# Cost-benefit analysis: Lecture note

**SOK-2031 (UiT)**

Tapas Kundu

# Contents

# 1 Section 1

## 1.1 Introduction

Cost-benefit analysis (CBA) is a policy assessment method that quantifies the value of a policy to all members of society in monetary terms.

Suppose we examine if a specific project should be implemented. This project can benefit certain groups in society but it also involves allocating valuable resources, such as land, labor, capital, etc., which could otherwise be used for alternative productive purposes. Our objective is to determine whether it is sensible to implement the project.

To arrive at an informed decision, we need to address several essential questions:

1. First and foremost, who is conducting the evaluation? Is it the society, an organization, or an interest group, for instance?
2. What are the potential alternatives? If resources aren't allocated to the project under consideration, how will they be used? How will the affected parties respond in the presence or absence (while taking the alternatives into account) of the project?
3. Who will be impacted by the project? Who stands to benefit, and who might incur costs? In the scenario of winners and losers, how should we assess the distributional effects?
4. How will we measure the benefits and costs? What if the effects of the project extend to non-marketed goods?
5. If the project is likely to have long-lasting effects, how should we incorporate these dynamic impacts into the analysis? In a scenario where the future is uncertain, how can we effectively quantify the overall impact?

The objective of a CBA is to establish a framework for assessing the relative efficiency of policy alternatives. This framework is constructed to address the above questions in a manner consistent with our understanding of economic principles, particularly those of welfare economics.

> Consider the following examples.
>
> a. Building a transportation project, for example, construction of a highway, or extension of railways, aimed at improving and shortening travel between two regions.
>
> b. Developing a mine for extracting minerals for industrial production and export.
>
> Discuss questions 1-5 within the contexts of these two projects.

## 1.2 A framework of CBA

We will develop a framework for CBA from the perspective of a social planner. This approach implies, among other considerations, that our focus is on assessing the changes in social welfare resulting from the project implementation. It is not just about revenue and financial cost of a project but also about how the benefits and costs are distributed to different groups of society. Furthermore, we aim to convert these factors into a welfare-based metric for policy evaluation.

Evaluating the benefits and costs requires a comparison with alternative uses of resources. For instance, construction of a highway would necessitate the allocation of land, labor, concrete, construction machinery, etc. These resources could otherwise be employed in producing other valuable goods and services for the population. While evaluating costs, we will emphasize the concept of opportunity costs, rather than merely focusing on the price of an input. The opportunity cost of employing an input is its value in its most valuable alternative use. It quantifies the value of what society foregoes in order to utilize the input for implementing the project.

To start with, a project can be regarded as a perturbation in the economy that influences the well-being of individuals. This perturbation will be realized in different markets by different groups.

### 1.2.1 Impacted markets

A project employs inputs and transform them into output. Consider a market that will be impacted due to the project. Examine the diverse sectors (such as consumers, producers, taxpayers etc.) within the affected market and quantify the combined effects across sectors. Subsequently, aggregate the outcomes across all potentially affected markets. Pay attention if you are double counting the effects if a group appears in different markets with different roles.

Which are the affected markets? Here is a broad classification.

1. Output market
2. Input market
3. Secondary market
4. (Market) of non-traded goods

   Consider the following example and identify various impacted markets.

   Suppose a city builds a new railway line.

> Then the output market is the market for railway trips.
>
> Building the subway requires inputs, such as the land the stations occupy and the rail passes through, the materials used to build the stations and, of course, labor. These goods or services are called factor inputs. Markets for these inputs are "upstream" of the output markets, and are referred to as input markets.
>
> The project may also have indirect effects that occur "downstream." For example, a new rail station may affect the market for housing near to the station or may affect the market for gasoline if some commuters switch from driving to commuting via train. Such markets are called "secondary markets." In addition, the output of the project can be an input in another market. For example, suppliers of other goods can use the railway which may reduce their costs of transport considerably. This is also an example of an affected secondary market.
>
> In addition, we must also consider the externalities, for example, effects on landscape, clean air, climate, or even safety levels, associated with the project even though there are no markets for these "goods."

### 1.2.2 A simple policy approval rule

We will introduce some notations here. Let us denote the social surplus that can be accrued from various affected markets by SS. We will elaborate on this concept in the next section.

$SS = SS_O + SS_I + SS_S + SS_E$ (indices O, I, S, and E representing the various markets, output, input, secondary, and external non-marketed goods, respectively).

The change in social surplus, which will be denoted by a notation $\Delta$, is

$\Delta SS = \Delta SS_O + \Delta SS_I + \Delta SS_S + \Delta SS_E$.

If we can quantify these changes, then we can think of a simple intuitive policy approval rule: A policy can be approved if it is associated with a positive $\Delta SS$, and if there many such alternatives with positive $\Delta SS$, then choose the one with the largest margin.

Two important observations to note here.

Firstly, there will be actors who might participate in multiple types of markets. For instance, a producer in the output market also functions as a consumer in the input market. It's worth noting that occasionally (though exceptions should be kept in mind), the output and input markets of the focal primary product are often collectively referred to as primary markets. This is because these markets realize the direct effects of the production activities associated with the project.

Secondly, there can be situations in which government may play a more active role in the market, for example as one of the suppliers or purchasers. It will then be useful to examine the net budget impacts on the government.

We will, however, have to deal with various critical issues before we come out with a simple quantified measure of $\Delta SS$. For example, how do we measure the social surplus in different types of markets. For a long-lasting project, how do we aggregate effects over time? How do we quantify future effects when there is uncertainty?

We will now focus on these questions, and here is a tentative plan for how we will proceed in this course. I will begin by discussing how to measure changes in social surplus within the output market. Later, we will explore measuring changes in social surplus in secondary markets and input markets. This will be followed by analyses of time-related effects, externalities, and risk and uncertainty.

## 1.3 Social surplus and allocative efficiency

Why do we focus on the social surplus?

As mentioned at the outset, CBA serves as a framework for evaluating the relative efficiency of policy alternatives. However, at this juncture, it remains unclear how we can quantify the efficiency of a policy.

Social surplus, if it can be measured, serves two objectives. First, it quantifies benefits and costs. Second, if represented in monetary terms, it can be aggregated across groups and markets and comparing policies becomes easier.

However, how does a measure of surplus reflect market efficiency? Since our focus is primarily on the well-being of individuals, groups, and organizations, the concept of collective efficiency must be tied to the preferences of those involved.

Characterizing preference of a producer (of an entrepreneur or of a factor owner) is somewhat easy - a policy is preferred to another if the former results in higher net benefits. A measure of surplus typically quantifies benefits minus costs. We can measure benefits in terms of revenue from sales and costs in terms of the value of the next best alternative uses of inputs.

Establishing a connection between surplus and preference of consumers becomes more intricate. How can we determine whether a diverse group of consumers would be better off under a specific policy compared to an alternative?

Adding to the complexity, economists generally hold that consumers derive utility from consumption. Utility isn't quantified in monetary units, and to make the matter worse, it is not even observable. Moreover, when various consumers undergo distinct degrees of change, how can we express the collective preference of the entire group?

### 1.3.1 Pareto efficiency

A simple and intuitively appealing definition of efficiency, referred to as Pareto efficiency, underlies modern welfare economics and CBA. An allocation of goods is considered Pareto-efficient if no alternative allocation can make at least one person better off without making anyone else worse off.

Conversely, an allocation is inefficient if an alternative allocation exists that would make at least one person better off without making anyone else worse off.

However, the requirement of Pareto efficiency does not significantly narrow down the choices among alternatives—several policies can be Pareto-efficient while still reflecting conflicting preferences among individuals.

### 1.3.2 Welfare function

To capture the collective preference, we can define a general social welfare function. To keep things simple, consider a society comprises of two consumers.

$W = W(U_1, U_2)$ where $U_1$, $U_2$ are individual utilities.

An example of a welfare function can be

$W(U_1, U_2) = U_1 + U_2$.

A change in welfare can then be expressed as

$dW = \sum_{i=1}^{2} \frac{\partial W}{\partial U_i} dU_i$

A direct application of the utility-maximization principle from consumer theory gives us that

$dU_i = \sum_{j=1}^{m} \frac{\partial U_i}{\partial M_i} p_j dx_{ij}$,

where consumer i choice set consists of m goods, and $x_{ij}$ is i-th consumer's utility-maximizing demand for j-th good.

Together, we can then express the change in welfare as

$dW = \sum_{i=1}^{2} \sum_{j=1}^{m} \frac{\partial W}{\partial U_i} \frac{\partial U_i}{\partial M_i} p_j dx_{ij} = \sum_{i=1}^{2} \sum_{j=1}^{m} \beta_i p_j dx_{ij}$

This observation is important for two reasons. First, it establishes a link between the change in welfare and the change in demand.

Second, it shows that the change in welfare is a weighted sum of the monetary value of the change in demand. These weights affect the marginal impact of a change in income on welfare. Since society comprises individuals with varying levels of wealth, it is expected that these weights should not be the same for everyone.

We will however follow a simplified path: assigning the same weights. We will come back to this issue later in the course while discussing distributional concerns.

Now, we will establish a connection between observed demand behavior and another useful concept: willingness to pay.

### 1.3.3 Willingness to pay (WTP)

Individual demand curves slope negatively due to diminishing marginal utility. The market demand curve is the horizontal sum of individual demand curves and also slopes downward.

The inverse market demand, where the price is represented as a function of quantity , can be interpreted as a Marginal Benefits (MB) curve: it indicates the maximum amount someone is willing to pay for an additional unit of a good – the marginal unit.

Willingness to Pay (WTP) is a general concept that can be applied not only to a single consumption good but also in the context of broader policies. For example, it is the payment that one would have to make or receive under the policy so that one would be indifferent between the status quo and the policy with the payments.

The algebraic sum of the WTP values is the appropriate measure of the net benefits of the impacts of a policy. Only if the aggregate net benefits of the policy (as measured by WTP of affected individuals) are positive, then there exists a set of contributions and payments that create a Pareto improvement over the status quo.

We can therefore approximate the change in welfare by the net benefits, the sum of the WTP values. This approach is not without controversy; however, it is a pragmatic one. Additionally, there are strong arguments in favor of evaluating policies based on net benefits.

#### 1.3.3.1 Kaldor-Hicks Criteria

Adopt only policies that have positive net benefits. Reasons for adopting it:

- It is feasible.
- Society maximizes aggregate wealth.
- If different policies have different winners and losers, then, in aggregate, costs and benefits will average out over the entire population.
- It is possible to do redistribution wholesale rather than within each separate policy.

#### 1.3.3.2 Limitations of WTP

- It can lead to intransitive social orderings of policies (Arrow 1951).
- Dependence on the distribution of wealth: If the distribution of wealth of society changes, then individual WTP changes, and perhaps, the ranking of alternatives could change.

### 1.4 Consumer surplus

Now that we have to decided to measure benefits in terms of WTP, which can be represented by the inverse demand curve, we can measure the consumer surplus as follows.

The area under the market demand curve from the origin to $X^*$ (see Figure 1 below) measures the gross benefits to society of consuming $X^*$ units of the good.

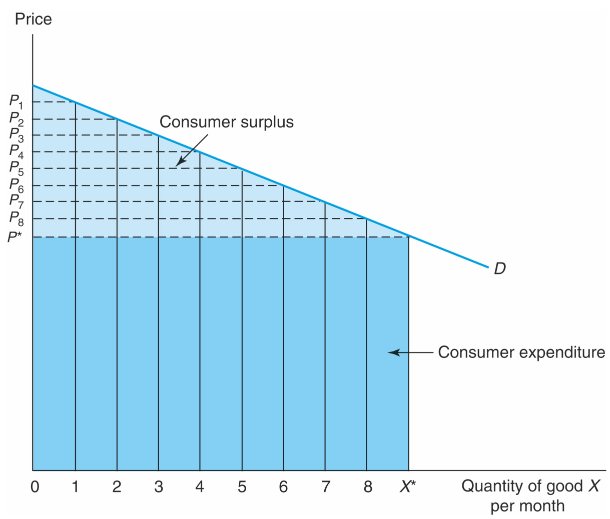



Figure 1: Consumer surplus and consumer expenditure

If one has to pay $P^*$ for $X^*$ units of the good, then the rectangle bounded by $P^*$ and $X^*$ is the aggregate cost to consumers. The net benefits to consumers are the gross benefits minus the consumers' costs.

The net benefits are called the consumer surplus (CS): it is the area between the demand curve and the $P^*$ line—the light shaded area in Figure 1.

Consumer surplus is important in CBA because changes in CS can be viewed as close approximations of the WTP for (the benefits of) a policy change.

### 1.4.1 Equivalence of consumer surplus and compensating variation

(Recall from microeconomic theory) We deal with two types of utility-maximizing demand:

1. Uncompensated or Marshallian demand - Utility-maximizing demand for a given fixed level of income.
2. Compensated or Hicksian demand - Expenditure-minimizing demand for a given fixed level of utility.

For CBA purposes, as an approximation of WTP, it is important to measure the change in consumer surplus on a Hicksian compensated variation demand curve.

Why?

The maximum amount of money that consumers would be willing to pay to avoid a price increase is the amount required to return them to the same level of utility that they enjoyed prior to the change in price, an amount called compensating variation. If consumers had to spend any more than the value of their compensating variation, then they would be worse off paying to avoid the increase, rather than allowing it to occur.

If they could spend any less, then they would be better off paying to avoid the increase, rather than allowing it to occur. Hence, for a loss in consumer surplus resulting from a price increase to equal the consumers' WTP to avoid the price increase, it has to correspond exactly to the compensating variation value associated with the price increase.

To get the compensating variation, analysts often estimate the market demand by directly asking consumers about their willingness to pay for a change.

The Marshallian demand curve, however, is sometimes the only one that is usually available. Computing consumer surplus on a Marshallian demand curve will be different than a Hicksian compensated variation demand curve because the income effect will be inappropriately included (if price increases, CS is smaller on a Marshallian than on a Hicksian demand curve; and if price decreases, it is larger).

The difference is usually small, however, and can be ignored unless there is a large income effect, i.e., large prices changes in key goods (housing, leisure, etc.) are being considered.

Equivalent variation, which is an alternative to compensating variation, is the amount of money that, if paid by a consumer, would cause him or her to lose just as much utility as a price increase.

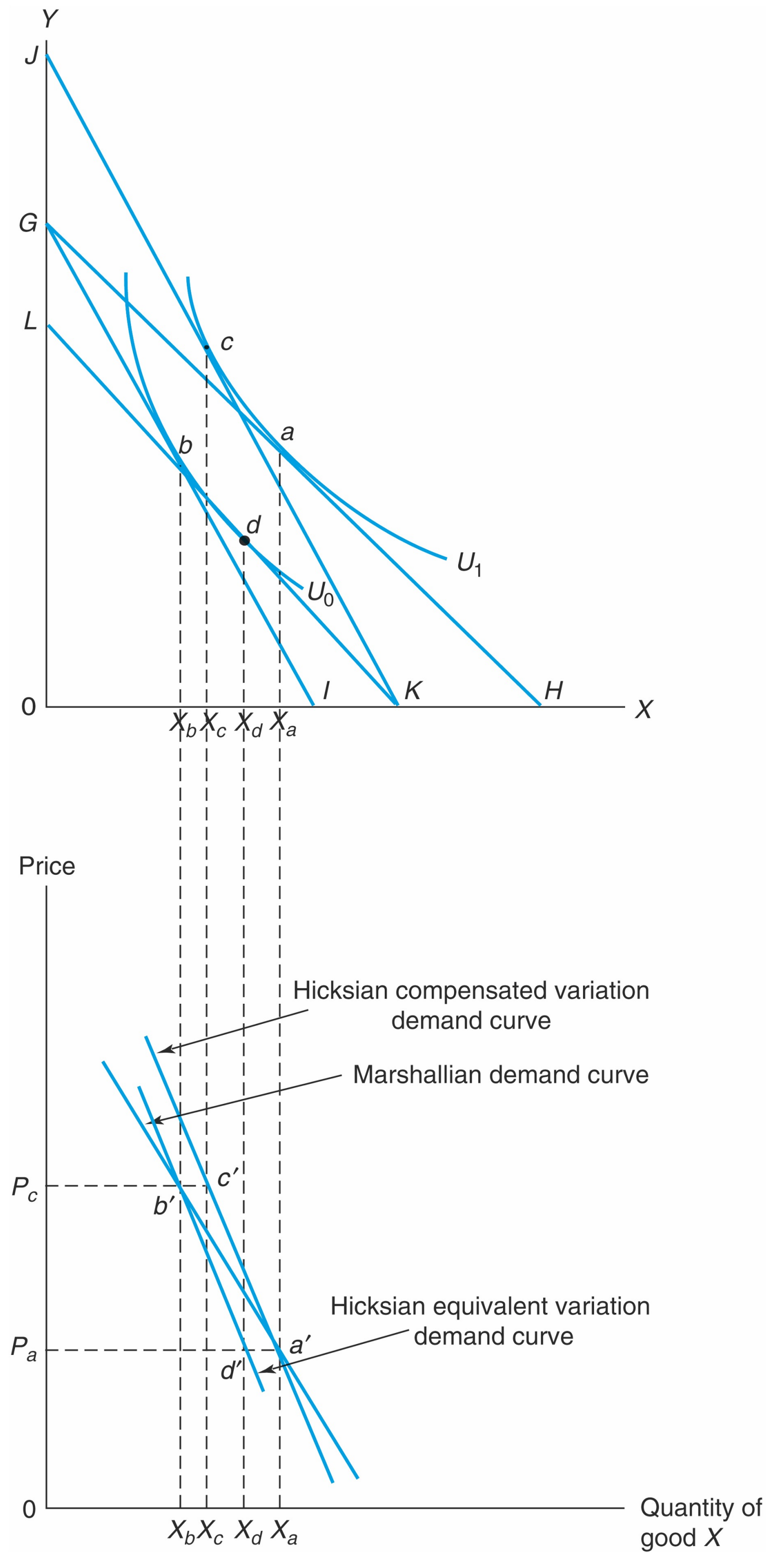



Figure 2: Compensated and uncompensated demand

### 1.4.2 Quantifying the change in consumer surplus

Suppose that a policy results in a price decrease, as in Figure 3. Let $\Delta P = P1 - P^* < 0$ denote the change in price and let $\Delta X = X1 - X^* > 0$ denote the change in the quantity of good X consumed. If the demand curve is linear, then the change in consumer surplus, $\Delta CS$, can be computed by using the formula:

$\Delta CS = -(\Delta P)(X^*) - (1/2)(\Delta X)(\Delta P)$

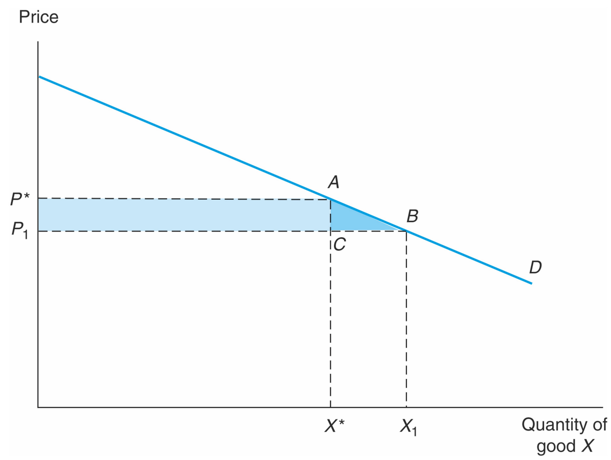



Figure 3: Change in consumer surplus

Sometimes we may not know the shape of the demand curve and, therefore, may not know directly how many units will be demanded after a price change, but we may know the (own) price elasticity of demand, $E_d$.

The price elasticity of demand is defined as the percentage change in quantity demanded that results from a 1 percent increase in price.

