AFRICAN ECONOMIC RESEARCH CONSORTIUM (AERC)

COLLABORATIVE MASTERS DEGREE PROGRAMME (CMAP) IN ECONOMICS FOR SUB-SAHARAN AFRICA

JOINT FACILITY FOR ELECTIVES

Teaching Module Materials

ECON 536 - Environmental Economics I

(Revised: August, 2020)

Facebook Twitter Website Email Website

Our mailing address is:
African Economic Research Consortium (AERC)
3rd Floor, Middle East Bank Towers, Jakaya Kikwete Road
P. O. Box 62882
00200 Nairobi
Kenya

Module 1.1: Nature and Evolution of Environmental Economics (4 Hours)

Learning Outcome

This Module introduces the student to the world of natural and environmental resource economics. It discusses the various classifications of natural and environmental resources and the evolution Environmental and Natural Resource Economics as a sub-discipline. It also exposes readers to the nature of interactions between human activities and the environment and the issues that flow from such interactions. After going through this Module, you should be able to

- ✓ relate with the subject matter of environmental and natural resource economics and its evolution as a sub discipline.
- ✓ know what environmental and natural resources are, their classification and the role they play in society.
- ✓ appreciate the interactions between human/economic activities and environmental and natural resources.
- ✓ know the limits imposed by environmental and natural resources on human activities and the effect of the latter on the former.

Outline

1.1 .1 Definition and Classification of Environmental and Natural resources

1.1.1.1 Classifying Environmental and Natural Resources
1.1.1.2 Algebra of Resource Classification

1.1.2 Evolution of Environmental Economics

1.1.3 Paradigms and Basic Concepts related to the Interaction between Environmental processes and Economic management

1.3.1 Environmental functions and Ecosystem Services

1.3.2. Economic Activity and the Environment

1.1.4. Efficiency, Optimality, Sustainability. Ethics, and Discounting

Summary

Discussion/Review Questions and Exercises

Materials used for the Lecture

1.1 .1 Definition and Classification of Environmental and Natural Resources

1.1.1.1 Classifying Environmental and Natural Resources

Natural Resources are those resources that occur naturally, that is, they are not man-made. Such resources can be classified in various ways: by **source of origin**, the **state of development**, or **renewability** of the resources. A more useful classification scheme incorporating these features distinguish between "Stock" and "Flow" resources.

- **Stock resources**: are resources for which today's use has implications for tomorrow's availability (e.g. plant and animal populations and mineral deposits). As such they can be considered as assets yielding flows of services or benefits over time. For such resources, the pattern of use over time matters as much as use at a point in time.
- **Flow resources** are those resources for which today's use does not affect tomorrow's availability (e.g. solar radiation, the power of the wind, tides and flowing water).

Flow resources are basically renewable while stock resources could be renewable or non-renewable. A resource is **renewable** when it has the capacity to grow in size over time, through biological reproduction, or tend to be perpetual in nature. Renewable resources are usually **biotic resources** (resources that come from living and organic materials and the materials that can be obtained them), including flora (plants) and fauna (animals). Basic examples include cereal grains (food), fish, forests, and animals. They also include flow resources which are by nature perpetual (e.g., solar radiation, the power of the wind, tides and flowing water etc.). These kind of resources (particularly **biotic resources)** have natural thresholds**.** Thus, they are **exhaustible** and could be depleted if the rate of exploitation exceeds their regeneration capacities.

Resources that do not have the capacity to grow over time are said to be **non-renewable**. This means there is no natural reproduction, *except on geological timescales*. Consequently, more use now necessarily implies less future use. Most **abiotic resources** (resources that come from non-living and non-organic materials, such as land, fresh water, heavy metals: gold, iron, copper, silver, etc.) are non-renewable. Some abiotic resources originated from biotic resources, that is, they are formed from organic matter that has decayed. Examples are fossil fuels, such as coal and petroleum.

It is important to note that the distinction between renewable and non-renewable natural resources depends on the nature of the link between current use and future availability, that is. the timescale over which they develop. In practice, since natural resources are products of natural cycles, they are in that sense renewable. But the timescale over which some are renewed may be too long, so that from the human standpoint they are considered non-renewable. Some authors consider all stock resources to be non-renewable and take all renewable resources as flows. In this sense stock resources are considered to be natural resources that have taken millions of years to form and so from a human perspective are fixed in supply, even though the limits may not actually be known. In contrast, flow resources are naturally renewed within a sufficient time span to be relevant to mankind.

For some renewable natural resources (RNR), the continuation and volume of their flow (availability) depend crucially on human activities. This is true of biotic resources and materials from them, **biological resources** (resources which depend on biological reproduction, e.g. birds, fish etc.), and others, such as forest, soils, aquifers, etc. Resources in this category are sometimes called **critical zone resources** (CZR). For some other renewable natural resources, the flow is independent of human activities. In other words, the amount consumed by one generation does not reduce the amount that can be consumed by subsequent generations. Flow resources, which are perpetual by nature, generally fall into this category. They can be regarded as **non-critical zone resources** (NCZR).

The stock of a CZR increases if in any period, use of the resource is less than natural growth. Stock remains the same if the rate of use is the same as natural growth rate, in which case, the resource can be maintained indefinitely. The use or harvest rate that maintain the stock of a CZR is often referred to as the '**sustainable yield**'. Stock declines if use or harvest rates is in excess of sustainable yield. For example, soil can effectively be transformed by human use from a renewable resource to a short-lived stock. Soil erosion, salinization and desertification arising from overuse makes soil irrecoverable either naturally or through planned remedial programme within time scales of relevance to human activity. Sometimes where this could be done within a foreseeable future, the cost could be prohibitive! In the same way, Aquifers may be exploited to exhaustion with no hope of recovery for hundreds of years. Questions 1 requires you to examine two case studies illustrating these possibilities.

NCZRs are not totally free from misuse also! Some can be temporally affected by overuse. For example, river flows can be reduced by over-pumping. So also, the capacity of water bodies to degrade waste can be ruined by too high levels of effluent and sewage discharge, while the quality of local air resource can fall due to polluting emissions. However, flow and quality levels are naturally and speedily restored once the rate of exploitation is controlled within the regenerative or assimilative capacity.

Non-renewable resources (NRRs) are sometimes referred to as **exhaustible resources**, or **depletable resources** because of the absence of a positive constant rate of use that can be sustained indefinitely – eventually the resource stock must be exhausted. However, the term can be confusing because, as noted above, renewable resources are also exhaustible under certain conditions. Some NRRs could be **replenishable**, e.g. water. However, the natural rate of replenishment is so low that it does not offer a potential for augmenting the stock in any reasonable time frame, so they are still depletable. Some NRRs are also **recyclable**. A recyclable resource is one that is although currently being used for some particular purpose, exists in a form allowing its mass to be recovered once that purpose is no longer necessary or desirable. For example, copper wiring from an automobile can be recovered after the car has been shipped to the junkyard. Other examples include lead (mineral) recycling, paper, bottles, e-waste. The current reserves of both replenishable and recyclable non-renewable resources (NRRs) can be augmented by economic replenishment or recycling. Prices and technology play a big role in this possibility. Higher prices and technological progress act as stimulants. Depletion rate is affected by the demand for and the durability of the products built with the resource, and the ability to reuse the products.

Current use of depletable, non-recyclable resources, precludes future use. This raises questions on how they should be shared among generations. Recycling and reuse make the useful stock of depletable resources to last longer, *ceteris paribus*. However, the cumulative useful stock is finite, and current consumption patterns still have an effect on future generations (100% recycling is not possible!). The challenge for non-renewable stock resources involves allocating dwindling stocks among generations while meeting the ultimate transition to renewable resources. In contrast, the challenge for managing renewable stock resources involves the maintenance of an efficient, sustainable flow.

We will be looking at the issues involved in the management of renewable and non-renewable resources in details in Module 3. For now, it is important to note that environmental problems are often associated with the use of renewable resources (whether they are stock or flow resources). Problem arises when one or more of such resource stands in danger of exhaustion or are used in a manner that alters their basic properties. However, environmental problems can also arise in the use of non-renewable resource when the process of use affects the sustainability of renewable resources. For example, when extraction of mineral resources leads to increased air or water pollution.

Figure 1.1a and 1.1b help to summarize our discussion on the classification of environmental and natural resources. Figure 1.1a divides natural resources into three classes (perpetual, renewable and non-renewable) while Figure 1.1b classifies them into Stocks and Flows.

.

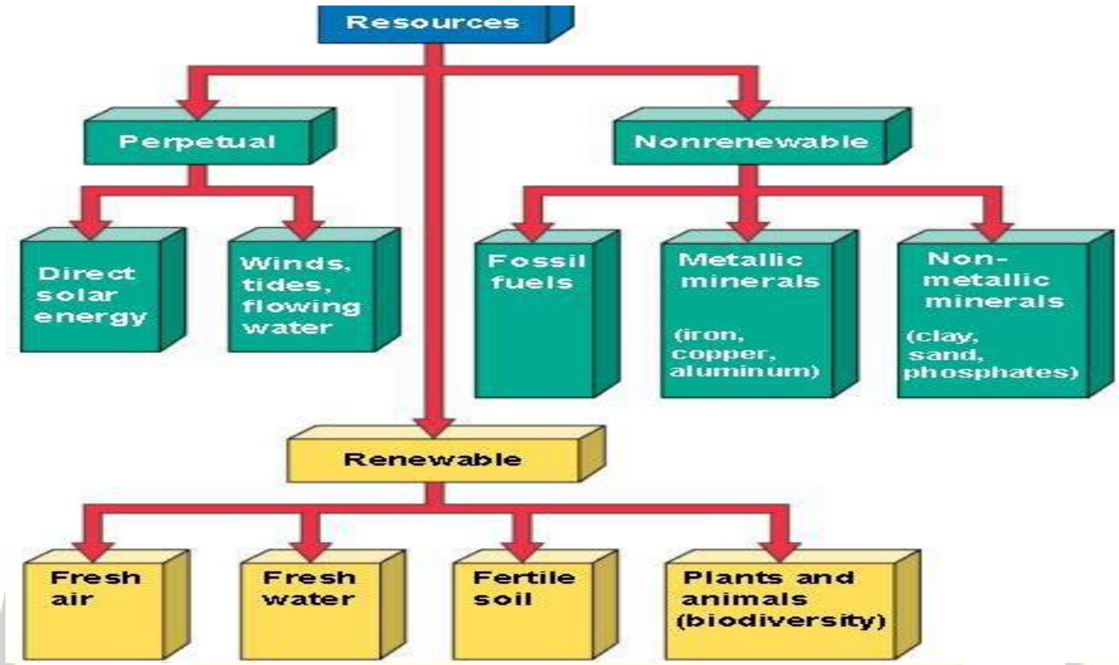


Fig. 1.1a. Taxonomy of Resources.

