the State Council of Nature Protection concludes:

> *Poland contrasts with the rest of Europe and North America as far as dead wood is concerned. Its adequate inventory has been considered a key element in forest ecosystems now. Despite declaring that the problem is appreciated, Polish foresters keep its inventory at the level much lower than what is necessary for biodiversity protection.*

Perhaps because of aesthetic reasons, foresters try to "clear" the forest of everything that seems to disturb timber production. Dead wood is removed in order to constrain insect populations. The efficiency of these activities is questionable since money spent on struggling with pest infestation does not translate into sufficiently higher revenues from selling timber. Alleged preferences of people who visit forests for recreational purposes were an argument against keeping dead wood. Once it turned out that people actually appreciate forests with fallen mossy trunks, this argument failed. Hence foresters' reluctance to welcome survey results. Finally, let me offer an explanation of the phrase *caeteris paribus*. The words mean "everything else unchanged". We use this expression whenever we want to trace consequences of a single factor (assuming that nothing else changes – in this case it means that we talk about the same forest practice, except that the amount of the dead wood can be higher or lower).

11.8 Calculating WTP based on declared prices makes use of OE or DC questions. Which format seems to be more natural for respondents?

It depends on how you buy or sell market goods you are accustomed to. In many European stores prices are not negotiable: you see a price tag attached to the product, and you take a DC type of a decision. The decision is "yes", if you think the price is acceptable, or "no" otherwise. Thus a DC format of WTP questions seems more natural. But in some stores – for instance in the Middle East – there are no price tags. If you ask for a price, the sales person asks you about what you are willing to pay, or (often by quoting a very high price) invites you to start negotiations. If this is what you are used to, then perhaps an OE format of WTP questions seems more natural.

11.9 Welfare improves if a dirty lake located next to a town is cleaned-up. How do economists interpret this phenomenon? Hint: please analyse picture on page 86 (ERE-11-4).

Recreation has an impact on the economic value of a lake near a town. As long as the lake is polluted, the demand for recreation is D_0 (see picture on page 86, or in my overhead ERE-11-4). If the lake is cleaned-up, the demand will increase to D_1. We interpret this picture in the following way. If the effort required to visit the lake is as it was earlier (the vertical coordinate is the same), then the number of visits will increase (the horizontal coordinate will be higher). This shift of the demand curve results in higher economic surplus caused by the environmental improvement. The shaded area illustrates the increase of economic surplus (a welfare improvement measure).

11.10 Does not Robert Costanza see any difference between forests in Croatia and Sweden?

He does. Both are boreal forests. Differences between various types of the boreal forest are smaller than differences between boreal and tropical forests. The value of food production in a tropical forest is estimated at 32 $/ha, and at 50 $/ha in a boreal forest, but the value of other raw materials (mainly timber) is estimated at 315 $/ha in the former, and only 25 $/ha in the latter. If one takes into account other "services", their total value is estimated at 2007 $/ha in the former (tropical), and only 302 $/ha in the latter (boreal); the difference is thus almost sevenfold. In other words – all benefits considered – the difference between a Swedish forest and a Croatian forest is smaller than between a German one and a Brazilian one.

11.11 Why does the table on page 88 (ERE-11-6) contain so many empty cells?

Because the University of Berkeley library did not have publications with values estimated for some ecosystem "services" in some biomes. For example the row corresponding to tundra is completely empty, even though something takes place there too. In many cases an empty cell means that a given biome does not provide a given "service", or provides it at a negligible level. For instance, in the desert row the column #13 (food production) is empty, because indeed there is nothing to eat. However the column #10 (pollination) is empty as well, even though some plants exist there and some insects pollinate them. Yet the scale of this process is negligible.

11.12 Please give an example of an ecosystem "service" consumed by us without the necessity of carrying out any market transaction?

Perhaps everybody picked berries to be eaten immediately. It is a different story if we buy berries in the market (this consumption is included in GDP – at least theoretically). Eating berries just picked is an example of consumption that can be carried out without making any market transaction.

11.13 Why are wetlands so precious?

There are not many ecosystem "services" that wetlands do not contribute to. The table on page 88 (ERE-11-6) lists 10 "services" they contribute to, and 7 where they do not. In particular, column #2 (climate regulation) is empty, even though we know that wetlands serve as an important carbon sink. Perhaps in the early 1990s there were no publications on this. Despite that (despite these omissions) the value of ecosystem "services" provided by wetlands was estimated at 14,785 $/ha, that is many times more than that provided by grasslands. It is also more than provided by forests (969 $/ha). The value of wetlands results from biodiversity, among other things. Please note that the value of their recreation "services" was estimated at 574 $/ha (that is much more than in the forest). People who are reluctant to think of visiting such places for recreation purposes (because of mosquitos) have to acknowledge that wetlands are a favourite place for many bird species, and these in turn are eagerly observed by numerous ornithologists-amateurs.

11.14 Why are ecosystem "services" provided by tropical forests higher than those provided by boreal forests?

In table on page 88 (ERE-11-6) ecosystem "services" provided by tropical forests are estimated at 2007 $/ha, and those provided by boreal forests at 302 $/ha (see also question 11.10 above). This difference is partially caused by the fact that in several columns tropical forest cells are filled, while boreal forest ones are left empty (probably because there were no publications in the early 1990s). But irrespective of that tropical forests are much denser and more biodiversity-rich than boreal forests. Consequently one can expect them to have more of everything per unit of area.

12. Environmental policy

Earlier lectures – especially the first one (ERE-1) and the fifth one (ERE-5) – emphasised the need to estimate benefits and costs of environmental protection before taking a decision whether to make environmental policy stricter or not. There are 4 fundamental concepts used in order to judge whether a policy is adequate or not. The same concepts are used to judge the adequacy of projects, that is minor elements any policy is composed of. These 4 concepts are:

- Effectiveness = reaching a goal
- Efficiency = maximizing the positive difference between benefits and costs (costs are justified in terms of benefits; advantage of benefits over costs is maximised)
- Cost-effectiveness = reaching a goal at the least cost
- Equity = making the distribution of costs proportional to the distribution of benefits

Effectiveness means achieving an objective. If a project is aimed at lake restoration, its effectiveness implies that the lake is cleaned-up. We do not ask whether it could have been done cheaper. Neither do we ask whether it was equitable (fair). Checking the effectiveness is easy since it does not require economic valuations. Neither does it require analysing who gained and who lost. It requires just making sure that the objective has been achieved.

Economists ask more difficult questions. They ask if the objective was a reasonable one, that is, whether its benefits were higher than costs. If a project was to restore the lake, it is good to check whether the lake was restored indeed, but also whether the costs were justified by the benefits. If an important reason to restore the lake was to boost tourism, but in the area there is also another (unpolluted) lake serving as a tourist attraction, the project is not very convincing. It would have been more convincing if the lake to be restored was a unique one.

Maximising the advantage of benefits over costs is the essence of *efficiency*, as understood by economists. Not every effective policy is efficient. Despite effectiveness, it may not be efficient if what it achieved was too expensive. For instance, if benefits from relaxing very sharp drinking water standards would have decreased just a little bit, but compliance cost would have been lowered significantly, then policy of enforcing these ambitious standards is not efficient. If it delivers its objective, it is effective. Yet it is too expensive to be considered efficient. Society would be better off by relaxing the standards. And conversely, for instance, if benefits from improving air quality would have increased immensely by adopting a stricter standard, while the cost of it would have been moderate, then it is worthwhile to do it. Policy of keeping a weak standard is not efficient in this case.

But the previous lecture (ERE-11) demonstrated problems linked to valuating the environment. Estimating benefits of environmental protection is difficult, since the evaluation of non-market goods is cumbersome. If someone claims that environmental benefits of certain activities are equal to some amount, then this statement can be questioned easily. Perhaps the benefits are higher or lower. Hence referring to the efficiency criteria is often considered controversial and hence avoided.

When valuation results are difficult to defend, many people prefer to use the effectiveness rather than efficiency criterion. Nevertheless there is a way to escape hardships of efficiency without being left with effectiveness only. Economists invented the criterion of *cost-effectiveness* (sometimes called *cost-efficiency*).

If the lake restoration project is to be referred to again, its cost-effectiveness requires that it is done as cheaply as possible. The restoration does not have to be justified by its benefits. This would be a matter of efficiency rather than effectiveness. However, cost-effectiveness differs from effectiveness, because it is not sufficient to reach an objective (to restore the lake in this case). This must be achieved at the minimum cost possible. Let us assume that the benefit from lake restoration was estimated at 1 million €. Depending on the technology applied, the cost of its restoration can be 4 million €, 3 million € or 2 million €; it cannot be cheaper than 2 million €. Irrespective of whether the method selected costs 4 million € or less, all three cases are effective, but inefficient. According to the definition above, cost-effectiveness is achieved in the third case though.

The concepts explained so far – effectiveness, efficiency, and cost-effectiveness – do not exhaust criteria of policy assessments. Social critics are interested in *equity* (fairness) understood as consistency of cost and benefit distributions. For example, let us look at landscape protection. The protection results in imposing certain constraints on economic activities. Their cost is born by people who own real estate. Their benefits are enjoyed by themselves, but also by some others who did not have to pay the cost. Equity requires that both distributions are proportional, that is all those who benefit pay the cost proportionally. In the landscape protection case it is usually achieved through the budget: beneficiaries are taxed, while those who pay the cost are subsidised (to lower the burden for them).

Environmental policies are assessed according to these four criteria. These assessments are often difficult, since cost and benefit quantification is controversial. It is easier to list typical mistakes that the policies should avoid (but very often they do not). Typical policy failures can be classified in the following way:

- Lack of cost-effectiveness
  - Use of uniform standards
  - Politically motivated exemptions

- Compromising enforcement
  - Over-ambitious targets
  - Lack of monitoring capability

They will be explained in detail.

Using uniform standards provides the most typical example of violating cost-effectiveness. A simple calculation illustrates the problem. Some cars drive 1000 km per year, and some drive 100,000 km per year. All of them have to be equipped with a catalytic converter. Without the device the emission of nitrogen oxides would be 2 g/km, and with the device it is 0.2 g/km (catalytic converters abate 90% of nitrogen oxides). Thanks to the catalytic converter the car with the low annual mileage (1000 km) avoids the emission of 1.8 kg (90% of 2 kg), and the one with the high annual mileage (100,000 km) – 180 kg (90% of 200 kg) every year. The cost of the device is identical in both cases, but in the second case it is 100 times lower per unit of avoided emission (also per km driven). Let us check its practical consequences. For the sake of calculations, we will assume that the cost of the catalytic converter is 360 € (actually it varies for various car types). We will also assume that damages caused by the emission of 1 kg of nitrogen oxides (mainly health damages) are 10 € (the number adopted in many European reports). Thus if the car drives 20,000 km per year, the cost of its catalytic converter corresponds to the benefits (avoided health damages) recorded in just one year. But the device can work even longer, so its installation is justified even in cars with a lower annual mileage. Besides, it abates some other pollutants as well.

Yet cost effectiveness of air protection requires that cars with a higher annual mileage (hence emitting more per year) are equipped with more effective converters (to eliminate even more than 90% of emission). Those with a lower annual mileage can install a less sophisticated converter. But this does not happen, because regulations require that every car must be equipped with similar devices. The regulations violate cost effectiveness. Perhaps equity considerations ("everybody pays the same") justifies such a requirement. In addition to that, administrative complexity explains why other alternatives (like requirements based on the number of kilometres driven) were not implemented.

More examples of cost-effectiveness violations can be found. If environmental funds finance projects characterised by the unit costs of abatement effects (such as e.g. tonnes of pollution avoided) that differ from each other by an order of magnitude, they violate cost effectiveness obviously. To see this, it is sufficient to note that better results could have been achieved, if the money was redirected from the more expensive projects to less expensive ones.

Another frequent mistake is to tolerate politically motivated exemptions from standards. In particular, it happens that standards are enforced against small polluters, while large ones are being exempted. Large polluters are often exempted, because politicians do not want to be lobbied; small polluters lack sufficient political power to represent a significant threat. Yet such exemptions violate cost effectiveness, because large polluters are more likely to have cheap abatement options.

Neglecting enforcement is also a typical mistake of environmental policies. In the 1980s Poland had the limit of sulphur dioxide (SO_2) concentration in the air of 32 $\mu g/m^3$. At the same time the Swiss limit was 80 $\mu g/m^3$. Did this mean that the air was better protected in Poland than in Switzerland? It was not. The standard adopted in Poland was too ambitious to be taken seriously. The government absolved itself from the lack of enforcement, since the standard was not realistic. In contrast, the Swiss standard was realistic and taken seriously.

Adopting unrealistic requirements is not energising. On the contrary, it demoralises. The lack of enforcement of some regulations may result in indulgence with respect to other regulations and thus compromised environmental protection. Adopting unrealistic requirements is a mistake motivated by ideological reasons. People think that it helps environmental protection, but in fact this is the other way around.

Likewise, ambitious standards are adopted with respect to various pollutants, that cannot be enforced because of insufficient monitoring. For instance Polish air protection regulations list more than 60 contaminants, even though inspectors can perform routine checks against few of them only. Theoretically, they can control everything by commissioning appropriate analyses to be carried out by scientific institutes. But such sophisticated analyses cost money and require a lot of time. Polluters know this and in many cases they are left emitting excessively, but unpunished.