$E_d = (dX/dP)/(X/P)$

Then the following can be an approximation for the change in consumer surplus for demand curves which exhibit (close to) constant elasticity:

$\Delta CS = -X^*\Delta P - E_d X * (\Delta P)^2/2P^*$

> Exercise 1: A person's demand for gizmos is given by the following equation: $q = 6 - 0.5p + 0.0002I$, where $q$ is the quantity demanded at price $p$ when the person's income is $I$. Assume initially that the person's income is $60000EUR$.
>
> - At what price will demand fall to zero? (This is sometimes called the choke price because it is the price that chokes off demand.)
> - If the market price for gizmos is $10EUR$, how many will be demanded?

- At a price of $10EUR$, what is the price elasticity of demand for gizmos?
- At a price of $10EUR$, what is the consumer surplus?
- If price rises to $12EUR$, how much consumer surplus is lost?
- If income were $80000EUR$, what would be the consumer surplus loss from a price rise from $10EUR$ to $12EUR$?

### 1.5 Producer surplus

Producer surplus is the supply-side equivalent of consumer surplus. It is the difference between total revenues (a rectangle bounded by $P^*$ and $X^*$) and total variable costs. Diagrammatically, it is the area between the price and the supply curve; see Figure 4.

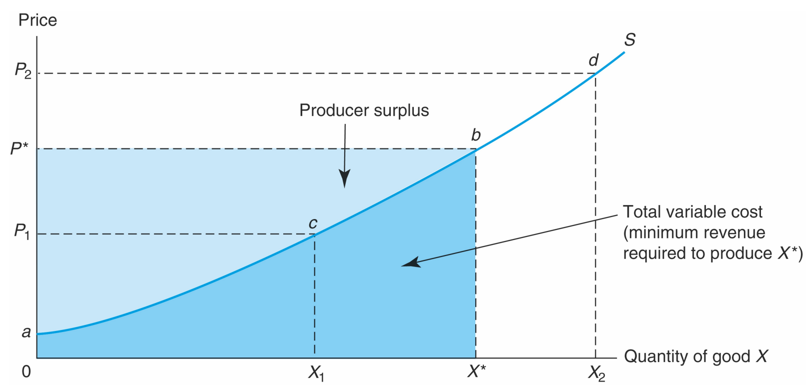



Figure 4: Producer surplus

### 1.6 Social surplus

Consumer surplus plus the producer surplus equals social surplus. Graphically (see Figure 5), social surplus is the area between the demand and supply curves to the left of the equilibrium point.

Since demand reflects MB and supply reflects MC, net social benefits (social surplus) is maximized where the supply and demand curves intersect. Because this equilibrium price $P^*$ and output level $X^*$ come about under perfect competition, we see that perfect competition maximizes social surplus. This point is a Pareto optimum point and it is allocative efficient.

In a perfectly competitive market, anything that interferes with the competitive process will reduce allocative efficiency.

Suppose, for example, government policy restricts output to $X_1$, due, for example, to output quotas. At least some people will be worse off relative to output level $X^*$. The loss in social

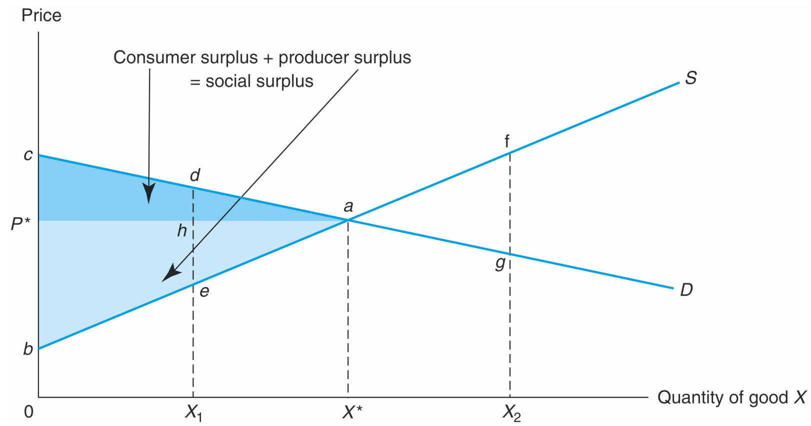



Figure 5: Social surplus

surplus at $X_1$ would equal the triangular area $dae$ – the area between the demand curve (MB) and the supply curve (MC) from $X_1$ to $X^*$.

Similarly, the loss in social surplus at $X_2$ would equal the triangular area $afg$ – the area between the demand curve and the supply curve from $X^*$ to $X_2$. These deadweight losses reflect reductions in social surplus relative to the competitive market equilibrium (at $X^*$). A government policy that moves the market away from the perfectly competitive equilibrium increases deadweight loss and reduces social surplus.

### 1.6.1 profits and factor surplus

The formula for producer surplus is not always satisfactory. Sometimes, a change in policy would not only affect the profit of the producer of the final product but can also change prices of the inputs used in production.

We will come back to this issue later in more detail. For the time being, we can simply say that a more useful description of the surplus generated at the production stage could be a sum of the final producer's surplus (sometimes called profits or surplus to those holding entrepreneurial capital) plus the surplus to the factor owners.

$SS = CS + \Pi + FS$

The incremental net social benefit $\Delta SS$ of a change in policy is then given by:

$\Delta SS = \Delta CS + \Delta\Pi + \Delta FS$

### 1.6.2 Budget impact

Thus far, we have only considered the effects of policy changes on consumers and producers. Public projects are financed with public funds managed by the government, so changes in government revenue or expenditure actually come back to individual investors or consumers.

Often in projects subjected to CBA, a government can bear all the costs of a project, while none of the financial benefits accrue to them. For example, the government may cover all the expenses of constructing rent-free housing for disabled individuals. To simplify matters and align with the assumption of perfect competition, it's reasonable to assume that there's no change in producer surplus in such situations. The project's benefit is the increase in consumer surplus, the cost is the net government expenditure, and the net social benefit equals the benefits minus the costs.

Now, let's contrast this with a scenario where the government builds the same housing but charges a market rent. Is it still reasonable to assume that there's no change in producer surplus? Indeed, it is. We can measure the benefit as the change in consumer surplus and the cost as the change in government expenditure (i.e., construction costs plus operating costs). The rent paid is essentially a transfer – a cost to consumers but a benefit to the government. The net effect of a transfer is zero. Thus, it can be disregarded in the calculation of net social benefits.

> Consider the following example.
>
> Assume that initially the perfectly competitive market shown in Figure 6 is in equilibrium at a price of $P^*$ and the quantity $X^*$. Then suppose that a new policy is enacted that guarantees sellers a price of $P_T$. Generically, these are known as *target pricing policies*.
>
> At a target price of $P_T$, sellers desire to sell a quantity of $X_T$. However, buyers are willing to pay a price of only $P_D$ for this quantity, so this becomes the effective market price. Under target pricing, the gap between $P_T$ and $P_D$ is filled by subsidies paid to sellers by the government.
>
> As the marginal cost of producing $X_T$ exceeds marginal benefit for this quantity of good X, a social surplus loss (deadweight loss), corresponding to area *bde*, results from the policy.

### 1.6.3 Marginal cost of public funds

Most public policies and projects are financed using public funds, which are raised through taxation. However, taxes often lead to distortions in production efficiency and are associated with deadweight losses.

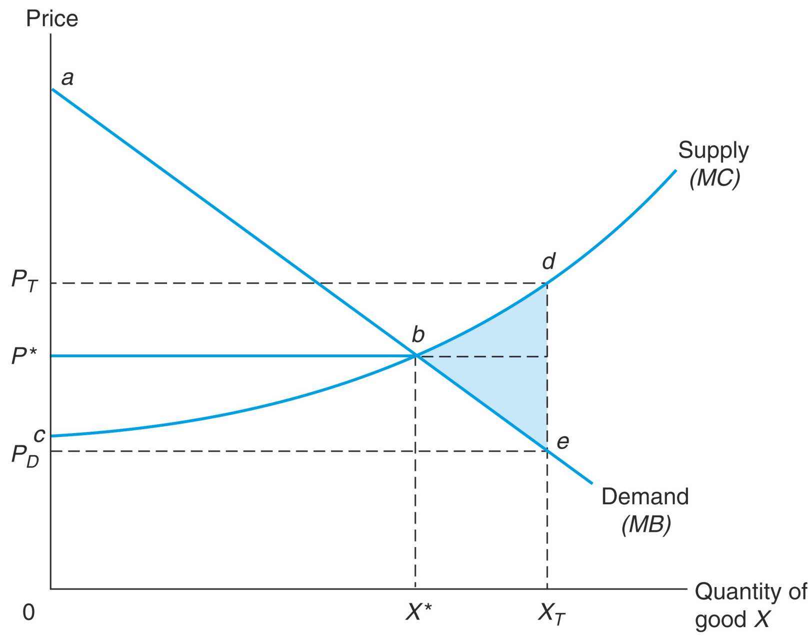



Figure 6: Deadweight loss from target pricing

We refer to the portion of each tax or subsidy that contributes to a deadweight loss as a potential source of leakage or an excess tax burden. In the case of the target pricing example, for instance, this leakage or excess tax burden can be represented by the area $bde/P_T deP_D$. TThe increase in deadweight loss resulting from raising an additional dollar of tax revenue is termed the *marginal excess tax burden* (METB).

The social cost of generating monetary units through taxes is equivalent to 1 + METB, which is called the *marginal cost of public funds* (MCPF).

Numerous empirical estimates exist regarding the magnitude of METB. Additionally, there are direct estimates available for MCPF. These estimates vary for several reasons, including the type of tax, as mentioned earlier.

For instance, in the case of Norway, NOU 1997:27 (Nyttekostnadsanalyser – Prinsipper for lønnsomhetsvurderinger i offentlig sektor) discussed the appropriate size of the marginal cost of public funds. They recommended setting the MCPF in CBA at 1.2. For a detailed discussion on this matter, refer to Holtsmark and Bjertnæs (2015).

A program that requires government expenditure funded by taxation incurs a social cost equal to the amount spent multiplied by the MCPF.

### 1.6.4 Reading materials for Section 1

1. Boardman et al. Chapters 2 and 3.
2. "Veileder i samfunnsøkonomiske analyser" section 3 and 4.

# 2 Section 2

## 2.1 Quantifying effects in the impacted market

So far, we have considered a general framework to quantify the impact of a policy by measuring the changes in social surplus within a market.

Ignoring the non-price goods for the time being, we can broadly classify the impacted markets into three categories: the output market, the input market, and the secondary market.

Measuring the change in surplus critically depends on whether we are dealing with an efficient or an inefficient market.

Recall that a market is allocatively efficient if the output is produced up to the level at which the marginal benefit is equal to the marginal cost. Consequently, the social surplus is maximized at the equilibrium production level.

If the market is efficient, then there will be no deadweight loss before or after the policy/project is implemented. Further, prices match the (social) marginal cost of production, thereby accurately mirroring the genuine opportunity cost of production. In contrast, in an inefficient market, prices fail to accurately represent the true opportunity costs. To quantify the cost of a project, it is customary to introduce the notion of a shadow price—a concept we will introduce later in our discussion. We will begin with a preview of the main findings/rules that we will derive.

- Changes in output markets
  - Value benefits as WTP for the change and costs as WTP to avoid the change.
  - Value change as net change in social (i.e., consumer and producer) surplus plus (less) any increase (decrease) in revenues.
- Change in input markets
  - Value costs as the opportunity cost of the purchased resources.
  - When the market is efficient: If supply schedule is flat, value cost as direct budgetary expenditure.
    If supply schedule is not flat, value cost as direct budgetary expenditure less (plus) any increase (decrease) in social surplus in market.
  - When the market if inefficient: Value costs as direct budgetary expenditure less (plus) any increase (decrease) in social surplus in market.
- Change in secondary markets
  - Most impacts in secondary markets can be valued in primary markets

- If the market is efficient: If prices do not change in secondary market, ignore secondary market impacts.
  If prices do change, but benefits in primary market are measured using a demand schedule with other market prices held constant, then social surplus changes in the secondary market will always represent reductions in social surplus that should be subtracted from changes in the primary market.
  However, if benefits in the primary market are measured using a demand schedule that does not hold other prices constant, ignore secondary market impacts.
- If the market is inefficient: Costs or benefits resulting directly from increases in the size of the distortion should, in principle, be measured.
  Other impacts in secondary market should be ignored if prices do not change.

## 2.2 The output market

A project or policy has the potential to shift the demand or supply curve in the output market. Such shifts in the demand and/or supply curves can sometimes lead to changes in prices (resulting in a price effect), while at other times, they may leave the price unaffected (resulting in a case without a price effect).

However, to accurately assess the impact of these shifts, it's crucial to comprehend the origins of the changes. For example, a shift in the supply curve might stem from alterations in the underlying production costs or the introduction of additional (free) goods. Similarly, a shift in the demand curve might arise from changes in underlying benefits or simply from the demand of an additional set of consumers. These diverse sources of shifts in demand and supply curves have implications for how we measure changes in social surplus.

We will now illustrate these issues using two examples. These examples pertain to efficient markets, where prices accurately reflect the marginal cost of production.

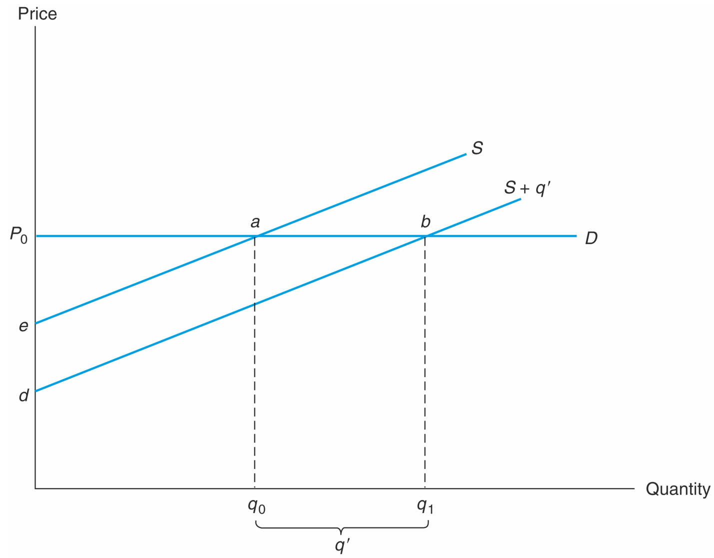
### 2.2.1 Example 1

Project (a): Suppose the government possesses certain good in excess quantities that it can supply to the market at zero cost.

Then the supply curve shifts to the right, resulting in an increased availability of goods to consumers at various price levels. However, if this increase in supply is relatively insignificant compared to the overall market supply, price will not drop.

Let's initially examine the scenario where price remains unaffected. If the government chooses to sell the additional units of the goods at the prevailing market price, it will be treated akin to other competitors in an efficient market. In this context, it will encounter a horizontal demand curve labeled as $D$.

Consider Figure 1, illustrating a rightward shift in the supply curve while maintaining the price at $P_0$.




Social surplus change (ignoring costs of project inputs to the government):
Project (a): Direct increase in supply of $q'$—gain of project revenue equal to area of rectangle $q_0abq_1$
Project (b): Supply schedule shift through cost reduction for producers—gain of trapezoid *abde*

Figure 7: Shift of supply with no effect on price

What will be the extent of the surplus change resulting from a shift in the supply curve?

Given our exclusive focus on the output market, we can identify the affected parties as the consumers and producers within this market.

The demand curve being horizontal, the change in consumer surplus is zero. The interpretation is that even if these additional quantities being sold, what the consumers pay are the same as what they are willing to pay.

The change in producer surplus is also zero; albeit with a more intricate rationale. Despite the rightward shift observed in the supply curve, there is no reduction in the marginal costs of production for the units sold earlier.

The government receives revenue equal to $P_0$ times $q'$, the area of rectangle $q_0abq_1$. The revenues received by the government are the only benefits that accrue from the project selling $q'$ units in the market.

Suppose the government instead of charging the price $P_0$, offers the good for free. Then, the rectangle $q_0abq_1$ will constitute a positive change in consumer surplus with zero government revenue. Thus, whatever price government charges, the net gain from this shift of supply curve will remain unchanged and is measured by the area $q_0abq_1$.

Here we ignore any additional costs of inputs to supply the surplus good to the market. Had there been any costs, it should be included in the calculation as well.

Project (b): Consider an alternative possibility in which the shift of the supply curve happens not due to the supply of some surplus goods, but because of an investment that reduces marginal costs of production. Such a reduction will move the supply curve downward and can be illustrated by a similar rightward shift as depicted in Figure 1.

In this case, measure the change in surplus for producer will be different. It will be represented by the trapezoid $abde$. No change in consumer surplus as before, due to the perfectly elastic horizontal demand curve at $P_0$.

### 2.2.2 Example 2

Next, consider a possibility similar to the project (a) above but assume that price moves downward due to a shift of the supply curve. Figure 2 illustrates this scenario with a downward sloping demand curve $D$.

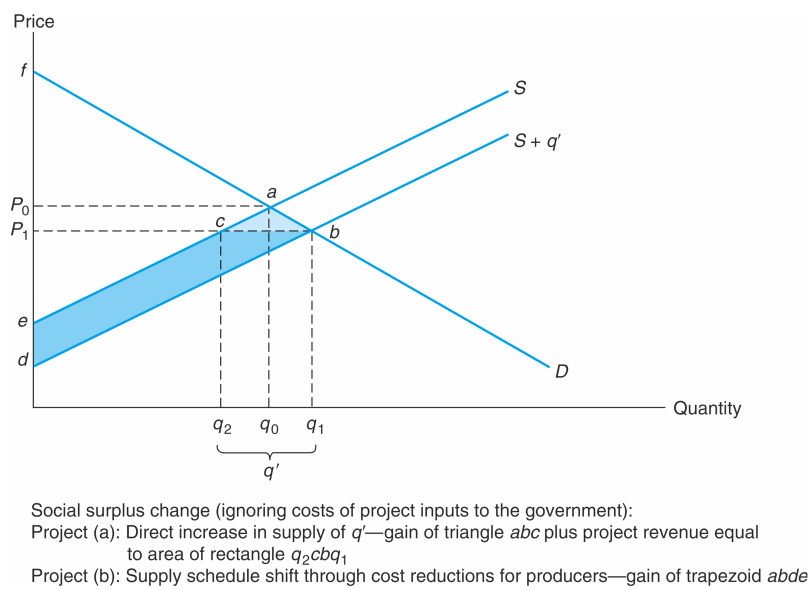



Figure 8: Shift of supply curve resulting in price reduction

Due to drop in price, there is an increase in consumer surplus, measured by the trapezoid $P_0abP_1$.

The new equilibrium demand is $q_1$. The production level comprises of two kinds of good, one was originally produced with an upward sloping marginal costs curve, of volume $q_2$, and the surplus volume that the government supplies of $q_1 - q_2 = q'$ units.

As the shift of supply curve is not due to a reduction of the marginal costs, it implies that producer surplus must be going down on the earlier production volume $q_2$, and this reduction is measured by the trapezoid $P_0acP_1$ .

And, there is an additional government revenue from selling the surplus good, measured by the rectangle $q_2cbq_1$.

Adding these three changes, we can find that the total change in surplus will be the area of $q_2cabq_1$.

As before, we assume that the surplus volume was available and could be supplied at zero costs. Otherwise, the cost of producing/supplying the additional goods should also be taken into account.

Project (b): We can also consider an alternative possibility in which the shift of the supply curve happens because of an investment that reduces marginal costs of production. Such a reduction will move the supply curve downward and can be illustrated by a similar rightward shift as depicted in Figure 2.

In this case, measure the change in surplus for producer will be different. The producer surplus before the project is the triangle $P_0ae$ and the producer surplus after the project is the triangle $P_1bd$. Therefore the change in producer surplus is the trapezoid $ecbd$ - the trapezoid $P_0acP_1$.