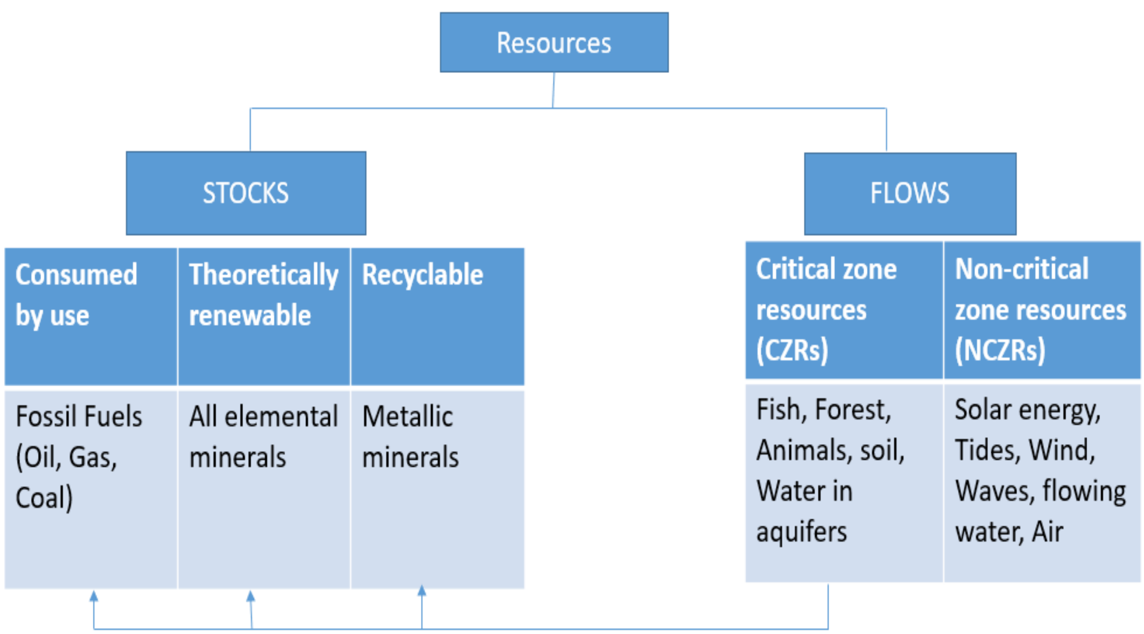


Fig. 1.1b Alternative taxonomy of resources (**Source**: based on Judith Rees, 1990, p15)

1.1.1.2 Algebra of Resource Classification

The following equation can be used to classify natural and environmental resources. Assume two periods, the present (denoted by 0 and the future denoted by 1). Then we can define the equation

$$S_1 = S_0 - Q_0 + \Delta S \qquad (1.1)$$

where S_1 = amount of resource available in period 1

S_0 = amount of resource available in period 0

Q_0 = amount of resource used in period 0

ΔS = increment to the resource in period 0

The critical factor in the equation is ΔS.

- If $\Delta S = 0$, we have a non-renewable resource.
- If $\Delta S > 0$, then the resource is renewable.
- If $Q_0 > \Delta S$, the stock of resource decumulates over time
- If $Q_0 < \Delta S$, the stock of resource accumulates over time
- If $Q_0 = \Delta S$, resource stock remains constant.
- If $S_1 = \Delta S$, resource is characterized by a non-accumulating regenerative process (e.g., free flowing river, incoming stream of solar energy).

For a recyclable resource,

$$S_1 = S_0 - Q_0 + \alpha Q_0 \qquad (1.2)$$

Where α = recycling ratio: the percentage of first period use recovered through recycling (recycling ratio)

1.1.2 Evolution of Environmental Economics

Environmental economics is the subset of economics which deals with the integration of economics and the environment. It applies the principles of economics to address environmental challenges and in a way that balances environmental, economic and social goals. The origin of environmental economics is traceable to the post-war boom in the United States. The emergence of the United States after World War II as a dominant economic power, characterized by an emergent middle class, unrivalled industrial economy, a robust international trade position, was accompanied by serious environmental consequences. By the 1950s, increasing accumulation of industrial wastes became a focus of concern among scientists and policy makers. This development prompted the establishment, in the United States, of the Resources for the Future (RFF). As a think-tank, the RFF emerged as an independent research body, which employed developed and applied economics to address a wide array of environmental issues. Its first focus, prompted by the U.S. President's Materials Policy Commission (the Paley Commission), relates to natural resource scarcity. The Paley Commission was mandated to review the future supply of minerals, energy and agricultural resources in the light of the robust demands made on these resources during World War II. Thus, RFF's early work focused on these issues, culminating in the influential and widely cited study, *Scarcity and Growth*, by Barnett & Morse in 1963. The study was set against the backdrop of the first environmental revolution initiated by Rachel Carson's *Silent Spring* in 1962, which focused on the detrimental impacts of agrochemicals on the environment. During that era, economists elaborated the cost-benefit analysis of economic activity, particularly as it relates to the environment. They determined that the costs, among other things, take the form of "external effects," (externalities) in this case alleged loss of biological diversity, an emerging issue that linked the theory of external effects with an economic interpretation of the rising tide of environmentalism.

Environmental economics did not just emerge from a vacuum. Like all sub-disciplines in economics, it has borrowed thoughts of its precursors. For example, the challenges associated with externality, a detrimental (or beneficial) effect to a third party for which no price is exacted, flowed from the work of Pigou in the 1920s. Early studies in environmental economics acknowledged that pollution damage fitted neatly into this framework, as polluters often cause damage to third parties, while they are not required to pay damages. As market economies did not account for externalities, they could not maximize human well-being, in a development that often leads to market failure. However, intervention in some firms to internalize the externality (i.e. getting the third-party effect included in the internal costs of the polluter) has emerged as a popular solution to the challenges of externality. Also, Dupuit, in the Nineteenth Century elaborated policies, which could be evaluated in terms of their costs and benefits, with costs and benefits defined in terms of human preferences and willingness to pay.

In a separate stream of intellectual development, Gray, in the early twentieth century, and later, Hotelling had established the idea that any natural resource had some optimal rate of use. Initially, these optimal use theorems were confined to natural resource economics rather than environmental economics. The distinction between the two concepts was that while the former was mainly concerned with the rates of exhaustible resource depletion and the determination of optimal harvest rates of renewable resources; the latter, on the other hand, is focused on pollution. However, the distinction largely broke down as soon as it was recognized that theorems from the former could be applied to the latter, particularly where pollutants were cumulative and also in the context of the theory of optimal growth. Indeed, mathematical models of economics with single exhaustible natural resources were stimulated by emergent world issues. For example, the 1973 oil price shock, driven by the Organization for Petroleum Exporting

Countries (OPEC), prompted concerns about the stability of fossil fuel-dependent economic systems.

The evolution of environmental economics is hardly complete without its interrelationship with welfare economics. For example, the latter provided the analytical foundations for determining the optima in economic systems. Within welfare economics, externalities were transformed from being fairly minor deviations from the optimum to being pervasive, central and potentially large; which may cause economic systems to become significantly inefficient. This development led to the emergence of economic growth models incorporating resource endowments and purporting to show that an optimal world might require significant intervention, but may also be unsustainable without technological transformation (ELC, 2007; Halcos, 2011)

While the modern sub-discipline of natural resource economics arose mainly from neoclassical growth economics, environmental economics takes its roots from welfare economics and the study of market failure, though earlier contributions to the two sides can be identified, both effectively date from the early 1970s. Environmental and Natural Resource economics is a fusion of the two sub-disciplines. Related to this is the relatively new, interdisciplinary field of Ecological economics, which emerged in the 1980s through the works of a number of economists and natural scientists (mainly ecologists). The motivation was to gain greater understanding of environmental problems through a multidisciplinary approach in order to more effectively address them. While economics had mainly focused on efficiency and optimality, the central burden here was on sustainability. Understanding nature's housekeeping (ecology) and human housekeeping (economics) and the relationship between the two is considered crucial to solving the sustainability problem. The main points raised include the following

- The economic system is part of the larger system, that is, the planet earth. The economic and environmental systems are interdependent, and need to be studied together in the light of principles from the natural sciences.
- There is need to work towards a more holistic discipline that would integrate natural-scientific and economic paradigms.
- The sustainability problem requires nothing less than a fundamental change in social values as well as a scientific reorientation.

However, unlike mainstream economics, the impact of ecological economics on environmental and natural resource economics has been limited. Little progress has been made in the direction of interdisciplinary cooperation and integration of analytical methods. Some economists see no need to go beyond the application of neoclassical techniques to environmental problems. They stress the importance of constructing a more complete set of quasi-market incentives to induce efficient behaviour and also reject the idea that existing social values need to be questioned. They seem to have great faith in the ability of continuing technical progress to ameliorate problems of resource scarcity and promote sustainability. In contrast, ecological economists tend to be more skeptical about the extent to which technical progress can overcome the problems that follow from the interdependence of economic and environmental systems.

Nevertheless, there is a lot of common ground between economists working in these two areas. None actually believes that the economy's relationship to the natural environment can be left entirely to market forces. At the same time, hardly anybody now argues that market-like incentives have no role to play in that relationship. In terms of policy, the arguments are about how much governments need to do, and the relative effectiveness of different kinds of policy instruments. All of these issues will come up as you go through this course. Figure 1.2 reveals the schools of thought leading to the emergence of environmental economics and the interactions between them.

