# Lecture notes on common-pool externalities*

Matti Liski

April 12, 2019†

**Abstract**

These notes present simple models to analyze the basic common-pool externalities relevant for resources that publicly owned. We identify where and how exactly the inefficiency arises in common-pool situations. We also formalize the solutions to the externality problems and explicit bargaining solutions are introduced. We also discuss transaction costs and introduce the reasons that prevent the market from achieving efficiency through bargaining.

*Key words*: Natural resources, pollution, common-pool problem, public goods, commitment, market power, technology, game theory

*Section 2 builds on "Games and Resources", Bård Harstad & Matti Liski (2013), In: Shogren, J.F., (ed.) "Encyclopedia of Energy, Natural Resource, and Environmental Economics", Vol. 2, pp. 299-308 Amsterdam: Elsevier.

†Edited and abridged by Iivo Vehviläinen in 2020.

# 1 Introduction

The purpose of these notes is to be explicit about the nature of strategic interaction among agents engaged in exploiting a common resource. Game theory is a natural tool for this purpose. However, reading these notes does not require prior knowledge of game theory, as we will go through a sequence of relatively simple strategic situations, keeping the focus on the substance matter and on explicit solutions. Throughout the text, references are made to resource use, activity level, output, and production. In all cases, these terms refer to the private activity contributing to the exploitation of the common resource. This can be, for example, private output increasing the pollution stock that reduces the common resource such as clean air.

# 2 Public Resources

The publicly owned resource is measured by $s \geq 0$ and the set of users is $\mathcal{I} = \{1, 2, ..., n\}$. Each user, $i \in \mathcal{I}$, decides on the private action $z_i$ impacting the total resource stock. In particular, $i$'s private utility is function:

$$U(x_i, z_i, s) = u(x_i, z_i) - d(s), \text{ and} \tag{1}$$

$$s = \sum_{i \in \mathcal{I}} z_i \tag{2}$$

Variable $x_i$ represents other decisions made by user $i$, assumed to not generate any direct externalities. It can be interpreted as a private valuation parameter or a technological choice, for example, investments in technologies that increase the private value of using the resource. We assume $u(\cdot)$ is concave in $x_i$ and that $d(s)$ is convex.

**Example 1:** We can call $s$ as the total pollution stock, and $z_i$ as emissions from agent $i$ (e.g., country, firm, or individual). Then, it is natural to define $d(s)$ as the pollution damage suffered by each agent. Assume that the damage is quadratic,

$$d(s) = \frac{d}{2}\Big(\sum_{i \in \mathcal{I}} z_i\Big)^2 \tag{3}$$

where $d > 0$ is a constant. Then, the marginal damage per individual from increasing the stock, $d'(s) = ds$, is linear in the stock. To complete the example, we may assume that the private marginal benefit is linear in $z_i$,

$$u(x_i, z_i) = u(x_i) z_i. \tag{4}$$

We can think that $u(x_i)$ is, for example, the monetary value of selling outputs produced. The choice of technology $x_i$ can impact this value.

**Example 2:** In some situations it is reasonable to think that the marginal environmental damage is linear. Then, we would assume

$$d(s) = D\Big(\sum_{i \in \mathcal{I}} z_i\Big) \tag{5}$$

so that $d'(s) = D > 0$ is a constant for all stock levels. For example, in climate change, this a reasonable approximation at least for a range of stock levels for greenhouse gases.

**Example 3:** For example, $s$ may represent the stock of clean air rather than stock of pollution if $z_i$ is $i$'s pollution level, reducing this stock. Then, we should think $s = s_0 - \sum_{i \in \mathcal{I}} z_i$, so that agents extract from the resource. The environmental harm from pollution is then given by the convex function $c\left(\sum_{\mathcal{I}} z_i\right) \equiv v\left(s_0\right) - v\left(s_0 - \sum_{\mathcal{I}} z_i\right)$, where $v(s)$ is a concave function of the benefits from having a clean environment, and $v\left(s_0\right)$ is simply a constant.