There will be a gain in consumer surplus, measured by the area of the trapezoid $P_0abP_1$.

Adding all of them, the net change (gain) in surplus will be measured by the area $eabd$.

The above two examples illustrate the effects of a shift in the supply curve, either due to the supply of surplus volume or a reduction in marginal costs, within an efficient market.

In cases of market inefficiency, determining the changes in surplus can become more intricate. This complexity arises from the presence of deadweight loss at the market equilibrium, both prior to and following the implementation of a policy.

As we covered in the previous lecture, taxes or subsidies can lead to allocative inefficiency (resulting in deadweight loss) in an otherwise efficiently functioning perfectly competitive market.

Additionally, other factors can contribute to market inefficiency, such as monopoly, information asymmetry, and externalities, among others.

---

## 2.3 The secondary markets

A project's effects can be extended to markets beyond the primary output market. Suppose that undertaking a project results in a lower price for a good.

It is expected that the demand for a complement product would increase. If this complementary product is in perfectly elastic supply, there will be no change in its price.

However, since the demand curve for the product has shifted right, the area of consumer surplus measured under the demand curve must have increased.

Should this increase in consumer surplus be measured and included in CBA?

In theory, the answer depends on whether the secondary market is efficient (such that the price there equals the marginal costs of production) or distorted.

Consider first the case of an efficient secondary market.

We refer to Figure 3(a), in which we consider a project that reduces marginal costs of production in the primary market resulting in a drop of price. In Figure 3(b), we observe a rightward shift of the demand curve of a complementary good.

### 2.3.1 Efficient market with no price change

Suppose the secondary market has a perfectly elastic supply curve, and therefore there will be no price change even after the demand shift. The question is whether we should include the apparent gain in consumer surplus, measured by the area $cdfe$ in our analysis?

The answer is no.

The shift in demand within the market for the complementary good reflects consumers reallocating their expenditures to capitalize on the lower price of the good provided by the project. The advantage of this lower price is comprehensively quantified by the change in consumer surplus, as measured in the project's output market.

To illustrate this point (see Chapter 7 for further discussion), let's consider the example of stocking a lake near a city with fish, which subsequently reduces fishing costs for residents. As depicted in Figure 3(a), this leads to an increase in the number of fishing days for residents and results in a gain in consumer surplus.

The market for fishing equipment constitutes a secondary market and has witnessed an upswing in demand, as indicated in Figure 3(b). However, accounting for this surplus gain in the secondary market would result in duplicating the benefits.

To understand why, let's recognize two types of consumers that could potentially comprise the additional consumer base in the primary market: those who already possess the equipment and thus would not contribute to the supplementary demand for equipment, even if their valuation

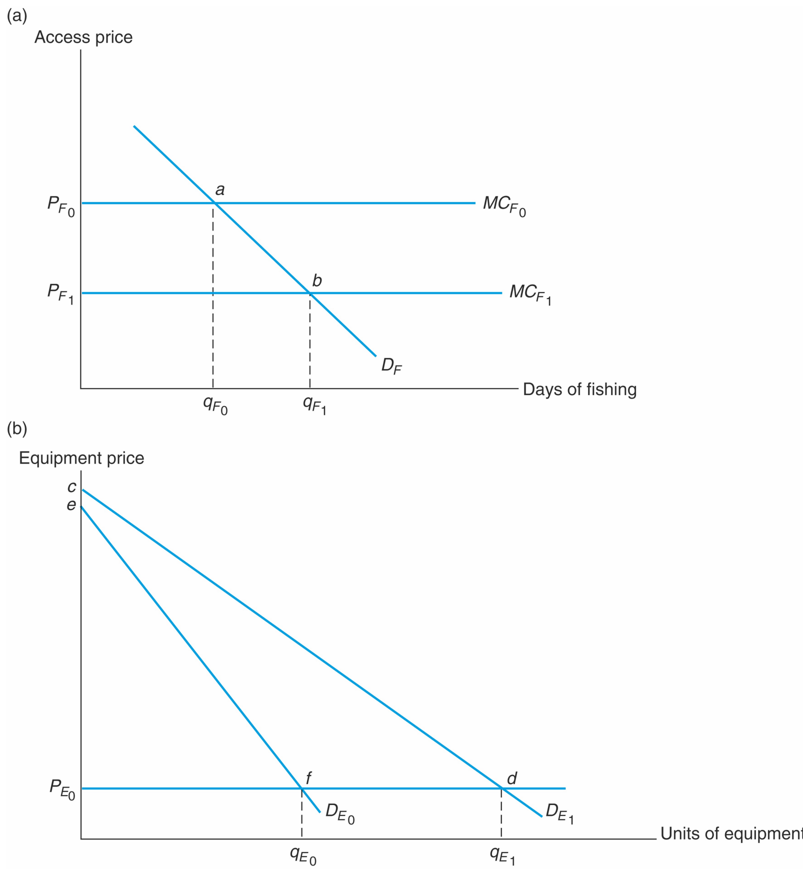



Figure 9: Efficient secondary market with no price effect

for it has increased; and those who previously lacked the equipment and could now join the augmented demand in the secondary market.

Nonetheless, for the second type of consumer, their willingness to pay for fishing days in the primary market has already factored in the potential benefits and costs of acquiring the equipment in the secondary market. Consequently, their net benefits are encompassed in our calculation of the change in consumer surplus in the primary market.

### 2.3.2 Efficient market with price change

Suppose the price changes in the secondary market. Can we still disregard the change in surplus there to avoid double counting? The answer is yes, but the argument is somewhat more intricate.

Figures 4(a) and 4(b) build upon the previous example illustrating the cost reduction in the primary market for fishing. This extension considers a potential secondary market involving golfing activities, which can be viewed as a substitute for fishing activities.

To understand these figures, first observe the effect in the primary market. Stocking the lake with fish reduces the marginal cost of fishing, and a movement along the demand curve $D_{F_0}$ from the point $a$ to the point $b$. This movement is equivalent to a reduction of price from $P_{F_0}$ to $P_{F_1}$ in the primary market.

A consequence of the price reduction in the primary market implies a leftward/downward shift of the demand curve for the substitute good, golfing activities. This movement is captured in Figure 4(b) by the shift of demand curve from $D_{G_0}$ to $D_{G_1}$. Assuming a upward-sloping supply in this secondary market implies a drop in price there, in particular, from $P_{G_0}$ to $P_{G_1}$.

What is the measure of change in surplus in the secondary market? And, should we include that in our CBA?

As we discussed in the previous example (no price change), had there been no changes to the price, the downward shifting of the demand curve would not constitute any additional loss to consumer surplus that is not captured in the calculation of change of surplus in the primary market.

However, in this case, the drop in price in the secondary market implies existing consumers there would have a gain in surplus (they are paying less than what they were paying previously) and a drop in producer surplus (they are receiving less than what they were getting before).

But some of these changes are simply transfer between producers and consumers and should not constitute as a change of social surplus. To have a better estimate of the change, observe first what consumers might be gaining: the area $P_{G_0}efP_{G_1}$ - this is because the existing $q_{G_1}$ consumers are experiencing the lower price (in addition to a lower willingness to pay).

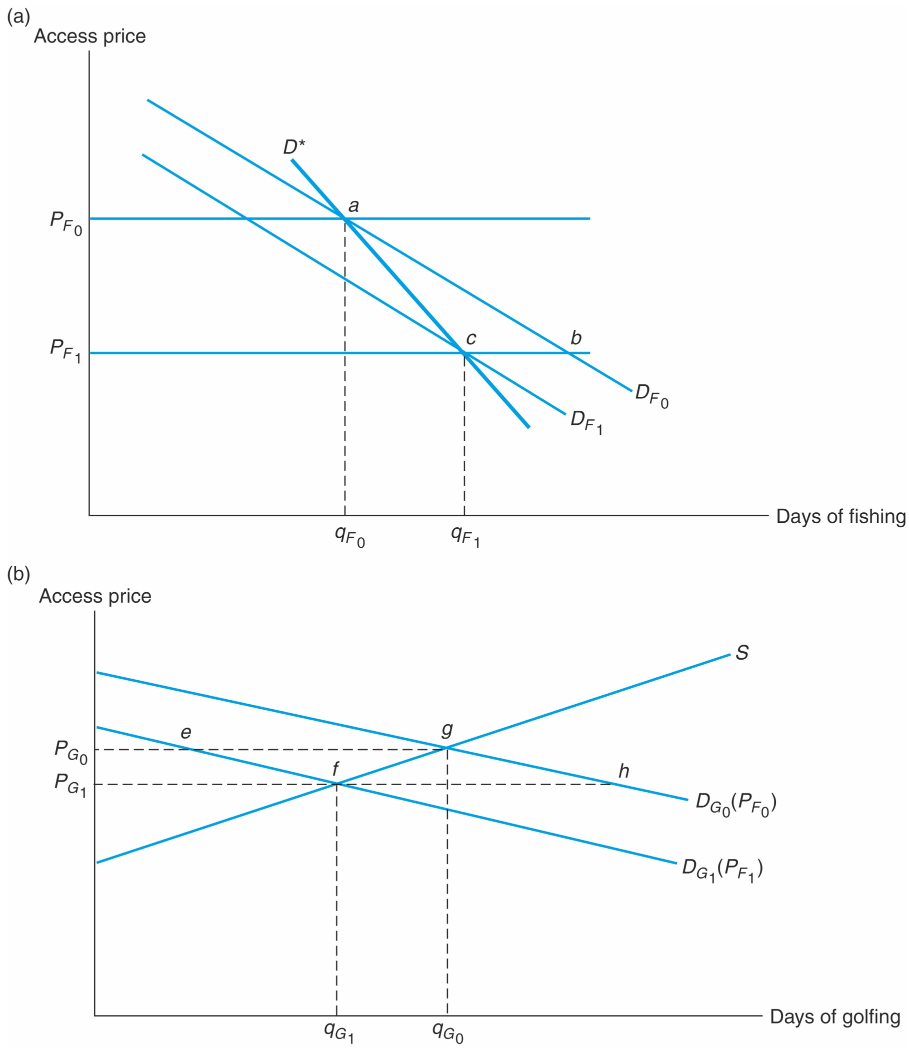



Figure 10: Efficient secondary market with price effect

In contrast, the producers would losing the area $P_{G_0}gfP_{G_1}$ - this is because the price drops and the consumer base is also shrinking.

The net effect is a loss of surplus the area $efg$ in the secondary market, arising from the change in price and shrinking of the consumer base.

Should we take it into account in CBA?

This effect is solely a first-order price effect in the secondary market. Any alteration in the price of the substitute good within the secondary market should also impact the primary market. How do we assess the second-order feedback effect in the primary market?

A decline in price for the substitute good indicates a corresponding decrease in demand for fishing activities. This, in turn, results in a downward or leftward shift in the demand curve, leading to a reduction in prices within the primary market.

Given that we can anticipate a series of feedback effects between these two markets due to price changes, let's consider, for the purpose of our analysis, that the equilibrium price in the secondary market stabilizes at $P_{F_1}$, causing a corresponding shift in the demand curve to $D_{F_1}$ in Figure 4(a). It's important to note that these demand schedules signify the demand for fishing activities while keeping prices in other markets constant. In particular, $D_{F_0}$ and $D_{F_1}$ represent the demands in the primary market when the price in the secondary market is held at $P_{G_0}$ and $P_{G_1}$, respectively.

The initial equilibrium in the primary market is at point $a$ and after the project is implemented, it moves to point $c$ in Figure 4(a). Upon connecting these two points, we arrive at what we commonly refer to as an observed demand schedule $D^*$—representing the demand curve in the primary market without the assumption of constant prices in other markets.

From an empirical perspective, we frequently estimate the observed demand schedule rather than the demand schedule that presumes constant prices in other markets. However, quantifying surplus based on the observed demand schedule under-represents the change in consumer surplus in the primary market. It's measured by the area $P_{F_0}acP_{F_1}$ instead of the area $P_{F_0}abP_{F_1}$, which signifies the surplus gain in the primary market. The disparity is depicted by the triangle area $abc$ —a measure of the underestimation when estimates are drawn from the observed demand schedule $D^*$.

The underestimation of surplus in the primary market, as determined by the area $abc$, is often a practical approximation for not estimating the loss in producer surplus within the secondary market, represented by the triangle area $efg$.

In fact, this approach of estimating surplus change based on the observed equilibrium demand schedule in the primary market, while disregarding the corresponding change in the secondary market (assuming the secondary markets are efficient), is recommended. This recommendation stems from the empirical difficulties involved in identifying all conceivable substitutes or complements for the primary good.

Taking into consideration our observation from the scenario where no price changes occur, as previously discussed, we can formulate a guiding principle here: *Irrespective of whether price adjustments occur in the secondary market, when measuring change of surplus in the primary market through empirically derived demand schedules that were estimated without holding prices constant in secondary markets, it is advisable to disregard effects in efficient secondary markets.*

### 2.3.3 Distorted secondary markets

If the secondary market has distortion or inefficiency, meaning that the price does not accurately represent the (social) marginal cost, the estimation of the surplus change in the primary market even using the observed equilibrium demand schedule will unfortunately overlook certain pertinent effects within the secondary market. This is because the changes in deadweight losses before and after the project will not be reflected in the measure of change in surplus in the primary market.

The following example, which is an extension of the previous example of stocking a lake with fish, can illustrate this possibility.

Suppose in the secondary market for fishing equipment, there is some negative externality from production, because of which the social marginal cost is higher than the private marginal cost of production. In this case, even in absence of the project, there will be excessive (socially costly) production. However, now with the implementation of the project, as the demand curve shifts further to the rights (because the market for fishing equipment faces a higher demand due to a drop in cost of fishing at the lake), there will be an additional deadweight loss, measured by the shaded area in Figure 5.

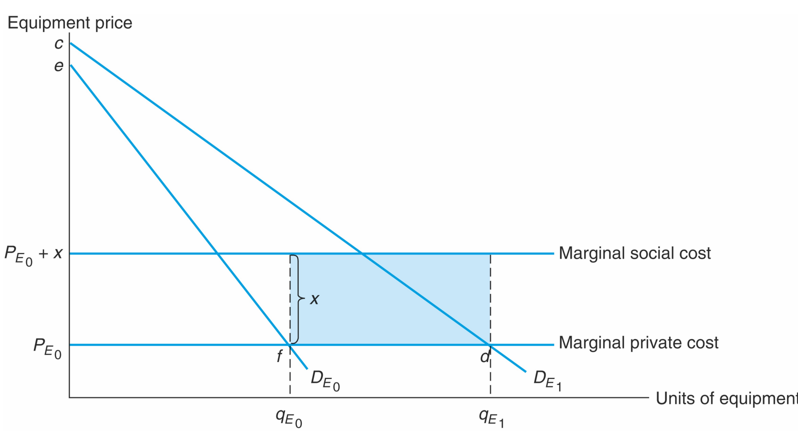



Figure 11: Secondary market with negative externality

Similar distortionary effects can also be found in the secondary markets, due to government intervention, such as taxes, subsidies, quotas etc.

However, in practice, obtaining an accurate estimate of the surplus change in the secondary market, beyond what we can capture through our measurements in the primary market, is typically challenging and empirically complex. Estimation issues arise when attempting to gauge the extent of demand shifts in secondary markets, as well as when measuring the magnitude of these distortions.

If price fluctuations in secondary markets are expected to be minor, it is reasonable to anticipate that significant demand shifts are unlikely to occur in those markets. Consequently, even in cases where secondary markets experience distortions, disregarding these markets may lead to relatively minimal bias in CBA.

### 2.3.4 Other sources of indirect effects

In the previous examples, the indirect effects in secondary markets stem from shifts in the demand for complementary or substitute goods due to price changes in the primary market. Additional sources of indirect effects may also be present; for example, a shift in the cost curve within markets where the primary good is used as an input.

While this effect differs in nature, a similar question arises: Can the change in surplus within these indirectly affected markets be adequately captured by concentrating solely on the change in surplus within the primary market?

In both scenarios, the recommendation remains consistent. If the indirectly affected markets are free from distortions, then we can disregard the indirect effects. However, in cases of distortion, it is important to acknowledge them, although we can anticipate that neglecting these markets might introduce relatively minimal bias, particularly if substantial price changes are not observed within those markets.

The argument unfolds as follows: Let's consider a situation where a policy intervention in the primary market leads to a reduced cost of input in another market. Naturally, this circumstance would likely increase producer surplus in the second market. Nevertheless, the degree to which producers can retain this surplus hinges on the competitive pressure within the second market. When competition is intense, producers are inclined to transfer this surplus to consumers by lowering prices (or if the goods are intermediary, it will raise consumer surplus in subsequent markets as well). However, as we've previously discussed in examples of demand shifts in secondary markets, such indirect increases in consumer surplus are already captured within the direct surplus gain achieved in the primary market. Our argument, however, relies on the efficient operation of secondary markets.

## 2.4 Quantifying effects in the input market

Before we introduce issues related to measuring changes in surplus in the input market, we briefly discuss the concept of a shadow price, which will frequently be used in our analysis.

Remember that in an inefficient market, prices fail to precisely represent the real opportunity costs.

As an example, public universities have recently introduced student fees for non-EEA students. However, as of now, the exact amount is not determined through a market clearing mechanism. It's highly unlikely that this fee accurately represents the true incremental cost of education for the marginal applicant or their benefits. In some instances, we might not even encounter a fee or price – consider, for instance, the valuation of an industrial worker's accident at the workplace.

When prices are not readily observed, or when observed prices fail to accurately capture the social value of a good, we assign explicit values or adjust observed prices to align them as closely as possible with the correct social values. These adjusted prices are called shadow prices. These shadow prices, for instance, the value of a statistical life saved or the social cost of different pollutants, are essential elements of CBA and various methods are being used to determine them.

In some of our analyses below, particularly regarding the measurement of input costs in distorted markets, we assume the explicit utilization of shadow prices.

The central question of interest here is: *What is the cost of a project*?

The implementation of a project requires resources. Once resources are assigned to the project in question, they become unavailable for other productive activities. Consequently, every project entails an opportunity cost, which is the focal point of our measurement.

When assessing the opportunity cost of a project, it is important to consider whether the resources are sunk. In the scenario where a project is already underway and certain resources have been committed to the process, calculating the opportunity cost of the project at an interim stage requires identifying the most advantageous alternative utilization of those resources. For instance, if the optimal alternative involves selling those resources in the secondhand goods market after incurring scrapping costs, then the opportunity cost would be the difference between the price of the used goods and the scrapping cost. This might even result in a negative value.

In conceptual terms, this opportunity cost is equivalent to the value of the most valuable alternative uses of these resources. However, we might only have knowledge of the current budgetary expenditures we are incurring, such as the price times the volume of resources used. When does this expenditure accurately represent the opportunity cost, and when does it fall short of doing so?

The answer depends critically on how the input market functions and whether it is distorted or not.

As previously, we will first illustrate the case of an efficient market.

### 2.4.1 Measuring input costs in an efficient market

Assume that the input market is efficient, indicating that the factor price accurately represents the genuine marginal cost of the input. We can explore two distinct scenarios.

The first involves no price effect: the implementation of the project brings about no alteration in the factor price.

The second scenario deals with a price effect: the implementation of the project results in a change in the factor price.

#### 2.4.1.1 Efficient input market and no price effect

An example of this case arises when the project increases the demand for an input, yet its price remains unaffected due to a perfectly elastic supply curve. This situation is more probable when the rise in demand for the input isn't substantial in relation to the overall national demand for that input.

Consider Figure 6. Assume that the project creates an additional demand for input, shifting the demand curve in the input market from $D$ to $D+q'$. As the input supply is perfectly elastic, there is no change of price, the equilibrium point moves from $a$ to $b$. The area under the supply curve represents the total opportunity costs of the factor. Consequently, the opportunity costs of using the additional $q'$ units of inputs to the project is simply $P_0.q$, as measured by the shaded area $abq_1q_0$.