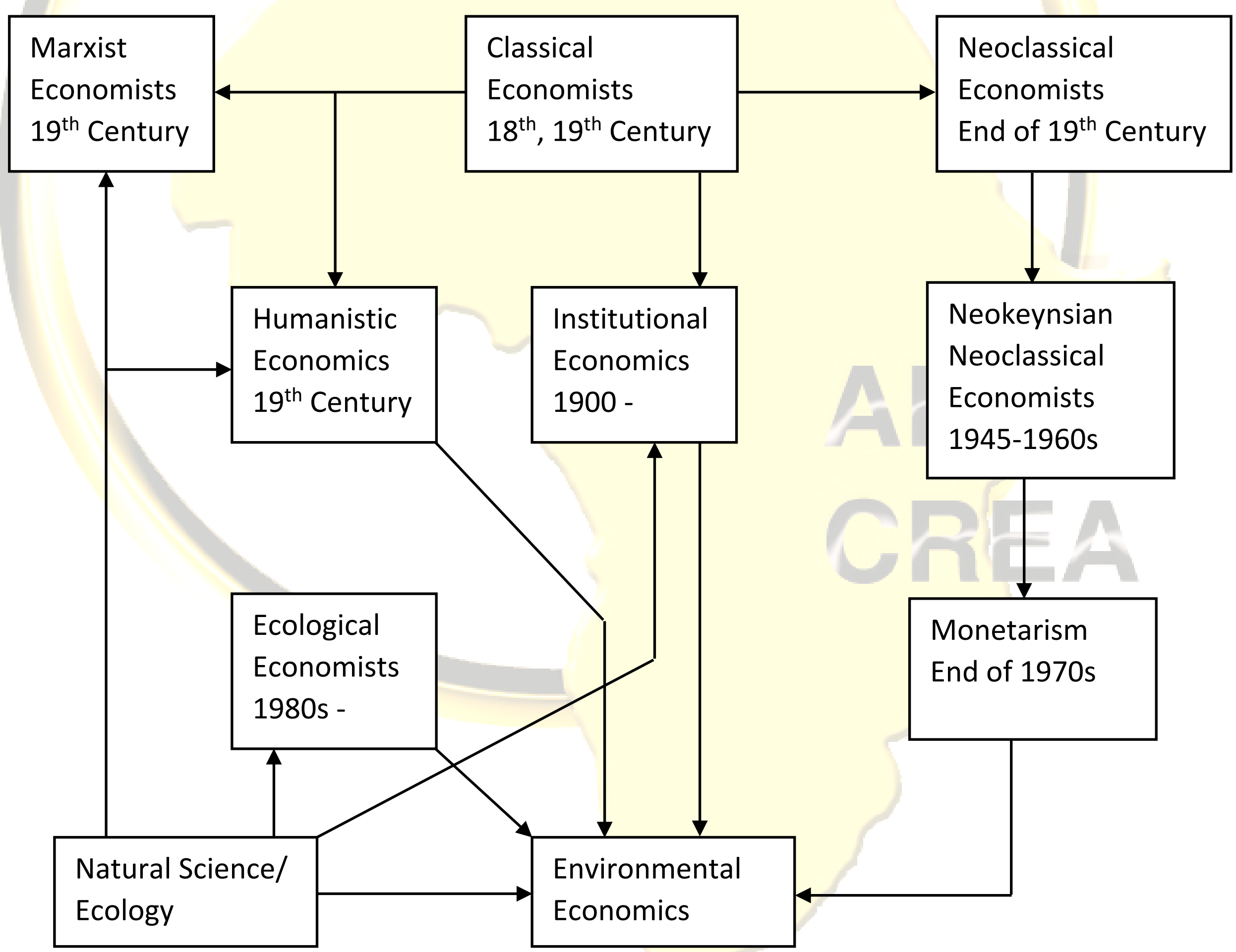


Fig. 1.2: Schools of Economic Thought and the Environment. **Source:** Halcos, 2011

1.1.3 Paradigms and Basic Concepts related to the Interaction between Environmental processes and Economic Management

1.1.3.1 Environmental functions and Ecosystem Services

The natural environment or **Environmental resource base (**ERB), is considered as the paradigm of environmental resources. It is viewed in economics as a composite (and very special) asset of nature providing a variety of services. It is also called societies' natural capital. It includes stocks of natural resources (renewable and non-renewable, including land) and the totality of ecological systems (called ecosystem for short).

Ecosystems are the dynamic and complex living system consisting of biological communities that interact with the physical and chemical environment in both time and space (see Figure 1.3). They constitute earth's most important regenerative resource.

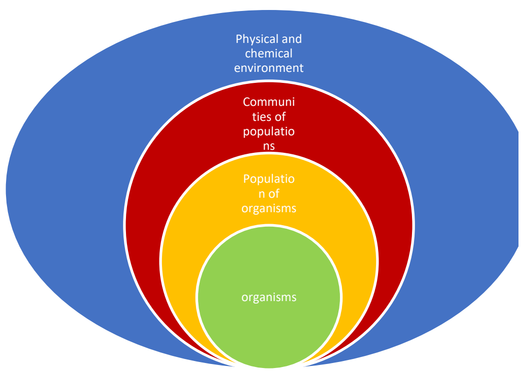


Fig. 1.3 An Ecosystem

Ecosystems perform some ecological services many of which are indispensable and vital to human existence. They

- provide various sources of water, animal and plant food, e.g., through the process of photosynthesis.
- operate the hydrological cycle.
- maintain the gaseous composition of the atmosphere.
- sustain the process that preserves and regenerate the soil, recycle nutrients, control floods, filter pollutants.
- help to assimilate waste.
- help to pollinate crops.
- help to maintain a genetic library

The interactions between the populations of organisms and the physical and chemical environment, which characterizes ecosystems, are often non-linear, implying that the state of the ERB base can display thresholds effect, leading to discontinuity in the flow of services. There are two useful concepts that flow from this reality: resilience and carrying capacity.

Resilience refers to the capacity of an ecosystem or the ERB to recover from perturbations, shocks and surprises, that is, the capacity to absorb disturbances (stress) without undergoing a fundamental change or losing its basic properties. This is often dynamic and complex as the resilience level of an ecosystem may change over time. Large reductions in resilience may make an ecosystem to break down even in the face of minimum shock. In addition, changes in the resilience of an ecosystem or the ERB are not easily observable until the effect becomes pronounced. There may be lots of uncertainties as to the current state of an ecosystem such that the breaking point may not be known practically.

Carrying capacity this refer to the maximum stress an ecosystem or the ERB can absorb without undergoing fundamental alteration. Since ecosystems evolve continuously, carrying capacity may change in an unpredictable manner.

A more useful approach towards classifying ERB functions involves grouping these functions into three broad types.

- **Resource functions:** many goods and services are derived directly from converting natural resources into other forms: gasoline from crude oil, timber from forest resources, fishes from ocean resources, etc.
- **Sink functions:** many of the activities involved in production and consumption give rise to waste products, or residuals, which are discharged into the natural environment. Exhaust gases from combustion or chemical processing, waters used to bathe or clean products, discarded packages or goods no longer wanted etc. These waste products are vented into the air, water, or buried in landfill sites. These locations in the natural environment that absorb such wastes (air, water, landfills) are called *sinks.*
- **Service functions**: the environment provides some services to mankind that are directly independent of productive (or conversion) activities on the part of man. Some of these functions are critical to life on earth while others are not, even though they are desirable. Thus, the service functions of ERB may be divided into amenity and survival functions.

  (i) *Amenity functions*: services provided by the environment that are not essential to human survival but nevertheless improve the quality of life. The biosphere has natural gardens, waterfalls, rivers, landscapes, species of animals, which provides humans with recreational facilities and other sources of pleasure and stimulation. For example, swimming in an ocean beach does not require productive activity to transform an environmental resource into a source of human satisfaction, wilderness recreation is defined by the absence of other human activity. Also, some people like simply lying out of doors in sunshine. If these services were not available, life may not be as interesting for many persons! (Consider the experience of many persons during the lockdown imposed by the COVID-19 pandemic. Why do you think many nations were eager to ease the lockdown and open beaches etc.? It was not just all about the economy!).

(ii) *Survival function*: functions provided by the ERB without which life on earth becomes impossible. Over and above serving as resource base, waste sink and amenity base, the biosphere currently provides the basic life-support functions for humans. It creates the natural habitat for all living beings. For example, there are specific requirements in terms of breathable air, temperature range and water intake, etc. that man can survive on. Box 1.1 provides an illustration of the survival function of the ERB.

As with other assets, we wish to enhance, or at least prevent undue depreciation of the ERB so that it may continue to serve its purpose. In the next subsection we will examine a general conceptual framework used in economics to approach environmental problems, the nature and

Box 1.1 Survival functions of the ERB: Solar radiation, human activities and climate change

Consider solar radiation as an environmental resource flow (an element of ERB). For some people sunbathing is an environmental amenity service. In fact, solar radiation as it arrives at the earth's atmosphere is harmful to humans. There it includes the ultraviolet wave- length UV-B, which causes skin cancer, adversely affects the immune system, and can cause eye cataracts. UV-B radiation affects other living things as well. Very small organisms are likely to be particularly affected, as UV-B can only penetrate a few layers of cells. This could be a serious problem for marine systems, where the base of the food chain consists of very small organisms living in the surface layers of the ocean, which UV-B reaches. UV-B radiation also affects photosynthesis in green plants adversely.

Solar radiation arriving at the surface of the earth has much less UV-B than it does arriving at the atmosphere. Ozone in the stratosphere absorbs UV- B, performing a life-support function by filtering solar radiation. In the absence of stratospheric ozone, it is questionable whether human life could exist. Currently, stratospheric ozone is being depleted by the release into the atmosphere of chlorofluorocarbons (CFCs), compounds which exist only by virtue of human economic activity. They have been in use since the 1940s. Their ozone-depleting properties were recognized in the 1980s. However, policy to reduce this form of pollution is now in place. Thanks to Environmental and Natural resource Economics! (**Source**: Perman et al. 2003.P. 19)

severity, and the basis for designing effective policies to deal with them.

1.1.3.2. Economic Activity and the Environment

Figure 1.4 below presents a framework illustrating the interactions between human (economic) activities and the environment. The environmental resource base (ERB) is represented in the outer heavy black lined box. The economic system is located within the environment and involves human actions manifested through production and consumption activities.

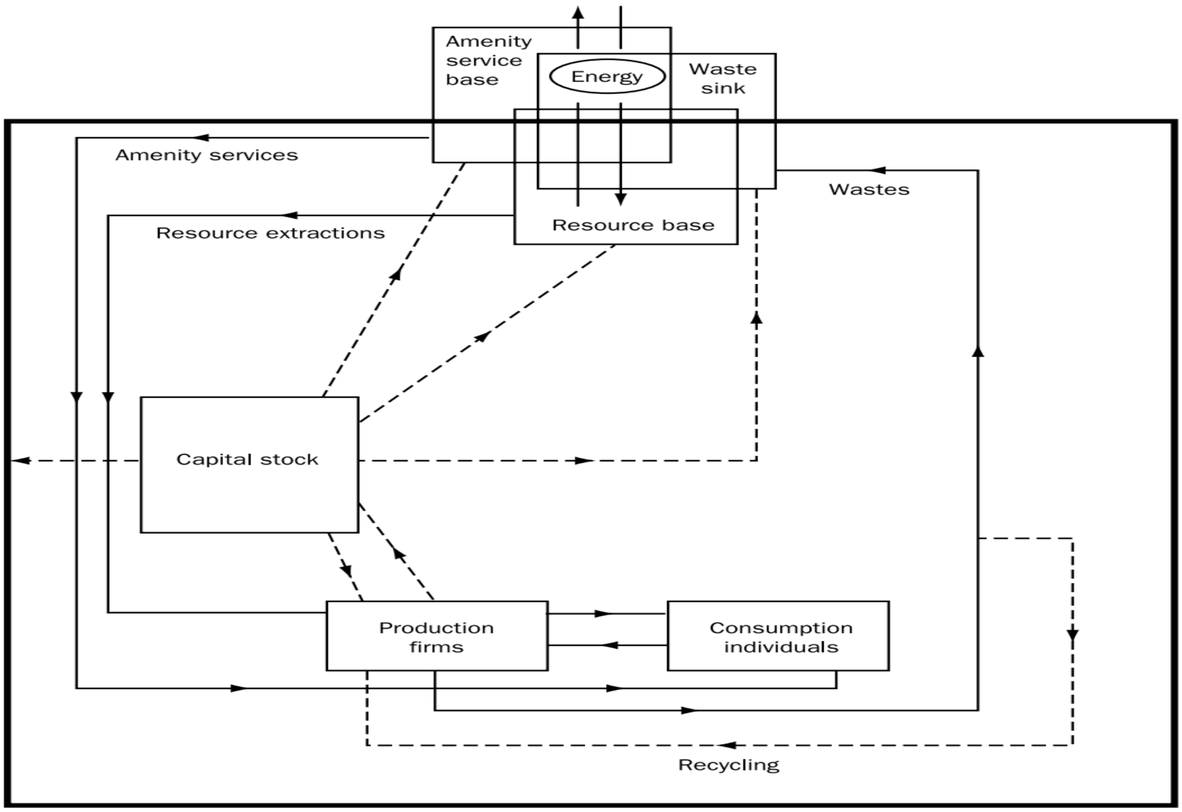


Fig. 1.4. Economic activity and the Environment. **Source:** Commons, 1995. P.17.