Detailed analyses of environmental policies are complicated and arduous. Instead, it is often expected that some simple principles these policies are assumed to comply with will allow a success. Three of them seem to be the most popular ones:

Polluter Pays Principle

- *Sensu largo*: the polluter is responsible for whatever harm its activity can cause
- *Sensu stricto*: the polluter is responsible for meeting environmental regulations required by the law

Subsidiarity Principle

- Decisions should be taken by the lowest possible level of government adequate for the nature of damages (adequate for the scope of externalities involved)

Precautionary Principle

- Measures applied should take into account the worst possible outcome

Polluter Pays Principle (PPP) is the best known of them. If the Minister of Environment (in any country) was asked about his (or her) policy, he (or she) would indicate the PPP as the cornerstone of it. Yet the understanding of the principle would turn out to be doubtful if additional questions were asked.

The example below explains why PPP is difficult to enforce. Many people know that an agricultural chemical called *alar* (commercial name of the chemical substance *deminozide*) is a poison. Individuals are poisoned by it from time to time. An investigation can prove that the victim was poisoned after eating an apple. The first idea is to identify the producer of *deminozide* as the polluter. There is a handful of its producers in the world, and they can be caught easily. But producers can defend themselves convincingly that they just provide pesticide manufacturers and other suppliers of agricultural inputs (*alar* is not a pesticide; it makes fruits hang on trees until they are picked by somebody) with chemical substances that they need. Hence they are not polluters. The polluters are manufacturers of products sold to farmers (not to other factories). The number of such manufacturers is higher but they can be tracked as well. These smaller suppliers defend themselves by saying that they mark their products with appropriate warnings, and they inform about the necessity of waiting periods after spraying. Many farmers do not read (they are illiterate) and they do not necessarily comply with waiting periods, but the warnings have been printed indeed. Should farmers be considered polluters? This is possible, but not many politicians are courageous enough to say so. Perhaps groceries where apples are sold should be considered polluters? Or maybe we ourselves are polluters (because we are not careful and suspicious enough)?

Another example of problems in interpreting PPP is provided by the American *Love Canal* scandal. The name has nothing to do with any romantic connotations. The Love Canal is just a small town in north-eastern United States. In the 1920s there was a landfill used by inhabitants of the Niagara Falls municipality and a local manufacturing plant. Later on it was recultivated and the area was designated for housing purposes. Not many people remembered its earlier purpose. Monitoring of air contamination did not reveal any violations. In the 1980s paediatricians were puzzled by the occurrence of child diseases. Children are sick quite often, but if it happens too frequently, paediatricians try to find a more general cause. In this case, after a while, they linked the problem to the operations of the landfill several decades earlier.

Who was the polluter that could be expected to pay? The most obvious choice was the firm which collected and landfilled the waste. But it had not existed any more at that time. Perhaps households whose garbage was collected by the firm should be considered polluters? But it would be difficult to make households liable for something after several decades. The government could not identify the polluter convincingly. Nevertheless damages were unquestionable, so it was necessary to pay for them while violating the PPP.

In 1972 OECD (*Organisation of Economic Cooperation and Development*, a club of some 40 countries that are considered economically developed) adopted a broad definition of PPP making the polluter responsible for all the damages. In this sense households that located their waste in the Love Canal landfill in the 1920s were responsible for damages that occurred in the 1980s. Yet their responsibility could not be enforced. In 1974, as a result of arguments similar to those emerging in the *Love Canal* case, OECD adopted a narrower definition of PPP making the polluter responsible for meeting legal regulations. The Love Canal landfill operated legally, and households paid for their waste collection, so PPP (*sensu stricto*) was not violated.

PPP is sometimes interpreted as a ban on subsidising environmental protection: "let every polluter pay for complying with regulations". If a damage occurred when everything was legal, then the government should pay for it, since regulations turned out to be not quite effective. This is somewhat controversial. The continuity of governments can be deduced somehow, but the continuity of taxpayers cannot be deduced. In the case of the *Love Canal* scandal taxpayers from the 1980s were supposed to pay for what the local government in the 1920s was responsible for. This may sound odd, but perhaps there is no better solution.

Identifying the PPP as a ban on subsidies is not considered fully realistic. PPP precludes subsidising what the polluters are required to do. Nevertheless OECD added certain conditions which lift an absolute ban on subsidies. If the following three conditions hold, subsidies are allowed:

- The subsidy is granted for a determined time period, not longer than 5 years
- The subsidy does not favour specific economic agents (any agent who meets certain criteria is entitled)
- The subsidy does not lead to serious distortions in international trade

Virtually all countries subsidise environmental protection, and they violate PPP even in its weaker version (*sensu stricto*). Everywhere there exist or existed subsidies to municipal sewage treatment plants. Even though subsidies flow to municipalities rather than firms they may distort international trade indirectly, since local production is made easier (and hence cheaper). Subsidies to sewage treatment plants may have another negative effect (in addition to trade distortions). They encourage municipalities to scrutinise costs less carefully and expose citizens to high maintenance costs that could have been lower.

Sometimes the principle is used in its plural form – *Polluters Pay Principle*. When strict financial responsibility of a polluter is difficult or impractical to enforce, governments occasionally apply this alternative version to charge polluters at large with the environmental protection costs. If an abatement facility – like a municipal sewage treatment plant – is financially self-sufficient, then it can be said that the *User Pays Principle* applies, i.e. the users pay for its operation. We would object to calling us "polluters", but we are not offended by being called "users".

Subsidiarity Principle owes its strange name to church organisation patterns suggesting that a higher level does not try to solve lower level problems if these can be addressed adequately there (bishops should not interfere with what can be tackled by parish priests). In the European Union there is a rule not allowing the European Commission to establish regulations that can be adopted successfully by member countries.

In general, in environmental policy the subsidiarity principle states that environmental protection should be undertaken by the lowest level appropriate given the nature of externalities. This level has to be determined on a case by case basis. For instance, noise regulations do not require the United Nations Organisation to refer to. The problem can be tackled effectively at a local level. Solving the European "acid rain" requires a continent-wide cooperation. If the problem was left to national governments, solutions would be like they used to be 100 years ago in Europe. England "solved" the problem of burning fossil fuels by increasing the height of the stacks; local depositions in England were lower, but Norway complained. Protecting the global climate cannot confine even to a continent (like in the European "acid rain" case), because carbon dioxide mixes in the atmosphere well, and – irrespective of where emitted – it hurts everybody. Hence a solution needs to be worked out at the United Nations level (not at a lower level).

Precautionary principle seems to be consistent with common sense: when we do something, we should take into account the worst possible outcome, and act accordingly. But it is consistent with common sense only partially. It is known that an airplane flight can end up with a catastrophe. Everybody is aware of that. What does need to be done in order to prevent this bad outcome? One should not fly airplanes, of course. Nevertheless we violate the precautionary principle routinely, because we attach a very low probability to this worst outcome, and fly airplanes if the benefits from this uncertain activity are very attractive, and therefore they are estimated high enough. In practice we adhere not to the Precautionary Principle, but to a decision making process consistent with rules identified in the first lecture (ERE-1) and repeated in the fifth lecture (ERE-5).

Questions and answers to lecture 12

12.1 Why do economists go beyond the effectiveness criterion?

Because effectiveness looks only at whether an objective was achieved. It does not answer the question whether a given project makes sense, or whether a given policy is worth to be supported.

12.2 Why does not cost effectiveness of an activity guarantee its efficiency?

A cost effective activity does not have to be efficient, because even its cheapest variant may cost more than the benefits achieved. Let me remind you that in order to check the cost effectiveness, we do not estimate the benefits of the objective; we are supposed to make sure that the objective (whatever it is, and irrespective of whether it is "worthy") is achieved at the least cost.

12.3 Is the following statement correct? In one plant reducing emission by 1 tonne costs 1000 €, and in another one – 100,000 €, but emissions should be abated in both of them.

This is an opinion often voiced by technical experts. Indeed, in some instances abatement cannot be cheap; very expensive measures have to be applied. Nevertheless economists are interested in total expenditures a society needs to pay in order to achieve some objective, say, to decrease the emission by x tonnes. If there were x plants with cheap options to abate 1 tonne each then by spending x times 1000 €, the objective could be achieved. If however the number of "cheap" plants is lower than x, the society needs to reach for more expensive options. Perhaps even the plant where abating 1 tonne costs 100,000 € has to be taken into account. Abatement in plants where unit cost of reducing emission is different can be reconciled with cost effectiveness only in the case when abatement in all "cheap" plants is not sufficient to achieve the objective. A frequent practice of reducing emission by both cheap and expensive options violates cost effectiveness (since not all cheap options were exhausted). Therefore the statement is not correct in general.

12.4 Europeans expect Brazilians to protect Amazonia more effectively. Brazilians claim that this is their internal issue. Who is right?

It is difficult to say. Teaching Brazilians by Europeans what they should do is inappropriate indeed. On the other hand, the magnificent nature of Amazonia is valuable not only for the Brazilians. Efficiency requires that the Brazilian nature is protected better than it is protected now (since benefits – for all of us – are higher than the necessary costs). But efficiency is not identical with equity. It would contradict equity if the cost of protecting Amazonia was to be paid by Brazilians alone, since they are not the only beneficiaries. Europeans should participate in the cost of protecting Amazonia. But even if this happened Brazilians are supposed to take the decision about its protection. Countries and institutions offering money for certain objectives can state specific conditions. If the beneficiaries are rational, they will accept a financial offer to cover the deficit caused by the costs higher than domestic benefits (the costs exceed domestic benefits, not the global ones). Of course beneficiaries may act emotionally rather than rationally, and reject the offer. My impression is that this is why less wealthy countries protest against attempts of more wealthy countries to give them hints how to manage their natural capital. Nevertheless a more frequent situation is when wealthy countries expect the less wealthy ones to appreciate the economic value of their natural capital irrespective of whether the money is offered or not.

Not only Brazilians, but other beneficiaries should participate in the cost of Amazonia protection too. Some of us expect that our governments – using the taxpayers' money – would persuade the Brazilian government to enhance the protection. But there is also a decentralised option available. There are environmental NGOs that "buy" the protection. If you trust such an NGO, you may support it with your money, and in this way contribute to protecting the Amazon forest.

12.5 In many countries all newly registered cars have to be equipped with catalytic converters. Even though this contradicts the cost effectiveness principle, it is applied almost universally. Why?

As mentioned in the lecture, regulations require that every newly registered car is equipped with a catalytic converter irrespective of its annual mileage. Some people believe that political lobbing of catalytic converter producers can be a factor. Other explanations refer to a broader definition of costs.

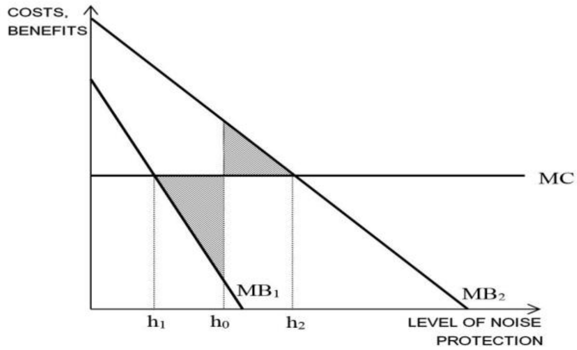
The cost of a regulation does not confine to the cost a device required. The cost of enforcement is important too. One can envisage a principle that requires installing a better catalytic converter in a car with a higher annual mileage, and allowing a less effective device in a car with a lower annual mileage. This cannot be based on *ex ante* expectations, but requires an *ex post* check. Such checks can be performed rather easily (falsification of the mileage is difficult). Nevertheless administrative costs required – especially when compared to a simple check whether the device is installed – imply that the unified standard can be cheaper than more sophisticated solutions (assuming a broader definition of the cost) and perhaps this is why such standards have been implemented in many countries.

12.6 Environmental ministers claim that they apply the Polluter Pays Principle. Do they know what it means and what it implies?

I am afraid that sometimes they do not. First, not always are they aware of the two versions of the principle (*sensu largo* and *sensu stricto*). Second, the identification of a "polluter" is difficult. Presumably PPP is identified with the ban on subsidies, even though in reality subsidies are tolerated subject to certain conditions. PPP seems to be consistent with common sense, but it is not that obvious.

12.7 There are several European Directives that regulate noise emitted by lawn mowers. This seems to contradict subsidiarity principle. Why are there such high-level regulations adopted against noise?

It can be demonstrated easily that adopting a uniform standard is not economically justified. Let us assume that there are two countries – 1 and 2 – characterised by very different preferences with respect to noise protection. Let 1 be a country where noise is not considered very harmful, and 2 – a country where people are willing to spend more in order to protect themselves against the noise. MB_1 and MB_2 are their respective "demand" curves for noise protection. Country #1 prefers a low level of noise protection h_1, and country #2 – a higher one h_2. If the level is to be unified, then h_0 is the average of h_1 and h_2. This average ("compromise") level does not satisfy anybody. The shaded triangles illustrate welfare loss caused by adopting this level (it is assumed in the picture that the marginal cost MC of noise protection is constant; if it is not, conclusions do not change). Country #1 prefers h_1. If forced to adopt a higher one (h_0), it records a loss caused by paying costs that are higher than benefits (left shaded triangle). Country #2 prefers h_2. If forced to adopt a lower one (h_0), it records a loss caused by benefits lost that were not compensated by avoided costs (right shaded triangle).



Adopting average ("compromise") standards for noise emitted by lawn mowers is not only inefficient, but also controversial from the point of view of the subsidiarity principle. Why does the European Commission do it? Although unjustified from the point of view of environmental protection, uniform standards for noise emitted by lawn mowers make sense from the point of view of the unified European market. If a manufacturer of a lawn mower had to certify its product in every country, then the certificate obtained in country #1 would be insufficient in country #2. In contrast, if a common level for the noise protection is valid everywhere, then a single certificate is sufficient. Therefore it can be argued that adopting a uniform standard for noise protection does not violate the subsidiarity principle when understood not only in the environmental context.