Therefore, in this case, the budgetary expenditure perfectly reflects the opportunity costs.

However, when the project's input has a perfectly inelastic supply curve, even if the input market functions efficiently, presents a different situation. Consider, for example, the government acquires land for the project, which could otherwise be traded in the private market where consumers could purchase it for developing buildings/houses.

In such situations, government often decides to pay a 'fair' market based price to the seller to acquire the land via eminent domain. Consider Figure 2, which presents a similar case. Here, the government expenditure is given by the area $PbAO$.

Does it also reflects the social opportunity costs of the input? No, it underestimates the costs. The reason is that the potential private buyers of the land lose consumer surplus, measured by the triangle $aPb$ in Figure 7, as a result of the government taking away their opportunity to purchase land. This loss is not included in the government's purchase price.

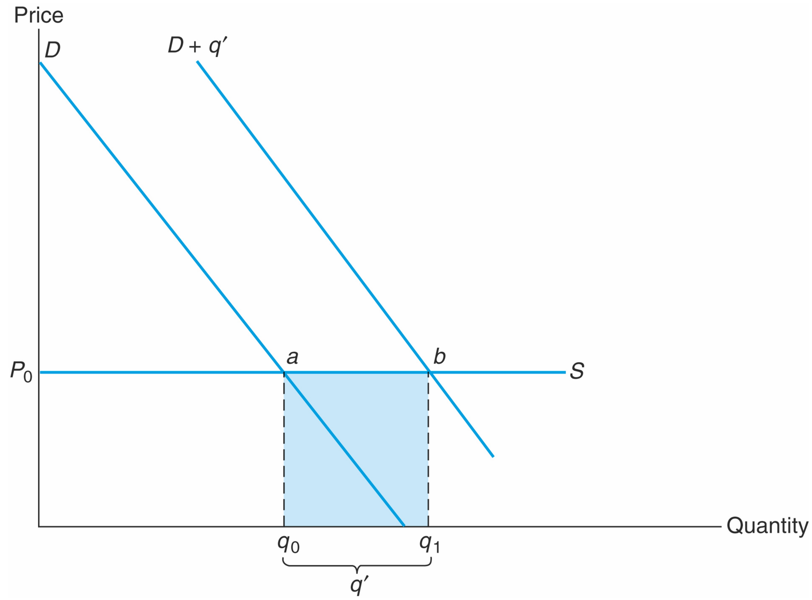



Figure 12: Efficient input market with perfectly elastic supply curve

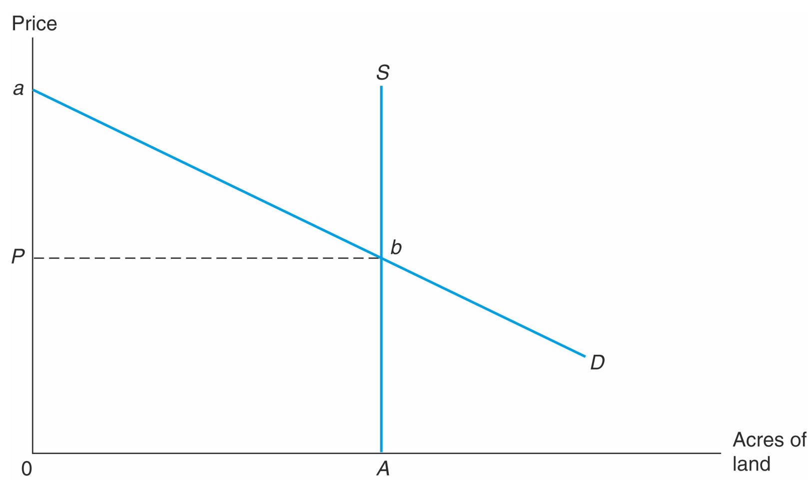



Figure 13: Efficient input market with perfectly inelastic supply curve

#### 2.4.1.2 Efficient input market and significant price effect

When a large quantity of an input is purchased, its price may increase, even when purchased in an efficient market. In such a scenario, the project encounters an upward-sloping supply curve for the resource.

The price increase will have an adverse impact on the original buyers in the input market. Refer to Figure 3: The purchases made by the original buyers decrease from $q_0$ to $q_2$, while the total purchase of the input rises to $q_1 = q_2 + q'$.

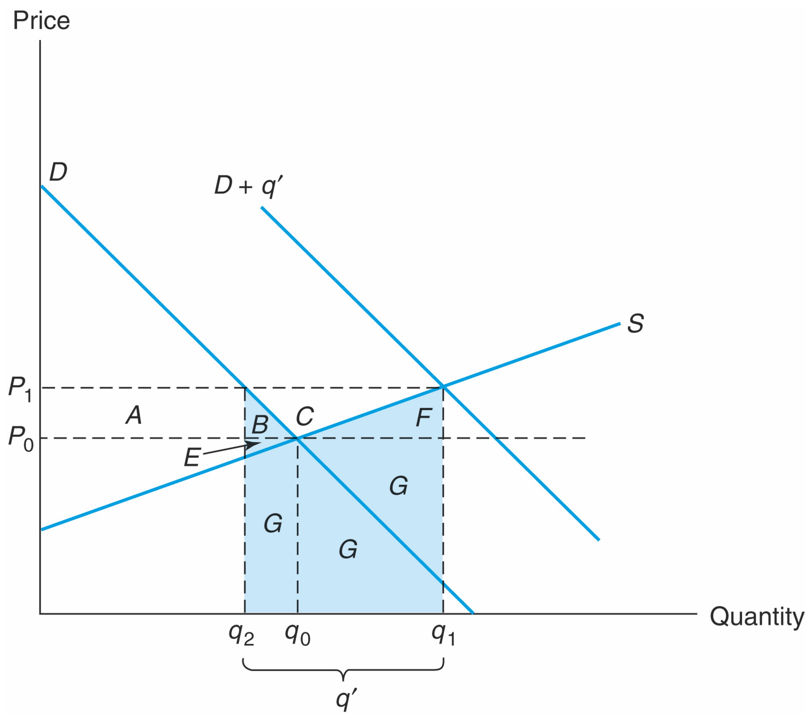



Figure 14: Change of price in an efficient input market

The $q'$ units of the resource purchased by the project come from two distinct sources:
(1) units bid away from their previous buyers and
(2) additional units sold in the market.

For (1), there will be a transfer of surplus between consumer and producer, measured by $A$, and a loss of surplus, measured by $B + E$.

For (2), a part of total expenditure is a transfer to the producer surplus, measured by $B+E+C$, and the rest is the total cost of supplying $(q_1 - q_2)$ units of input, which is the area below the supply curve, measured by $G + F$.

However, when we measure the cost of the project simply by the budgetary expenditure, which is $P_1.q'$, we cover the whole area $B+C+E+G+F$, it does not exclude a potential transfer from the government in the form of producer surplus, measured by $C$. The budgetary expenditure therefore overestimates the true opportunity costs of the project.

Also, note that the true opportunity cost in this case actually exceed the real cost of supplying the input, which is measured by $G + F$.

Sometimes, instead of paying a fixed price (the cost of the marginal unit) for every unit, the government estimates the total cost of supply and purchases the aggregate volume at its cost value. In Figure 8, this implies that the government expenditure in that case will be $G + F$. However, even then, the expenditure will not perfectly reflect the opportunity cost of the project, as it ignores the loss of surplus measured as $B$, which constitutes a part of the opportunity costs of the project.

The fundamental point here is that when prices change, budgetary outlays do not equal social costs. Unless the increase in prices is substantial, however, the change in social surplus will be small relative to the total budgetary costs. This suggests that in many instances, budgetary outlays will provide a fairly good approximation of the true social cost. If prices do increase significantly, though, budgetary costs need to be adjusted for CBA.

How do we make this adjustment in practice?

We often measure the opportunity cost by multiplying the amount purchased by the average of the new and old prices – that is, by $\frac{1}{2}(P_1 + P_0)$ times $q'$.

The average of the new and old prices here serves as a shadow price; it reflects the social opportunity cost of purchasing the resource more accurately than either the old price or the new price alone.

### 2.4.2 Measuring input costs in a distorted market

A variety of circumstances can lead to inefficiency, including the absence of a functioning market, market failures (e.g., public goods, externalities, monopolies, and information asymmetries), and distortions resulting from government interventions (such as taxes, subsidies, regulations, price ceilings, and price floors).

Any of these distortions can arise in factor markets, making the estimation of opportunity costs more complex. Shadow pricing becomes necessary to accurately measure the opportunity cost of inputs in such cases.

We will now discuss some of these scenarios:

1. The government makes purchases of an input that is in fixed supply.
2. The government hires labour from a market in which there is unemployment.
3. The government purchases inputs for a project from a monopolist.

#### 2.4.2.1 Inputs with inelastic supply

When the supply of an input is fixed (e.g., due to an import quota), then a project will increase the market price of the input, thereby reducing the consumption of it by current consumers.

See Figure 4 and consider the pre-project input demand as $D_b$ and post-project input demand as $D'_b$. Because of the project, price increases, and the government's budgetary expenditure will be $P_2.(Q_1 - Q_2)$.

Does it actually reflect the opportunity cost of the project? No.

To see this, note that a section of old consumers will no longer remain as consumers, measured by $Q_1 - Q_2$, which is replaced by the project's input demand.

Further, the existing consumers face an increase in price and thereby a loss of surplus. However, this loss of surplus for the current consumer is simply a transfer, and therefore, a gain in the form of producer surplus and would not appear in CBA.

For the outgoing consumer, the loss of consumer surplus is given by the triangular area below the demand curve $D_b$ and above the price line $P_1$ for quantities between $Q_1$ and $Q_2$. The gain in producer surplus due to purchase of input for the project is the rectangle between the price lines $P_2$ and $P_1$ and for quantities between $Q_1$ and $Q_2$. The sum of two changes results in a positive net surplus in the form of producer surplus. Thus, a part of the government's budgetary expenditure is a transfer in the form of a producer surplus, which we would not have seen if the project was not undertaken.

To obtain an estimate of the opportunity costs, we would need to subtract this potential transfer from the government expenditure. As mentioned earlier, a common practice in this context is to determine the cost of the project by using the average of the old and new market prices as a shadow price.

#### 2.4.2.2 Hiring unemployed labours

When the unemployment rate is high, a project is likely to draw a significant proportion of its workers from the unemployed labour force. If a portion of the labour input would have been unemployed in the absence of the project, how can we measure the opportunity cost of the project?

Consider Figure 10. Let's assume that, for certain reasons, such as union pressure or wage laws, there is a wage floor set at $P_m$. This implies that the employment level remains at $L_d$, while $L_s - L_d$ segments of the labour force are willing to work at the wage level of $P_m$ but remain unemployed.

As a result of the project, let's suppose that labour demand shifts to$D + L'$, leading to the employment of an additional $L_t - L_d$ workers. The budgetary expenditure amounts to $P_m.(L_t - L_d)$; does this accurately reflect the opportunity cost of the project's use of labour?

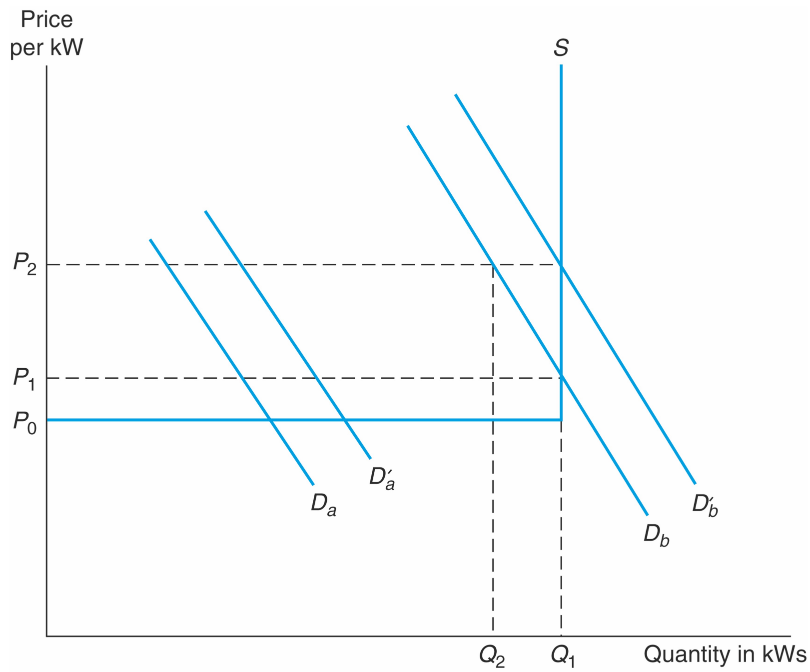



Figure 15: Input in fixed supply

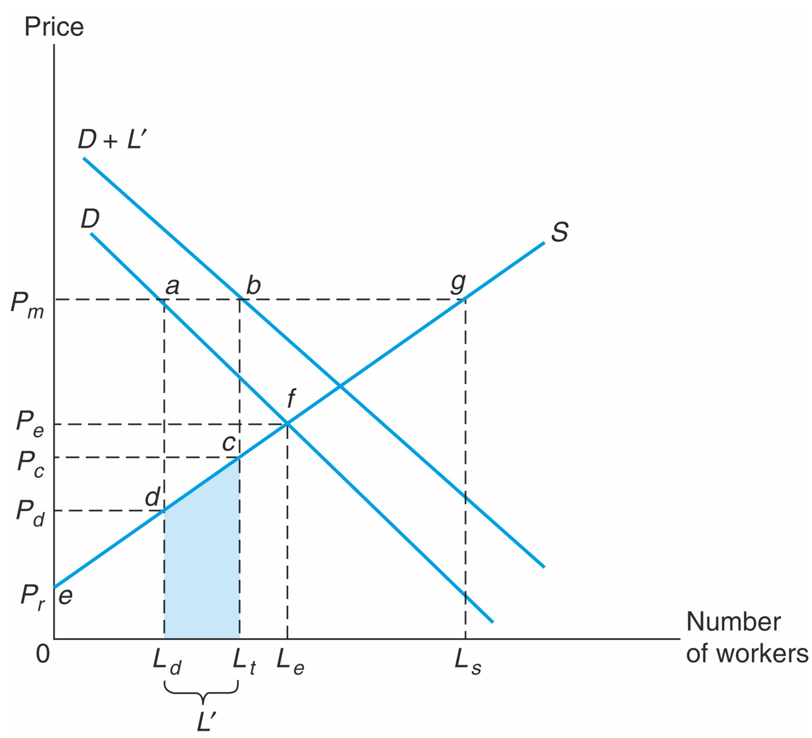



Figure 16: Labour supply with unemployment

Various solutions have been proposed in different scenarios, yet the issue remains quite complex. First, can we regard the opportunity cost as zero because these workers would be unemployed otherwise? Not necessarily, as doing so would imply treating the unemployed as if their time is valueless. This approach is inappropriate for two reasons.

First, many unemployed individuals are engaged in productive activities such as job searching, childcare, and home improvements. Second, even if the unemployed were entirely at leisure, leisure itself holds value for those enjoying it.

So, one possibility could be to estimate the wage at which they are willing to accept a job offer. In other words, this would provide us with a measure of their opportunity cost of labour, which is given by the shaded area in Figure 5.

However, this measure can still be inappropriate for the following reasons: The labour supply curve presents workers in the order of their respective willingness to accept a wage offer. When the project recruits workers, it may not necessarily be in the same order. Even those already employed are not necessarily the workers who are willing to accept lower wages. Therefore, the shaded area hardly captures the opportunity cost of labour for the new workers.

It is common practice to consider a shadow price, typically calculated as an average of $P_m$ and the reservation wage $P_r$. However, there are empirical challenges associated with estimating the reservation wage, as it is likely to vary depending on the nature of work and the heterogeneity of the workforce. At times, a compromised solution is to use a shadow wage of $\frac{1}{2}P_m$.

#### 2.4.2.3 Input market with a monopoly supplier

Consider a scenario where inputs are procured from a monopoly, as illustrated in Figure 11. Recall that the monopolist supplies up to the point where marginal revenue equals marginal cost.

In Figure 11, before the project, the monopolist supplies up to $Q_1$, and the price remains at $P_1$. With the project being implemented, the demand curve shifts to the right, resulting in an increase in both price and quantity sold to $P_2$ and $Q_2$, respectively.

The government's budgetary expenses is $P_2.(Q_2 - Q_3)$. Does it accurately reflect the opportunity costs? No, it is an overestimate.

To illustrate this, consider that the monopolist's producer surplus increases as it sells more at a higher price. Initially, a portion of the original buyers exits the market, resulting in a loss of surplus equivalent to area $C$. The remaining consumers, apart from the project, also experience a loss in surplus due to the high price, measured by $B$. However, the loss of $B$ effectively transfers to the monopolist as an increase in her producer surplus. In addition to $B$, the monopolist also realizes an additional gain in her producer surplus, measured by the area $C + E + G$.

When we sum up the changes in the surplus of the existing consumers and the monopolist, we find a net positive gain, represented by the area $E+G$. The monopolist realizes this gain only because the project is implemented. Therefore, to obtain an accurate estimate of the project's opportunity cost, we must deduct this gain from the government expenditure, resulting in a true measurement represented by the area $C+A$.

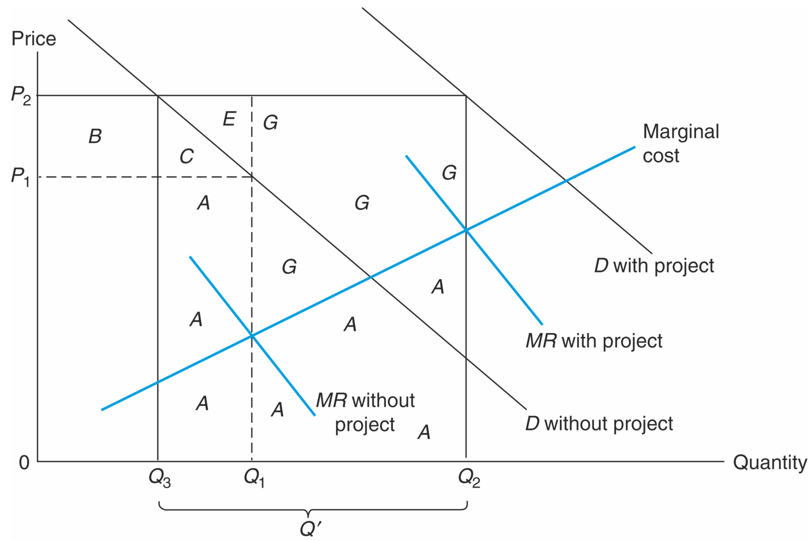



Figure 17: Input with a monopoly supplier

To rectify the overestimation of the opportunity cost, the price should be adjusted downward using shadow pricing. There is no fixed formula for this adjustment, but it's important to consider several contributing factors to the overestimation. These factors include the steepness of the demand and supply curves, the extent of price change, and more.

For instance, the error resulting from the use of unadjusted budgetary expenditures may not be significant if the price charged by the monopolist is not substantially higher than the marginal cost, indicating that the monopoly lacks significant market power.

#### 2.4.2.4 Other distortions

There can be other sources of distortion in the input market, for instance, taxes or subsidies, and externalities.

If supply of an input is taxed, direct expenditures overestimate the opportunity cost; if subsidized, expenditures underestimate the opportunity cost.

If there are positive externalities of the input supply, expenditures overestimate the opportunity cost; in presence of negative externalities, expenditures underestimate the opportunity costs.

An essential part of CBA is to understand the extent to which the budgetary expenditure accurately represents the opportunity costs of the project. As demonstrated by the previous examples, there are several reasons why it might fail to reflect the true cost. While the concept of using a shadow price helps in correcting possible over- and under-estimations of the project's costs, determining an appropriate shadow price is also a complex exercise.