The earth's environment is considered a thermodynamically closed system because it exchanges energy (but not matter) with its external environment.[1] It receives inputs of solar radiation, some of which is absorbed while some is reflected back into space. This is represented by the arrows crossing the heavy black line at the top of the diagram. Balance between energy absorption and reflection (shown in the in and out arrows at the top part of Fig. 1.4) drives environmental processes and determines how the ERB performs its resource, sink, amenity and survival functions. Resource, sink and amenity functions are represented in the three boxes at the top part of the diagram, while the survival function is represented in by the heavy black lined box itself.

The interdependencies between economic activity and the environment are pervasive and complex. The complexity is increased by the existence of processes in the environment that ensures that the ERB functions interact with each other. For example, a river estuary serves as resource base for the local economy (commercial fishery) and as a waste sink in that urban

[1] Appendix 1.1 provides a background examination of these underlying ecological concepts and how they relate to Environmental sustainability.

sewage is discharged into it. It also serves as the source of amenity services, being used for recreational purposes such as swimming and boating. Finally, it contributes to life-support functions being a breeding ground for marine species which are not commercially exploited, but which play a role in the operation of the marine ecosystem. This reality is illustrated in the diagram by having the three boxes representing the resource, sink and amenity functions of the ERB intersect one with another and jointly with the heavy black line representing the life-support function.

Production and consumption activities take place in the economic system and are represented in the two boxes at the bottom part of Fig. 1.4. Production activities involve inputs of man-made capital and labour (represented in the “capital stock” box). This process is illustrated in the dashed line flowing from the capital stock box to the production box. Production also uses resource inputs from the ERB (the line flowing from the resource box at the top to the production box at the bottom). Consumption derives from what is produced (the line from the production box to the consumption box) and from amenity services provided directly by the ERB (the line flowing from the amenity box at the top to the consumption box at the bottom). Not all of produced goods are consumed. Some are added to the man-made, reproducible, capital stock (the dashed line flowing from the production box to the capital stock box). Capital is accumulated when output from current production is not used for current consumption. Human capital also increases when current production is used to add to the stock of knowledge. This is the basis for technical change. However, for technical change to impact on economic activity, it generally requires embodiment in new equipment. Knowledge that could reduce the demands made upon environmental functions does not actually do so until it is incorporated into equipment that substitutes for environmental functions!

Many of the activities involved in production and consumption give rise to waste products, or residuals. These are discharged into the sinks in the natural environment (the line flowing from the production box at the bottom to the waste box at the top). The ERB is considered as having an ‘assimilative capacity’ for the wastes from production activities. At rates of sewage discharge equal to or below the assimilative capacity of the ERB, all its four functions (resource, sink, amenity, survival) can coexist. However, a pollution problem arises if the rate of sewage discharge exceeds assimilative capacity. There is no pollution if the residual flow rate is equal to, or less than, assimilative capacity.

Pollution is not only a problem by itself, it impairs all the other estuarine functions of the ERB. Consider the river estuary earlier mentioned. Pollution will interfere with the reproductive capacity of the commercially exploited fish stocks, and may lead to the closure of the fishery, and reduce the capacity of the estuary to support recreational activity or even (as in the case of swimming, for example) totally destroy the possibility. Pollution will also impact on the non-commercial marine species, and may lead to their extinction, with implications for marine ecosystem functions. Another example can be garnered from human activities and climate change, an issue we shall be learning about in Module 7.2. The bottom line here is that if the residual flow rate from economic activities is persistently in excess of the assimilative capacity of the ERB, the latter's resilience and carrying capacity may be altered.

In practice, some waste discharges from economic activities could be recycled (this involves interception of the waste stream before it reaches the natural environment, and the return of some part of it to production). In Figure 1.4, this possibility is represented in the dashed line flowing from the production box at the bottom to the line that connects to the waste box at the top. Recycling substitutes for environmental functions in two ways. First, it reduces the demands made upon the waste sink function. Second, it reduces the demands made upon the resource base function, in so far as recycled materials are substituted for extractions from the environment.

It may also be possible to directly substitute some services performed by the ERB with services from man-made capital. This possibility is reflected in the four dashed lines running from the capital box to the three boxes and the heavy black line representing environmental functions. Substitution may take various forms. For example, in relation to the sink function performed by the river estuary, various levels of treatment of the sewage prior to its discharge into the river are possible. These will reduce the demand made upon the assimilative capacity of the estuary for a given level of sewage. Capital in the form of a sewage treatment plant can also substitutes for the natural environmental function of waste sink to an extent dependent on the level of treatment that the plant provides. Similarly, for a given level of human comfort, the energy use of a house can be reduced by the installation of insulation and control systems. These add to total capital stock. Take note, however, that the insulation and control systems are themselves material structures, the production of which involves extractions, including energy, from the environment).

Similar fuel-saving substitution possibilities exist in productive activities. In relation to amenity services performed by the ERB, an individual who likes swimming can do this in a river or lake, or from an ocean beach, or in a manufactured swimming pool. The experiences involved are not identical, but they are close. In addition, the capital equipment in the entertainment industry means that it is possible to see wild flora and fauna without leaving an urban environment. Apparently, it is envisaged that computer technology will, via virtual reality devices, make it possible to experience many of the sensations involved in being in a natural environment without actually being in it.

Note that the implications of any given substitution may extend beyond the environmental function directly affected. For example, a switch from fossil fuel use to hydroelectric power reduces fossil fuel depletion and waste generation in fossil fuel combustion. But it also impacts on the amenity service flow when it leads to flooding of a natural recreation area. Furthermore, as we shall see in Module 1.2, the extent to which man-made capital can substitute for functions of the ERB is a thorny issue in Environmental and Natural Resource Economics and at the heart of the differences in perception between core economists and ecologists in this field.

The interactions between the economic system and the environment displays both positive and negative feedback loops. **Positive feedback loops** are those in which secondary effects tend to reinforce the basic trend, whether such trend is desirable or undesirable. They make a process to be self-reinforcing. For example, new investment generates greater output, which when sold, generates profits. These profits can be used to fund additional new investments. Similarly, scientists believe that the relationship between emissions of methane and climate change may be described as a positive feedback loop. Because methane is a greenhouse gas, increases in methane emissions contribute to climate change. The rise of the planetary temperature, however, could trigger the release of extremely large quantities of additional methane currently trapped in the permafrost layer of the earth; the resulting larger methane emissions would further increase temperature, resulting in the release of more methane, etc.

Human activities or responses can also intensify environmental problems through positive feedback loops. When shortages of a commodity are imminent, for example, consumers typically begin to hoard the commodity. Hoarding intensifies the shortage. Similarly, people faced with shortages of food may be forced to eat the seed that is the key to more plentiful food in the future. Situations giving rise to this kind of downward spiral are particularly troublesome and highlight some of the issues that may create environmental problems in poor counties.

A **negative feedback loop** is self-limiting rather than self-reinforcing. Perhaps the best-known planetary-scale example is provided in the *Gaia hypothesis,* a theory advanced by the English scientist James Lovelock and named after the Greek concept for Mother Earth. It suggests that the earth is a living organism with a complex feedback system that seeks an optimal physical and chemical environment. Deviations from this optimal environment trigger natural, nonhuman response mechanisms that restore the balance.

The presence of positive and negative feedback loops raises questions on the degree to which our economic and political institutions serve to intensify or to limit emerging environmental problems. They also help shape differences in opinions among economists and ecologists as it relates to the use of environmental resources. While some group believe that current use of environmental resources could still be sustained, others are concerned that the scale of human housekeeping is now such that it threatens the viability of nature's housekeeping in ways which will adversely affect future generations of humans. The following questions (drawn from Titenberg and Lewis, 2012, p10) highlight some of the general concerns.

- Does the earth have a finite carrying capacity? If so, how can the carrying-capacity concept be operationalized? Do current or forecasted levels of economic activity exceed the earth's carrying capacity?

- How does the economic system respond to scarcities? Is the process mainly characterized by positive or negative feedback loops? Do the responses intensify or ameliorate any initial scarcity?
- What is the role of the political system in controlling these problems? In what circumstances is government intervention necessary? What forms of intervention work best? Is government intervention uniformly benign, or can it make the situation worse? What roles are appropriate for the executive, legislative, and judicial branches?
- Many environmental problems involve a considerable degree of uncertainty about the severity of the problem and the effectiveness of possible solutions. Can our economic and political institutions respond to this uncertainty in reasonable ways or does uncertainty become a paralyzing force?
- Can the economic and political systems work together to eradicate poverty and social injustice while respecting our obligations to future generations? Or do our obligations to future generations inevitably conflict with the desire to raise the living standards of those currently in absolute poverty or the desire to treat all people, especially the most vulnerable, with fairness?
- Can short- and long-term goals be harmonized? Is sustainable development feasible? If so, how can it be achieved? What does the need to preserve the environment imply about the future of economic activity in the industrialized nations? In the less industrialized nations?

Box 1.2 features the difference in emphases and methodological approach between Environmental and Ecological Economics on some of these issues

1.1.4. Efficiency, Optimality, Sustainability. Ethics, and Discounting

As we saw in the preceding section, efficiency, Optimality and Sustainability are three themes that pervade discussions on Environmental and Natural Resource Economics. Two kinds of efficiency are often considered in Economics: technical and allocative efficiency. **Technical** (also called physical) **efficiency** occurs when resources are used in a technically efficient way, requires goods to be produced in the cheapest way. This is often assumed in economics (no rational agent produces a good in a more costly manner when there are cheaper ways of producing it!). Thus, the concern is more on allocative efficiency. **Allocative efficiency** is efficiency from the viewpoint of society and not the individual. Technical efficiency may allow an economic agent to use a resource that creates some losses to society because it is the cheapest source of production but allocative efficiency considers what is cheapest taking all costs into consideration. Thus, a resource allocation choice may result in allocative inefficiency despite being technically efficient! For example, electricity can be generated in a technically efficient way through the burning of coal or gas. Coal is cheaper than gas and hence will be chosen by profit-maximizing firms. However, the pollution associated with use of coal is far much higher.

Box 1.2 Some Differences in Emphases and Methodological Approach between Environmental and Ecological Economics

Standard environmental economics and ecological economics take different perspectives, although these are not always precisely defined and there is room for overlap in a broad spectrum of analytical approaches. But there are clear differences of emphasis in terms of important concepts such as value, the role of the market, balancing present and future impacts, economic growth, and long-term sustainability.