## 2.1 The Static Common Pool Problem

We refer to the *first-best*, denoted below by superscript $FB$, as the outcome preferred by a social planner maximizing the sum of the users' payoffs.[1]

**Definition 1.** *The (interior) first-best is an allocation of* $\{x_i, z_i\}_{i \in \mathcal{I}}$ *maximizing the sum of utilities:*

$$\max_{\{x_i\},\{z_i\}} \quad \sum_{i \in \mathcal{I}} U\left(x_i, z_i, s\right)$$

$$\Leftrightarrow$$

$$\underbrace{\frac{\partial u\left(x_i, z_i\right)}{\partial z_i}}_{\text{private gain}} = \underbrace{n \frac{\partial d\left(s\right)}{\partial s}}_{\text{private \&public damage}} \quad \text{and} \quad \frac{\partial u\left(x_i, z_i\right)}{\partial x_i} = 0.$$

In the first best, the private gain from increasing the activity level of $i$ is balanced against the full social loss, that is, the marginal damage that $i$ suffers and the damage that all others suffer. This means $n$ times the marginal damage is full the social marginal cost.

However, in the game referred to as the *problem of the commons*, the variables are not chosen by a social planner, but by the individual users. When all the users act simultaneously, the natural equilibrium concept is Nash equlibrium. We use superscript $NE$ to denote the Nash Equilibrium when convenient.

**Definition 2.** *A Nash equilibrium is a set of choices,* $\{x_i, z_i\}_{i \in \mathcal{I}}$ *, such that, when user* $i$ *takes as given the other users' choices,* $\{x_j, z_j\}_{j \in I \backslash i}$ *, then* $\{x_i, z_i\}$ *maximizes private payoff* $U\left(x_i, z_i, s\right)$.

The static version of the game is the one where all users act simultaneously and only once. Later, we will look at situations where agents can first choose technologies and after this decide on the use of the common pool.

**Proposition 1** *Consider the static common-pool problem: (i) Each user's choice* $z_i$ *is larger than the first-best:*

$$\frac{\partial u\left(x_i, z_i\right)}{\partial z_i} = \frac{\partial d\left(s\right)}{\partial s} < n \frac{\partial d\left(s\right)}{\partial s}.$$

[1] The second-order conditions for the first best problem are satisfied when $|u_{12}| \leq \sqrt{u_{11}\left(u_{22} - n^2 d''\right)}$ since $\partial^2 u\left(x_i, z_i\right) / \partial q_i^2 - n^2 d'' < 0$ and $\partial^2 u\left(x_i, z_i\right) / \partial x_i^2 < 0$.

*(ii) For any given $z_i$, the equilibrium $x_i$ is socially optimal:*

$$\frac{\partial u(x_i, z_i)}{\partial x_i} = 0.$$

Since user $i$ does not internalize the negative externality on the other users, $i$ extracts from the resource too much. In other words, $i$ extracts more from the common resource than what a social planner would prefer and as specified by the first-best. Note that the difference between the equilibrium extraction and the first-best increases in the number of users, $n$. Since choice-variable $x_i$ was assumed to create no externality, it is optimally chosen. This fact is natural in the static setting but it might no longer hold in dynamic settings, as we soon will learn. The following quotation describes the negative externalities in the Nash Equilibrium:

> The Tragedy of the Commons. Garrett Hardin. Science, New Series, Vol. 162, No. 3859 (Dec. 13, 1968), 1243-1248:
> "The tragedy of the commons develops in this way. Picture a pasture open to all. It is to be expected that each herdsman will try to keep as many cattle as possible on the commons. Such an arrangement may work reasonably satisfactorily for centuries because tribal wars, poaching, and disease keep the numbers of both man and beast well below the carrying capacity of the land. Finally, however, comes the day of reckoning, that is, the day when the long-desired goal of social stability becomes a reality. At this point, the inherent logic of the commons remorselessly generates tragedy.
>
> As a rational being, each herdsman seeks to maximize his gain. Explicitly or implicitly, more or less consciously, he asks, "What is the utility to me of adding one more animal to my herd?" This utility has one negative and one positive component.
>
> 1) The positive component is a function of the increment of one animal. Since the herdsman receives all the proceeds from the sale of the additional animal, the positive utility is nearly +1.
>
> 2) The negative component is a function of the additional overgrazing created by one more animal. Since, however, the effects of overgrazing are shared by all the herdsmen, the negative utility for any particular decision-making herdsman is only a fraction of 1.
>
> Adding together the component partial utilities, the rational herdsman concludes that the only sensible course for him to pursue is to add another animal to his herd. And another; and another... But this is the conclusion reached by each and every rational herdsman sharing a commons. Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit–in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in a commons brings ruin to all."

### 2.1.1 Working through Example 1

We work out now Example 1 above. Let us assume $u(x) = u > 0$ so that the marginal gain from using the resource is a constant. The choice of $x_i$ can be thought of as made in the past, so we consider it fixed here in this Section.

**Nash equilibrium:** We want to explicitly obtain the first best and Nash equilibrium outcomes to be able to quantify the welfare loss from the externality. Consider the Nash outcome first. Taking the emissions of the other agents as given, $\{z_j\}_{j\in\mathcal{I}\setminus i}$, the marginal gain from increasing $z_i$ for agent $i$ is

$$\begin{aligned}\frac{\partial U(z_i,s)}{\partial z_i} &= \frac{\partial U(z_i, z_i+\sum_{j\in\mathcal{I}\setminus i} z_j)}{\partial z_i}\\ &= u - d\frac{\partial\big(z_i+\sum_{j\in\mathcal{I}\setminus i} z_j\big)}{\partial z_i}\big(\sum_{i\in\mathcal{I}} z_i\big)\\ &= u - d\big(\sum_{i\in\mathcal{I}} z_i\big)\end{aligned}$$

Note that all players are symmetric, and therefore it is natural to focus on symmetric equilibria where all players choose $z_i = z^{NE} > 0$ (we will look at asymmetries below). By this, we must have for all $i$,

$$\begin{aligned}\frac{\partial U(z_i,s)}{\partial z_i} &= 0\\ \Rightarrow u = d(nz^{NE}) \Rightarrow z^{NE} &= \frac{u}{dn}.\end{aligned}$$

From this, obtain the equilibrium payoff as

$$\begin{aligned}U(z^{NE}, nz^{NE}) &= uz^{NE} - \frac{d}{2}\big(nz^{NE}\big)^2\\ &= \big(1-\frac{n}{2}\big)\frac{u^2}{dn}.\end{aligned}$$

We may immediately note that $U(z^{NE}, nz^{NE})$ is negative for all $n > 2$. Yet, this does not mean that player $i$ could do better by choosing $z_i = 0$: this would only leave the negative externality coming from other player's actions but no private gain.

**The first best:** The first best can be obtained by choosing

$$
\begin{aligned}
\max_{\{z_1,..,z_n\}} \sum_{i\in\mathcal{I}} U\left(z_i, s\right) &= \max_{\{z_1,..,z_n\}} \sum_{i\in\mathcal{I}} \Big(uz_i - \frac{d}{2}\big(\sum_{k\in\mathcal{I}} z_k\big)^2\Big) \\
&\Rightarrow \\
\forall i, \frac{\partial U(z_i,s)}{\partial z_i} &= u - d\frac{\partial\big(z_i + \sum_{j\in\mathcal{I}\backslash i} z_j\big)}{\partial z_i}\big(\sum_{k\in\mathcal{I}} z_k\big)\underbrace{(1+...+1)}_{=n} \\
&= u - d\big(\underbrace{\sum_{k\in\mathcal{I}} z_k}_{=nz^{FB}}\big)n \\
&= u - d\big(nz^{FB}\big)n \Rightarrow z^{FB} = \frac{u}{dn^2}
\end{aligned}
$$