### 2.4.3 Reading materials for Section 2

1. Boardman et al. Chapters 5, 6 and 7.
2. "Veileder i samfunnsøkonomiske analyser" section 3 and 4.

---

# 3 Section 3

## 3.1 Valuation of non-marketed goods

Estimating the change in social surplus relies on knowledge regarding the impact of a policy, such as the number of affected individuals, and the marginal social benefit or cost associated with an additional unit of the affected good or service. In an ideal market scenario, the market price aligns with both the marginal social cost and the marginal social benefit of an extra unit of the good or service. When market conditions are distorted, we search for a shadow price that can better represent the true marginal social cost.

However, in many CBA, markets for specific goods, such as human life or leisure, do not exist. When no market exists for the particular item in question, two primary methods for estimating shadow prices come into play. It's plausible that the shadow price can be indirectly inferred from the market for a related commodity. By examining this related market, we can extrapolate the value of the non-marketed good. This estimation method relies on observing actual behavior, albeit of a different yet related good traded within a market, and typically falls within the realm of *revealed preference* methods.

The second approach to estimating a shadow price involves using contingent valuation through surveys and belongs to methods grounded in *stated preference*. Note that the key distinction between these two methodologies depends on the use of observations of real behaviors. For instance, the survey method can also be employed to gather information about actual behavior, aligning it more closely with revealed preference rather than stated preference.

### 3.1.1 Revealed preference methods

We will begin with discussing the revealed preference methods.

#### 3.1.1.1 Market analogy method–Estimation based on information of an analogous good

Governments often provide private goods such as housing, campsites, university education, among others. However, they may provide these services at substantially lower prices than those prevailing in the market. The price paid for these publicly provided private goods may not align with the market supply curve and only represents a single point on the demand curve. It may, however, be possible to estimate the true demand curve using data from a similar good provided by the private sector.

#### 3.1.1.1.1 Using the market price of or expenditures on an analogous good

The market price of a comparable good in the private sector provides a good estimate of the value of a publicly provided good if it equals the average amount that users of the publicly provided good would be willing to pay (WTP).

Where the government provides a good or service at a lower than market price, the price paid by occupants would generally underestimates the benefit of this service because users would be WTP at least this amount; some might pay more.

Let's consider a local government project that offers apartments to 100 households at a monthly rent of 5.000 kr. The government revenue amounts to 500.000 kr. How can we gauge the full extent of the benefits derived from this project? One straightforward estimate might be to solely look at the revenue figure. However, this approach tends to underestimate the actual benefit, as many residents might be willing to pay more for these apartments.

Now, if we take comparable apartments in the private market, which charge a rent of 10.000 kr per month, and consider 10.000 kr as a shadow price, the estimated benefits amount to 1 million NOK.

It's crucial to consider the target demographic of consumers in the public project. For instance, if these apartments are primarily occupied by well-off individuals, then the revenue figure falls significantly short of representing the true measure of their Willingness to Pay (WTP).

#### 3.1.1.1.2 Using information about an analogous private-sector good to estimate the demand curve for a publicly provided good

Rather than focus on the average amount that users of a publicly provided good are willing to pay, it is conceptually better and easier to think about the demand curve for the good. We can use private-sector data to help map out the demand curve for a publicly-provide good if the goods and their markets are similar. Of course, using expenditures alone underestimates total benefits because it ignores consumer surplus.

Consider a scenario where the government provides a service, such as access to a public wifi, at no cost, and it attracts 300.000 users daily. This observation represents just one specific point on the demand curve. However, if we can identify a comparable location where the same service is offered at a positive price, we can have better estimate of the potential demand curve. For instance, in another area, a similar service is available at a price of 5 kr, and it attracts 100.000 users per day. See Figure 1. Assuming linearity in demand, we can estimate it using the line *abc*, and the consumer surplus would correspond to the area *oac*.

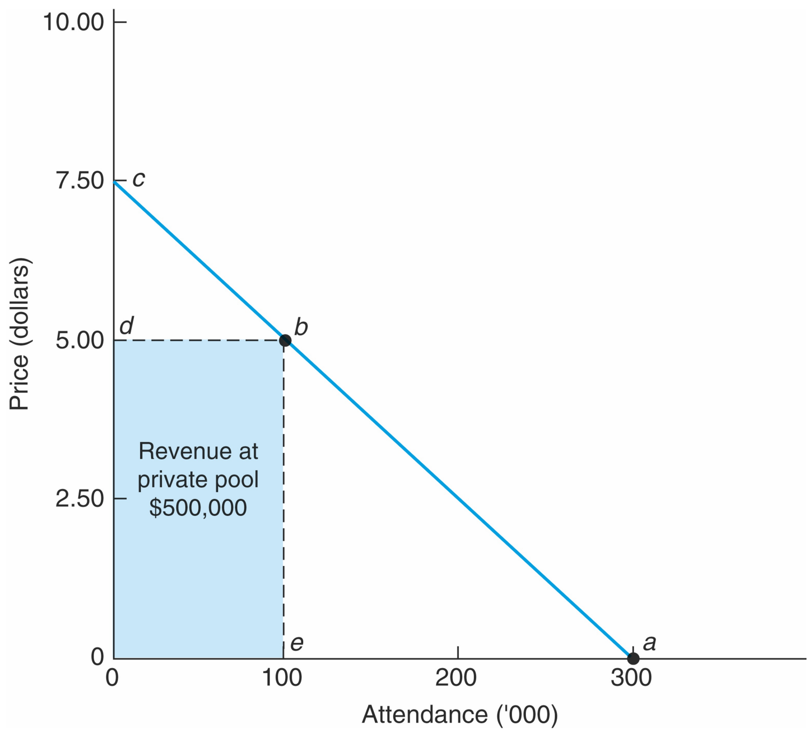



Figure 18: probable demand curve

#### 3.1.1.2 Trade-off method–Estimation of shadow prices based on trade-offs

We can use the opportunity cost, the value we give up to get a certain good, as a measure of its value. For example, time saved could be valued using the after-tax wage rate. Similarly, the trade-off that people make between changes in fatality risk and wages can be used to value a statistical life.

##### 3.1.1.2.1 Value of time saved

When a public project impacts individuals' time, such as alterations in travel duration or waiting in queues for government services, we frequently seek to determine the shadow price associated with the time saved or expended.

The labor market serves as a clear parallel to valuing time saved. In cases where market imperfections are absent (meaning individuals can freely choose their work hours, and there is no unemployment), the wage rate is equivalent to the marginal value of time. Nonetheless, several challenges arise when attempting to use the wage rate as a measure for valuing time saved:

- Wages ignore benefits, which are also a form of compensation for work.
- It should take account of taxes as one's decision to work typically depends on the after-tax wage rate (plus benefits).
- People value different types of time differently.
- The wage rate may not be appropriate due to rigidities in the market or market failures. For example, people may not be able to easily adjust the number of hours they work.

#### 3.1.1.3 Value of a statistical life

The valuation of human life remains a subject of debate. We allocate substantial resources to rescue trapped miners or provide heart transplants to certain individuals, but may not invest in initiatives aimed at improving mine safety or reducing the prevalence of heart disease. In practice, to efficiently allocate resources and assess the advantages of life-saving projects, we must establish a monetary value for a saved life.

##### 3.1.1.3.1 Forgone earnings method

This method suggests the value of a life saved equals the person's discounted future earnings. It generates higher values for young, high-income males than old, low-income females. For retired people, the resultant value of life may be negative. Conceptually, most problematically, this method does not reflect what people are WTP for a small reduction in risk of their death.

#### 3.1.1.3.2 Willingness to purchase safety

This method estimates the value of life by observing how much people pay for life-saving devices, such as safety belts. If people are indifferent between paying an extra 1000 kr to reduce the probability that they will die by 1/10000, then they value their life at 10 million; Consider the indifference condition: $(\rho + \omega)V - 1000 = \rho V$.

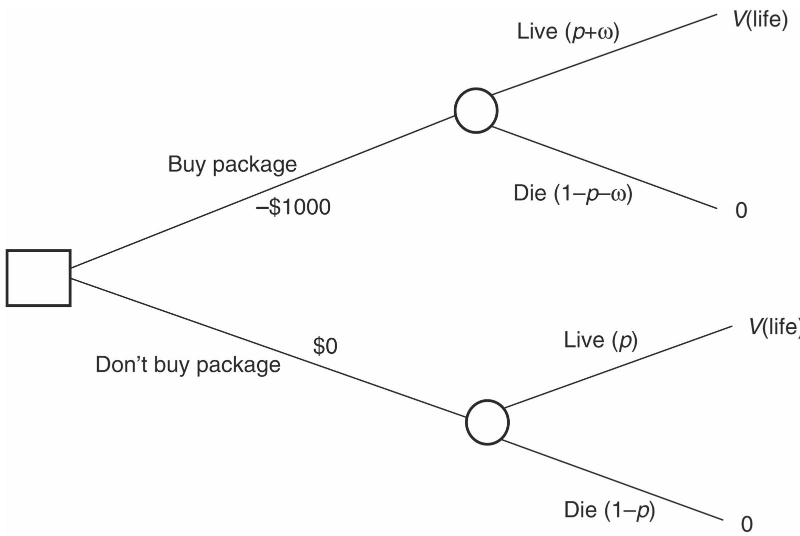



Figure 19: Decision tree for safety package purchase

#### 3.1.1.3.3 Willingness to avoid danger

Similarly, if a person is willing to forgo an extra 50000 kr per year to increase the probability that he will not have a fatal on-the-job accident by 1/1000, then he values his life at 50 million; considering the indifference condition: $(1/1000) * V = 50000$.

There are several problems with these simple approaches:

- These methods assume workers and consumers fully understand the risks, which they may not.
- Are the participants true representative of the population? For example, people who take risky jobs may be more likely to take risks which would lead to a relatively small gap in the salary between risky and less-risky jobs.
- Can we accurately measure the risks?
- The WTP to reduce fatality risk (and therefore the estimated VSL) depends on both the level of risk and the change in the risk level due to the policy. People probably have diminishing marginal utility for safety.

#### 3.1.1.4 Valuation methods–Valuation of related goods

##### 3.1.1.4.1 Estimation based on prices of capital goods

The impacts of a project or policy can be inferred from changes in the price for certain capital goods. For example, the value of noise can be inferred from comparing the price of a house in a noisy neighborhood to the price of a similar house in a quiet neighborhood. Changes in the market values of firms following a regulatory change can be used to estimate the change in producer surplus of the new regulations.

An advantage of using market prices is that information is quickly and efficiently capitalized into prices so that the price differential provides a good estimate of the value of the policy change.

##### 3.1.1.4.2 Non-traded intermediate good

If a project produces an intermediate good that is not sold in a well functioning market, then its value can be imputed by determining the value added to the *downstream activity*.

For instance, consider an irrigation project that supplies water to farmers. If water is traded within a well-functioning market, it becomes feasible to estimate the market demand curve. However, if such a market does not exist, we have to infer the shadow price. We can consider to what extent the level of income changes in the downstream industry.

*Annual Benefit = NI(with project) – NI(without project)*

where, NI = net income of downstream business. The total benefit of a project can be computed by discounting these annual benefits over the project's life. This method can be used to value improvements in human capital, such as training programs, by comparing the average incomes of those in the program to those who are not.

The method assumes the difference in income captures all of the benefits. The assumption is questionable when there may be additional consumption benefits.

##### 3.1.1.4.3 Critique of the simple valuation methods

All of the methods discussed above suffer potentially from the *omitted variable problem* (we assume that only the price of the comparable good has changed but in practice the project could affect numerous goods and we may not control for all potential affected markets) and *self-selection bias*.

#### 3.1.1.5 Hedonic price method

The hedonic price method is a tool for appraising an attribute when its value is reflected in the pricing of assets such as houses or salaries. This method addresses potential challenges arising from omitted variables and self-selection bias and comprises two distinct steps.

Suppose we are trying to assess the worth of a scenic view. First, we estimate the impact of an improved scenic view on the value of houses, represented by a slope parameter within a regression model. Consider, for instance, the following regression equation.

$$lnP = b_0 + b_1 ln(area) + b_2 * ln(view) + b_3 * (closetocity) + b_4 * (otherchac.) + error$$

This equation is commonly referred to as a *hedonic price function* or an *implicit price function*.

Note that the coefficient $b_2$ measures the price elasticity with respect to the variable measuring the quality of *view*.

The change in a house's price resulting from a unit change in a specific attribute, often referred to as the slope, is known as the hedonic price, implicit price, or rent differential associated with that attribute.

In a well-functioning market, the hedonic price can naturally be understood as the extra cost incurred when purchasing a house that offers a slight improvement in a particular attribute.

In the context of the above model, the hedonic price of view can be described as

$$h_v = \frac{\partial P}{\partial view} = b_2 * \frac{P}{view}$$

Next, we can estimate the WTP for scenic views, after controlling for individual taste, which can be proxied by income and other socioeconomic factors. For example, we can estimate the following WTP function (inverse demand function) for scenic views:

$$h_v = W(view, Y, Z)$$

where, $h_v$ is estimated from the hedonic price function, $Y$ is household income, and $Z$ is a vector of household characteristics that reflects tastes. The estimated $W-$function reflects an inverse demand function for scenic views, and can be used to measure the changes in consumer surplus for a certain change in quality of scenic views.

There are several potential problems with hedonic models. First, we implicitly assume that people understand the implications of the attribute that is being valued. For example, people should know the level of pollution at the property they buy and know the expected effect of this level of pollution on their health. In addition, the econometric approach is vulnerable

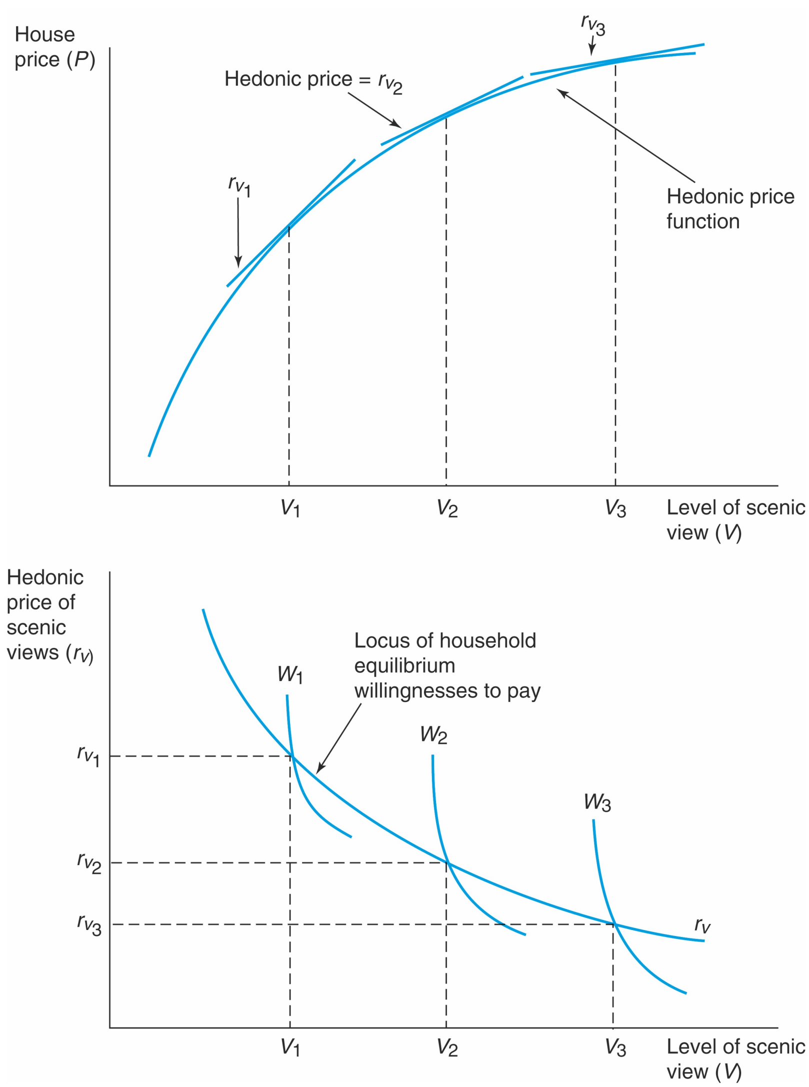



Figure 20: Hedonic pricing

to measurement error and specification error. Further, markets may not necessarily adjust to changes in the attributes of interest and to all other factors in the short run.

#### 3.1.1.6 Travel cost method

Most applications of the travel cost method (TCM) have predominantly focused on assessing the value of recreational sites.

When we seek to estimate the value of a specific recreational site, we anticipate that the number of visits to a recreational site depends on various factors, including the *true* cost to visit, the price of substitutes, income level of the targeted population, and other potential preference-related characteristics.

The TCM acknowledges that the true cost incurred by individuals for a visit to a recreational site encompasses more than just the admission fee. It encompasses expenses associated with traveling to and from the site. Within these travel costs lie elements like the opportunity cost of time spent traveling, the operational costs of vehicles used for travel, expenses for accommodations during overnight stays while traveling or visiting, and parking fees at the site, among others. The sum of these costs constitutes the total cost of visiting the site.

The ingenious aspect of the TCM lies in recognizing that, although admission fees are generally uniform for all individuals and, in some cases, even nonexistent, the total cost encountered by each person varies due to disparities in the travel cost component. Consequently, usage patterns also differ, allowing us to estimate a demand schedule for the site.

Conceptually, estimating the demand schedule involves several steps. First, a random sample of households residing within the market area of the site is selected. Second, these households are surveyed to determine the frequency of their visits to the site over a defined period, encompassing all costs incurred during these visits, expenses related to visiting alternative sites, their incomes, and other characteristics that may influence their demand. Third, a functional form for the demand schedule is specified, and estimation is performed using the survey data.

However, the empirical approach has its limitation in terms of measurement error, selection bias (endogeneity issues), and omitted variable bias.

### 3.1.2 Stated preference methods

For some public goods, there are no obvious ways to determine preferences through observation of behaviors. In these cases, there may be no alternative to asking a sample of people questions about their valuations. These surveys are typically referred to *contingent valuation* surveys.

The primary use of contingent valuation surveys is to elicit information about WTP for changes in the quantity of a good. Valuation by contingent valuation surveys of goods that are directly consumed by potential consumers is relatively non-controversial. However, valuation of passively-used or non-used good (for example, assessment of preserving an archaeological site) with such a survey is more controversial. The uses of these contingent valuation surveys however, are rapidly growing.

Typical steps in conducting a contingent valuation survey are as follows.

1. Identify a sample of respondents from the population.
2. Ask respondents questions about their valuations of a good.
3. Estimate respondents' WTP for the good using information from the survey.
4. Extrapolate the results to the entire population.

#### 3.1.2.1 Direct elicitation method

*Open-Ended method*

Respondents are simply asked to state their maximum WTP for a good or policy being valued.

*Close-Ended Iterative Bidding Method*

Respondents are asked if they would pay a specified amount for the good or policy. If yes, then the amount is increased incrementally and they are asked again (until there is a response of no). If no, then the amount is lowered and they are asked again (until there is a response of yes). This method is however, found to be too sensitive to the initially presented value.

*Contingent Ranking Method*

Respondents are asked to rank specific feasible combinations of the good being valued and monetary payments. An example would be low water quality and low taxes vs. high quality and high taxes, including several combinations in between. This method makes it is easier for the respondent to answer (an ordinal procedure). The WTP must be inferred from the rankings, however, rather than being directly elicited. Also, responses tend to be sensitive to the order in which the alternatives are given.

#### 3.1.2.2 Dichotomous-choice method

Respondents are asked whether they would be willing to pay a particular amount, or bid price, to obtain a good or policy. The range of bid prices are chosen by the analyst. Because many respondents are surveyed, accept/reject probabilities can then be calculated for each bid price. Data can then be plotted in a histogram (number of yes responses versus bid price). The curve fitted to the histogram can be viewed as the demand curve of a randomly drawn member of the sample.