Question	Viewpoint of Environmental Economics	Viewpoint of Ecological Economics
How is the value of the environment determined?	Using economic value, based on people's willingness to pay	Economic value may be useful, but also recognize inherent values
How are values measured?	Convert all values to monetary terms if possible	Some values, particularly inherent value, cannot be expressed in monetary terms
Advocate market-based solutions to market failures?	Yes, in the majority of cases	Perhaps, but micro-level market solutions may fail to address macro-level issues.
Consideration given to future generations?	Some, with weights inferred from market activity	More weight given to future generations based on ethical considerations
Is value neutrality desirable?	Economics aims to be value neutral (objective)	Values are acceptable in a pluralistic framework
What is sustainable development?	Maintaining the well-being of humans across time	Maintaining ecological functions across time
Are there ultimate limits to economic growth?	Perhaps not, at least in the foreseeable future	Very likely, based on the limited availability of natural resources

As it turns out, externalities (a concept examined in greater details in Module 2 is often a primary source of allocative inefficiency. Allocative inefficiencies occur "naturally" in the use of natural and environmental resources (due to their peculiar characteristics). One of the goals of Environmental economics is to determine how society can avoid this kind of inefficiencies in the allocation of natural and environmental resources.

Optimality implies allocative efficiency but goes beyond it. It considers what the objective should be for society and uses this to determine which allocation is desirable from societal viewpoint. This is done through the use of a social welfare function. A resource-use choice is socially optimal if it maximizes the objective (as reflected in the social welfare function), given any relevant constraints that may be present. In Module 2 we will be taking a closer look at efficiency and optimality withing the framework of Welfare Economics.

Optimality is vital in the management of environmental resources, but it is also not sufficient. Sustainability has to do with ensuring that future generations also benefit from existing environmental and natural resources. For example, an optimal allocation may not necessarily benefit posterity, even though it maximizes the benefits to society today. Thus, the pursuit of optimality will need to be constrained by a sustainability requirement. The idea is that since future generations cannot speak for themselves, the current generation must speak for them. The concept of sustainability and the associated criteria is a major part of the issues examined in Module 1.2.

Associated with sustainability in the use of natural and environmental resources, is the issue of **ethics**. The ERB is a gift of nature, a bequeath to all generations of mankind. Thus, there are ethical issues involved in their use and perceptions about what is right or wrong. Environmental ethics is a philosophical discipline that examines the moral and ethical relationship of human beings to the environment. It aims to determine what, if any, moral obligation does man has to the preservation and care of the non-human world. Ethical issues on the environment are hardly a new phenomenon. It has always been the subject of debates for centuries. However, environmental issues did not emerge as a philosophical discipline until the 1970s, driven by the increasing awareness of how the rapidly growing world population was impacting the environment, as well as the environmental consequences that accompany increasing use of pesticides, technology and rapid industrialization.

Environmental ethics help define human's moral and ethical obligations towards the environment. Indeed, human values are things that are important to individuals that they can then use to evaluate actions or events. These values are unique to each individual because not everyone places the same importance on each element of life. For example, an impoverished individual living in a developing country may find it morally acceptable to cut down the forest to prepare the land for farming, where he can grow food to feed his family. However, another individual in a developed country may find this action morally unacceptable because the destruction of the forest undermines carbon sequestration and increases emission of toxic compounds into the atmosphere, which negatively impacts the environment. Thus, environmental ethics, along with human values, make for challenging philosophical debates about man's interaction with the environment. In contemporary times, water and air pollution, depletion of natural resources; loss of biodiversity; destruction of ecosystem, as well as the global climate change have all emerged as part of the environmental ethics debate. Consequently, environmental ethics has become a controversial issue, accompanied by tough ethical decisions for individuals, communities, the public and private sectors, researchers, development partners, as well as policy makers.

At the heart of environmental ethics is a question about what has value and the proposition that the environment, or parts of it, has intrinsic value quite distinct from functions it provides. This leads us to the concept of **discounting**. Since natural and environmental resources offer a steam of services over time, we need to be able to determine the present value of future service flows. What discount rate should be applied to do this is one of the contentious issues in environmental management, especially given also the moral and ethical dimensions.

Summary

- Natural Resources are those resources that occur naturally, that is, they are not man-made. They can be classified into Stocks and Flows.

- Stock resources are resources for which today's use has implications for tomorrow's availability. As such they can be considered as assets yielding flows of services or benefits over time. Flow resources are those resources for which today's use does not affect tomorrow's availability.

- Flow resources are basically renewable while stock resources could be renewable or non-renewable.

- For some renewable resources, the continuation and volume of their flow (availability) depend crucially on human activities. These are called critical zone resources (CZR). For some others, called non-critical zone resources, NCZR, the flow is independent of human activities.

- The challenge for non-renewable resources involves allocating dwindling stocks among generations while meeting the ultimate transition to renewable resources. In contrast, the challenge for managing renewable resources involves the maintenance of an efficient, sustainable flow.

- Overuse of a renewable resource beyond its regenerative capacity creates environmental problems. Such problems can also arise in the use of non-renewable resource when the process of use affects the sustainability of renewable resources.

- Environmental and natural resource economics is the subset of economics which deals with the integration of economics, natural resource management, and the environment. It applies the principles of economics to address environmental challenges and in a way that balances environmental, economic, and social goals.

- While the modern sub-discipline of natural resource economics arose mainly from neoclassical growth economics, environmental economics takes its roots from welfare economics and the study of market failure.

- The natural environment or environmental resource base **(**ERB), including the stocks of natural resources (renewable and non-renewable) and the totality of ecological systems) is considered to be the paradigm of environmental resources. It is viewed in economics as a composite (and very special) asset of nature providing a variety of services.

- Services provided by the environmental resource base includes resource functions, sink functions, amenity and survival functions.

- The interdependencies between economic activity and the environment are pervasive and complex. This complexity is increased by the existence of processes in the environment that ensures that the ERB functions interact with each other.

- The interactions between the economic system and the environment displays both positive and negative feedback loops which raises questions on the degree to which our economic and political institutions serve to intensify or to limit emerging environmental problems. They also help shape differences in opinions among economists and ecologists as it relates to the use of environmental resources.

- While some group believe that current use of environmental resources could still be sustained, others are concerned that the scale of human housekeeping is now such that it threatens the viability of nature's housekeeping in ways which will adversely affect future generations of humans.

- Efficiency, Optimality and Sustainability are three themes that pervade discussions on environmental and natural resource economics.

- Allocative inefficiencies occur "naturally" in the use of natural and environmental resources due to their peculiar characteristics. One of the goals of environmental economics is to determine how society can avoid this kind of inefficiencies in the allocation of natural and environmental resources.

- Optimality implies allocative efficiency but goes beyond it to consider what society's objective should be in relation to the use of resources. Optimality is vital in the management of environmental resources, but it is also not sufficient.

- Sustainability has to do with ensuring that future generations also benefit from existing environmental and natural resources.

- Associated with sustainability in the use of natural and environmental resources, is the issue of ethics. The environmental resource base is a gift of nature, a bequeath to all generations of mankind. Thus, there are ethical issues involved in their use and in perceptions about what is right or wrong.

- Since natural and environmental resources offer a steam of services over time, we need to be able to determine the present value of future service flows. What discount rate should be applied to do this is one of the contentious issues in environmental management, especially given also the moral and ethical dimensions.

Discussion/Review Questions and Exercises

1. Consider and evaluate the following case studies in the light of what you have learnt in this Module

Case study I: **Restoring Ogoni land (Cross Rivers State, Nigeria)**. Available links: Nigeria Launches $1 Billion Ogoniland Clean-up and Restoration Programme, https://www.unenvironment.org/news-and-stories/story/nigeria-launches-1-billion-ogoniland-clean-and-restoration-programme; Environmental assessment of Ogoniland, https://postconflict.unep.ch/publications/OEA/UNEP_OEA.pdf

Case study II: **What has happened to aquifers underlying the Sahel?**

2. Classical economists placed great emphasis on natural resources (land) as determinant of wealth. How important are natural resources in the wealth of nations today?

   Some useful materials:

(a) *Where is the Wealth of Nations, Measuring Capital for the 21st Century,* World Bank, 2006.

(b) *The Changing Wealth of Nations, Measuring Sustainable Development in the New Millennium*, World Bank 2011

(c) *The Changing Wealth of Nations*, Building a Sustainable Future, World Bank 2018.

3. Has the intensive programme of the United Nations (UN) to combat desertification been able to control the soil degradation process or recover significant areas of damaged land? (Useful link : https://www.unccd.int/).
4. Can you think of an example in your country involving the transformation of a CZR into a vanishing stock?
5. What factors do you think can influence the recycling ratio? Why has recycling been slow in many parts of Sub-Saharan Africa?
6. Does the divergence in views and approaches among economists in relation to environmental problems have any effect on the progress of the discipline and ability to solve the environmental challenges facing the world today?
7. What are the merits in each of the two sides of the debate on the future of society? Can you identify historical evidences to support each view? How can we reconcile the two views? What role can economics play?
8. Waste discharge into the environment is not necessarily an environmental problem. Do you agree? Why or why not? Support your answer with a strong theoretical argument.
9. Is waste management of any long-run importance in addressing environmental challenges faced by cities?
10. Using relevant diagram and equation, illustrate the importance of recycling in natural resource conservation.
11. Is the rapid depletion (over-extraction) of a non-renewable resource, such as gold or crude oil, an environmental problem? Compare this with over-harvesting of fishes.
12. Explain and different between the three concepts of efficiency, optimality and sustainability, highlighting their roles in Environmental & Natural Resource Economics. Which branch of economics contributed the most to the development of each concept?

Materials used for this Lecture

Freeman A.M. (1993). **The Measurement of Environmental and Resource Value**s. Washington DC, RFF

Jonathan M. Harris and Brian Roach (2017), **Environmental and Natural Resource Economics** 4th Edition, Routledge.

Judith Rees)1990). **Natural Resources: Allocation, Economics and Policy**, Second, Routledge

Pearce D. and R. Turner (2004). **Economics of Natural Resources and the Environment.** Harvester Wheatsheaf. London.

Perman, R., Ma Y., McGilvray J. and Common M. (2003). **Natural Resource and Environmental Economics,** 3rd edition**.** Edinburgh, Longman.

Titenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics

.

Appendix 1.1

This Appendix introduces the reader to the scientific (ecological) background to understanding the interactions between economic activities and the environment.

1.1.1 Thermodynamics and Implications for Environmental and Natural Resource Use[2]

Thermodynamics is the science of energy. Energy is defined as the potential to do work or supply heat. It is a characteristic of things, rather than a thing itself. Work is involved when matter is changed in structure, in physical or chemical nature, or in location.