12.8 Why does the subsidiarity principle recommend taking decisions at the lowest possible level?

If a decision is taken by a high level (say, national government) then people living at a lower level (say, municipality) may feel alienated and they do not identify themselves with the decision. That is why it would be good to have decisions taken at the lowest possible level. A possibility that such a decision can impose an external cost on somebody else may justify letting the decision be taken by a higher level. This does not violate subsidiarity.

13. Environmental policy instruments

The previous lecture (ERE-12) introduced four criteria to assess environmental policies: effectiveness, efficiency, cost-effectiveness, and equity. In addition, it defined three principles these policies are supposed to comply with: PPP, subsidiarity, and precautionary. Other analyses of environmental polices look at instruments they apply. At least four groups of instruments are listed:

- Price (tax) instruments
  - Pigouvian taxes (with or without thresholds)
  - Sub-Pigouvian pollution charges
  - Non-compliance fees
- Quantity regulations

  - Standards
  - Non-tradable permits
  - Zoning

- Tradable permits

- Voluntary instruments
  - Moral suasion
  - Eco-labels
  - Voluntary agreements

Especially the first two groups are analysed and compared to each other. In order to characterise them, one needs to recall the first lecture and one of its illustrations (the lower picture on page 5, or ERE-1-6). The graph identifies the point (q^*,p^*) at the intersection of the MAB and MAC curves. The point corresponds to the environmental protection level q^* (for example related to water or air quality) such that the marginal cost of abatement is equal to its marginal benefit (p^*). In the second lecture a Pigouvian tax was defined (page 16, ERE-2-7) as a means to let the polluter see as an internal one the external cost it imposes. As a result of the Pigouvian tax the polluter is motivated to undertake abatement up to the level q^*. The same outcome can be achieved by simply requiring the polluter to abate up to this level. So-called non-tradable permits force polluters to do this.

The same picture (page 5, or ERE-1-6) illustrates two alternative approaches to environmental protection policy. Economic theory indicates that (q^*,p^*) is the optimum solution. One approach is to require the polluter to move towards q^*. It is called "quantity regulation" (it looks at the horizontal – quantity – axis). Another approach is to impose a Pigouvian tax in order to move towards p^*. It is called "price regulation" (it looks at the vertical – price – axis). Even though totally different methods are used, both approaches motivate polluters to do the same thing, namely to abate up to the socially desirable level. Pigouvian taxes and non-tradable permits are the most important examples of the first two groups of instruments, but they do not exhaust them.

The group of price (tax) instruments includes sub-Pigouvian taxes, i.e. moderate environmental fees. For instance the Polish sulphur dioxide charge is roughly 120 €/tonne. In contrast, the Pigouvian rate is many times higher. Does this mean that the Polish fee does not provide any incentives to abate? No. It does give such incentives, but they are insufficient to let polluters abate up to what is justified by the external damages their emission causes. In order to bring them up to this level, non-tradable permits are used. They force polluters to abate more than what would have been implied by the environmental fee.

There are also non-compliance fees in many countries (including Poland). They differ from environmental fees. The latter are charged for permitted emission. The former are charged when emission is higher than permitted. In Poland non-compliance fees are multiples of "regular" environmental fees. Sometimes they are high enough to resemble a Pigouvian tax. But their enforceability is low, as apparent from the statistics. Almost 100% of environmental fees are actually paid. At the same time much less than 50% of non-compliance fees are paid. In other words, non-compliance fees seem to be "dangerous", but only theoretically; polluters know that violating emission permits does not imply serious consequences always; it implies serious consequences rarely.

Apart from non-tradable permits, the group of quantity regulations includes standards and zoning. The difference between a standard and a non-tradable permit is that the former applies to any polluter of an indicated type, while the latter is issued individually for a specific polluter. For instance the emission of car exhaust gases is regulated by standards rather than non-tradable permits. Emission from an electricity plant is regulated both by standards and non-tradable permits. For instance, if the plant uses a given technology, then it must comply with standards adopted for this technology. At the same time every plant is located in a specific place. Its non-tradable permit allows – at least theoretically – to emit only what can be tolerated by the neighbourhood. In the case of a power plant located in a densely populated area the non-tradable permit can be stricter than the standard. On the contrary, if the plant is located in the middle of nowhere, then its non-tradable permit can be weaker than the standard.

Zoning belongs to the quantity regulation group. Unlike standards and non-tradable permits, it simply introduces a ban for placing certain objects in certain areas. For example zoning may establish a ban for siting manufacturing plants in certain locations. Or it can establish a ban on car traffic. The advantage of zoning is its simplicity. At the same time, its disadvantage is the lack of flexibility. Perhaps some small emission is justified, while zoning precludes any emission from a given source type.

Some environmental economics textbooks indicate only two groups of policy instruments: price (tax) instruments, and quantity regulations. Newer textbooks emphasise that tradable permits do not fit either of the groups. Discussion of tradable permits will make it clear why we added the adjective "non-tradable" in earlier analyses. Everybody understands what is a non-tradable permit. Tradable permit means that if a polluter uses its part only, then the unused part can be sold to somebody else. And *vice versa*, one can emit more if one buys an unused part of the permit from somebody else.

Tradable permits have certain characteristics of quantity regulation, because their sum is equal to what the government considers acceptable. But they have certain characteristics of price (tax) instrument, since they are flexible. They do not oblige polluters to abate to a certain level, but they give them incentives. The incentives are twofold. Those who emit less than the permit allows can sell its unused part. Those who find their marginal abatement cost to be high may try to buy additional permits from those who managed to abate at a lower cost.

Tradable permits are so important that a separate (next) lecture will be devoted to them.

As the name suggests, voluntary instruments are non-obligatory – they simply inform and/or suggest.

Moral suasion is the first instrument listed in this group. Some people think that it is too vague to be considered a policy instrument. Nevertheless waste management demonstrates that moral suasion can be effective. Municipal garbage is a difficult problem and it has not been solved satisfactorily anywhere in the world. Of course there are price lists that motivate us to segregate the waste and throwing them to wrong containers is sometimes detected and punished. A total success is impossible though, because it is not possible to control our behaviour fully. Yet moderate errors do not lead to damages. Some time ago I learnt that a mixture composed of scrap paper in 90% only (that is up to 10% "contaminated" with something else) can be processed in paper mills as if it was 100% "pure". Let us assume that we urge people – by referring to their honesty, since not everything can be checked – to throw scrap paper only to appropriately marked containers. Some people will turn out to be stubborn. But if the share of the honest is 90%, the system will work. Hence moral suasion can prove to be sufficient and it is not expensive.

Eco-labels work similarly. They do not force to do anything. They simply provide information. Refrigerators have alternative labels attached: A+, A, B, C, or D. A+ ones are characterised by the lowest energy use. Unfortunately they are often the most expensive ones. But if there are people who wish to demonstrate – even without any external pressure – that they care for energy conservation, they can do so. Eco-labels provide quick information which can be obtained otherwise, although at a higher effort. Economists analyse circumstances where so-called information asymmetry (sellers and buyers have different information about what is to be sold and bought) leads to non-optimal choices. Eco-labels are supposed to inform potential buyers about product characteristics that are not universally known, yet are relevant for environmental protection.

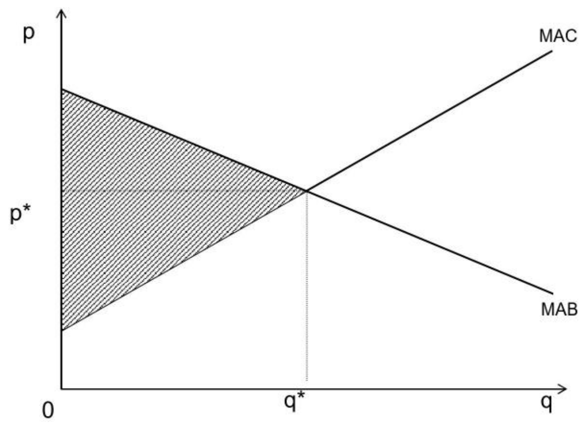
Sceptics raise the credibility problems. For instance in Finland (considered by many people a country of high environmental awareness) there were two systems of wood product certification. These two systems required complying with certain criteria in order to call their products (e.g. boards) "green" or "sustainable". One organisation claimed that certificates awarded by the other were not reliable, and *vice versa*. Therefore an average buyer did not know if the product he (or she) buys was produced according to principles he (or she) expected, or not.

As indicated earlier (in the second lecture), as well as today, quantity regulations are prevalent. Pigouvian taxes are praised in economic textbooks, but almost no politicians were ever convinced to implement them. In other words, virtually all environmental improvements achieved over the last century are due to quantity regulations. Why did environmental policy makers hesitate to apply price (tax) instruments?

The most popular explanation of this phenomenon are controversies about externality monetisation. The Pigovian tax rate is equal to the marginal external cost calculated for the social optimum ($MPT=MEC(q^0)$, where q^0 satisfies $MSC(q^0)=MSB(q^0)$; see the second lecture, ERE-2). Imposing the Pigouvian tax requires monetisation of environmental outcomes, and this may prove to be controversial. Politicians prefer to avoid the problem by pretending that environmental policies are based on some "objective" criteria rather than looking at monetary values.

Here is how it works. The government wants the environmental protection to enjoy the abatement of q^* (see picture below which replicates the lower part of what you saw in the first lecture on page 5). It can alternatively impose a Pigouvian tax of p^* to motivate polluters to abate up to the level of q^*. Both methods are equivalent. The explanation is different though. In order to justify q^* (for instance 1000 tonnes) the government can claim that this is what the doctors say, this is what ecologists say, or this is what engineers say. In order to justify p^* (for instance 1000 €/tonne) the government would have to admit that by imposing the tax (justified by external damages), it is expected that polluters will abate up to what is optimal. Quoting the tax rate and admitting that it was based on some monetary valuations invites journalists, environmentalists, business leaders, and everybody else to undermine the credibility of the policy. Reactions will be less violent if monetisation is not there.

Both approaches are equivalent, and they are based on the same argument, and on the same graph. Nevertheless their political appeal is different. If one looks at the vertical axis, then controversies linked to economic valuation are unavoidable. If one looks at the horizontal axis, then economic valuation problems can be hidden, despite the fact that they are important for both approaches. As emphasised in the lecture on uncertainty (ERE-5), zero pollution objective is inappropriate, and yet it is the only one that does not refer to any economic context. If the objective is to keep the pollution at some non-zero level, then the question can be raised why not to increase it or decrease it by a little bit. The reason why a specific (non-zero) pollution target is adopted is that this is where MAB=MAC, and changing the target would not be justified. Economic valuation can be hidden, but it cannot be ignored.

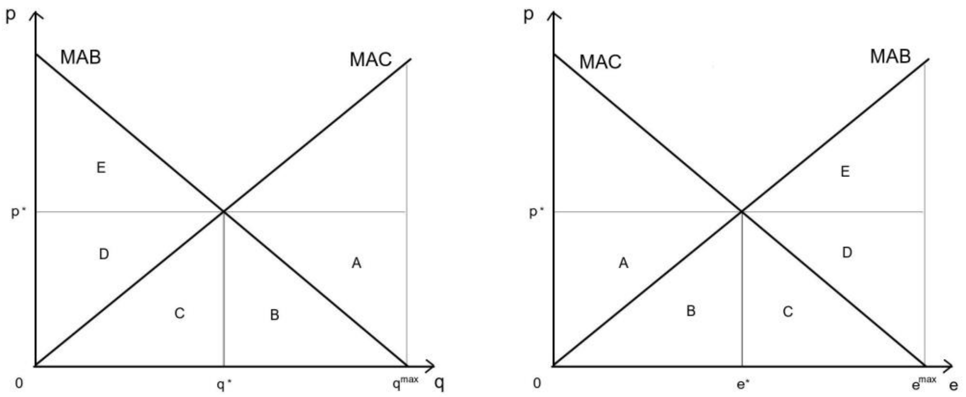


Alleged avoidance of monetisation – in fact hiding the monetisation – is perhaps the most important reason why quantity regulations are much more popular than price (tax) instruments. But there are other reasons for circumventing the latter.

Numerous externalities created by car traffic can be addressed by price (tax) instruments. The externalities are not fully determined by the car vintage. Neither are they proportional to the fuel tanked. Thus they cannot be taxed uniformly at the moment of car registration or at the moment of buying the fuel; they have to be taxed according to what the car (and its driver) does. This, however, would require an enormous administrative effort, and that is why car-related externalities are tackled by quantity regulations (such as catalytic converters – see ERE-12).

An additional reason why price (tax) instruments are not popular is uncertainty as to the outcome. When one looks at the picture above, one obvious uncertainty is that a Pigouvian tax can be miscalculated, and it will motivate polluters to an abatement effort different from q^*. This will not happen with quantity regulations: regulators may be surprised with abatement costs imposed, but at least the abatement effort will be as expected.

Another uncertainty is about the incentive effect of the tax. Even if its rate is calculated accurately, tax revenues are to be paid back to the economy. Unless they are paid in a way totally uncorrelated to how they were raised (economists call it "flat rate"), they are going to change their incentive characteristics. If for instance, electricity plants are subject to a Pigouvian tax, the electricity price goes up, and the demand for electricity goes down. But if the tax revenues are used to subsidise old-age pensioners, then the beneficiaries of this programme are more wealthy and they will demand more electricity. In the end the incentive effect of the Pigouvian tax will be weaker than expected. This is an additional reason why in many environmental protection programmes quantity regulations rather than price (tax) instruments are relied on.