The demand curve shows the probability that an individual would be willing to pay for the good or policy at each price. The area under this curve provides an estimate of the individual's WTP.

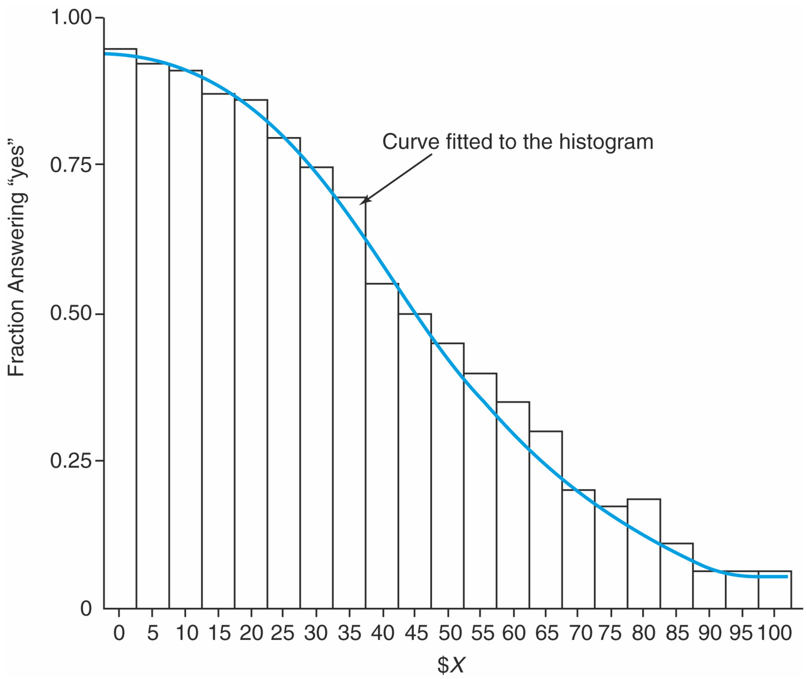



Figure 21: Histogram of dichotomous choice responses

The method has several potential challenges.

First, the issue in consideration can be quite complex. Problem arise in defining exactly what the good or policy is (especially for non-used goods). Second, can a question be phrased to elicit a neutral response? Third, the respondents may have non-commitment bias - they do not have to pay their stated willingness to pay. In addition, in iterative bidding methods seem to quite sensitive to initial bidding value and order to bidding.

Judgmental biases may arise in response to certain questions—i.e. is the question framed as WTP or willingness to accept (WTA)? If consumers act rationally and markets operate efficiently, the distinction between WTP and WTA for most goods should have minimal impact. However, empirical evidence indicates that WTA amounts are often higher than WTP amounts, possibly attributed to loss aversion.

#### 3.1.3 Reading materials for Section 3

1. Boardman et al. Chapters 15
2. "Veileder i samfunnsøkonomiske analyser" section 3.4.

---

# 4 Section 4

## 4.1 Discounting future impacts

We will now address several practical considerations that are essential for calculating the Net Present Value (NPV) of a project. First, we assume that a discount rate has been provided to us. Subsequently, we will introduce the theoretical aspects related to choosing an appropriate discount rate.

### 4.1.1 Basics of discounting

Discounting takes place over periods not years. However, for expositional simplicity, we assume that each period is a year. Consider an investment that lasts for one year and yields return at an annual interest rate of $i$.

The future value in one year of an amount $X$invested at interest rate $i$ is: $Y = X(1+i)$.

Then the present value (PV) of this future aggregate fund of $Y$ must be the same as $X$, which gives us a formula for PV as

$PV \ = \ X = Y/(1+i)$

We can extend this line of argument to investments that yield returns over multiple years.

The present value, $PV$, of an amount $Y$ received in $T-$th years, with interest compounded annually at rate $i$ is:

$PV = \frac{Y}{(1+i)^T}$.

This formulation presents a case of exponential discounting where the $T$-th period discount factor is given by $\delta^T = 1/(1+i)^T$.

The present value for a stream of benefits $\underline{B} = (B_0, B_1, ..., B_T)$ and costs $\underline{C} = (C_0, C_1, ..., C_T)$ over $T$ years (here 0 denoting the current year) *is*:

$PV(\underline{B}) = \sum_{t=0}^{T} \frac{B_t}{(1+i)^t}$ and $PV(\underline{C}) = \sum_{t=0}^{T} \frac{C_t}{(1+i)^t}$

### 4.1.2 Net Present Value (NPV)

Net present value (NPV) of a project with a stream of benefits $\underline{B} = (B_1, ..., B_T)$ and costs $\underline{C} = (C_1, ..., C_T)$ over $T$ years is

$NPV(\underline{B}, \underline{C}) = \sum_{t=0}^{T} \frac{B_t}{(1+i)^t} - \sum_{t=0}^{T} \frac{C_t}{(1+i)^t} = \sum_{t=0}^{T} \frac{B_t - C_t}{(1+i)^t}$.

See, for example, Figure 5 and Figure 6, which illustrates the timeline of a project with streams of benefits and costs accrued over 5 years, and consider an interest rate of 0.04.

| Year | Event | Annual benefits | Annual costs | Annual net social benefits |
|---|---|---|---|---|
| 0 | Purchase and install | 0 | 500,000 | –500,000 |
| 1 | Annual benefits and costs | 150,000 | 25,000 | 125,000 |
| 2 | Annual benefits and costs | 150,000 | 25,000 | 125,000 |
| 3 | Annual benefits and costs | 150,000 | 25,000 | 125,000 |
| 4 | Annual benefits and costs | 150,000 | 25,000 | 125,000 |
| 5 | Annual benefits and costs | 150,000 | 25,000 | 125,000 |
| | *PV* | 667,773 | 611,296 | 56,478 |

Figure 22: Streams of benefits and costs of a project

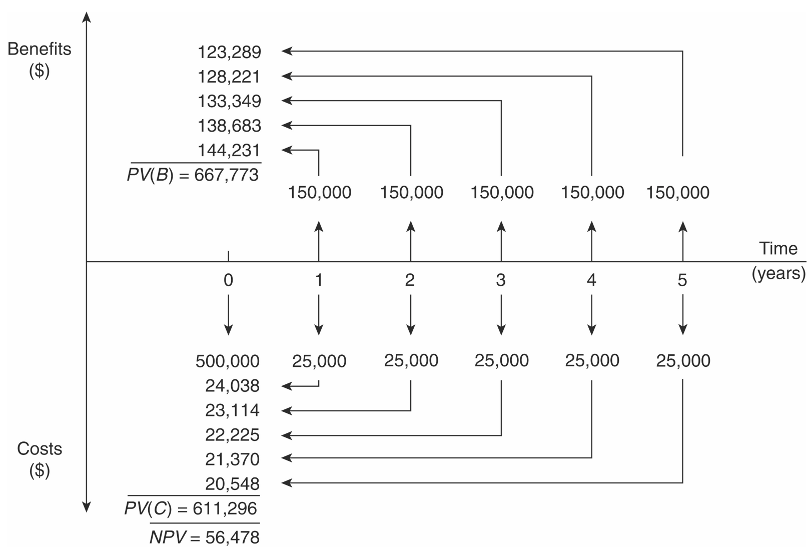



Figure 23: A project with streams of future benefits and costs

Inspecting the formula for NPV, we can see that for a project that typically yields benefits in later years and incurs higher costs in the current year, NPV can decrease with the interest rate $i$. If we set $i = 0$, the formula simply adds the net benefits (benefits minus costs) over the years.

For example, the project described in Figure 6 will have an aggregate (non-discounted) benefit of $750,000$ and an aggregate (non-discounted) cost of $625,000$, thus generating a net figure of $125,000$. For large values of $i$, the future benefits are less valuable in today's terms and will reduce the NPV.

### 4.1.3 Internal Rate of Return (IRR)

The Internal Rate of Return (IRR) of a project equals the interest rate at which the project's NPV equals zero. Sometimes, IRR provides a good indication of the desirability of a project. IRR is the highest interest rate that leaves the project profitable.

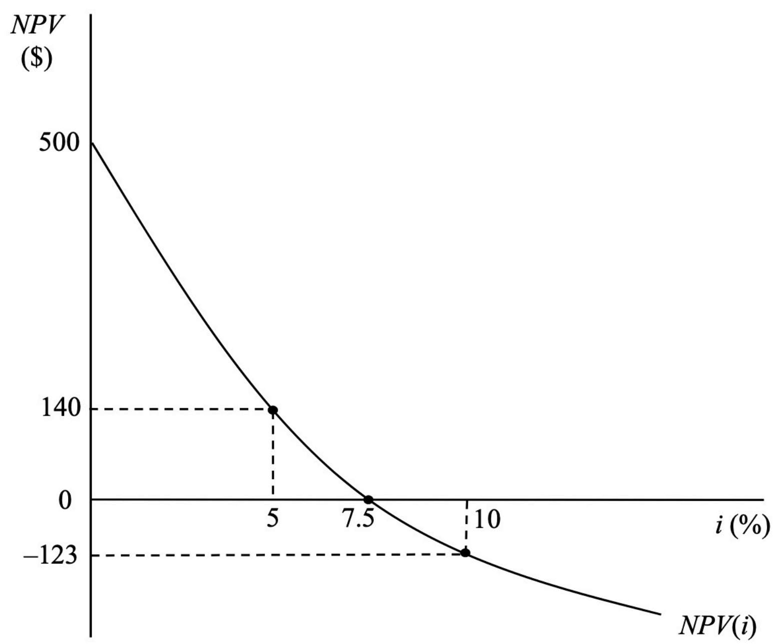



Figure 24: Internal rate of return

### 4.1.4 Compounding over multiple sub-periods

All the previous formulas considered the length of the periods to be one year and the rate of interest/discount to be an annual one. However, it might be possible that benefits or costs are realized at different intervals, such as semi-annually or quarterly, etc. In such cases, the interest rate needs to be adjusted accordingly.

The conversion formula is given by

$(1+i_s)^s = (1+i)$,

where $s$ is the number of sub-periods (for example, 2 in the case of half-yearly, 4 in the case of quarterly), and $i_s$ denotes the corresponding interest/discount rate.

Solving for $i_s$ gives

$i_s = (1+i)^{\frac{1}{s}} - 1$.

### 4.1.5 Some useful formulae

If future returns are constant across periods, then we can use the following formula to represent the sum of a finite geometric series.

$a + ad + ad^2 + ... + ad^n = a\frac{1-d^{n+1}}{1-d}$ where $a > 0$, and $d \neq 1$.

Also note that as $n$ approaches infinity, the absolute value of $d$ must be less than one for the series to converge. The sum then becomes

$a + ad + ad^2 + ... = \frac{a}{1-d}$ where $0 < d < 1$.

Using the formula for the sum of a finite number of terms in a geometric series

$\sum_{t=0}^{T} \frac{V}{(1+i)^t} = V[\frac{1+i-(1+i)^{-T}}{i}]$, and

$\sum_{t=1}^{T} \frac{V}{(1+i)^t} = V[\frac{1-(1+i)^{-T}}{i}]$.

Observe that the difference between the two expressions above stems from the different the number of periods considered in the two series, one starts from period 0 and the other starts from period 1.

### 4.1.6 Annuity and perpetuity

An annuity is an equal, fixed amount received (or paid) each year for a number of years. A perpetuity is an indefinite annuity. Many CBAs contain annuities or perpetuities. Fortunately, there are some simple formulas for calculating their PVs.

The present value of an annuity of $V$ per annum (with payments received at the end of each year) for$T$ years starting from year 1 with interest at $i$percent is given by

$PV = \sum_{t=1}^{T} \frac{V}{(1+i)^t} = V[\frac{1-(1+i)^{-T}}{i}]$

and the present value of the present value of an amount $V$ received at the end of each year (starting from year 1) in perpetuity is

$PV = \sum_{t=1}^{\infty} \frac{V}{(1+i)^t} = \frac{V}{i}$.

### 4.1.7 Equivalent Annual Net Benefits (EANB)

It is sometimes useful to convert the NPV of a project in terms of an Annutiy; for example, find $V$ such that a project with a stream of benefits $\underline{B} = (B_0, B_1, ..., B_T)$ and costs $\underline{C} = (C_0, C_1, ..., C_T)$ satisfies the following:

$NPV(\underline{B}, \underline{C}) = \sum_{t=1}^{T} \frac{V}{(1+i)^t} = V \sum_{t=1}^{T} \frac{1}{(1+i)^t} = V[\frac{1-(1+i)^{-T}}{i}]$.

The value of $V$ refers to the annuity payment (also called *Equivalent Annual Net Benefits* or EANB) and the term $(1-(1+i)^{-T})/i$ is called the $T-$period annuity factor, which equals the present value of an annuity payment of 1kr over $T$ periods at an annual interest rate of $i$. Thus,

$EANB = \frac{NPV}{annuityfactor}$

### 4.1.8 Comparing returns in nominal versus real terms

A relevant question when quantifying future benefits and costs is whether these measurements should be expressed in real or nominal rates.

Fortunately, as long as we maintain consistency in expressing benefits/costs and the discount rate in the same manner, both methods will yield the same comparative assessment. We should either measure benefits and costs in real dollars and discount them using a real interest rate, or measure them in nominal dollars and discount using a nominal interest rate.

Expressing benefits and costs in real terms is sometimes more intuitive because it takes into account the potential effects of inflation on future benefits and costs. However, when doing so, we must convert the nominal interest rate to a real discount rate using the following formula, where $i$, $r$, and $m$ represent the annual nominal interest rate, the annual real interest rate, and the annual inflation rate, respectively:

$(1+r)(1+m) = (1+i)$,

which gives

$r = \frac{i-m}{1+m}$.

Expressing figures in real terms would also grant us greater flexibility when selecting variable inflation rates for future periods. Future inflation rate forecasts are typically obtainable from statistical bureaus.

## 4.2 Project selection

After discounting the benefits and costs, we require a decision criterion for accepting or rejecting a project and for selecting one if there are multiple acceptable projects. NPV can serve as a key indicator for this assessment purpose.

When faced with the decision of whether to accept or reject a single project, it is logical to apply the decision rule:
Accept if $NPV > 0$; reject if $NPV < 0$.

This is because we have already expressed the benefits and costs of the public project under consideration in monetary terms. Consequently, we can interpret the project as an investment project. If we could invest funds in a project that would yield returns at the same interest rate, then its NPV would be precisely zero. We can regard such a project as the opportunity cost of the public project.

A project with a positive NPV implies that the funds invested in the project generate benefits exceeding its opportunity cost. This not only enables the recovery of the invested funds along with interest payments but also yields additional benefits.

This NPV-based decision rule can also be applied to choose among multiple projects. If the impacts of multiple, mutually exclusive alternative projects are calculated relative to the status quo, one should choose the project with the *highest NPV*, as long as this project's $NPV > 0$. If the $NPV < 0$ for all projects, one should maintain the status quo.

Another indicator for assessment is the IRR. Recall that IRR is the value of the interest rate that makes $NPV$ equal to zero.

A commonly used decision rule, when there is only one alternative, is to invest in the project if its IRR is greater than the discount rate.

Similarly, when comparing two projects with different IRRs, one might prefer the one with the higher IRR, especially when one project consistently dominates the other in terms of NPV.

Although IRR conveys useful information, there are some problems associated with using IRR as a decision rule. For instance, IRR may not be unique; that is, there may be more than one interest rate at which the NPV is zero. Second, IRRs are expressed as percentages (ratios), not in monetary values. Therefore, they should not be used to select one project from a group of projects that differ in size.

### 4.2.1 Comparing projects with different time frames

Choosing one project over another solely based on the NPV of each project is problematic if the two projects have different time frames. Such projects are not directly comparable.

Two appropriate methods to evaluate projects with different life spans are:

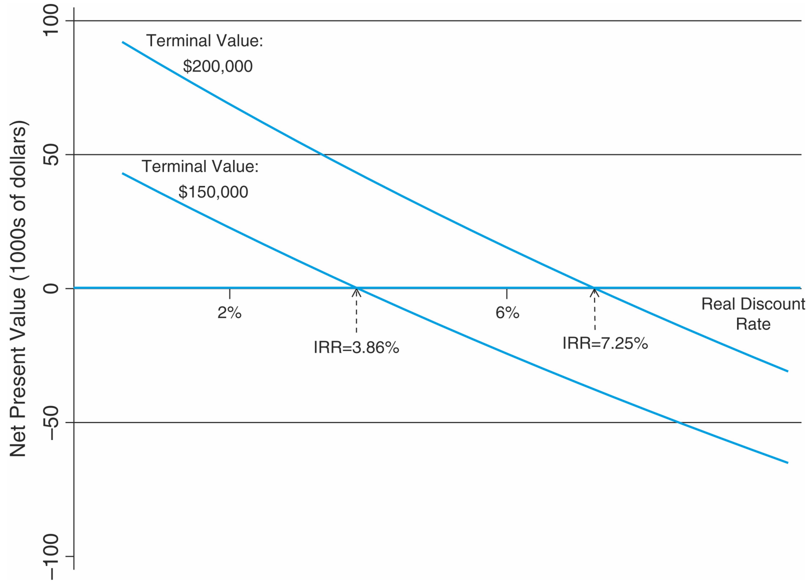



Figure 25: Two projects with different IRR

#### 4.2.1.1 Roll over method

If project $A$ spans $n$ times the number of years as project $B$, then assume that project $B$ is repeated $n$ times and compare the $NPV$ of $n$ repeated project $B$s to the $NPV$ of (one) project $A$.

For example, if project $A$ lasts 30 years and project $B$ lasts 15 years, compare the $NPV$ of project $A$ to the $NPV$ of 2 back-to-back project $B$s, where the latter is computed:

$NPV = x + \frac{x}{(1+i)^{15}}$,

where $x$ is the $NPV$ of one 15-years lasting project $B$.

#### 4.2.1.2 Equivalent Annual Net Benefits Method

Recall that $EANB = NPV/(annuityfactor)$, which is the amount received each year for the life of the project that has the same $NPV$ as the project itself.

Comparisons can be drawn based on EANBs, which provide more comparable per-year figures. However, it does not take into account that the two projects have different lifespans.

### 4.3 Selection of discount rate

So far, we have presented our analysis for a given discount rate. This discount rate effectively assigns varying weights to the benefits and costs accrued in different periods. When assessing government policies or projects, we must determine the suitable weights to apply to policy impacts occurring in different years.

The choice of the appropriate discount rate is equivalent to determining the proper set of weights, also known as social discount rates (SDR).

Various approaches have been employed to determine the Social Discount Rate (SDR). Two prominent methods are the Rate of Time Preference method and the Opportunity Cost of Capital method.

### 4.3.1 An Individual's Marginal Rate of Time Preference (MTRP)

An individual's marginal rate of substitution (MRS) between periods measures how much additional consumption she would accept in the future to be willing to postpone one unit consumption in the current year. Her marginal rate of time preference (MRTP) is $MRS - 1$.

### 4.3.2 Equality of MRTPs and market interest rate in perfect capital market

In a perfectly competitive capital market, an individual's MRTP equals the market interest rate $i$, as shown in Figure 9.

The observation of the equality between MRTP and the market interest rate is significant. It implies that even if individuals have varying time preferences, their consumption choices may differ, but their MRTPs will remain identical to the prevailing market interest rate. We can therefore uniquely characterize the common market interest rate as the social rate of time preference.

### 4.3.3 Equality of the Social Rate of Time Preference and the Return on Investment (ROI) in perfect capital market

Economists have extended the analysis to to incorporate production within a closed economy. The analysis shows that if there are no market failures or taxes, and that there are no transaction costs associated with borrowing and lending, then the MRTP and the market interest rate coincides with the marginal rate of return on investment (ROI). In such a closed economy (with the assumption of perfect capital market), the common interest rate would be the obvious choice for the SDR.