A system may be thermodynamically open, closed, or isolated. An 'open' system is one which exchanges energy and matter with its environment. An individual organism – a human being for example and indeed all living organisms – is an open system. Some energy input is necessary for it to maintain its structure and not become disordered (dead). A 'closed' system exchanges energy, but not matter, with its environment. Planet earth is a closed system in relation to its outer environment. An 'isolated' system exchanges neither energy nor matter with its environment. Apart from the entire universe, which can be considered as an isolated system, the concept is more of an ideal, an abstraction.

There are two important principles (called *laws*) in thermodynamics. The first law of thermodynamics states that *energy can neither be created nor destroyed; it can only be converted from one form to another*. This has implications for environmental and resource economics, that is, there is always 100% energy conservation whatever people do. Thus, the real issue is not conservation of energy but encouraging people to do the things that they do in ways that require less heat and/or less work, and therefore less energy conversion. The mass of materials flowing into the economic system from the environment has either to accumulate in the economic system or return to the environment as waste. When accumulation stops, the mass of materials flowing into the economic system is equal in magnitude to the mass of waste flowing into the environment. Thus, Insertions of materials into the environment as wastes are the necessary corollary of the extraction of material resources from it (the materials balance principle). To the extent, and only to the extent, that waste discharge gives rise to problems perceived by humans, to that extent can we say that there is a pollution problem.

The second law of thermodynamics states that *heat flows spontaneously from a hotter to a colder body, and it cannot be transformed into work with 100% efficiency*. All conversions of energy from one form to another are less than 100% efficient. Some energy is always lost during conversion, and the rest, once used, is no longer available for further work. As with the first, this

---

[2]Thermodynamics is not without its controversies, even among scientists. Classical thermodynamics involved the study of equilibrium systems, but the systems directly relevant to economic activity are open and closed systems which are far from equilibrium.

also has implications for environmental and natural resource use. Though energy can neither be created nor destroyed (the first law), not all of the energy of some store (a biotic or abiotic resource),is available for conversion. Energy stores vary in the proportion of their energy that is available for conversion. All energy conversions increase the entropy (the unavailable energy) of the system, hence the need to advocate for less energy conversion.

All energy conversions are irreversible. The fact that the conversion is less than 100% efficient means that the work required to restore the original state is not available in the new state. For example, fossil fuel combustion is irreversible, and of itself implies an increase in the entropy of the system which is the environment in which economic activity takes place. In the absence of new energy inputs, an isolated system must eventually use up its available energy. Since energy is necessary for life, life ceases when useful energy flows cease.

Since the planet is a closed, not an isolated, system, and is continually receiving energy inputs from its environment, in the form of solar radiation, this makes life possible. However, the entropy law reminds us that the flow of solar energy establishes an upper limit on the flow of available energy that can be sustained. Once the stocks of stored energy (such as fossil fuels[3] and nuclear energy) are gone, the amount of energy available for useful work will be determined solely by the solar flow and by the amount that can be stored (through dams, trees, etc.). Thus, in the very long run, the growth process will be limited by the availability of solar energy and our ability to put it to work.

The second law of thermodynamics has been described as the 'taproot of economic scarcity' or "the ultimate basis of economic scarcity" (Georgescu-Roegen, 1979). In his words *if energy conversion processes were 100% efficient, one lump of coal would last forever!*. Some have argued that complete recycling is physically possible if a sufficient amount of energy is available. Given available energy, there need be no scarcity of minerals. (This reasoning is behind the interest in nuclear power, and especially nuclear fusion, which might offer the prospect of a clean and effectively infinite energy resource: the "energetic dogma"). However, as argued by Nicholas Georgescu-Roegen, "matter matters as well", Even with enough energy, the complete recycling of matter is, in principle, impossible. Whatever the amount of energy that is available, such expenditure of energy required for 100% recycling would involve a tremendous increase in the entropy of the environment, which would not be sustainable for the biosphere' (Biancardi *et al*., 1993).

---

[3] Fossil fuels are accumulated past solar energy receipts, initially transformed into living tissue, and stored by geological processes. Given this origin, there is necessarily a finite amount in existence.

1.1.2. The Materials Balance Principle (MBP)

The MBP is used by economists to capture the law of conservation of mass, a derivative of the first law of thermodynamics. The law is based on two premises.

1. Economic activity essentially involves transforming matter extracted from the environment; it cannot, in a material sense, create anything, even though it transforms materials extracted from the environment into more valuable forms to humans.
2. All of the materials extracted from the environment must, eventually, be returned to it, albeit in a transformed state, even though some of the extracted material stay in the economy for a long time – in buildings, roads, machinery and so on.

The physical relationships implied by the MBP is illustrated in Fig. 1.5. It is much similar to our familiar circular flow of income and expenditures. It amplifies the picture of material extractions from and insertions into the environment provided in Figure 1.4. and abstracts from the lags in the circular flow of matter due to capital accumulation in the economy. Primary inputs (ores, liquids and gases etc.) are taken from the environment and converted into useful products (basic fuel, food and raw materials, etc.) by 'environmental' firms. These outputs become inputs into subsequent production processes (shown as a product flow to non-environmental firms) or to households directly. Households also receive final products from the non-environmental firms' sector. The materials balance principle implies an identity between the mass of materials flow from the environment (flow *A*) and the mass of residual material discharge flows to the environment (flows *B +C+D*).

Several other important identities are implied by the flow. Each of the four sectors shown by rectangular boxes receives an equal mass of inputs to the mass of its outputs leading to the following four identities:

- The environment: $A \equiv B + C + D$
- Environmental firms: $A \equiv A_1 + A_2 + C$
- Non-environmental firms: $B + R + E \equiv R + A_1 + F$
- Households: $A_2 + E \equiv D + F$

The Model yields several practical insights in relation to environmental management.

- In a materially closed economy in which no net stock accumulation takes place (that is, physical assets do not change in magnitude) the mass of residuals into the environment (*B +C+D*) must be equal to the mass of fuels, foods and raw materials extracted from the environment and oxygen taken from the atmosphere (flow *A*).

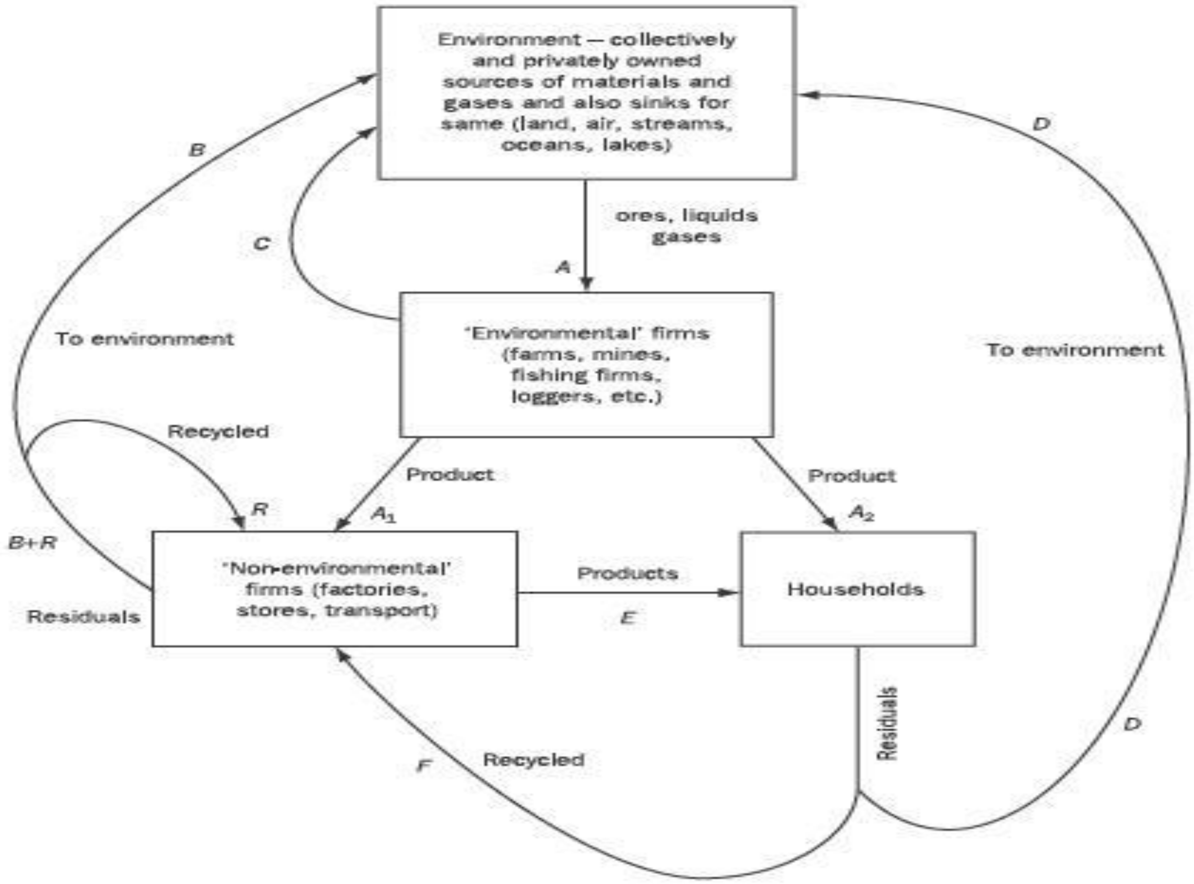


Fig. 1.5. A material balance model of environment economy interaction. **Source:** Perman et al., 2003:25)

- The treatment of residuals from economic activity does not reduce their mass although it alters their form (waste treatment does not get rid of residuals but waste management can transform residuals to a more benign form or change their location).
- Recycling is important. For any fixed magnitude of final output, *E*, if the amount of recycling of household residuals, *F*, can be increased, then the quantity of inputs into final production, A_1, can be decreased (look again at the identity $B + R + E \equiv R + A_1 + F$. and convince yourself that this is so!*)*. This in turn implies that less primary extraction of environmental resources, *A*, need take place.
- So, the total amount of material throughput in the system (the magnitude *A*) can be decreased for any given level of production and consumption if the efficiency of materials utilization is increased through recycling processes.

Module 1.2 Economic Development and the Environment (6 Hours)

Learning Outcomes

This Module examines further the nature human interactions with the environmental resource base. It introduces and discus the concept of sustainability development, its controversies and limitations. It also examines the factors that drive environmental degradation and the prospects for using policies to influence these factors as well as the concept of green growth. After going through this Module, you should

- ✓ appreciate the importance of economic growth and the notion that environmental resources set limits to growth,
- ✓ understand how growth itself sets limits on environmental sustainability,
- ✓ understand the concept of sustainability development, the various views on the concept and its operational difficulties,
- ✓ know the proximate drivers of the economy's impact on the environment and how they can be influenced by policies, and
- ✓ understand and appreciate the concept of green growth.