One can also check the political economy. Picture above explains who is charged how much, and who benefits depending on what policy instruments are applied. Please look at the left picture first. In the absence of any regulation there is no environmental protection at all (q=0). The government comes in and wants to regulate environmental protection so as to achieve the optimum (q^*,p^*), that is to imply the surplus of D+E. From now on, however, it is more natural to analyse the right picture which is simply a mirror image of the left one (more explanations are required in order to achieve mathematical precision). Please note that the horizontal axis is interpreted now as the level of emission (q=0, that is no protection corresponds to the maximum emission that the polluter chooses to release, i.e. e^{max}; q^{max} corresponds to no emission at all, i.e. e=0; q^* corresponds to e^*). If the Pigouvian tax p^* without any threshold is applied, then the polluter pays C (the cost of abatement) plus A+B (the tax). If the Pigouvian tax p^* with a threshold q^* is applied, then the polluter pays C (as before) and no tax. If the Pigouvian tax p^* with a very generous threshold e^{max} is applied, then the polluter pays C (as before), and is being paid (subsidised) C+D. That is the net gain is D. If the quantity regulation is applied then the polluter pays C (the cost of abatement) and nothing else. Thus under Pigouvian taxation the polluter either pays or gets a subsidy depending on the threshold applied. Nevertheless, as observed earlier (especially see question 2.8), it is usually assumed under Pigouvian taxation that there is no threshold and the polluter pays the cost of abatement plus the tax on the residual emission (i.e. A+B+C).

That is why quantity regulations are more prevalent. They are considered cheaper by the polluters, even though they can require paying exactly the same as under the tax regime (if an appropriate threshold is applied). If the tax applies a more generous threshold (and thus it works as a subsidy), it becomes even more attractive than a quantity regulation, but nobody thinks of price (tax) instruments in this way. Picture above can also illustrate the PPP *sensu largo* (the polluter is responsible for whatever damage the pollution causes). Under such (quantity) regulations – in addition to paying the abatement cost (C) the polluter has to compensate for the damages the residual pollution causes (B).

This discussion of environmental policy instruments needs more comments on voluntary instruments. Earlier examples referred to providing information. Concerns can be raised whether voluntary instruments make a fair solution. The doubts are caused by the fact that the response in not obligatory: some people will choose a better option, but some will not (and they will not be penalized). It may be considered unfair indeed, but – as explained earlier – the society may wish to have a problem solved cheaply rather than making sure that the good ones are rewarded, and the bad ones are punished.

Voluntary agreements point at a totally different aspect of voluntary instruments. The term is used for a voluntary compliance with regulations contemplated (but not enacted) by the government. Polluters declare that they will comply with what the government contemplates sooner and/or better, and urge the government not to regulate. Theoretically this sounds convincing. The Dutch "VOC 2000" programme aimed at reducing emissions of volatile organic compounds by chemical manufacturers has been considered a success story. Yet, when the University of Warsaw asked the Dutch government for more details, it turned out that neither the abatement cost saving (an important aspect raised by industry), nor even the quantity of the abatement had been recorded. Voluntary agreements are mentioned as environmental policy instruments, but one has to be cautious when trying to check if improvements promised are in fact delivered.

Questions and answers to lecture 13

13.1 Why do environmental economists claim that price (tax) instruments and quantity regulations can achieve the same outcome?

Both solutions can achieve the same optimum outcome (q^*,p^*). They are based on the same graph (page 5 or ERE-1-4 and ERE-13-2). In addition, both solutions can require the polluter to pay exactly the same amount (see page 106, or ERE-13-4). Both solutions require the polluter to pay its abatement cost, that is C. In addition, Pigouvian tax implies the payment of the tax: $PT(q)=MEC(q^0)\ (q-q_{threshold})$, where $MEC(q^0)$ stands for the marginal external cost calculated for the economically justified level of externality, that is q^0 (please note that what is understood in the definition of a Pigovian tax on page 16 or in ERE-2-7, i.e. q^0, corresponds to q^* in pictures referred to here). As a result of the tax rate, the polluter has an incentive to set its pollution exactly at the economically justified level of externality, that is $q=q^0$ (i.e. $q=q^*$). If you substitute this amount as the threshold into the PT formula, you will see that $PT(q^0)=MEC(q^0)\ (q^0-q^0)=0$. Hence when this threshold is applied, then the polluter does not pay the tax in fact, and the burden is the same as in the case of quantity regulation.

13.2 By manipulating thresholds ($q_{threshold}$ in the definition of the Pigouvian tax) the government can change the burden of the tax for polluters. How can this be achieved?

By increasing thresholds – without changing the tax rate – one can make the tax payment arbitrarily low. In particular, it can be achieved that for the sector (not necessarily for individual polluters) it will be zero. The example below explains how this was achieved in Sweden.

In 1992 the Swedish government carried out a successful experiment with so-called nitrogen fee (imposed on the emission of nitrogen oxides) based on variable thresholds. The nitrogen fee of 40,000 SEK/tonne was to be paid by all commercial installations that produce electricity or heat (400 plants in Sweden). The fee was very large (it can be considered a Pigouvian tax), but it turned out to be politically feasible, since it was paid back to those who were taxed. If the tax is paid back, then nobody can claim that the government kills industry, but does such a fee make sense at all? If everybody received what was paid, then obviously the fee would not have any incentive effect. The key characteristics of the fee is that even though the entire group of taxpayers (heat and electricity producers) receive what they paid, some of them get more, and some – less. The payment is proportional to the emission of nitrogen oxides (at the beginning of the last decade it was approximately 4670 €/t), and the return is proportional to the energy produced (at the beginning of the last decade it was approximately 0.95 €/MWh). In other words, if the plant emits per unit of energy exactly what the average for the entire group is, the return is exactly what it paid. If the emission is higher than average, it pays in net terms, and if the emission is lower than average, it is subsidised in net terms. This corresponds to the Pigouvian tax defined as:

$$PT(q_i) = MEC(q^0)(q_i - q_{threshold}),$$

where q_i is the emission of nitrogen oxides per unit of energy produced in plant i, $MEC(q^0)$ – as in the question 13.1, or in the formula on page16 (or ERE-2-7) – is the uniform tax rate, $q_{threshold} = (\sum_i q_i e_i)/\sum_i e_i$ is the tax threshold, and e_i is the energy produced in plant i. As before, PT stands for a Pigouvian tax, and $MEC(q^0)$ – for the marginal external cost calculated for the economically justified level of externality q^0. One can check that the sum of budgetary revenues is zero (what the budget receives, it pays back):

$$\sum_i e_i PT(q_i) = \sum_i (e_i MEC(q^0)(q_i - q_{threshold})) = MEC(q^0) \sum_i (e_i\, (q_i - q_{threshold})) =$$
$$MEC(q^0)\, (\sum_i (e_i q_i) - \sum_i e_i q_{threshold}) = MEC(q^0)\, (\sum_i (e_i q_i) - q_{threshold} \sum_i e_i) =$$
$$MEC(q^0)\, (\sum_i e_i q_i - (\sum_i q_i e_i)/\sum_i e_i)(\sum_i e_i)) = MEC(q^0)\ 0 = 0$$

The formula demonstrates the logic of the Swedish nitrogen oxide charge. Intuitively it is very simple: plants have an incentive to abate nitrogen oxide emission up to the unit cost of 40,000 SEK/tonne. Those ones who arrive at the emission lower than the average are rewarded. Those ones, whose emission is higher than the average contribute to this reward. The budget is neutral; it neither pays nor gains from the scheme. The fee caused excellent results. In 1992 the average emission was higher than 0.4 kg NOx / MWh. In 2008 it was approximately 0.2 kg NOx / MWh. The results can be also phrased otherwise: the total annual emission from the plants covered by the fee decreased from 16,000 tonnes to approximately 14,000 tonnes, even though the annual electricity production increased from below 40 TWh to approximately 70 TWh.

So much for the Swedish nitrogen charge. Because of applying specially designed thresholds, it resulted in various plants paying different amounts or receiving various subsidies. The solution is a unique one (it was not replicated anywhere else). It demonstrates that by manipulating thresholds one can decide on the tax obligation without changing the tax rate. Higher tax thresholds $q_{threshold}$ lower the tax obligation while lower ones – increase it.

The reason that this Swedish tax was not replicated in other sectors is that other industries do not have a non-controversial criterion to repay the tax. In the case of power plants the criterion referred to electricity production. Other plants that emit a pollutant (nitrogen oxides or something else) have non homogenous outputs and apply various inputs; some employ a number of workers while some rely on sophisticated automated technologies. Thus it would be controversial to suggest any method of repaying the tax (think of paying in proportion to kilogrammes produced, in proportion to workers employed, in proportion to the capital invested, and so on). Electricity sector proved to be unique.

13.3 Which of the instruments – standard or a non-tradable permit – imposes stricter constraints for economic activities?

It depends. Typical standards require an abatement level allowed by an available technology. Let us think of an abatement level appropriate for an electricity plant. If a plant is located in the middle of nowhere (far from houses and far from ecologically valuable areas), then a relaxed non-tradable permit can be issued. It will probably be less strict than a standard. On the contrary, if a plant is located where damages can be significant (like in the middle of a densely populated area), the non-tradable permit should be strict. Perhaps it will be even impossible to comply with. This simply means that the plant cannot be located there.

13.4 *Voluntary Agreements* are a special case of voluntary instruments. A group of polluters can approach the government and urge it not to regulate them. They argue that they know better than the government who can do what, and how much this will cost. Can such "voluntary agreements" substitute environmental regulations?

The notorious Dutch VOC 2000 programme casts serious doubts on whether industry delivered what the government expected. Perhaps Voluntary Agreements are a good solution, but a specific objective has to be stated first. If voluntary agreements are to substitute a regulation, they need to achieve the objective that the regulation was to secure.

13.5 Why were eco-labels considered a bad instrument of environmental protection in some countries?

In many countries, discussions on eco-labelling were carried out in the 1980s and 1990s. As a result of these discussions, it was introduced in some places. Nevertheless analysts have warned that governments should not support them officially. Every eco-label should be endorsed by an NGO, but not by the government. The warning is based on the assumption that if a product is identified as an absolutely worse one, it should be simply banned. If it is legal then it can be worse according to some criteria, but better – according to other ones. If an NGO has information about the superiority of a product over alternative ones, let it disclose this information. However, this does not preclude a different choice by somebody else who attaches different weights to specific characteristics of the product.

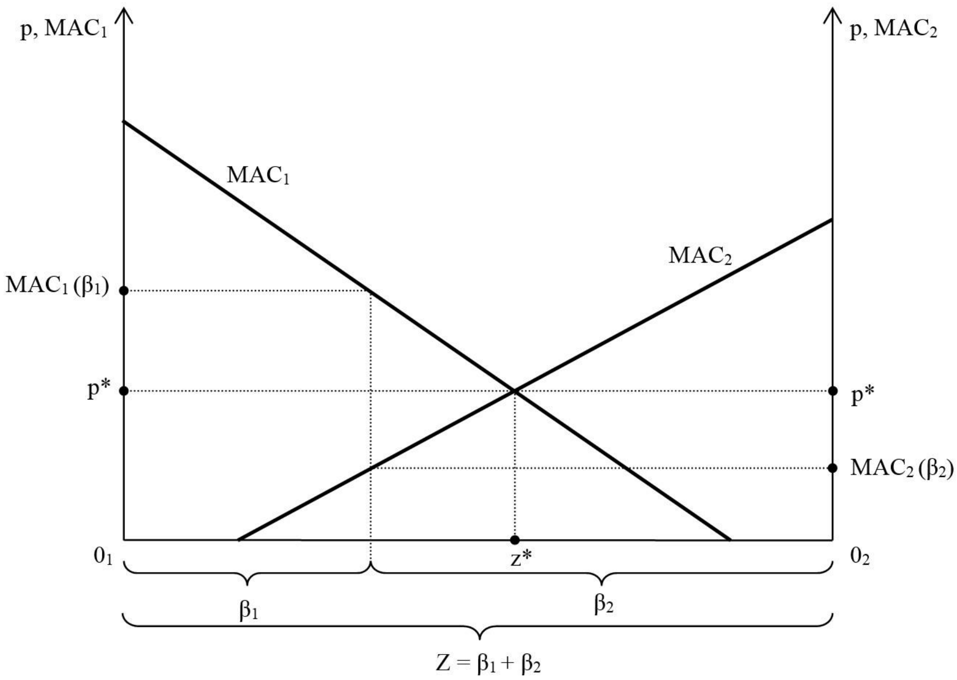
13.6 Mexican and Japanese fishermen effectively protested against the label "dolphin-safe" on tuna cans sold in the USA. Some fishing techniques are dangerous for dolphins that can be trapped in nets aimed at tuna fish; only some tuna fishermen (especially the American ones) apply the technique that is safe. Is the label "dolphin-safe" unnecessary?

Many people like dolphins, and do not want to consume a can containing tuna fish that was caught while killing a dolphin unwillingly. American fishermen persuaded the US government to include an appropriate information when selling tuna caught by a safe technology. This was sued by their foreign competitors as a hidden barrier to trade. As a result, voluntary eco-labels were introduced but thy were sued as well. The US government (who did not endorse these labels, but just tolerated them) lost this case too, since the label was found to be discriminatory and thus illegal. A question can be raised, whether the "dolphin-safe" information is a kind of discrimination. Indeed in principles adopted by the WTO (*World Trade Organisation*) there is a ban on letting the fate of sales depend on production technology. In other words, tuna sales should not depend on the harvest technology applied; apparently this technology was legal in the country where the fish was caught (even though it would not be legal in the USA). But on the other hand, some consumers are actually interested in the fate of dolphins, and they would not like to eat tuna from the dangerous harvests. Hence the label "dolphin-safe" does not have to be a matter of international trade discrimination, but a matter of a genuine care.