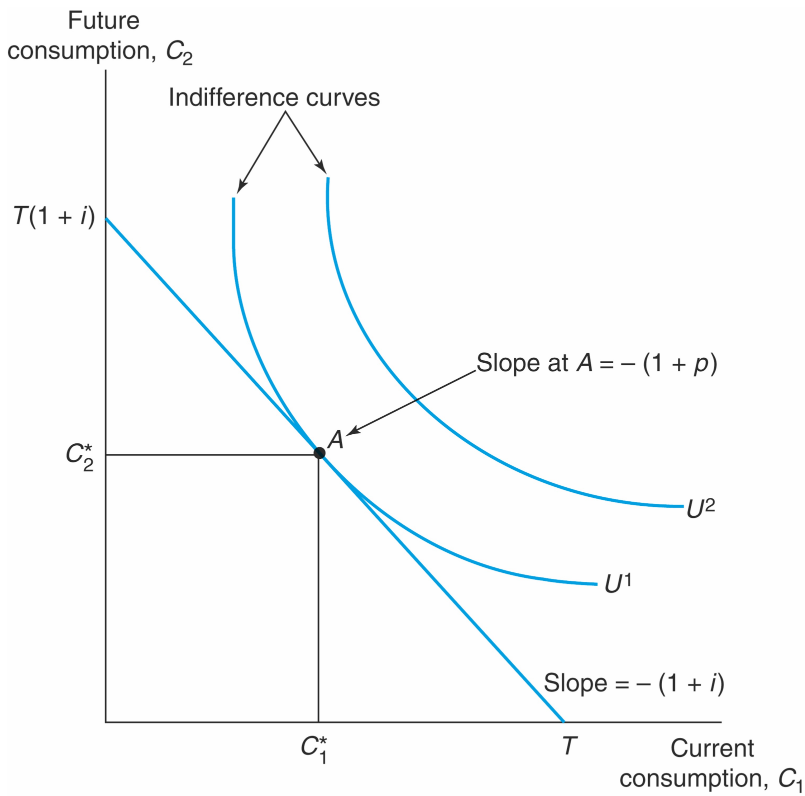



Figure 26: MRTP equals the market interest rate

### 4.3.4 Social opportunity cost of capital method

The social opportunity cost of capital (SOC) method builds on influential work by Arnold Harberger. Harberger analyzed a closed domestic market for investment and savings. He assumed that any new government project would be funded by domestic government borrowing which would raise the interest rate and result in a fall in private investment and an increase in savings.

Harberger argues that the social discount rate should reflect the social opportunity cost of capital, which can be obtained by weighting the marginal rate of return on private investment and the after-tax marginal return on savings (or foregone consumption) with respective weights given by the size of the relative contributions that investment and consumption would make toward funding the project.

### 4.3.5 Ramsey model: Social time preference method

Frank Ramsey suggested an approach for determining the SDR that does not rely on market rates of interest. He proposed a model with infinite periods in which society (or a single representative individual) attempts to maximize a social welfare function that reflects the values society places on per capita consumption over time.

Following this model, the society's *marginal rate of time preference* (STP), would equal the sum of two components: one that reflects the reality of impatience and the other that reflects society's preference for smoothing consumption over time:

$STP = \rho + \mu g,$

where $g$ is the percentage change in per capita consumption (i.e. consumption growth) and is the absolute value of the elasticity of the marginal utility of consumption with respect to changes in consumption, and $\rho$, $g$, $\mu \geq 0$. This equation is known as the Ramsey formula.

## 4.4 Discounting Inter-generational projects

So far we've discussed only constant (time-invariant) SDRs. Time invariant SDR may be appropriate for projects affecting the same group of people over different years. There are reasons, however, to suggest the use of a time-declining SDR for projects that might affects different generations.

Empirical evidence suggests that people use lower discount rates for events that occur farther into the future. Long-term environmental and health consequences have very small present values when discounted using a constant rate, often implying that spending a relatively small amount today to avert a costly disaster several centuries in the future is not cost-beneficial.

Constant rates do not appropriately take into account the preferences of future, as yet unborn, generations.

Constant rates do not appropriately allow for uncertainty as to market discount rates in the future.

To address these issues, a time-declining rate schedule is often recommended.

## 4.5 Current practice in Norway

Current discounting practices for public projects vary depending on the nature and scope of the projects; see the Norwegian guideline in the review committee's report on cost benefit analysis (NOU 2012: 16, pp. 66-67). Here I am quoting some relevant parts of it (from page 13):

> *Recommendation (pp. 13)*
>
> – In principle, the real risk-adjusted social discount rate should reflect the risk-free interest rate and the risk associated with the project and, consequently, reflect the project's risk adjusted opportunity cost of capital… A real risk-adjusted discount rate of 4 percent will be reasonable for use in the cost-benefit analysis of an ordinary public measure, such as a transportation measure, for effects in the first 40 years from the date of analysis.
>
> – Beyond 40 years, it is reasonable to assume that one will be unable to secure a long-term rate in the market, and the discount rate should accordingly be determined on the basis of a declining certainty equivalent rate as the interest rate risk is supposed to increase with the time horizon. A rate of 3 percent is recommended for the years from 40 to 75 years into the future. A discount rate of 2 percent is recommended for subsequent years.
>
> *Historical guidelines (pp. 66)*
>
> From 1967 to 1999, Norway used an approach where the discount rate was based on the Ramsey model. Based on a report by Leif Johansen in 1967, in which it was assumed that  = 1 percent, $\mu = 3$ and $g = 3$ percent, the discount rate was put at 10 percent in circular R-3/1975. In circular R25/78, the discount rate was changed to 7 percent … The recommendations in the NOU 1997: 27 Green Paper were based on the Capital Asset Pricing Model. Circular R-14/99 stipulates that a riskfree real rate of 3.5 percent should be assumed for cost-benefit analysis purposes. Three different risk classes were defined, with a risk-adjusted rate of 4, 6 and

8 percent, respectively. Specific calculation of the risk-adjusted required rate of return was recommended for large projects or groups of projects.

*Current guidelines (pp. 67)*

A new circular R-109/2005, replacing R-14/99, was issued following revision of the Ministry of Finance cost-benefit analysis guide in 2005. The risk-free rate for use in the cost-benefit analysis of central government measures was here put at 2 percent, and it was noted that a normal project will have a risk premium of 2 percentage points, and thus a risk-adjusted required rate of return of 4 percent. For measures where considerable systematic risk may reasonably be assumed, it is stated that a risk premium of 4 percentage points, and thus a discount rate of 6 percent, may be appropriate. Specific calculation of the riskadjusted required rate of return continued to be recommended for large projects or groups of projects. This applies, in particular, to projects falling within the scope of the system for quality assurance of major public projects ... The Ministry of Transport and Communications concluded by applying a discount rate of 4.5 percent to all projects within its area of responsibility, whilst for practical reasons not distinguishing between road, railway and aviation.

### 4.5.1 Reading materials for Section 4

1. Boardman et al. Chapters 9 and 10.
2. "Veileder i samfunnsøkonomiske analyser" section 3.5.

---

# 5 Section 5

## 5.1 Uncertainty and risk

A crucial aspect of conducting a cost-benefit analysis entails addressing uncertainties. The decision to approve or reject a project depends on the future costs and benefits. However, the future is inherently uncertain and so a cost benefit analysis must take into consideration these uncertainties.

We will focus on the following commonly used methods.

1. Expected NPV
2. Sensitivity and scenario analyses
3. Simulation analysis
4. Quasi-option values

### 5.1.1 Expected NPV (ENPV)

Following this approach, we model the future outcome as a set of *relevant contingencies* - a range of possible outcomes that might and will occur.

For instance, our decision to build a dam depends on the availability of irrigation water, which, in turn, relies on future rainfall. Nevertheless, uncertainty exists, and we can describe the nature of this uncertainty by categorizing it into three potential states: high, moderate, and low rainfall in a year. The list is exhaustive (ruling out other possibilities) and mutually exclusive (not overlapping possibilities).

In the context of CBA, the set of contingencies ideally should encompass the entire spectrum of conceivable variations in the net benefits of the policy. However, the potential number of variations is limitless. To address this practically, we should consider incorporating contingencies that offer a diverse range of scenarios, sufficient to represent the possible outcomes effectively.

After we specify a tractable, representative set of outcomes, we should assign probabilities to these outcomes. The probabilities can be objective, for instance, historically observed frequencies, or subjective assessments. Notice that this approach, which models uncertainty as a list of potential outcomes with associated probabilities, effectively transforms it into a risk model.

To illustrate this, let's assume, for the time being, that a project will yield returns for only one period. However, the outcomes (possible benefits and costs) during that period can be uncertain. The expected net benefit is calculated as the weighted sum of net benefits across all potential contingencies, with the weights determined by their associated probabilities.

$$ENPV = \sum_{k=1}^{K} p_k(B_k - C_k),$$

assuming there are $K$ possible contingencies, $p_k \geq 0$ denoting the probability of the $k-$th contingency, and $\sum_k p_k = 1$.

This formulation can be easily extended to a multi-period setting.

For example, if an investment of $I_0$ at time 0 in a project yields $ENPV_t$ at time $t = 1, 2, ..., T$, then we can express the expected NPV of the project as

$$ENPV = -I_0 + \sum_{t=1}^{T} \frac{ENPV_t}{(1+i)^t},$$

where $i$ denoting the discount rate.

In CBA, it is a common practice to treat expected values as if they were the precise values to be gained, even though the realized values seldom match these expected values. This approach is not conceptually accurate when assessing the Willingness to Pay (WTP) in situations where individuals confront uncertainty. Nonetheless, in practical terms, considering them as comparable is reasonable when either risk is spread across a set of policies or when risk is spread across a group of individuals affected by a policy. This practice tends to bring the actual realized values of costs and benefits closer to their expected values.

The above formulation for extending ENPV in a multi period setting works well as long as the probabilities assigned to various contingencies in a specific period are independent of the actions taken in previous periods. For instance, future rainfall is less likely to be influenced by whether a dam was constructed in the previous year. However, in certain situations, this assumption can be questionable. Take, for example, the decision to launch a vaccination project, which depends on the social costs to infection and the potential spread of infection in the future. In this context, whether the vaccination project has already started or not significantly influences the distribution of infection spread.

### 5.1.2 Decision trees and ENPV

Decision analysis can be thought of as an extended-form game. It has two stages.

*Specification of decision tree.* First, we specify the logical structure of the decision problem in terms of sequences of decisions and realizations of contingencies using a diagram (called a decision tree) that links an initial decision to final outcomes.

*Backward induction.* Second, we work backwards from final outcomes to the initial decision, calculating ENPV across contingencies and pruning dominated branches (ones with lower ENPV).

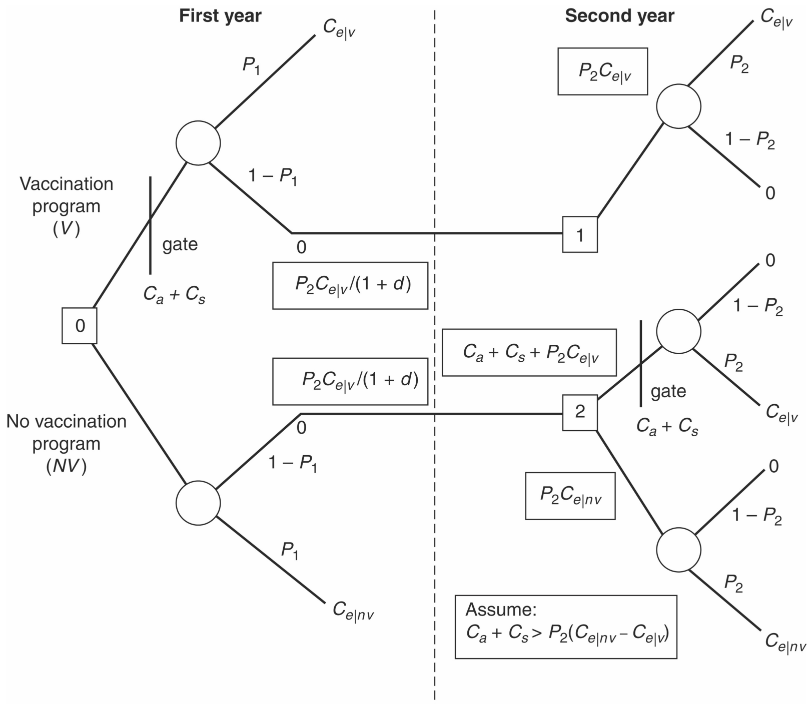



Figure 27: Decision tree and backward induction

Consider the vaccination example in Suppl 4; see Figure 1 (note the typo in the figure - in the NV-arm at the 0-node, conditional on the event of "no epidemic", the return is $\frac{P_2 C_e|_{nv}}{1+d}$.

Three types of costs:

$C_a$ : costs of administering the vaccination project

$C_s$ : costs associated with the side-effects of vaccination

$C_e|_v, C_e|_{nv}$ : costs of epidemic with vaccination and with no vaccination, respectively.

$P_t$ is the probability an epidemic breaking out in period $t = 1, 2$.

At decision node 1, the expected cost is $P_2 C_e|_v$.

At decision node 2, the expected cost of initiating the vaccination project is $C_a + C_s + P_2 C_e|_v$, and the expected cost of not initiating is $P_2 C_e|_{nv}$.

It is assumed in this context that $C_a + C_s > P_2(C_e|_{nv} - C_e|_v)$, which gives us that initiating vaccination at node 2 is a dominated action, and so we can conclude that the rational action to take at node 2 is not initiating the vaccination project and the associated expected cost is $P_2 C_e|_{nv}$.

Finally, coming back to node 0, we can calculate the expected cost (EC) of the vaccination program as $EC(V) = C_a + C_s + P_1 C_e|_v + (1 - P_1)\frac{P_2 C_e|_v}{1+d}$,

and the expected cost of no vaccination program as

$EC(NV) = P_1 C_e|_{nv} + (1 - P_1)\frac{P_2 C_e|_{nv}}{1+d}$.

As an illustration, suppose that we have data suggesting the following values for parameters:

$P_1 = 0.4$, $P_2 = 0.2$, $d = 0.05$;

$C_{e|v} = 0.5C_{e|nv}$ - the vaccination program cuts the costs of influenza by half),

$C_a = 0.1C_{e|nv}$ - the vaccination costs 10 percent of the costs of the influenza), and

$C_s = 0.01C_{e|nv}$ - the side-effect costs are 1 percent of the costs of the influenza).

For these values,

$EC(V) = .367C_{e|nv}$, and

$EC(NV) = 0.514C_{e|nv}$.

We can determine the ENPV of the vaccination program as the expected benefits (which is the costs it can avoid by implementing the project now, i.e., $EC(NV)$) minus the expected cost $EC(V)$, which gives us

$ENPV = 0.147C_{e|nv} > 0$.

### 5.1.3 Sensitivity analysis

There are several key ideas to sensitivity analysis:

- We face uncertainty about the predicted impacts and the values assigned to them.
- Most plausible estimates comprise the *base case*.
- The purpose of sensitivity analysis is to show how sensitive predicted net benefits are to changes in assumptions. (If the sign of net benefits doesn't change after considering the range of assumptions, then the analysis is robust and we can have greater confidence in it.)
- Looking at all combinations of assumptions is infeasible.

#### 5.1.3.1 Partial sensitivity analysis

Consider how net benefits change as one assumption varies, holding other assumptions constant. It should be used for the most important or uncertain assumptions.

The value of a parameter where net benefits switch sign is called the *breakeven* value.

A thorough investigation of sensitivity ideally considers the impact of changes in each of the important assumptions.

For the purpose of illustration, we will consider a more detailed specification of the costs relevant to the decision analysis of the hypothetical vaccination program; See Figure 2 and Figure 3, which specify various parameters used in the analysis and the formulae to computing the cost components.

| Parameter | Value [range] | Comments |
|---|---|---|
| County population ($N$) | 380,000 | Total population in the county |
| Fraction high risk ($r$) | .06 [.04, .08] | One-half population over age 64 |
| Low-risk vaccination rate ($v_l$) | .05 [.03, .07] | Fraction of low-risk persons vaccinated |
| High-risk vaccination rate ($v_h$) | .60 [.40, .80] | Fraction of high-risk persons vaccinated |
| Adverse reaction rate ($\alpha$) | .03 [.01, .05] | Fraction vaccinated who become high-risk |
| Low-risk mortality rate ($m_l$) [.000025, .000075] | .00005 | Mortality rate for low-risk infected |
| High-risk mortality rate ($m_h$) [.0005, .00015] | .001 | Mortality rate for high-risk infected |
| Herd immunity effect ($\theta$) | 1.0 [.5, 1.0] | Fraction of effectively vaccinated who contribute to herd immunity effect |
| Vaccine effectiveness rate ($e$) | .75 [.65, .85] | Fraction of vaccinated who develop |
| Hours lost ($t$) | 24 [18, 30] | Average number of work hours lost to illness |
| Infection rate ($i$) | .25 [.20, .30] | Infection rate without vaccine |
| First-year epidemic probability ($p_1$) | .40 | Chance of epidemic in current year |
| Second-year epidemic probability ($p_2$) | .20 | Chance of epidemic next year |
| Vaccine dose price ($q$) | \$9/dose | Price per dose of vaccine |
| Overhead cost ($o$) | \$120,000 | Costs not dependent on number vaccinated |
| Opportunity cost of time ($w$) | \$20/hour | Average wage rate (including benefits) in the county |
| Value of life ($L$) | \$10,000,000 | Assumed value of life |
| Discount rate ($d$) | .035 | Real discount rate |
| Number high-risk vaccinations ($V_h$) | 13,680 | High-risk persons vaccinated: $v_h rN$ |
| Number low-risk vaccinations ($V_l$) | 17,860 | Low-risk persons vaccinated: $v_l(1-r)N$ |
| Fraction vaccinated (v) | .083 | Fraction of total population vaccinated: $rv_h + v_l(1-r)$ |

Figure 28: Parameters and their base case values

Note the typos in the formula for $EC_{nv}$ - specifically, $C_{e|v}$ must be replaced by $C_{e|nv}$.

```
import matplotlib.pyplot as plt
import numpy as np

# Define the ENB function
def ENB(p1, p2, d, Vh, Vl, q, o, alpha, w, t, mh, L, r, i, ml, n, theta, v, e):
```

| Variable | Value (millions of dollars) | Formula |
|---|---|---|
| $C_a$ | 0.404 | $o + (V_h + V_l)q$ |
| $C_s$ | 9.916 | $\alpha(V_h + V_l)\ (wt + m_h L)$ |
| $C_{e\mid nv}$ | 147.3 | $i[rN(wt + m_h L) + (1 - r)N(wt + m_t L]$ |
| $C_{e\mid v}$ | 87.9 | $(\mathrm{i} - \theta ve)\{(rN - eV_h)(wt + m_h L) + [(1 - r)N - eV_l](wt + m_l L)\}$ |
| EC$v$ | 55.7 | $C_a + C_s + p_1 C_{e\mid v} + (1 - p_1)\ p_2\ C_{e\mid v}/(1 + d)$ |
| $\mathrm{EC}_{nv}$ | 76.0 | $p_1 C_{e\mid v} + (1 - p_1)p_2 C_{e\mid v}/(1 + d)$ |
| $E[NB]$ | 20.3 | $\mathrm{EC}_{nv} - \mathrm{EC}_v$ |