**Comparison & discussion:** We now compare explicitly how the total payoff depends on the number of agents. First, in the Nash equilibrium of the common pool game, we observe that

$$
\begin{aligned}
\sum_{i\in\mathcal{I}} U\left(z_i^{NE}, s\right) &= \Big(uz^{NE} - \frac{d}{2}\big(nz^{NE}\big)^2\Big)n \\
&= \Big(1 - \frac{n}{2}\Big)\frac{u^2}{d} < 0 \Leftrightarrow n > 2.
\end{aligned}
$$

Thus, the total payoff goes negative quickly with the number of agents. Second, in the first best we have

$$
\begin{aligned}
\sum_{i\in\mathcal{I}} U\left(z_i^{FB}, s\right) &= \Big(uz^{FB} - \frac{d}{2}\big(nz^{FB}\big)^2\Big)n \\
&= \Big(1 - \frac{1}{2}\Big)\frac{u^2}{dn} > 0, \forall n.
\end{aligned}
$$

Careful planning thus avoids externalities and keeps the payoff positive in total. Note that it declines with $n$ however: it is not possible to scale up the activity with the number of users since the total resource becomes scarce. It is important to notice that the increase in scarcity is different from the increase in negative externalities. For example, the illustration by Hardin above describes how a given number herdsmen overuse the resource by increase the number of cows. However, increasing the number of herdsmen, even if they can optimally manage the use of the commons, leads to similar problems: the resource to be used remains limited but the number of users increases. There is less and less of the resource per herdsman.

# 3 Solutions to the common pool problem

We consider now how the inefficiency from the externality could be removed. The most famous solution is the one proposed by Coase.

## 3.1 Coasian solution

Coase proposed that if the property rights are well defined, then the market will take care of the externality problem. Since this idea builds on the thinking that the market will set prices for externalities, let us first modify our definition of Nash equilibrium, given some (at this point) arbitrary payment that follows if individual $i$ generates externalities. Since the social cost of those externalities depends on how much the others producing them, the externality payment depends on both $z_i$ and $z_{-i}$.

**Definition 2.** *A Nash equilibrium with externality payment $p(z_i, z_{-i})$ is a set of choices, $\{x_i, z_i\}_{i\in\mathcal{I}}$ , such that, when user $i$ takes as given the other users' choices, $\{x_j, z_j\}_{j\in I\setminus i}$ , then $\{x_i, z_i\}$ maximizes private payoff $U(x_i, z_i, s) - p(z_i, z_{-i})$.*

**Proposition 2** *Consider the static common-pool problem with $p(z_i, z_{-i}) = (n-1)d(z_i + \sum_{j\in\mathcal{I}\setminus i} z_j)$: (i) Each user's choice $z_i$ coincides with the first-best:*

$$\frac{\partial u(x_i, z_i)}{\partial z_i} = n\frac{\partial d(s)}{\partial s}.$$

*(ii) For any given $z_i$, the equilibrium $x_i$ is socially optimal:*

$$\frac{\partial u(x_i, z_i)}{\partial x_i} = 0.$$

The result says that if user $i$ must pay the externality cost imposed on other agents, then it follows, by definition, that the individual choices will be socially optimal. The Coasian solution achieves this either by assigning property rights to those who suffer or to those who produce the externality. In the former case, the holder of the property right has the right to live without the damage from the resource use, so the agent that wants to use it must pay for the usage.