14. Tradable permits

Tradable permits are an instrument of environmental policy (see my previous lecture: ERE-13). They are so important and interesting, that they deserve a separate lecture. They were conceived more than a half century ago. Tom Crocker of the University of Wyoming is the first person who mentioned them in 1966. He emphasised that when analysing an economically justified level of environmental protection, the demand for it has to be matched by its supply understood as the availability of abatement projects (recall the first lecture, especially pages 2 and 5, and question 1.4). In some textbooks, John Dales, a professor of the University of Toronto, and his 1968 book are indicated as the first ones. In 1972, David Montgomery provided a mathematical framework for using the instrument in environmental policy.

The idea is quite simple. Polluters are given permits that can be traded: if one plans to emit more, the missing part of the permit can be bought; if one emits less than the permit allows, then its unused part can be sold. The total emission must be equal to what individual permits allowed in sum. This mechanism can be explained with the same picture that was used in the third lecture to prove the Coase theorem (see page 22, or ERE-3-5).



As illustrated above, there are two plants (1 and 2) that emit pollution. Their abatement costs are MAC_1 and MAC_2, respectively. Both abatement cost curves – MAC_1 and MAC_2 – are downward sloping: the more they emit, the less expensive abatement is (please note that the horizontal axis of the second plant runs from right to left!). The no emission variant is the most expensive one ($MAC_1(0)$ and $MAC_2(0)$ are the largest). If the plants are allowed to emit as much as they want, they will choose emission such that $MAC_1=0$ and $MAC_2=0$. The vertical axes measure economic variables (costs and prices). The horizontal axis measures emissions.

The left axis refers to the first plant and it has a standard orientation. The second plant is represented by a mirror image of a traditional picture. Thus its vertical axis is on the right, and the direction of the horizontal axis is from right to left.

If the two systems are to be included in the same picture and the horizontal axis is the same for both, one needs to determine how far from each other the vertical axes should be. It is assumed here that the distance is equal to the total emission from the two plants that we can tolerate, i.e. Z, the sum of what their permits allow.

Let us assume that the two plants are given emission permits β_1 and β_2, respectively. The first plant can emit β_1, and the second one: β_2. The sum of their permitted emissions is $\beta_1+\beta_2=Z$. The plants have to abate in order to emit what they were permitted. Their marginal abatement costs read $MAC_1(\beta_1)$, and $MAC_2(\beta_2)$, respectively. $MAC_1(\beta_1)>MAC_2(\beta_2)$. The first plant has an incentive to approach the second one and to ask that the second plant abates more (i.e. emits less) and sells the unused part of its emission permit. This corresponds to negotiating a different distribution of abatement tasks: the first plant can emit more than β_1, if the second one emits less than β_2. If the first plant buys the part of emission permit unused by the second plant, it can emit more and thus avoid some of its abatement cost. The first plant is willing to pay less than $MAC_1(\beta_1)$, and the second plant should be willing to accept anything more than $MAC_2(\beta_2)$. As $MAC_1(\beta_1)>MAC_2(\beta_2)$, they should agree on a mutually attractive price of the unused part of the emission permit. By looking at the picture, you see that they have incentives to trade until they hit z^*, i.e. an allocation such that their marginal abatement costs are identical (the first plant will emit z^*, and the second one: $Z-z^*$). The price of the emission permit is then p^*. This is how the mechanism works.

In the early 1990s, tradable permits were considered a novelty in Europe, even though they had been successfully tested in the USA for two decades. They were not banned in Europe, but they were treated warily.

The emission originating from two plants located in Chorzów (a town in Southern Poland) was much higher than what could be tolerated, and local administration sought methods to cut it. Unfortunately – because of the wary attitude – the administration could not fully implement tradable permits with respect to the two plants (the Kosciuszko steel mill, and the Chorzów power plant). Nevertheless, despite the unfriendly legal framework, the head of the Katowice regional environmental authority tried hard to issue a much stricter permit for the low abatement cost plant (the steel mill) so that together with the high abatement cost plant (the power plant) they emit jointly what the town can tolerate. The project succeeded, but more time was needed to change the attitude towards tradable permits in Europe.

There are several variants of tradable permits:

1. Absolute limit ("cap-and-trade", "harvest quota", etc.)
2. Credit system
3. Relative limit
4. Hybrid solutions

Picture on page 110 refers to what we call the "absolute limit". This absolute limit is given by the sum $\beta_1+\beta_2$. It is called cap-and-trade system: setting the limit (the cap) and allowing polluters to trade ("under the cap"). It does not confine to pollution abatement. Similar approach can be used in fisheries. The absolute limit for the harvest is determined first. It is allocated among the fishermen somehow. Then fishermen are allowed to depart from what they are permitted to do, but those who want to catch more should make sure that others will catch less. The fishermen who do not use their permits can sell their unused parts to somebody else.

The second variant differs from the first one by the lack of setting the limit explicitly. The trading mechanism is the same though: if one plans to increase the emission, a permit has to be bought from somebody who did not use it fully. Everything looks like in the first variant, except that abuses (i.e. fraudulent applications) are easier. A plant can prove a reduction of emission which was not accounted for by the administration (before it was reduced) and demand that the administration "credits" (recognises) this reduction so that it can be sold to somebody else. For instance, if emission of particulate matter starts to be monitored, it is not allowed to increase it unless a permit is bought from somebody who obtained an "emission reduction credit". But simply by pouring concrete on top of a gravel parking (which results in less dust around) an "emission reduction credit" could have been obtained and sold to a firm which built a power plant and looked for a particulate matter permit. The firm would not have been allowed to build this power plant without a permit, but as a result of this arrangement it is allowed.

Another fraudulent use of the second variant has been practised in the so-called *Clean Development Mechanism* (CDM) under the international climate convention. CDM allows using an additional permit for carbon dioxide emission in a country where there is a limit for this emission. This time the fraud is caused by the fact that "emission reduction credit" is sold by a country that does not have such a limit. As a result, the total emission does not have to be unchanged; it can go up, since the revenues from selling the "emission reduction credit" can be used to make an investment leading to an increase of the emission (if the selling country does not have a binding emission limit).

The third variant is more complicated. It resembles somewhat the mechanism of the Swedish nitrogen oxide charge analysed in my answer to the question 13.2. In some countries it is used in order to allocate carbon dioxide permits among electricity plants. Let us assume that electricity plants are allowed to emit 120 million tonnes of this gas. If they produce 100 TWh of electricity, then 1 TWh corresponds to 1.2 million tonnes of the emission. Yet if the production goes up to 120 TWh, then 1 TWh corresponds to only 1 million tonnes. The limit of 120 million tonnes is absolute (it cannot be exceeded by the electricity plants). The limit per unit of electricity production is a relative one; electricity plants do not know how much it will be per unit of electricity produced. If the total electricity production is 100 TWh, it will be 1.2 million tonnes per 1 TWh. But if the total production is 120 TWh, it will be only 1 million tonnes per 1 TWh. It is called a relative limit. Electricity plants do not know in advance what they will be allowed to emit per unit of production.

The fourth variant is based on combining absolute and relative limits. Hence it is even more complicated.

The trading mechanism is always the same (whoever did more than permitted can sell the effect of this effort to somebody who did less), but the initial allocation of permits is more complicated. In many applications – especially during their first decades – tradable permits relied on historical precedents. This is called the *grandfather principle* (or *grandfathering*). Permits are allocated according to historical emission. Hence if a polluter emitted 50 tonnes, it gets a permit for 50 tonnes, and if a polluter emitted 10 tonnes, it gets a permit for 10 tonnes. If the administration wishes to reduce the emission by 30%, the first one will get the permit for 35 tonnes, and the second one will get the permit for 7 tonnes.

This is an unfair principle since the polluter who has always emitted 50 tonnes will get a permit for 35 tonnes, and if the second polluter emitted also 50 tonnes in the past, but reduced it recently to 10 tonnes, it will get a permit for 7 tonnes only. If the polluters expect permits to be grandfathered then they hesitate to undertake environmental improvements, since they do not want to be penalised by a lower permit if they start abatement earlier. In water management this is called *use it or lose it*. If somebody has a permit to use 10 m^3 of water, he (or she) hesitates to use only 9 m^3. He (or she) is more likely to use 10 m^3, being afraid of getting a permit for 9 m^3 in the future. Despite these obvious disadvantages, grandfathering is preferred by many polluters, but also by administrators who do not want to face more embarrassing lobbying caused by auctions (lobbying is always practised).

Economists praise advantages of auctions. Auctions imply that initial allocation is zero to everybody. Polluters need to buy permits in order to emit anything. Administration is not accused of corruption and the lack of fairness like when it gives low permits to those who made improvements in the past, or gives high permits to those who are the loudest ones. Yet polluters do not like auctions.

In the 1980s and 1990s economists warned against pathologies caused by insufficient market competition. In particular, they were afraid of monopolistic behaviour and speculative permit purchases in order to hurt competitors. These fears turned out to be premature, since auction systems proved to be immune to such pathologies. On the contrary, prices reflect the supply and demand correctly, and there are no signs of harmful speculations.

So-called "banking" is allowed in many systems. "Banking" means keeping own unused permits in order to use them in the future. If a firm expects that future permit prices will be high, then it does not sell its unused permits, but keeps them – either for itself or for sale (at a higher price). Speculation of this type is not damaging; it is a normal strategy in competitive markets.

Tradable permits have numerous advantages. These can be summarised by the following statements that will be explained in a moment:

1. Ideal implementation of the cost effectiveness
2. Separation of scale and allocation decisions
3. Freeing administration from the necessity to take into account individual circumstances of economic agents (polluters)
4. Possibility of reaching environmental objectives gradually

Cost effectiveness was defined in an earlier lecture (ERE-12). It requires that the objective is achieved as cheaply as possible. Let us look at the picture from page 110. If the objective is to achieve the emission of $\beta_1+\beta_2$ then tradable permits let achieve it at the minimum abatement cost. As long as the polluters know that their marginal abatement costs are not the same ($MAC_2(\beta_2)<MAC_1(\beta_1)$ or $MAC_2(\beta_2)>MAC_1(\beta_1)$), they have incentives to trade permits. Incentives disappear when marginal abatement costs are the same ($MAC_2(\beta_2)=MAC_1(\beta_1)$), that is when their sum cannot be made lower. Similar logic applies to the situation when there are more than two polluters (but drawing a diagram for three plants in the two-dimensional space would be difficult).

The second of the advantages listed above requires introducing a distinction between scale and allocation decisions. An appropriate quotation from Herman Daly (an American economist) is in place.

> *In loading a boat we also have the problems of allocation and scale – allocating or balancing the load is one problem (a microeconomic problem), and not overloading even a well-balanced boat is another problem (a macroeconomic problem). To avoid overloading and sinking even a well-balanced boat we have a Plimsoll line defining an absolute scale limit. But the boat can be well or badly balanced even when the water line is far below the Plimsoll mark. And if the water line is above the Plimsoll mark, rearranging the load will be only a small help. Economists who are obsessed with allocation to the exclusion of scale really deserve the environmentalists' criticism that they are busy rearranging deck chairs on the Titanic.*

Tradable permits make a clear distinction between scale and allocation decisions. The scale is decided by the environmental administration who issues permits and by doing so it sets the absolute limit of pollution. But afterwards the allocation is determined by the market, and every polluter can emit less or more than allowed by the permit. If the permit price is lower than its marginal abatement cost, it will benefit from doing less than required and buying a missing part of the permit. If the permit price is higher than its marginal abatement cost, it will benefit from doing more than required and selling the unused part of the permit. Even if someone has a large amount of money to spend on permits, what can be bought is equal to what somebody else abated; the sum of emission cannot be higher than what the administration decided (scale). At the same time administration does not interfere with what the market can do better (allocation).

This relates to the third advantage too. In non-tradable permit systems, environmental administration gives permits that are ultimately binding for the polluters. Often the administration is afraid of imposing strict environmental protection requirements. It has to judge whether the abatement cost required (implied by the permits) is prohibitive or not. But this abatement cost – which requires installing protection devices, changing the technology, or lowering the production – is hidden by the polluter. Every polluter blackmails the administration with shutting down the plant, firing workers, losing market competition to foreign companies, and so on. The administration is afraid of such results and gives up occasionally. Sometimes it is bribed by the polluter, and sometimes it is convinced that the abatement cost would be prohibitive or simply it does not have an opportunity to carry out a deeper analysis, and issues the permit just to have peace of mind. In the tradable permit system it does not have to study individual circumstances, because the requirements are not final – they can be changed by polluters themselves.

The fourth advantage will be explained in greater detail later on, in the story of phasing out lead additives to the American gasoline. Here I would like just to mention that tradable permits are excellent if one plans for a long run. If one plans to reduce some emission by 80% in 20 years, the abatement should progress 4% per year on average. It would be more reasonable to cut even more than 4%, but not every year; only when some modernisation takes place. Those who do not reduce the emission at a given moment can comply with the 4% requirement by buying an unused permit (from themselves – if they reduced the emission more than 4% earlier – or from somebody else).