Outline

1.2.1. Economic Development and Sustainability

1.2.1.1 Economic Growth, the Environment and Growth Limits
1.2.1.2 Sustainable Development: Definitions, Basic Concepts and Operationalization
1.2.1.3 The Basic Framework
1.2.1.4 Weak and Strong Sustainability

1.2.2. Factors responsible for Environmental Degradation

1.2. 2.1 Economic activities, Market Failure and Government Failure

1.2.2.2 Environmental Impact of Economic Activities: the IPAT Identity

1.2.2.3. IPAT Identity and Determinants of Environmental Degradation

1.2.2.4 Economic activities and Environmental Degradation: Some Observation on Recent Trends

1.2.3. Green Growth

Summary

Discussion/Review Questions and Exercises

Materials used for the Lecture

1.2.1. Economic Development and Sustainability

1.2.1.1 Economic Growth, the Environment and Growth Limits

Given economists' idea about the effectiveness of growth, for many years, it was thought that the eradication of poverty required well-designed development programmes that were largely independent of considerations relating to the natural environment. The goal of economic and political debate was to identify growth processes that could allow continually rising living standards. Economic development and 'nature conservation' were seen as quite distinct and separate problems. Concerns for the natural environment was interpreted by some as a rather selfish form of self- indulgence on the part of the better-off.

The 1970s witnessed a shift in perspectives. While the pursuit of economic growth and development continues, it was recognized that the maintenance of growth has an important environmental dimension. Concern for sustainability began to appear on the international political agenda, most visibly in the proceedings of a series of international conferences. The common theme of these debates was the interrelationship between poverty, economic development and the state of the natural environment. The first major challenge to the claim to growth as solution to the world's problems questioned the sustainability of growth itself. In its book, *The Limits to Growth* (Meadows et al., 1972), the Club of Rome posits that environmental limits would cause the collapse of the world economic system in the middle of the twenty-first century (see Box 1.3 and Figure 1.6).[4] The arguments, though roundly condemned by most economists, stimulated the re-emergence of interest in the management of natural resources among economists.

[4] Some authors have also identified "social limits to growth", that is, the desirability of growth itself, rather than its feasibility. They believe that aggregate growth has "self-cancelling effects on welfare". They also emphasize the corrosive effects on moral standards of the very attitudes that foster growth, such as glorification of self-interest and a scientific–technocratic worldview (Daly, 1987). According to Hirsch (1977), the process of economic growth becomes increasingly unable to yield the satisfaction which individuals expect from it, once the general level of material affluence has satisfied the main biological needs for life-sustaining food, shelter and clothing. As the average level of consumption rises, the satisfaction that individuals derive from goods and services depends in increasing measure not only on their own consumption but on consumption by others as well. (Hirsch, 1977, p. 2). For example, the satisfaction a person gets from the use of a car depends on how many other people do the same. The greater the number of others who use cars, the greater is the amount of air pollution and the extent of congestion, and so the lower is the satisfaction one individual's car use will yield. Also, the utility from expenditure on education in an attempt to raise one's chances of securing sought-after jobs declines as an increasing number of others also attain that level of education. The bottom line is that beyond a certain level, growth does not deliver the increased personal satisfactions that it is supposed to.

Box 1.3 The Limits to Growth

The Limits to Growth reported the results of a study in which a computer model of the world system, World3, was used to simulate its future. World3 represented the world economy as a single economy, and included interconnections between it and its environment. Five major trends of global concern were investigated: accelerating industrialization, rapid population growth, widespread malnutrition, depletion of non-renewable resources, and a deteriorating environment. The goal was to understand the causes of these trends, their interrelationships and their implications as much as one hundred years in the future. Environmental limits are defined to include (a) a limit to the amount of land available for agriculture; (b) a limit to the amount of agricultural output producible per unit of land in use; (c) a limit to the amounts of non-renewable resources available for extraction; (d) a limit to the ability of the environment to assimilate wastes arising in production and consumption, which limit falls as the level of pollution increases.

Based on these limits and given current behaviour of the economic system, on the basis of a number of simulations, the modeling team came to the following conclusions

1. If the present growth trends in world population, industrialization, pollution, food production and resource depletion continue unchanged, the limits to growth on this planet will be reached sometime within the next 100 years. The most probable result will be a sudden and uncontrollable decline in both population and industrial capacity (see Figure 1.6).
2. 2. It is possible to alter these trends and to establish a condition of ecological and economic stability that is sustainable far into the future. The state of global equilibrium could be designed so that the basic material needs of each person on earth are satisfied and each person has an equal opportunity to realize his or her individual human potential.
3. If the world's people decide to strive for this second outcome rather than the first, the sooner they begin working to attain it, the greater will be their chances of success.

Source: Meadows et al., 1992, see Perman, et al. 2000. pp44-46)

In the basic Limits to Growth (LTG) model, resource depletion and an increase in pollution will force a reduction in food, industrial output, and population by the middle of the twenty-first century. The LTG authors also model a "sustainable world" scenario where average family size is two children, modest limits are put on material production, and society invests heavily in sustainable technologies, resulting in human population stabilizing at less than 8 billion, with industrial output also stabilizing and pollution eventually decreasing.

Some of the criticisms raised by economists against the propositions from *The Limits to Growth* states that the feedback loops in the model on which the Book was based were poorly specified in that they failed to take account of behavioural adjustments operating through the price mechanism. Changing patterns of relative scarcity would alter the structure of prices, inducing behavioural changes in resource-use patterns. Given a well-functioning market mechanism, limits to growth would not operate in the way envisaged. Furthermore, in the absence of properly functioning markets or any market at all, proper policy responses can be used to address emerging problems.

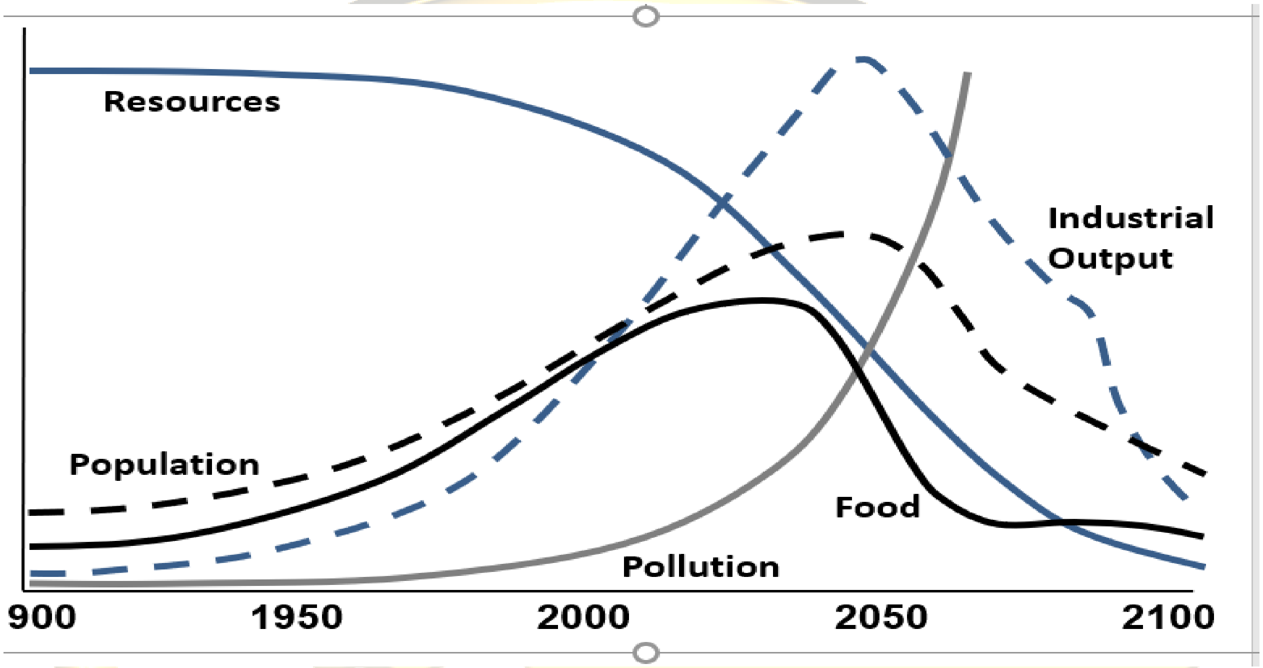


Figure 1.6 The Limits to Growth Model, Business-as-Usual Scenario. *Source*: Meadows, et al., 2004. Jonathan and Roach, 2017)

A sequel to *The Limits to Growth*, written by the same team and entitled *Beyond the Limits* and published in 1992 provided an update of the numerical values used in estimation but came up with much the same conclusions.

1.2.1.2 Sustainable Development: Definitions, Basic Concepts and Operationalization

The paradigm shift in relation to economic growth and environmental sustainability came with the publication of the Brundtland report (*Our Common Future*, 1987) by the World Commission on Environment and Development (WCED)[5] The Report provided much information about "the

[5] WCED was established four years earlier by the United Nations with mandate to

- re-examine the critical environment and development issues and to formulate realistic proposals for dealing with them;
- propose new forms of international cooperation on these issues that will influence policies and events in the direction of needed changes and
- raise the levels of understanding and commitment to action of individuals, voluntary organizations, businesses, institutes and governments.

sustainability problem". It sets out the nature of economy–environment interdependence, identified a number of potential environmental constraints on future economic growth, and argued that current trends cannot be continued far into the future. Environment and development are not separate challenges; they are inexorably linked. Development cannot subsist on a deteriorating environmental base; the environment cannot be protected when growth does not account for the costs of environmental protection. Accordingly, attempts to maintain social and ecological stability through old approaches to development and environmental protection will increase instability.

The Brundtland Report defined sustainable development as a *development that meets the needs of the present, without compromising the ability of further generations to meet their own needs*. It recognizes that the problems of poverty and underdevelopment cannot be solved unless we have a new era of growth in which developing countries play a large role and reap large benefits. Growth must be revived in developing countries because that is where the links between economic growth, the alleviation of poverty, and environmental conditions operate most directly. Yet developing countries are part of an interdependent world economy; their prospects also depend on the levels and patterns of growth in industrialized nations. The medium-term prospects for industrial countries are for growth of 3–4 per cent, the minimum that international financial institutions consider necessary if these countries are going to play a part in expanding the world economy. These growth rates could be environmentally sustainable if industrialized nations can continue the recent shifts in the content of their growth towards less material and energy-intensive activities and the improvement of their efficiency in using materials and energy. (the green growth revolution to be discussed later). Much of resource and environmental economics is about the right kind of policy instruments for achieving this goal. In recent times the concept of sustainable development is being used to set the agenda for development as reflected in the United Nations Sustainable Development Goals (SDCs) and more recent Millennium Development Goals (MDGs).

The concern for sustainability derives from an ethical concern for future generations together with an appreciation of the fact that such concern needs to be incorporated into current decision making. Unfortunately, the definition of sustainable development provided by the Bruntdland report was left vague with no specificity about what this implies in terms of policy and actions. Concern for future generations can take a variety of expressions, and does not translate into a single simple constraint on current planning. Consequently, the literature is replete with various definitions, meanings and interpretations of what sustainable development is (at least 50 as at 1989!). Table 1.1 provides six popular definitions.