**Victims hold the rights**. Going back to our Example 1 with two agents $i = 1, 2$, assume that only agent 1 has some private gain from production: $u(x_1, z_1) = uz_1$ with $u > 0$ while $u(x_2, z_2) = 0$. Yet, both agents suffer from output as before: $d(s) = \frac{d}{2}(s)^2$ is the individual damage per agent. In the Nash equilibrium, choices are

$$z_1^{NE} = \frac{u}{d}$$
$$z_2^{NE} = 0$$

since agent 1 cares only about his own damage, and agent 2 has no gain from production at all. Now, if agent 2 owns the property right to have no damage, agent 1 must compensate agent 2 for giving up those rights. If the compensation equals the full damage imposed by 1 on agent 2, then we arrive at:

$$p(z_1, z_2) = p(z_1, 0) = \frac{d}{2}(z_1)^2.$$

Facing this payment, agent chooses the activity level to solve

$$\frac{\partial}{\partial z_1}\left(uz_1 - \frac{d}{2}(z_1)^2 - p(z_1)\right) = 0 \Rightarrow u - \underbrace{dz_1}_{\text{own marginal damage}} - \underbrace{dz_1}_{\text{marginal payment to 2}} = 0$$

$$\Rightarrow z_1^{NE} = z_1^{FB} = \frac{u}{2d}$$

We observe that agent 1 chooses the first-best outcome, and compensates agent 2 exactly for the losses. This is what the proposition predicts. However, remember that agent 2 owns the property right, and it is not the case that the agent suffering has to follow exactly this pricing rule for selling her rights. The efficient sale of the rights to pollute can be obtained in different ways. For example, agent 2 could set a unit price $u/2$ so that the total payment received from the polluter is

$$p(z_1) = \frac{u}{2} z_1.$$

When facing this linear payment in $z_1$, agent 1 solves

$$\frac{\partial}{\partial z_1}\left(u z_1 - \frac{d}{2}(z_1)^2 - \frac{u}{2} z_1\right) = 0 \Rightarrow u - d z_1 - \frac{u}{2} = 0$$
$$\Rightarrow z_1^{NE} = z_1^{FB} = \frac{u}{2d}$$

This gives the first best incentives as well. Why is this? Note that here agent 2 is fixing one price for all levels $z_1$, and this price equals the marginal damage on agent 2 at activity level $z_1^{NE} = z_1^{FB}$. For this reason, agent 1 pays a high price for all units: the price is higher than the actual marginal damage for all $z_1 < z_1^{FB}$. As a result, trading leaves some extra surplus for agent 2:

$$\underbrace{p(z_1) - \frac{d}{2} z_1^2}_{\text{compensation - damage}} = \frac{u}{2} z_1 - \frac{d}{2} z_1^2 = \frac{u^2}{4d} - \frac{u^2}{8d} > 0$$

What do we learn from this? The Coasian solution says that it is in principle possible to trade with rights to achieve efficiency when the property rights are well defined – it does not tell how exactly the parties decide to transact.

**Externality-generating agents hold the rights**. The key proposition made by Coase was that it does not matter for efficiency how the property rights are allocated, as long as they are well defined. To see this, consider the other extreme where the user of the resource has all the rights. For example, the polluter has the right to pollute, the herdsmen have the right to use the pasture, etc. How can the market set a price for the externality in this case? Now, the victims have to compensate the externality-generating agents for reducing the level of the externality causing activity: the externality payment that the producer faces must be negative, $p(z_i, z_{-i}) < 0$, for choices $z_i < z_i^{NE}$. That is, the payment must compensate the agent for not generating the externality level associated with the Nash equilibrium where, as we have seen, the agents ignore how much costs they inflict on others.