Apart from these advantages, tradable permits have several traps environmental administration should avoid. Four of them are listed below:

1. Inappropriateness of combining with some other instruments
2. Possibility of cheating in the credit system
3. Possibility of discrediting market mechanism
4. Possibility of discrediting any financial mechanism

Tradable permits should not be applied jointly with subsidies for reducing emission that the permits relate to. Unfortunately, environmental administrators often fall into this trap. European Trading System (ETS) combined with subsidies for renewable energy provides an example of the trap. If a power plant that participates in ETS burns fossil fuels (and hence emits carbon dioxide) wants to increase its production, it may buy an unused permit from another plant that produces electricity in a wind mill, i.e. from a renewable source. Investor in this wind mill decided to build the turbine knowing that an unused carbon dioxide permit can be sold (since wind mill does not emit carbon dioxide). In addition, it received a subsidy dedicated for renewable sources. At the same time, some analysts are surprised that despite subsidies, the emission of carbon dioxide does not go below the limit defined by the ETS. It does not decrease, because – by definition – the emission is determined by the sum of permits allocated. Subsidies to those who do not use these permits do not translate into an additional emission reduction; they just provide them with additional financing, and let some other plants emit.

A possibility to cheat on emission reduction credits was already mentioned. Here I draw your attention to the fact that it does not refer to cap-and-trade systems. Only programmes where the total cap was not indicated are affected.

Implementations of tradable permits do not succeed always. Every failure provides an opportunity to ridicule the instrument, even when the failure was caused by something else (like in the example of combining them with subsidies). In recent years, a general disillusionment with financial mechanisms has taken place. People are angry at the financial sector, and they do not trust any mechanisms that have something to do with money.

Applications of tradable permits started several decades ago. The list below includes historical ones (mainly in the USA), contemporary ones, and future ones.

- Historical
  - USA: "bubbles" and "offsets"
  - USA: un-leaded gasoline
  - USA: "acid rain"

    - etc.
- Contemporary
    - ETS
    - China
    - harvest quota
    - etc.
- Potential
    - PL: "traditional" pollutants
    - World: climate protection

The first applications in the USA were in the 1970s. They were called "bubble policy", since the total emission under a hypothetical "bubble" covering many polluters was to be controlled. If the "bubble" covered several stacks of the same plant, and the plant was to decrease its emission by, say, 30%, then individual stacks did not have to abate 30% uniformly. Their total emission was supposed to decline by 30%. Moreover, the emission originating from some stacks could even increase, provided that other stacks decreased it by more than 30%, so that the overall target was reached.

If the "bubble" covered a region, then – like in Chorzów (a city in south-eastern Poland) – environmental administration required that the aggregate regional emission was controlled, not an emission from individual firms. So-called "offsetting" programme in the USA was based on this principle. If air concentration standards were not attained, no new plant could obtain an emission permit. However, in order to allow economic development, a regulation was passed that a new plant in a non-attainment region could obtain a permit for *x* tonnes, if it found an existing plant and decreased the emission in that plant by *x* tonnes (or financed the abatement and thus claimed an emission reduction "credit").

The greatest success story is the phasing out of leaded gasoline. In 1980 the American government decided to eliminate (in 6 years) the leaded gasoline. This was produced by 300 refineries. A traditional approach would be to announce the target and to oblige refineries to stop the production after 1986. In many countries such an approach has failed. When something was expected to happen, say, in 6 years, the government checked if it happened after 6 years, and discovered that nobody did anything. If nobody did anything, the government could not shut down all the plants. The only step the government could take was to express its disappointment, and to re-announce the programme.

In order not replicate this pattern, the US government decided to divide 100% into 6 (roughly 16.7% per year), and to announce that it would check every year if refineries decreased the production of leaded gasoline by the required rate: by 16.7% after the first year, by 33.3% after two years, by 50% after three years, and so on. After 6 years only the lead-free gasoline was to be produced. Those of you who know something about engineering are aware of the fact, that changing a production technology is not a gradual process. Either a refinery invests in a new technology, and the old production regime can be eliminated in 100%, or it does not invest, and the old production regime does not change at all. Some installations approach the end of their lives, and they will be substituted by new ones; if this is the case then it would be better to wait until this happens rather than to invest in some transitional solutions.

Aware of such circumstances, the American government announced that every year each refinery would get a permit to use 16.7% less lead additives, and they would check if it complied. However, the permits are tradable. Thus if a refinery did not use its permit, it can sell the unused part. A refinery which plans to change the technology sometime later can buy missing permits from other refineries that decreased the production of leaded gasoline faster.

The programme was a great success. In 6 years the production of leaded gasoline was eliminated completely. Several hundred million dollars were saved when compared to an alternative with non-tradable lead additive permits.

In 1990 the US Clean Air Act was amended. Sulphur dioxide permits (sulphur dioxide is the most important factor of the "acid rain") were made tradable. It was planned that there would be two separate markets divided by the Rocky Mountains. Finally the two markets were merged, and polluters located anywhere in the United States were allowed to trade with each other.

An interesting example of applying the mechanism in water management is provided by the Fox river, a small river in the state of Wisconsin (USA). It was contaminated by a dozen of paper mills and two towns. Scientists calculated what was the maximum pollution load to be tolerated by the river. The head of the local environmental administration invited directors of all the plants and mayors of the two towns, announced the total pollution limit, asked them to allocate this limit, gave them several hours to discuss, and promised to respect the allocation by giving the tradable permits equal to what they agree upon. Polluters participated in a conclave-like meeting, and after a time they declared the agreed allocation. By the way, over the next several years only one transaction took place (one of the paper mills transferred its permit to the municipality it got connected to). Some analysts look at this lack of trading as the failure of the mechanism. Others presume that perhaps the allocation agreed upon was optimal from the outset (i.e. it minimised the total abatement cost), and no later corrections were necessary. Whatever the truth is, the Fox river example demonstrates that administration did not have to be concerned about the predicament of individual polluters; they could delegate the allocation, and concentrate on making sure that the total abatement is consistent with what the ecosystem tolerates.

Contemporary applications are dominated by carbon dioxide markets. The largest one functions in Europe. This is so-called European Trading Scheme (ETS). It covers about 50% of the European emission, especially from power plants. It does not cover mobile sources, especially from cars. It was established in 2005. Initially the price of 1 tonne of CO_2 permit was 25 €. Then for almost 10 years it was close to zero. More recently – as a result of confiscation of permits for almost 900 million tonnes in 2014-2016 – it exceeded 25 € again (in 2022-2023 it was roughly 100 €/tonne). In other words, European polluters had an incentive to abate up to the cost of 100 €/tonne.

In China an experimental trade in carbon dioxide emission permits was carried out in 2018-2019. The total emission did not decline (since 2008 China has been the largest CO_2 emitter in the world), but this results from the very definition of the instrument. It is not quite clear how the Chinese market functioned. It cannot be linked to the ETS, since China – as opposed to the European Union – does not have an emission limit, and it does not wish to have any. Nevertheless it is important that the Chinese administration registered the emission from some plants (plants that were covered by the programme), and some market price emerged, and it informs economic agents that carbon dioxide emission is not free.

Tradable permits are applied not only in environmental protection. In some fisheries so-called tradable harvest quota have been established. They start with monitoring of populations living in certain aquifers. Based on these findings, it can be determined how much can be harvested sustainably (see the ninth lecture, ERE-9). The total limit is then allocated among the fishermen who are allowed to trade quotas they received (in order to adjust their allocation to harvested quantities).

Transferable Development Rights (TDR) have been advocated as a solution to equity disputes triggered by landscape protection. Let us assume that in order to protect the historical landscape of a town, there is a ban on erecting tall buildings. Owners of the land affected by this ban expect to be compensated for losses caused by constraints on their investment opportunities. The government offers subsidies that can be considered too low by the owners, and too generous by others. There is no good solution to this dilemma, unless TDRs are applied.

Here is how the mechanism can work. If the number of storeys consistent with landscape protection in the historical centre is, say, three, then the government can establish a general regulation for the area that the number of storeys is four; this affects both the historical centre and more distant locations. Owners of real estate in the historical centre are constrained by three storeys, so they cannot use the permit to erect a four-storey building. But they can sell the unused part to somebody else. If there is somebody in a more distant location who wants to erect, say, a seven-storey building, he (or she) has to buy additional permits over what he (or she) is entitled to, that is over three storeys. Through this arrangement the government is freed from analysing what level of compensation is fair. If the local economy is stagnant, investment opportunities are not very profitable, and compensations can be low. On the contrary, if the local economy is vibrant, investment opportunities are attractive, and compensations should be high. Under a TDR regime, governments do not have to decide on relevant disbursements; the market does it better.

Analyses of future applications are dominated by the carbon dioxide emission. Nevertheless in some countries – e.g. in Poland – environmental policy can use this instrument for regulating "traditional" pollutants, such as sulphur dioxide, nitrogen oxides, or particulate matter. Water management can use it too.

Climate change, and the emission of the carbon dioxide is a matter of global concern. For the time being, the emission is subject to some constraints in a small group of countries only (so-called Annex I to the *United Nations Framework Convention on Climate Change*, UN FCCC – a group of 40 countries considered economically developed). If the global emission – estimated at 48 billion tonnes per annum – is divided into 8 billion people (the world population), the emission per head would be 6 tonnes. If every country utilised such a per capita limit, the total emission would be exactly what we have now, i.e. 48 billion tonnes. In many wealthy countries the present emission is larger than this per capita limit, and in many poor countries – it is much lower; in China it is moderately higher than the hypothetical limit. If the limits were announced tradable, wealthy countries would have to buy from poor ones the unused parts of their permits. China would not have to buy too much in this market. Yet neither China, nor wealthy countries do not advocate for such a solution, since they are afraid of excessive money transfers implied by the global tradability of permits.

Tradable permits make a promising instrument of environmental policy. It is not considered a novelty any more, and it started to be applied in many countries. It has excellent characteristics: namely it relies on the market mechanism where markets perform very well, and leaves other decisions to be taken independently. Moreover it realises cost effectiveness in a perfect way.

Questions and answers to lecture 14

14.1 Are tradable permits a new controversial instrument of environmental protection?

No. They have existed for a half of a century, and they should not be seen as a novelty. Their alleged controversy may be caused by some proposals of strange applications. Probably the weirdest one was made by Garret Hardin (an American biologist) in 1968 who suggested that each person has the right to have one child; married couples would be entitled to have two children. Somebody who did not use his (or her) right could sell it to somebody else. And conversely, if a person wants to have more children, he (or she) has to buy the right(s) from somebody else. This is a very strange idea and impossible to implement, but – perhaps – for many people this is where the controversy comes from. In environmental protection tradable permits are not controversial at all.

14.2 Why did the Chorzów project refer to tradable permits metaphorically rather than literally?

The emission originating from two plants located in Chorzów was much higher than what could be tolerated. In 1991 local administration sought ways to cut it by 50%. Unfortunately – because permit trading was not allowed in Poland at that time at all – the administration could not fully implement this approach with respect to the two plants (the Kosciuszko steel mill, and the Chorzów power plant). A standard solution would be to write uniform emission permits for 50% of what they emitted in the past, irrespective of what cost this implies for either of the plants. This makes little sense, because the plant where the abatement cost is very high (in this case: the power plant) would be required to reduce the emission by the same rate – i.e. 50% – which is expected from the other one where it is cheaper (in this case: steel mill). It would have been better to write a much stricter permit for the low abatement cost plant, and a more generous one for the high cost one, while the resulting emission is what the town can tolerate (i.e. 50%). Yet – because of the legal framework – this was difficult.

"Emissions trading" was incompatible with the existing legal framework. In order to avoid protests, the low-cost plant (steel mill) was to receive a "compensation" from the high-cost plant (power plant). However direct payments *à la* "emissions trading" were impossible. The administrator used the regional environmental fund to make an appropriate financial transfer. If environmental regulations allowed "emissions trading", the permits would have gone where they can be used cost-effectively without any intervention. But regulations did not allow for that. Hence the regional administrator had to justify the strict permit for the steel mill, the generous one for the power plant, and – with the help of the regional environmental fund – transferred the money serving as a compensation. Hence the experiment involved "emissions trading" in a metaphorical sense only. It did not involve "emissions trading" literally.

14.3 Where does the motivation to protect the environment come from in tradable permit systems?

The polluter who abates more than required by the permit can sell its unused part. The polluter has a motivation to do it, if the own abatement cost is lower than for a potential buyer. And even if there are no other polluters, but permits are issued for several periods to a single polluter, it can sell to itself, if it wishes to abate more than required in the beginning, and less – later on. The polluter can buy from itself (from its future allowance), if it plans to abate less than required in the beginning, and more – later on. Tradable permits let a target (the emission in an area or in a period of time) be achieved at the minimum cost; they are not introduced in order to minimise the pollution in the short run. In the long run, however, they may encourage setting more ambitious targets. Increasing the level of ambition is often difficult since people are afraid of excessive abatement cost. If this cost – thanks to tradable permits – turns out reasonable for an economy or a region, people are not so afraid of increasing the level of ambition.

14.4 What is the difference between cap-and-trade and credit systems?

Cheating in the cap-and-trade systems is more difficult, since there is a clear baseline for emission reductions. In the credit system – if the overall cap is not determined – emission reduction is more difficult to calculate (see also the answer to question 14.7).

14.5 Car manufacturers have to comply with certain standards for CO_2 emission per kilometre driven. What is the difference between this regulation and tradable permit system with a relative limit?

If car manufacturers were to comply with limits for actually produced cars, or – even better – for actually driven kilometres, then the regulation would be equivalent to tradable permit system with a relative limit. But they are to comply with something else. Namely every car manufacturer has to make sure that its engine combusts fuel with certain efficiency, so that driving 1 km causes the emission of, say, 120 g of CO_2, or something like that. This limit is uniform irrespective of whether the number of such cars produced is 10,000 or 100,000. Moreover, it is identical irrespective of whether the car has an annual mileage of 10,000 km or 50,000 km. In the first case a single car will emit 1.2 t CO per annum, and in the second – 6 t CO_2. This has nothing to do with tradable permits.