Figure 29: Formulas for calculating various costs components

```
    ca = o + ((Vh + Vl) * q)
    cs = alpha * (Vh + Vl) * ((w * t) + (mh * L))
    cenv = i * ((r * n * ((w * t) + (mh * L))) + ((1 - r) * n * ((w * t) + (ml * L))))
    cev = (i - (theta * v * e)) * ((((r * n) - (e * Vh)) * ((w * t) + (mh * L))) + ((((1 -

    ecv = ca + cs + (p1 * cev) + ((1 - p1) * p2 * (cev / (1 + d)))
    ecnv = (p1 * cenv) + ((1 - p1) * p2 * (cenv / (1 + d)))

    return ecnv - ecv

# Define parameter values for Scenario 1 and Scenario 2
S1 = {
    'n': 380000,
    'r': 0.06,
    'vl': 0.05,
    'vh': 0.6,
    'alpha': 0.03,
    'ml': 0.00005,
    'mh': 0.001,
    'theta': 1,
    'e': 0.75,
    't': 24,
    'i': 0.25,
    'p2': 0.2,
    'q': 9,
    'o': 120000,
    'w': 20,
    'L': 10000000,
    'd': 0.035,
```

```
    'Vh': 13680,
    'Vl': 17860,
    'v': 0.083
}

S2 = {
    'n': 380000,
    'r': 0.06,
    'vl': 0.05,
    'vh': 0.6,
    'alpha': 0.03,
    'ml': 0.00005,
    'mh': 0.001,
    'theta': 1,
    'e': 0.75,
    't': 24,
    'i': 0.25,
    'p2': 0.2,
    'q': 9,
    'o': 120000,
    'w': 20,
    'L': 5000000,
    'd': 0.035,
    'Vh': 13680,
    'Vl': 17860,
    'v': 0.083
}

S3 = {
    'n': 380000,
    'r': 0.06,
    'vl': 0.05,
    'vh': 0.6,
    'alpha': 0.03,
    'ml': 0.00005,
    'mh': 0.001,
    'theta': 0.5,
    'e': 0.75,
    't': 24,
    'i': 0.25,
    'p2': 0.2,
```

```
    'q': 9,
    'o': 120000,
    'w': 20,
    'L': 10000000,
    'd': 0.035,
    'Vh': 13680,
    'Vl': 17860,
    'v': 0.083
}

# Define the range of p1 values
p1_values = np.linspace(0, 0.5, 100)

# Calculate ENB values for Scenario 1 and Scenario 2
enb_values_S1 = [ENB(p1, S1['p2'], S1['d'], S1['Vh'], S1['Vl'], S1['q'], S1['o'], S1['alph
enb_values_S2 = [ENB(p1, S2['p2'], S2['d'], S2['Vh'], S2['Vl'], S2['q'], S2['o'], S2['alph
enb_values_S3 = [ENB(p1, S3['p2'], S3['d'], S3['Vh'], S3['Vl'], S3['q'], S3['o'], S3['alph


# Create the plot for two different values of life
plt.figure(figsize=(10, 6))
plt.plot(p1_values, enb_values_S1, label='L = 10 million')
plt.plot(p1_values, enb_values_S2, label='L = 5 million')
plt.xlabel('Probability of Disease in First Year (p1)')
plt.ylabel('Net Benefits')
plt.title('Net Benefits vs. Probability of Disease (p1)')
plt.legend()
plt.grid(True)
plt.ylim(-10000000, 30000000)
plt.xlim(0, 0.5)
plt.show()

# Create the plot for two levels of immunity
plt.figure(figsize=(10, 6))
plt.plot(p1_values, enb_values_S1, label='Full Herd Immunity')
plt.plot(p1_values, enb_values_S3, label='Half Herd Immunity')
plt.xlabel('Probability of Disease in First Year (p1)')
plt.ylabel('Net Benefits')
plt.title('Net Benefits vs. Probability of Disease (p1)')
plt.legend()
plt.grid(True)
```

```
plt.ylim(-10000000, 30000000)
plt.xlim(0, 0.5)
plt.show()
```

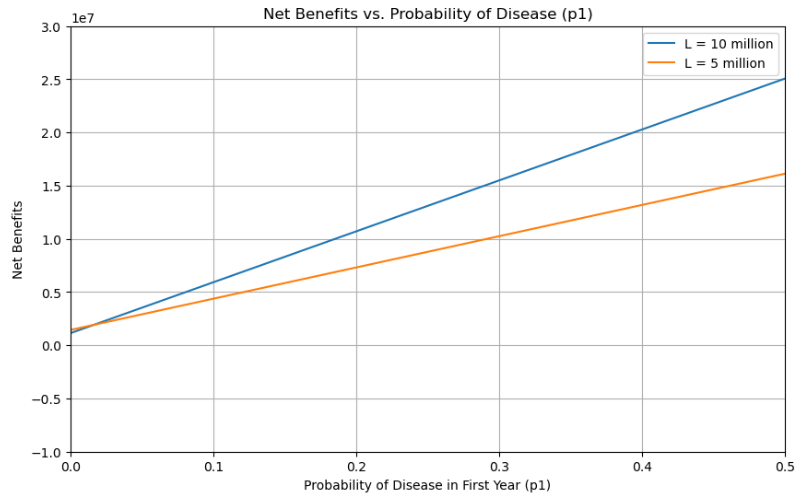



Figure 30: Expected net benefits against p1 for two values of life

Note that these figures were generated using the same parameter specification suggested in the example included in Suppl 4. However, the plots turn out to be qualitatively similar but not identical to the ones reported in the book (Figures 11.3 and 11.4). Check the code and formulas carefully.

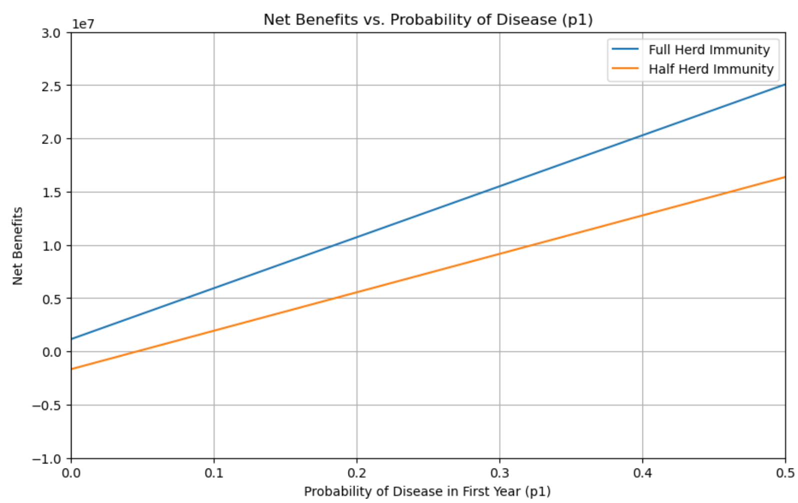



Figure 31: Expected net benefits against p1 for two levels of herd immunity

#### 5.1.3.2 Best-case and worst-case scenario

We can compare the base case with the best and the worst cases possible.

*Base Case*: Assign the most plausible numerical values to unknown parameters to produce an estimate of net benefits that is thought to be most representative.

*Worst Case*: Assign the least favorable of the plausible range of values to the parameters.

*Best Case*: Assign the most favorable of the plausible range of values to the parameters.

Worst case analysis is useful as a check against optimistic forecasts and for decision-makers who are risk averse.

### 5.1.4 Simulation analysis

Partial and best/worst case sensitivity analyses have two limitations.

First, they may not take into account all the available information about the assumed values of parameters (i.e., the worst and best cases are highly unlikely).

Second, these techniques do not directly provide information about the variance of the statistical distribution of realized net benefits (i.e., one would feel more confident about an expected value with a smaller variance because it has a higher probability of producing net benefits near the expected value).

The essence of a simulation analysis is to elicit a distribution of outcomes. The basic steps are as follows.

1. Specify the probability distributions for all the important uncertain quantitative assumptions. If no theoretical or empirical evidence suggests a particular distribution, you can use a uniform distribution if all values are equally likely, or a normal distribution if a value near the expected value is more plausible.
2. Execute a trial by taking a random draw from the distribution for each parameter to arrive at a specific value for computing realized net benefits.
3. Repeat the trial many times. The average of the trials provides an estimate of the expected value of net benefits. An approximation of the probability distribution of net benefits can be obtained by creating a histogram. (As the number of trials approaches infinity, the frequency will converge to the true underlying probability.)

Figure 6 presents a histogram comprising 10,000 replications of random draws based on the bracketed assumptions listed in Figure 2. The assumed distributions are all uniform, with the exception of hours lost $t$, which follows a normal distribution.

Note that the simulation analysis may involve calculation of the expected values of complex expressions (for example, ratios of variables) which will make the analysis quite complicated.

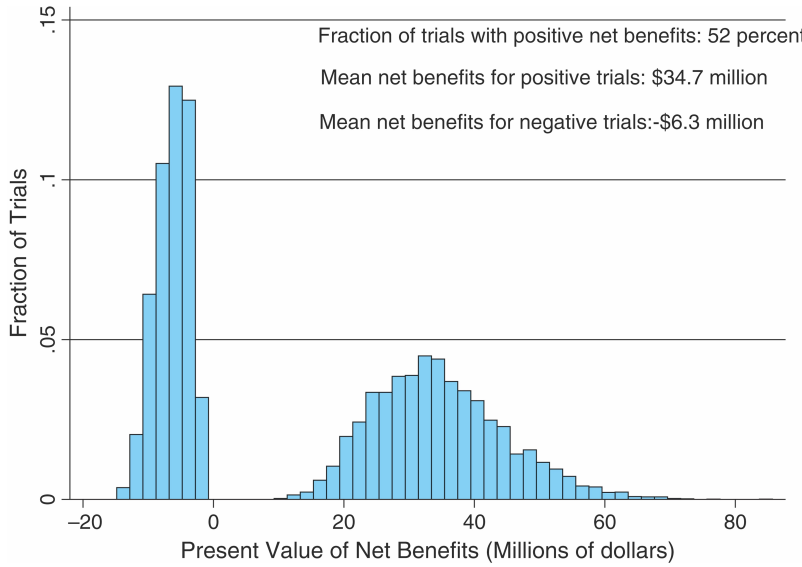



Figure 32: Distribution of the present value of net benefits

If the calculation of net benefits involves sums of random variables, then it is easy - expectation of the sum is the sum of the expected values of each variables.

If the calculation of net benefits involves products of random variables, using the product of expected values as the expected value of the product works only if the random variables are uncorrelated.

In the simulation analysis, correlations can be taken into account by drawing parameter values from either multivariate or conditional distributions rather than from independent univariate distributions.

If the calculation involves ratios of random variables, then even independence does not guarantee that the ratio of the expected values would be the same as the expected value of the ratio.

### 5.1.5 Information and quasi-option

#### 5.1.5.1 Value of information

We often encounter decision problems involving the allocation of resources, such as time (for example, delaying a project), money and effort (for example, use of a larger sample size), aimed at reducing uncertainty in the values of parameters used to calculate net benefits.

In such cases, the investment of resources are meant to release valuable information that could reduce uncertainty, and the investment can be justified if newly available information turns out to be sufficiently valuable.

How can we measure the value of information? One way to do it is to find how much an information would increase the expected net present value of the project.

#### 5.1.5.2 Quasi-option value

Quasi-option value is the expected value of information gained by delaying an irreversible decision.

Consider for example, a project with investment expenditure of NOK 18 million and that there is basically an equal probability that the operating profit in period 1 and all future periods will be either NOK 500000 or NOK 1500000. Also assume that by postponing the investment to the next period, the decision maker will know with certainty whether the operating profit will be NOK 500000 or NOK 1500000 in each period. The calculation interest is set equal to 4 percent.

In this case, the decision maker is faced with two possible options.

Alternative 1: The project is implemented immediately with probability 1, but with an uncertain operating profit. The expected net present value of alternative 1 is 7000000.

$$E_0(NV_1) = -18\,000\,000 + \frac{0{,}5\, \times\, 500\,000 +\, 0{,}5\, \times\, 1\,500\,000}{0{,}04} = 7\,000\,000$$

Alternative 2: The project is postponed by one year and only happens if the annual operating profit is 1,500,000. The probability that the project will be started is therefore 0.5. The expected net present value of alternative 2 is 9375000.

$$E_0(NV_2)\ =\ 0{,}5\,(\frac{-18\,000\,000\ +\dfrac{1\,500\,000}{0{,}04}}{1{,}04})\ =\ 9\,375\,000$$

If an investment leads to irreversible effects, a positive net present value is no longer a sufficient criterion to determine whether the measure should be implemented immediately. Starting immediately will only be correct if the present value is sufficiently large to also cover the lost *quasi-option value* by postponing the measure.

For example, even if the expected present value of alternative 1 seems to appear as 7000000, it does not take into account the opportunity cost in the form of lost option value.

$$E_0(NV_{1\,\mathrm{Korrekt}})\ =\ E_0(NV1)\ -\ E_0(NV2)\ =\ 7\,000\,000\ -\ 9\,375\,000\ =\ -\,2\,375\,000$$

It is therefore not profitable to implement the project immediately.

When there are more than one feasible actions in every period, we can use simple decision trees to analyze these problems. A decision is also helpful in analyzing situation where the possibility of learning might depend on actions taken in previous periods. The above example, in contrast, illustrates a case of exogenous learning, in which the possibility of learning does not depend on actions taken previously.

*Exogenous learning*: Learning is revealed regardless of the chosen option. After the first period, we discover with certainty which of the two contingencies will occur. Quasi-option value is the difference in expected net benefits between the learning and no-learning case.

We can illustrate this case with the following example. The society is considering a development program which can adversely impact the surrounding wild habitat. There are two levels of development: Full (FD) and Limited (LD); and it is also possible not to pursue (ND) the development program.

It is uncertain whether the development program will ultimately be benefiting or costly. If the program turns out to benefit, then the benefit under FD is higher than that under LD. If the program turns out to be costly, then the cost under FD is higher than that under LD. Specifically, $B_F > B_L > 0$ and $C_F > C_L > 0$. Assume ND results in zero returns.

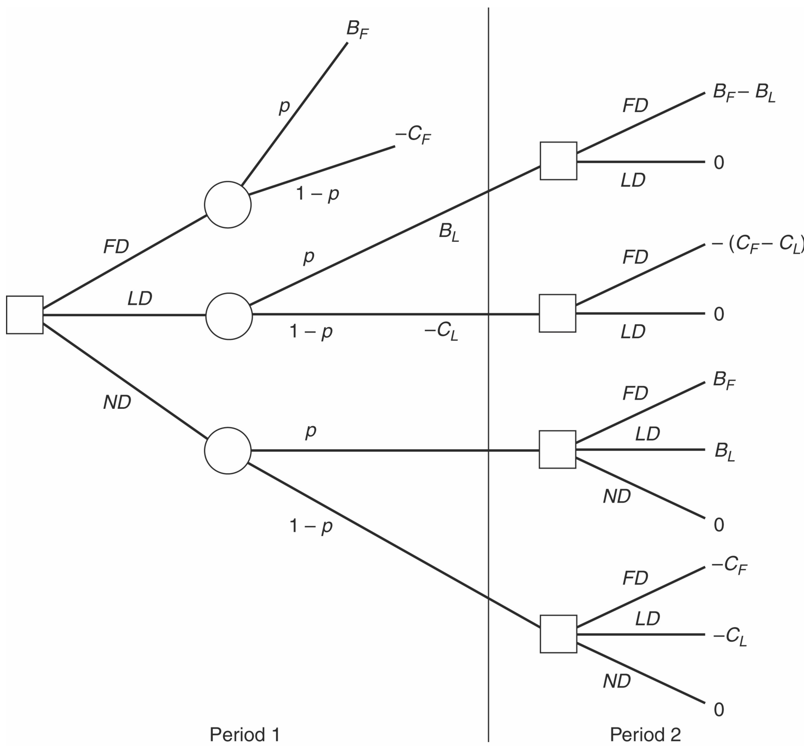



Figure 33: Exogeneous learning

Assume FD is an irreversible option in the sense that if it is taken, then the effects can never be adjusted at a later stage. If LD (or ND) is followed, then at a later stage, say at period 2, it can be adjusted to FD (or FD or LD). Between the two periods, there will be learning about the project's effect and the uncertainty will be perfectly mitigated. We assume exogenous learning in the sense it does not depend on what specific action is taken in period 1.

Doing backward induction, we can determine the quasi-option values, as reported in Figure 8.

| | No learning | Exogenous learning | Endogenous learning |
|---|---|---|---|
| $E[FD]$ | $pB_F - (1-p)C_F$ | $pB_F - (1-p)C_F$ | $pB_F - (1-p)C_F$ |
| $QOV$ | | 0 | 0 |
| $E[LD]$ | $pB_L - (1-p)C_L$ | $p[B_L + (B_F + B_L)/(1+d)] - (1-p)C_L$ | $p[B_L + (B_F - B_L)/(1+d)] - (1-p)C_L$ |
| $QOV$ | | $p(B_F - B_L)/(1+d)$ | $p(B_F - B_L)/(1+d)$ |
| $E[ND]$ | 0 | $pB_F\,(1+d)$ | 0 |
| $QOV$ | | $pB_F/(1+d)$ | $= 0$ |

Figure 34: Determination of quasi-option values

*Endogenous learning*: Information is generated only through the activity itself (whatever the program is). This leads to Exogenous learning yielding large no-activity results (i.e., a decision to hold off) and endogenous learning yielding large limited-activity results (i.e., a decision to pursue a limited program).

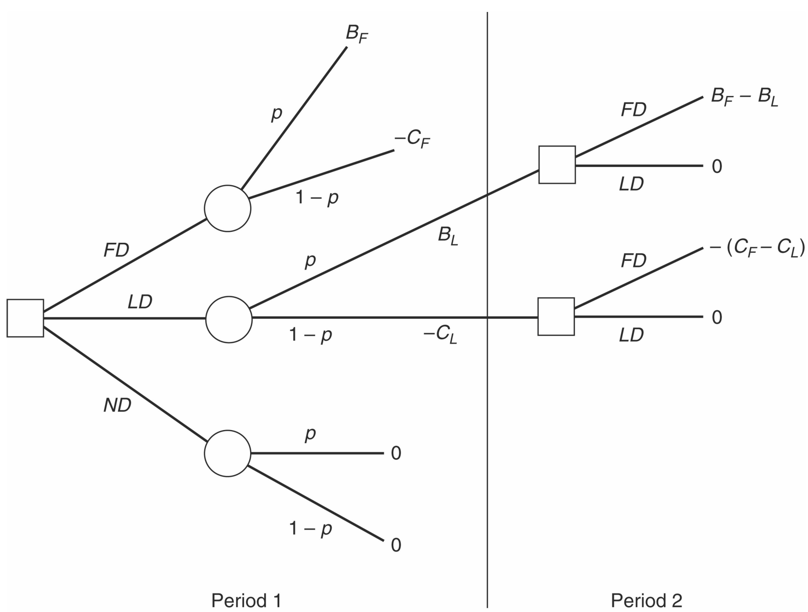



Figure 35: Endogenous learning

The expected values and the quasi-option values under endogenous learning are reported in Figure 8.

Figure 10 reports a numerical illustration of the expected values for a specific set of parameter values:

$(B_F = 100, B_L = 50, p = 0.5, C_F = 80, C_L = 40, d = 0.8)$

| Expected value | No learning | Exogenous learning | Endogenous learning |
|---|---|---|---|
| $E[FD]$ | 10.00 | 10.00 | 10.00 |
| $E[LD]$ | 5.00 | 28.15 | 28.15 |
| $E[ND]$ | 0.00 | 46.30 | 0.00 |

Figure 36: A numerical illustration of the quasi-option values

In this case, the optimal choice of development plan is sensitive whether learning is exogenous or endogenous or if there is no scope of learning. If there is no learning, then FD has the largest expected net benefits. If there is a possibility of exogenous learning, then ND has the largest expected net benefits. If there is a possibility of endogenous learning, LD has the largest expected net benefits.

### 5.1.6 Reading materials for Section 5

1. Boardman et al. Chapters 11.
2. "Veileder i samfunnsøkonomiske analyser" section 3.4 and 3.6.