Once again go back to our Example 1 with two agents $i = 1, 2$, and assume that only agent 1 has some private gain from production: $u(x_1, z_1) = u z_1$ with $u > 0$ while $u(x_2, z_2) = 0$. Yet, both agents suffer from output as before: $v(s) = -\frac{d}{2}(s)^2$ is the individual damage per agent. In the Nash equilibrium, activity levels are

$$z_1^{NE} = \frac{u}{d}$$
$$z_2^{NE} = 0.$$

Now, if agent 1 owns the property right to choose whatever activity level it pleases, agent 2 must compensate agent 1 for giving up those rights. If the compensation equals the full willingness to pay for reductions in damages, then

$$p(z_1, z_2) = p(z_1, 0) = \underbrace{\frac{d}{2}z_1^2 - \frac{d}{2}(z_1^{NE})^2}_{\text{difference in agent 2 damages}} = \frac{d}{2}z_1^2 - \frac{1}{2}\frac{u^2}{d}.$$

Notice that agent 1 is collecting money from reducing its activity level below the Nash equilibrium level. Facing this payment, agent 1 chooses the activity level to solve

$$\frac{\partial}{\partial z_1}\Big(uz_1 - \frac{d}{2}(z_1)^2 - p(z_1)\Big) = 0 \Rightarrow u - dz_1 - \underbrace{dz_1}_{\text{marginal compensation from 2}} = 0$$

$$\Rightarrow z_1^{NE} = z_1^{FB} = \frac{u}{2d}$$

We observe that agent 1 chooses the first-best outcome! Again, we could consider other compensation schemes: as long as they give the right incentives to reduce the activity at the margin, the optimum can be achieved. The main difference is that parties end up achieving different total surpluses from the activity, depending on the pricing schedule used. Ultimately, the chosen effective price for the externalities depends on the bargaining power between the agents, which is an issue that we consider in detail below.

## 3.2 Transaction costs

Above, we sketched the core idea of the Coasian solution to the externality problem. In practice, it is challenging concept since it may be difficult for the parties involved to transact in an optimal way. These difficulties go under the concept of transaction costs, covering several potential sources of such costs.

**The hold up problem** can arise when the victims have the right to avoid the external cost. As we have seen above, those who generate the external costs must then compensate the victims. The hold-up problem arises when some victims hold up the rest of the group from transacting, unless a large compensation is made exactly to the agents holding up the process. To make a stark case, suppose that all agents who suffer have been paid compensations according to their damages, excluding one agent, denoted by $j$. Thus, all agents $i \in \mathcal{I}\backslash j$ have agreed on the compensation scheme that leads to the first-best, provided agent $j$ also comes on board. This agent could now ask for a different compensation from each of the polluter:

$$\underbrace{p(z_i, z_{-i})}_{\text{compensation paid by i}} = \underbrace{d(z_i + \sum_{k\in\mathcal{I}\backslash i} z_k)}_{\text{damage caused by i on j}} + \underbrace{T_i}_{constant}$$

The first part of the compensation is what we have seen before: $j$ is compensated for the damage it suffers from the activity of $i$. The second part is a pure transfer of money: agent $j$ could demand the full surplus that agent $i$ achieves from the activity through $T_i$.

Why is it that $j$ could collect such extra transfers from all agents $i \in \mathcal{I} \backslash j$ generating the externalities? This follows because $j$ has huge bargaining power: by refusing to sell its rights, it denies the activity, and all agents would get zero payoff (no activity, no payoff). The hold-up follows from the strict nature of rights; one cannot produce at all if there is at least one agent who declines the activity.

The hold-up problems becomes a sever issue, in particular, if there is a small number of firms who need to procure the rights from the victims. The victims have an incentive to wait instead of striking a deal, because the terms of bargaining improve when the victim is later in the sequence of deals between the firm and the victims.

**The free rider problem** can arise when the externality generating firms have the rights. Then, the victims must get their act together and compensate the firms for reducing the activity level. Suppose there is a meeting where all parties excluding one victim show up. In this meeting bargaining leads to the best possible outcome for the $n-1$ agent present in the room so that they end up choosing for all $i$ in the room:

$$\frac{\partial u\left(x_i, z_i\right)}{\partial z_i} = (n-1)\frac{\partial d\left(s\right)}{\partial s}.$$

Notice that this is the first-best outcome when the number of agents is $n-1$. This follows because, as we have shown, those suffering can compensate those who generate the externality to reduce the activity level and thereby reach the first best. However, there is one agent missing from the meeting, so the outcome is not fully optimal. The agent who does not show up obtains a considerable reduction in the activity level, and thus in the negative external cost. But it does not have to compensate anyone for reductions in the activity, which is the key to the free rider problem. When $n$ becomes very large, the negotiations give almost the first best outcome but no payments follow for the agent staying out. But since all agents have incentives to stay out, the negotiations may not happen in the first place.