14.6 Is "grandfathering" always more attractive for emitters than an auction-based initial allocation?

It is often assumed that emitters prefer "grandfathering" rather than auction, since in the first case they get the permits for free, while in the second case they need to buy them. This is not a universal rule though. If a firm wants to enter the market, and it did not emit anything earlier, then according to the "grandfather principle" it has a worse position than a firm that was there before already. If an auction is applied, then the old ones and the new ones are subject to the same requirements. Hence for new emitters, auctioning is better than "grandfathering".

14.7 Why is "emissions trading" within CDM environmentally disruptive?

CDM (*Clean Development Mechanism*) promotes emission growth disguised as its decline. For instance, we can look at an Indian entrepreneur who announces that he would build a refinery with the annual CO_2 emission of 2 million tonnes. Environmental activists protest since this will increase the emission from India by 2 million tonnes, and will aggravate the global problem. The entrepreneur states that climate protection is given a very high priority on his agenda, but – because of the planned technology – the emission cannot be lower. Activists say that this is not true, and argue that by applying a newer technology, the emission can be made as low as 1.5 million tonnes. The entrepreneur analyses this technology and estimates that the cost of such a refinery would be 1 million € more than he planned. The activists assure him that this additional cost will be paid by one of the Annex I countries of the United Nations Framework Convention on Climate Change (UN FCCC) – a group of countries considered developed in the 1970s (Poland and Bulgaria belong to this group, while South Korea does not).

Annex I countries are required to abate but they can comply with their obligations by buying "emission reduction credits" from a non-Annex I country (which – like India – does not have any binding commitment). One of the Annex I countries is likely to buy "emission reduction credit" of 0.5 million tonnes (2-1.5=0.5). Even if it was to pay, say, 4 €/t of CO_2 "abated" in this way, it will be cheaper than abating at home (the domestic marginal abatement cost is probably higher than 30 €/t). Both sides find this transaction profitable. An Annex I country pays 2 million € for "emission reduction" of 0.5 million tonnes (perhaps the abatement cost required at home would have been 15 million € or more). The Indian entrepreneur is also happy, since he will get an external financing of the incremental cost of a newer technology. The global emission will go up by 1.5 million tonnes (it will not decrease by 0.5 million tonnes!), and the whole project will be subsidised by an Annex I country (2 million €). "Emissions trading" makes sense only when all its participants have emission limits. CDM involves trading with countries that do not have such a limit, and therefore it is environmentally disruptive.

14.8 "Banking" means keeping unused permits (to make the future emission easier), but also borrowing permits from what is expected in the future (to make the present emission easier). Is there a symmetry between both variants of "banking"?

No. There is a symmetry between both variants only if the emitter exists all the time covered by the "banking". Nevertheless there is a risk of bankruptcy. A firm who "borrowed" a permit from its future allowance should give it back later. However, if it went bankrupt, then it will never give it back. Thus emitting more than allowed turned out to be cheaper, because the firm does not pay for additional permits (if it goes bankrupt).

14.9 How do tradable permits comply with Herman Daly's principle of separating scale and allocation decisions?

The principle was explained by a quote on page 114. Tradable permits make a clear distinction between scale and allocation decisions. The scale is decided by the environmental administration and the allocation is determined by the market. Every polluter can emit less or more than required by the administration, but the total emission cannot exceed the scale.

14.10 In tradable permit systems the administration does not have to investigate economic hardships resulting from giving an excessively sharp limit of emission. Why?

Because issuing a sharp permit does not necessarily imply bankruptcy, firing workers, etc. In tradable permit systems no tasks have to be considered ultimate and final ones. If a firm wants to emit more than requested by authorities, it can buy an additional permit from a polluter who managed to abate. Of course the price of such a permit may turn out to be excessive, and the firm anyway will go bankrupt, fire workers, etc. Nevertheless the government is not the only responsible for the problem. This reduces some of the psychological pressure on environmental administration.

14.11 Was the "bubble policy" safe for the environment?

In principle, yes. If the total emission in a given area was not allowed to grow, then the fact that one source emitted more while a neighbouring one emitted less should not do any harm. The only frauds – in the "credit system" – are possible when emission reduction took place indeed, but the original emission was not included in local environmental protection inventories (like dust emitted by a gravel parking lot).

14.12 In the American programme to fight "acid rain" after 1990, emissions trading was allowed over the entire territory of the United States. Hence emitters from the Pacific coast could sell unused permits to emitters on the Atlantic coast. Did not it threaten formation of excessive emission in the East coast (or conversely, if emitters from the Atlantic coast sold their unused permits to emitters in the West coast)?

People were afraid of such an outcome, and initially the country was divided into two parts: East of the Rocky Mountains, and West of the Rocky Mountains. Trading was allowed within each of the two parts only. Finally, however, trading was allowed over the entire territory. In theory, it was possible that a "hot spot" emerges somewhere, if an emitter from the East coast buys a large number of unused permits from the West coast (or *vice versa*), and plans a very high emission. This has never happened since every plant had to hold two permits: a tradable one (aimed at fighting "acid rain" in North America), and a non-tradable one (aimed at protecting human health and forests, preventing corrosion, etc.). Therefore even if an emitter bought a large number of tradable permits from distant places, its emission was also constrained by the local non-tradable permit.

14.13 Why are "tradable harvest quota" considered a variant of "tradable permits"?

Because the logic of separating scale and allocation decision is the same. The scale decision is taken by an organ who can answer the question of sustainable harvest. Allocation decisions are taken by the market (by trading individual harvest quota, fishermen decide who is going to catch how much, but the total catch has been decided already before).

14.14 *Backloading*, that is withdrawing from the market a part of carbon dioxide permits, was expected to counteract low ETS (*Emission Trading System*) permit prices. Did it respect ownership rights?

In 2013 the European Commission decided to reduce its permission for Member State governments to auction tradable permits by 900 million tonnes in 2014-2016. Every emitter participating in the ETS has an individual account where all the permits are recorded (taking into account all market transactions). If somebody wanted to nationalise, expropriate or steal these assets, property rights would have been violated. Yet "backloading" implies claiming of what governments were supposed to sell in auctions. Thus it does not interfere with private property; it affects public domain. The "backloading" decision was taken in a procedure appropriate for international negotiations. The European Commission asked for what was "owned" by governments rather than private entities.

14.15 What are the prospects for establishing a global tradable permit market for CO_2?

"Emissions trading" is fairly easy and obvious once the market has been established. But the global market for carbon dioxide has not been established yet. In order to establish it, two conditions must be satisfied. First, all signatories of the UN FCCC need to have binding commitments to reduce their emission below a certain level. Second, an allocation of these commitments has to be agreed upon. For the time being the world (197 member countries of the United Nations) is divided into two groups: Annex I, that is roughly 40 countries considered economically developed; and non-Annex I countries, i.e. the rest of the world. The first group has certain emission limits assigned, while the other one does not have any, and struggles not to accept any. Signed in 2015, the Paris Agreement makes the first (very modest) step towards breaking the impasse, and accepting the necessity to reduce emission in the second group too. Only when both conditions are satisfied, a global market can be established.

Either condition is difficult. The first one – requiring that everybody has to take a commitment – requires to break the impasse indicated in the previous paragraph. But once it is achieved, another controversy will come to the fore; the signatories will have to agree on an allocation principle. A possible principle – outlined in the lecture on page 118 – is to divide the present annual emission (i.e. roughly 48 billion tonnes of CO_2) into the world population (i.e. roughly 8 billion people). This means the allocation of roughly 6 tonnes per head. It seems fair, but it is not promoted (neither by the developing countries, nor the developed ones). Perhaps it is too abstract for developing countries to become politically attractive; they simply object to accepting any commitments. Annex I countries did accept commitments, but they prefer to allocate permits not only according to a per capita criterion – e.g. 6 tonnes per head – but also by taking into account GDP per capita. In other words, according to this philosophy, a wealthy country should be given a limit higher than 6 tonnes per head, since its citizens are used to a higher standard of living, and therefore to a higher CO_2 emission. Prospects for establishing a global tradable permit market are dim.

15. Environmental tax reform

The last lecture is devoted to a problem that has frustrated many environmentalists for several decades. They cannot tolerate the fact that taxes are imposed on goods that do not damage the environment, and products and activities that are ecologically harmful are subject to charges too low to provide sufficient incentives to protect the environment. The state budget has to have certain revenues obviously, but why not to base it on the principle: "Tax bads not goods!" This is the general idea of the so-called *Environmental Tax Reform* (ETR).

Total		505
Including taxes		*465*
Of which:		
	VAT	230
	Excise	80
	Including fuels	*34*
	PIT	68
	CIT	70

At the beginning we will look at the Polish state budget in 2022. In the table above all revenues are denominated in billion PLN. In Poland (2022), VAT, excise, CIT, PIT provided 230 BPLN, 80 BPLN, 68 BPLN, and 70 BPLN, respectively (the state budget revenues were 505 BPLN, of which tax revenues were 465 BPLN

The table indicates that the state budget – totalling 505 billion PLN – consists mainly of tax revenues. These in turn consist mainly of VAT (*Value Added Tax*), followed by excise tax. The latter is imposed on some commodities only. Motor fuels – gasoline and diesel oil – are among those. Paid by us PIT (*Personal Income Tax*) brings less revenues. Also CIT (*Corporate Income Tax*) provides less than 15% of the budgetary revenues. In many countries the structure of the budget is similar; VAT is its most important component.

VAT is imposed almost uniformly on virtually all commodities. In Poland its rate is 23%. Some commodities – especially those considered basic goods – are taxed lower. For example, in Poland the VAT rate for bread is 5%.

A key feature of VAT is taxing the added value only. Thus if some components (such as raw materials) were used in the production process, and they were already taxed, then their value is not to be taxed again. For instance, if the product sold for 300 PLN used something that was bought for 200 PLN then only 100 PLN is subject to the tax. Technically this is done by imposing VAT on everything what was sold, that is 69 PLN (23% of 300), and subtracting what was already collected, that is 46 PLN (23% of 200). This principle seems simple, but in fact it is difficult to implement in a fraud-proof way. A typical offence is to claim a VAT refund allegedly paid by somebody at an earlier stage of the production process even though it was not paid. Despite that, VAT revenues make the most important component of state budgets almost everywhere.

The excise tax differs from VAT in that it is not refundable. Its rate is usually higher than for VAT. Most often excise is imposed on commodities characterised by an inelastic demand (the demand which does not react to price changes strongly), on luxury goods, or on goods considered harmful. Energy demand is fairly inelastic, so it is often subject to excise taxes. Perfume is a prime example of a luxury good. Excise on perfume provides high budgetary revenues and governments are not afraid of hurting the poorest citizens. Tobacco and alcohol are subject to high excise taxes, because their consumption is harmful for human health. The limits to excise are set by the black market risk. Increasing excise taxes to a very high level provides incentives for illegal transactions.

The frustration of environmentalists is caused by the fact that tax rates ignore the "ecological friendliness" of commodities. At the same time environmental charges and non-compliance fees, that Poles are so proud of, provide roughly 2 billion PLN, i.e. much less than 1% of tax revenues. Can any tax revenues (other than emission charges) be identified as "environmentally motivated" somehow?

Excise on fuels plays such a role in state budgets. Its rates are not diversified according to a fuel's environmental friendliness, but it can be assumed that any fuel is environmentally disruptive in its "life cycle" (from extraction, to transport, and to combustion). In many countries – especially in the European Union – there is a concept of "environmentally motivated taxes". Excise on fuels is an important element of this concept. In Poland excise on fuels provides many times more than all environmental fees.

Car emission is not subject to environmental fees, but excise on fuels is considered an incentive for environmental protection since it motivates to lower the demand for transport. Yet the incentive is very weak, and besides taxing rates are not related to environmental harmfulness of fuels and vehicles that use them. In particular, the excise on diesel used to be lower than the excise on gasoline, despite the toxicity of exhaust gases. Diesel engines are more difficult to control and they often pollute more than engines burning gasoline. Despite such controversies, excise on fuels is usually the most important component of "environmentally motivated taxes".

In professional (economic) terms, "environmentally motivated taxes" are levied on externality generating activities and on extraction of natural resources.

This problem was analysed by the OECD. It turned out that the share of such taxes in state budgets is fairly low. Given that budgets typically correspond to 50% of GDP, their share in GDP is twice as low. Please note that "environmental motivation" is not necessarily reflected in the official terminology (an "environmentally motivated" tax does not have to be called "environmental"). Shares of "environmentally motivated taxes" in GDP are estimated at:

- Average for EU – 2%-4%
- In Poland – 3%
- In Denmark – 6%

The number for Poland is not far from what can be found in other EU countries. Denmark is an outlier. Its share of "environmentally motivated taxes" is unique. Governments are constrained in setting high excise taxes on fuels by so-called "tanking tourism": if the excise tax on fuels is higher than in a neighbouring country, then there is a risk that drivers will tank abroad. In Denmark this risk is low, since the border with Germany is fairly far from where most of the Danes live (besides, the German excise is not lower than the Danish one). One can also travel from Copenhagen to Sweden over a long bridge, but the difference in excise is not large: in Denmark it is 0.63 €/l, and in Sweden it is 0.62 €/l. I believe that very few inhabitants of Copenhagen make an effort to travel to Sweden in order to tank cheaper. There are other burdens – like huge registration fees – the Danish drivers are exposed to (but do not complain about). Moreover in all Scandinavian countries governments enjoy extremely high confidence. Hence citizens tolerate even extraordinarily high taxes. They are wealthy, reveal high willingness to pay for environmental protection, and they expect that if they pay a tax, the government will use the money more efficiently than if they spend the same amount themselves (privately).