Let us go back to our Example 1 with n agents $i = 1, 2, .., n$, and assume that only agent 1 has some private gain from production: $u(x_1, z_1) = uz_1$ with $u > 0$ while $u(x_i, z_i) = 0$ for all other $i$. Yet, all agents suffer from output as before: $v(s) = -\frac{d}{2}(z_1)^2$ is the individual damage per agent. In this case, the negotiations lead to

$$z_1 = \frac{u}{(n-1)d}$$

so that the free-riding agents saves the damages

$$\frac{d}{2}(z_1^{NE})^2 - \frac{d}{2}z_1^2 = \frac{u^2}{2d} - \frac{u^2}{2(n-1)^2 d} = \frac{u^2}{2d}\Big(1 - \frac{1}{(n-1)^2}\Big)$$

without contributing any resources to compensate firm 1.

**Bargaining frictions** are yet another source of transaction costs. These are related to difficulties that parties might have in coming up with a bargaining protocol that manages to obtain efficiency. These problems may be less sever if there is an outside authority such as government with coercive powers, helping the parties to materialize the gains from trade, for example, through a mechanism that determines actions and contribution by each party. But, in particular, between sovereign countries, there is no such outside authority, and problems discussed just below may arise.

There are three players a, b, c who can produce emissions reductions. When a and b form a group, we denote that coalition by ab, and so forth. If all agents remain separate, the payoff remains at zero. We assume that ab produce 12 for group ab. The agent outside the group, who is c in this case, receives 9. We assume that whenever an agent remains outside and two others form a coalition, then the outside agent receives 9. Outside agent obtains thus some free-riding benefit: this agent does nothing but yet has more payoff than if the two others remain separate as well. To complete the payoffs, we assume ac gives 13, bc gives 14, and abc is of value 24.

The bargaining protocol puts a microstructure on how any one of the possible coalitions could form, including the potential transfers between the agents. One protocol is such that agents enter a room in a particular oder, for example, a first, followed by b, and then c arrives. Thus, in this situation a waits for b, and makes an offer to b to form team ab. What are the options for b? First, b can stay outside and compete with a for c when agent c arrives the room. Second, b can join a, and then they can make an offer together to c. Which one dominates? Well, in the first case, b can offer at most the value of bc minus the value of b alone if ac forms: 14-9=5. Similarly, a can offer c as follows: 13-9=4. We observe that b will win the competition for agent c. But b will not pay 5 but 4 to c (you can think that b can offer to pay one cent more than a to outbid a). You see that c has no option other than to accept since c has does not have the free riding option, as there will be no coalition if bc does not form. Thus, along this path, agent b gets 14-4=10. Now, consider the decision by agent a when seeing b entering the room. Agent a has to offer b at least 10 to join agent a. Then, ab together would have to offer c at least 9 to prevent her from free riding. Team abc has value 24 so this leaves payoff 24-10-9=5 for agent a. Does this make sense to a? No! Agent can refuse to talk to b when first meeting b, so this forces b to form a coalition with c, allowing a then to free ride on this coalition. We have found the final outcome: a stays alone, bc forms, and the payoffs are (9,10,4) to a,b,c, respectively.

What do we learn? The vagaries of bargaining can create frictions that prevent the full co-operative outcome to emerge. It is essential for this outcome that there is an externality: a can commit to free ride, given the order of moves in the bargaining. This observation is quite general in the sense that commitment not to contribute to the common good is often exploited, for example, by locking into dirty technologies prior to climate change negotiations.

# Exercises (voluntary!)

1. During the lecture we derived payoff matrix below, with actions $(D, D)$ and $(C, C)$ referring to Nash equilibrium and cooperative equilibrium, respectively. Player X was country 1, and player Y was country 2.

   (a) Once again: identify the externality in this example when one player deviates from $(C, C)$

(b) Now, suppose the players are at liberty to propose and accept contracts. A contract specifies four transfers from one player to another, one transfer for each of the four cells in the matrix. What would be a contract that solves the externality problem? Contract should be such that it eliminates the incentive to deviate from $(C, C)$, and both players are better off by signing it. Explain how the contract changes the payoffs in the matrix. Don't worry about how this contract is negotiated (which can be difficult as you have seen in the previous problem).

| | | Player $Y$ | |
|---|---|---|---|
| | | $C$ | $D$ |
| Player $X$ | $C$ | $\frac{1}{8d}, \frac{1}{8d}$ | $-\frac{1}{8d}, \frac{1}{4d}$ |
| | $D$ | $\frac{1}{4d}, -\frac{1}{8d}$ | $0, 0$ |

2. During the lectures, we discussed transaction costs: what do they mean precisely? Quite simply, transaction costs prevent efficient outcomes to emerge in negotiations. In this problem, you will work out negotiation process explicitly and see how externalities are the source of transaction costs in negotiations. In this first part, there are no externalities and the outcome is efficient. The numbers are the same as in the lecture slides: bargaining over joint production. Players $A, B, C$ can produce as follows. $AB$ can jointly produce 15, $AC$ can produce 16, and $BC$ can produce 18. Under full collaboration, $ABC$ produces 24 in total. If each player is alone, each can produce 6.

   (a) Assume the same order of contracting as in the lecture notes: $A$ is first in the negotiations room, offering something to $B$. If $B$ accepts, then $AB$ forms. When agent $C$ enters the room, $AB$ will make an offer to $C$. If $C$ accepts, then $ABC$ forms. In contrast, if $B$ declines $A's$ offer, then $A$ and $B$ compete to form either $AC$ or $BC$. Think from the end to the beginning and show that, in fact, coalition $ABC$ will form.

   (b) What are the final payoffs for each party, that is, how do they share the total 24?

3. Continuing with the previous example, we will change now the numbers to introduce externalities. If each player is alone, each can produce 0. We assume: $AB$ produces 12, $AC$ produces 13, $BC$ produces 14, and $ABC$ continues to produce 24. We assume that whenever an agent remains outside and two others form a coalition, then the outside agent receives 9. Here is the externality: one can benefit from two others forming a coalition, which holds true with pollution externalities for example.

   (a) Assume the same order of contracting as above: $A$ is first in the room, $B$ enters, and then $C$. Can you show that it is not possible for the coalition $ABC$ to form?

   (b) What is the equilibrium coalition that follows from this bargaining protocol, and what are the final payoffs?

(c) The main lesson from this exercise is that externalities prevent efficient bargaining outcomes. Explain in words why?

# 4 Further Reading

The Coasian solution to the externality problem (1960) can be found from most textbooks but the exposition in Friedman (2000), including the discussion on transaction costs, is outstanding. Harstad and Liski (2013) extended to dynamic common-pool problems.

Coase, Ronald H. (1960): "The Problem of Social Cost," *Journal of Law and Economics* 3: 1-44.

Friedman, D, " *Law's order* ( 2000). Princeton University Press.

Fudenberg, Drew and Tirole, Jean (1991): *Game Theory.* MIT Press.

Harstad Bård & Matti Liski (2013)"Games and Resources", In: Shogren, J.F., (ed.) "Encyclopedia of Energy, Natural Resource, and Environmental Economics", Vol. 2, pp. 299-308 Amsterdam: Elsevier.