Denmark is unique having the 6% share of "environmentally motivated taxes" in GDP, that is the 12% share in budgetary revenues. This is much more than anywhere else in the world. The *European Environment Agency* (EEA) in Copenhagen acknowledged that even the 15% share of "environmentally motivated taxes" in budgetary revenues makes a "drastic" environmental tax reform. In other words, replicating the Danish experience even at a slightly larger scale is a challenge requiring governments to make massive efforts.

The scepticism of finance ministers who hesitate to implement an environmental tax reform is caused by their concern about the stability of budgetary revenues. No matter what objections one can raise against VAT, it gives comfortably stable revenues. People will always buy something, and thus they will always give some money to the budget. There is no risk that they will stop spending money and the minister of finance will announce that defence, education, police, health care – tasks that every government has to fulfil – are left without funding.

In contrast, environmental taxes are characterised by what economists call "self-erosion of the tax base". Please imagine a tax on mercury (Hg) which is a very toxic substance. In the past it was considered a necessary component of thermometers. But traditional thermometers are now substituted by electronic ones that do not require mercury. How does this translate into revenues from a hypothetical mercury tax? Initially they could be quite high. Later on, when mercury starts to be withdrawn from production, they would decline. If the unit rate of such a tax increases, the motivation to look for mercury substitutes becomes even stronger. Abandoning mercury completely – even when the tax rate is extremely high – leads to zero revenues.

Or, take the water example. Its withdrawals from rivers and other aquifers can be environmentally disruptive, and an appropriate tax will provide an incentive to reduce them. Will the water withdrawals decline as a result of such a tax? Probably yes – this is just a matter of cost. Even complete elimination of the water withdrawals – like in a spaceship, where water flows in a closed circuit – is possible. Closing circuits can be done on earth too, but it is very expensive. However, people may be motivated to do so, if a sufficiently large tax is imposed on water withdrawals. What would be the reaction of finance minister? He (or she) would have to look for some other tax in order to keep budgetary revenues as before.

Self-erosion of the tax base is the reason why finance ministers stick to VAT which lets revenue be stable. Almost every item linked to environmental protection leads to the erosion of tax base which makes environmentalists happy, but makes the finance minister concerned. Any raw material such as mercury or water can be eliminated or recycled. There is only one input that – because of the laws of physics – cannot be eliminated completely. This is energy. According to the second law of thermodynamics (that you should remember from your high-school physics), its use leads to wasting some of its amount, so that its full recycling is impossible. Taxing energy is basically the only direction of ETR which is not to be objected by finance ministers. But revenues from this source can finance only a portion of the budget. That is why EEA does not recommend an ETR leading to more than the 15% share of environmentally motivated taxes in budgetary revenues.

In 1990s the idea of a "double dividend" from environmental taxation emerged. Its main outcome (the "first dividend") is to provide incentives for environmental protection. Additionally (and this is the "second dividend") it lets achieve economic objectives. For instance – according to this logic – higher taxes on raw materials would allow for lowering labour taxation, and thus for a lower unemployment. Or higher taxes on imported fossil fuels can provide incentives for investing in renewable energy sources and hence improving energy security. The "first dividend" is a lower emission from burning fossil fuels, and the "second dividend" is independence of imports. There are also concepts of the "third dividend" as an additional reason for ETR.

Nevertheless state-of-art economic research using so-called general equilibrium models (models that look at all interconnections between economic sectors) raise doubts regarding these common sense conclusions. Tax systems that exist in the world are so complicated that changes introduced in one part may trigger unexpected repercussions somewhere else and eliminate additional "dividends".

Arguments raised in good faith are accompanied by opinions voiced by certain economic agents about inappropriateness of taxing of their activities. Economists call this "political economy". For instance, German car manufacturers – that run on diesel oil – lobbied effectively for a low excise tax on this fuel by arguing that low diesel oil prices support farmers who drive tractors. Or some people can argue against water taxation by arguing that poor and old citizens require protection (even though the low price of water benefits first of all those who consume large quantities of water, that is wealthy citizens), and so on.

The reasons for not introducing ETR can be summarised in the following way:

1. "Political economy"
2. Self-erosion of the tax base (energy is the only commodity without a risk of the total elimination of the tax base)
3. Doubts with respect to arguments about multiple dividends (double dividend, triple dividend, etc.)

For the time being, state budgets in all countries are based on taxes that – like VAT – do not take into account environmental protection. The essence of ETR is linking the tax rates to the environmental harmfulness. Deeper analyses demonstrate that this would destabilise budgetary revenues though. Consequently expectations of environmentalists to restructure the budget completely are not realistic. The only realistic variant of ETR is to increase the taxation of energy.

It should be emphasised that every technology of energy conversion is environmentally harmful. Of course burning fossil fuels leads to the well-known climate change effects. Yet renewable energy production is not problem-free either. Hydroelectricity plants disrupt river ecosystems, windmills are dangerous for birds, and photovoltaic panels require so-called rare earth metals, whose extraction – mainly in China – causes tragic results for people and landscape. Thus taxing energy from any source is motivated by environmental protection.

There is an additional argument for energy taxation. Energy is used by everybody. Wealthy people use more energy than poor ones. Thus – if energy is taxed higher – it is the wealthy rather than poor who bear a higher burden. It is true that in poor households' budgets energy expenditures have a larger share. Consequently increasing the energy taxation may turn out to be "regressive" ("regressivity" means that in percentage terms the poor pay more than the wealthy; governments try to avoid it). This is not inevitable though. Higher energy taxation can be combined with lower tax rates on items bought predominantly by the poor or simply combined with subsidies addressed to the poor. There are both environmental and social arguments for higher energy taxes.

Questions and answers to lecture 15

15.1 What components is a state budget composed of?

State budget revenues originate from taxes mainly. Some revenues come from tariff duties, and other sources (e.g. dividends from state owned enterprises), but they are of secondary importance. In Poland taxes give more than 90% of budgetary revenues. VAT provides a half of these revenues. Another important source of revenues is excise tax. In 2022 Personal Income Tax (PIT) gave less than 15% of revenues only, almost the same contribution as from

Corporate Income Tax (CIT). PIT revenues were 68 billion PLN. Taking into account that there are more than 20 million taxpayers, this corresponds to less than 3500 PLN per head annually, that is less than 300 PLN per month. We contribute to the state budget much more whenever we buy commodities, paying prices that include VAT.

15.2 Why are CIT and PIT revenues smaller than those originating from VAT?

Some people take advantage of tax exemptions, reliefs, allowances, etc. The more wealthy the taxpayer is, the more reasons he (or she) identifies to pay less. In the case of firms paying CIT rather than PIT, exemptions are particularly numerous. An important reason for exemption is spending profits on investment. In some countries almost all investment expenditures are subtracted from the taxable income. Practised by many firms, so-called "tax optimisation" means manipulating financial flows so as to make taxable items as small as possible. In the case of multinational firms this procedure – called *transfer pricing* – boils down to making transactions between subsidiaries located in various countries, such that lower income is recorded where tax rate is high, and a higher income is recorded where the tax rate is low. For instance, if a firm in country X produces something which requires an import from its subsidiary located in Y, then the price the former pays to the latter can be either too low or too high. It will be too high, if tax on profits in X is higher than in Y (since the subsidiary in X will show a lower profit, or perhaps even a loss, and the subsidiary in Y will show a higher profit; in total the firm will pay lower taxes). The price will be too low otherwise. All finance ministers try to fight against *transfer pricing*, but – as a rule – firms reveal greater creativity, and they are shrewder than administrators. Consequently revenues from CIT are small – almost as small as those from PIT.

15.3 Do environmentally motivated taxes include Polish environmental fees?

It is difficult to say. Environmental fees are not *sensu stricto* taxes. They are collected in special funds which are separate from the budget; the funds are more flexible than the budget, and the National Fund for Environmental Protection (where these fees go to) is a legal person (local branches do not have the same legal status). But nevertheless the state budget takes these revenues into account sometimes too. The President of the National Fund is appointed by a council whose members are nominated by the Minister of Environment. Therefore, in a sense, he (or she) should be considered an official government organ. The most important reason for excluding Polish environmental fees from "environmentally motivated taxes" is the rule that fees are considered production costs. Thus they decrease the profit. In contrast, the *sensu stricto* taxes are paid from profits. Polish environmental fees (around 2 billion PLN) cannot change the overall conclusion that environmentally motivated taxes account for 6% of the state budget.

15.4 Are Polish environmentally motivated taxes lower than in the European Union?

No. Both with respect to the state budget, and with respect to GDP, they are in the middle of intervals typical for the European Union countries.

15.5 What are the chances that countries decide to implement ETR unilaterally?

The chances are rather weak. ETR implies a higher taxation of energy. If a country does this unilaterally, its tax base is likely to erode, since energy intensive products will be cheaper when imported rather than produced domestically. No country has a motivation to do it unilaterally.

15.6 Why does the European Environment Agency (*EEA*) call a proposal to increase the share of environmentally motivated taxes in state budgets to 15% a "drastic" ETR?

The share of environmentally motivated taxes in the Danish budget is 12%, and this is much more than anywhere else in the world. EEA expects that exceeding the Danish pattern slightly is possible, but a more intensive substitution of traditional taxes with "environmental" ones is not realistic.

15.7 What is the "self-erosion of the tax base"?

People prefer to pay as little as possible. If a tax goes up, then they try to avoid making transactions involving this tax. Sometimes they manage to accomplish it by changing production (or consumption) patterns, sometimes they decrease the use of a material (affected by the tax), and sometimes they close circuits (they recycle). Irrespective of how they react to the increased tax, the demand for what is taxed more heavily goes down. Economists call it "erosion of the tax base".

15.8 Why does not energy taxation lead to the "self-erosion of the tax base" to the same extent as other environmentally motivated taxes?

The Second Law of Thermodynamics (that all high-school graduates should remember) implies this. The Law states that no energy conversion has 100% efficiency; some part of the energy is dissipated (it is called entropy) and cannot be utilised. For example in a coal fired power plant out of 100 units of energy stored in the fuel at most 60 can be transformed into electricity. The rest transforms into heat. Some of it can be used in order to distribute so-called "network heat". But at most 30 units can be captured in this way. The rest – that is at least 10 units – will be dissipated in the neighbourhood of the plant, and cannot be utilised. In hydropower plants the potential energy of water can be used either to produce electricity or to fill the upper reservoir (in pumped storage power plants), but 100% efficiency will never be achieved. A part of the original energy will be lost irreversibly in the form of friction, noise, or otherwise. Levying taxes on energy leads to a lower demand for it, but – contrary to other raw materials, e.g. water – it does not let close circuits completely. *Perpetuum mobile* is not possible. Storing energy is difficult too, even though new technologies allow us to do it better and cheaper now than a couple of years ago. Nevertheless, irrespective of solving the storage problems, the Second Law of Thermodynamics does not let achieve the 100% efficiency of energy conversions.

15.9 What does the "double dividend" of environmental taxes mean?

The direct effect of an environmental tax (the "first dividend") is a reduced pressure on the environment. But apart from that, the tax can make other commodities more attractive (in relative terms). For instance, a tax on fossil fuels makes the electricity produced out of these more expensive. As a result, the electricity produced out of renewable sources becomes more competitive. Renewable energy sources allow independence from traditional power plants and make the economy less "addicted" (to importers or other suppliers). Energy security can be identified as the "second dividend".

Sometimes environmental taxes promote exports, and this is considered the second dividend. Sometimes, they imply a higher employment (see also question 15.10), and this is considered the second dividend then.

15.10 Why does taxing environmentally harmful products make a chance for lowering labour taxes?

Taxing an environmentally harmful product (that was taxed less or was not taxed at all before) provides the budget with an additional revenue. If the budget is to have the same total as before, other taxes can be lowered. For instance, labour is taxed everywhere. If it becomes being taxed lower, then the demand for it goes up (something all governments would like to see).

15.11 What is the "political economy" of lobbying for or against taxes?

"Political economy" explains what forces act for or against a specific movement. Usually nobody recommends to be forced to pay more; on the contrary, everybody recommends that he (or she) should pay less. Since all economic agents are likely to convey the same message, the most effective ones are those who are the loudest ones, or who provide the most convincing arguments. The arguments against taxation are often disguised as opinions in favour of public interest rather than private one. A prime example is the argument used by Mercedes-Benz aimed at taxing diesel lower than gasoline. It did not refer to the well-being of car owners, but rather to the well-being of farmers who drive tractors. Even though this is a nonsensical argument, it used to be effective for many years.

15.12 Everybody knows that burning coal and hydrocarbons (oil and gas) damages the environment. But in fact, any energy conversion process is environmentally harmful. Why?

Environmental disruption caused by traditional power plants is obvious and well documented. At the same time, producing electricity from renewable sources is considered environmentally friendly. It has to be pointed out that it is environmentally harmful as well (of course to a lower extent). Hydroelectricity makes a good solution in some circumstances. There are not many countries (think of Norway as an exception), where the potential energy of water in mountain rivers can be easily used for electricity production at a significant scale. Most often hydropower requires constructing dams, building artificial reservoirs, etc. These necessary objects are environmentally disruptive. Windmills are considered environmentally friendly, but ornithologists inform that for birds – especially during the night – they are dangerous. Even photovoltaic panels, that seem to be completely safe, are responsible for unprecedented damages at the production stage; they require so-called rare earth metals, whose extraction causes enormous damages. Only the direct solar energy use (for instance drying your clothes outside) – difficult to tax, by the way – is free from any harmful impact on the environment.