---

while focusing on population growth, food security, biodiversity loss, energy, resource depletion and pollution, and urbanization

Table 1.1 Definitions and perspectives on Sustainable Development

	Sustainable development		Perspective
1	one in which utility (or consumption) is non-declining through time	Utility-based approach	Economics, view held by most environmental economists
2	one in which resources are managed so as to maintain production opportunities for the future	Opportunity-based approach	
3	one in which the natural capital stock is non-declining through time		Economics, view held by most ecological economists
4	one in which resources are managed so as to maintain a sustainable yield of resource services		Ecology, view held by most ecologists
5	one which satisfies minimum conditions for ecosystem resilience through time		
6	consensus-building and institutional development.		Governance, view held by Institutionists

Source: Adapted from Perman et al., 2003. p.86

Note that the six concepts/definitions in Table 1.1 are not mutually exclusive. For example, the first largely entails the second (if utility or consumption is not to decline, resources must be managed so that productive opportunities are maintained for sub- sequent generations). The fourth is also a particular case of the second, while the first seems to require the fifth if we take the view that production and consumption cannot be maintained over time in the face of ecosystem collapse.

In the light of the absence of a single and all-embracing definition of what 'sustainable development' actually means, some authors have suggested the idea of a "sustained' development or 'survivable' development (Pezzey, 1997). The distinctions and intuition behind the two concepts and that of sustainable development are illustrated as follows. Let

- U_t = utility at time t
- $\dot{U_t} = \frac{dU_t}{dt}$ = the rate of change of utility at time t

- U_{tMAX} = the maximum utility which can be held constant for ever from time t onwards, given production opportunities available at time t.
- U^{SURV}= the minimum utility level consistent with survival of the given population

Development is **sustainable** if $U_t \leq U_{tMAX}$ always, **sustained** if $\dot{U}_t \geq 0$ always and **survivable** if $U_t > U^{SURV}$ always. If we take utility to be a function of consumption(C_t) alone, development is **sustainable** if $C_t \leq C_{tMAX}$ always, **sustained** if $\dot{C}_t \geq 0$ always , and **survivable** if $C_t > C^{SURV}$ always.

Note that while the level of utility (or consumption) corresponding to survivability is taken to be constant over time (C^{SURV} carries no time subscript), C_{tMAX} includes a time subscript. Also, the highest level of constant, sustainable, consumption an economy can obtain from any point of time onwards does depend on which point in time we consider. For example, at the end of a prolonged and major war in which large stocks of resources have been consumed or irretrievably degraded, the maximum feasible level of sustainable consumption is likely to be smaller than it was before the war broke out.

These definitions establish three poplar criteria for operationalizing the concept of sustainability: **non-declining consumption**, **survivable development**, and **minimum condition**. A serious objection to the non-declining consumption as a way of operationalizing sustainable development is that it does not impose any requirements on how large the non-declining level of consumption should be. An economy is sustainable even if living standards are abysmally low and remain so, provided they do not get any lower over time. Survivable development introduces the idea of a minimum level of consumption consistent with biophysical survival requirements. This avoids the problems with the non-declining consumption constraint, but may not be considered as 'fair' enough to future generations. Setting a minimum consumption standard that is socially or culturally, rather than biologically, determined (ethically acceptable poverty lines) may help reduce the complexity of choices but also rules out better outcomes. Question 2 at the end of the Module tests your understanding of these issues. In what follows, we provide the basic framework used by economists in the effort to operationalize the concept sustainable development.

1.2.1.3 The Basic Framework

From the economist's viewpoint, providing future generations the same opportunities as that which is available to the present boils down to giving them the same consumption level. In other words, the opportunities that matter from the economists' perspective are consumption opportunities. Obligation to future generations lies in bequeathing a generalized productive capacity or, certain standards of consumption/living possibilities over time. Future generations will be interested in the consumption opportunities, not the stocks of resources, that they inherit.

These issues are explored within the framework of the simple optimal growth model, where production uses a finite non-renewable resource, along with physical capital; population is assumed constant, and labour's role in production is ignored. The goal is to maximize welfare over an infinite time horizon:

$$\text{Max} \quad W = \int_{t=0}^{t=\infty} U(C_t)\, e^{-\rho t} \text{dt} \qquad (1.3)$$

subject to the constraints

$$\dot{K} = Q\,(R_t, K_t) - C_t \qquad (1.4)$$

$$\dot{S} = -\, R_t \qquad (1.5)$$

$$\bar{S} = \int_{t=0}^{t=\infty} R_t \, \text{dt} \qquad (1.6)$$

where W= social welfare, C_t is as earlier defined, ρ is the utility discount rate (describing impatience, it shows how future utility from consumption is discounted to present value), Q is output, K_t and R_t are the quantities of physical capital and non-renewable resources used in producing Q at time t respectively, $\bar{S}$ is the finite initial stock of the non-renewable resource, and $\dot{S}$ is the change in the stock of the nonrenewable resource at time t.[6]

The fundamental question here is under what conditions is constant consumption forever possible, notwithstanding that production uses inputs of a non-renewable resource available only in finite total amount? The answer crucially depends on the nature of the production function, Q(R_t, K_t). For constant consumption forever to be possible where a non-renewable resource must be used as an input in production, three conditions are required.

- The technology of production must allow for some degree of substitution (though imperfect) between the inputs R_t, and K_t.
- Production technology must exhibit constant returns to scale, and
- The returns to K_t must be greater than that of R_t

Given these conditions, very high levels of output can be produced with very small levels of resource input, and there exists a programme of capital accumulation such that R_t, approaches but never actually becomes zero and consumption can be maintained constant forever.

The programme of capital accumulation that guarantees constant consumption over time, given the above conditions is summed up in what has been called the **'Hartwick rule**" (Hartwick, 1977, 1978). This requires that at every point in time, the total rent arising in the resource extraction industry should be saved and invested in reproducible capital. In the context of our model, the rule is that $\dot{K}$ must be equal to the total rents arisiing in the resource extraction industry. The unit

[6] For a review of the underpinnings of this model, read Perman, et al 2003, chapter 3.

rent is the difference between the price at which an extracted unit of the resource sells and the marginal cost of extraction. It is, essentially, the scarcity value of the resource. Total rent is simply unit rent times the number of units extracted.

Following the Hartwick rule means that the total value of the economy's stock of reproducible capital together with its stock of the non-renewable resource is held constant over time. As the value of the remaining stock of the resource declines, so the value of the stock of reproducible capital increases in compensating amount. Thus, the constant consumption level that goes with following the Hartwick rule can be thought of as being like the interest on this constant stock of total wealth. Module 5.1 explores in greater details the implication of the Hartwick rule foe resource accounting for environmental sustainability.

The Hartwick rule is necessary but not sufficient in itself to guarantee constant consumption over time. This will be realized under the Hartwick rule only if intertemporal efficiency conditions are satisfied (we shall examine this more closely in Module 3.3), and if sustainability as constant consumption is actually feasible; that is, if the substitution possibilities between capital and natural resources are great enough. The issue of substitutability between capital and natural resources is fundamental in the sustainability debate and is the focus of our discussion in the next subsection

1.2.1.4 Weak and Strong Sustainability

The possibility (or impossibility) or degree of substitution between environmental (and natural) resources and man-made capital (R_t and K_t) in the production of goods and services leads to two different ways of looking at the concept of sustainability: 'weak sustainability' and 'strong sustainability'. Practically, they refer to two different views about the conditions that need to be met for the realization of sustainability as constant consumption (or utility).

- **Weak sustainability**: that definition of sustainable development which allows for imperfect (and in some cases even perfect) substitution between environmental (and natural) resources and man-made capital (R_t and K_t) in the production of goods and services.
- **Strong sustainability**: that definition of sustainable development which allows for no (or in some cases very little) substitution between environmental (and natural) resources and man-made capital (R_t and K_t) in the production of goods and services.

Weak sustainability is based on a **capital approach** to sustainable development. This approach views all forms of capital resources (human made or natural) as assets and emphasize the possibility of substitution between them. In contrast, 'strong sustainability' is based on an **ecological approach** which sees sustainable development in terms of maintaining the ecosystems and their health (or properties). A third approach, the **three-pillar approach** emphasizes the three spheres of human needs (economic or material needs, social needs, and environmental needs) and argue that each is important on its own and none can take the place of the other. (Figure 1.7 illustrates the three-pillar approach to sustainable development).

Central to the weak versus strong sustainability debate is the concept of "natural capital". Production potential at any point in time depends on the stock of productive assets available for use. This stock can be classified into human labour and all other productive resources, which are grouped as capital. Capital in this sense will include

- **Natural capital**: any naturally provided stock, such as aquifers and water systems, fertile land, crude oil and gas, forests, fisheries and other stocks of biomass, genetic material, and the earth's atmosphere itself. It is the collectivity of the environmental assets from which all such services flow.
- **Physical capital**: plant, equipment, buildings and other infrastructure, accumulated by devoting part of current production to capital investment.
- **Human capital**: stocks of learned skills, embodied in particular individuals, which enhances the productive potential of those people.
- **Intellectual capital**: disembodied skills and knowledge. This comprises the stock of useful knowledge, which we might otherwise call the state of technology. These skills are disembodied in that they do not reside in particular individuals, but are part of the culture of a society. They reside in books and other cultural constructs, and are transmitted and developed through time by social learning processes

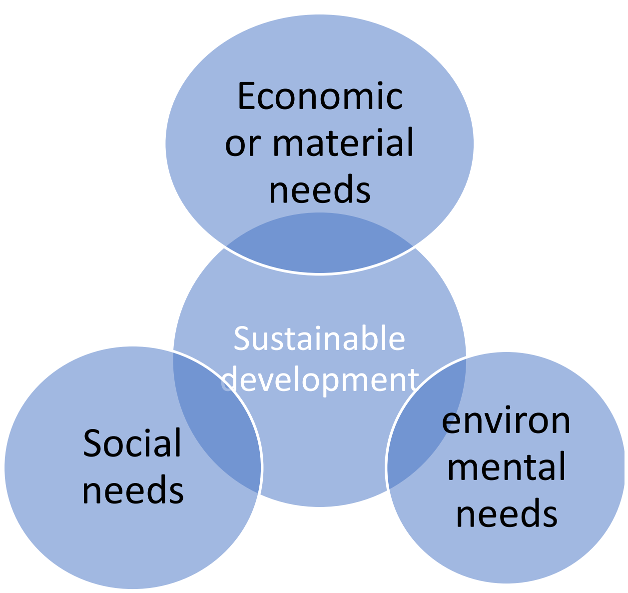


Fig. 1.7 The three-pillar approach to sustainable development

Using the above concepts, the economy's production function can be represented as

$$Q = Q\,(L,\, K_N,\, K_H) \qquad (1.\,7)$$

Where L = Labour.

K_H= human-made capital (also called **reproducible capital**) and is the sum of physical, human and intellectual capital, and

K_N= natural capital.

This formulation of the production function does not allow the function to change with changing technology (since technology is itself defined as part of reproducible capital).

The difference between weak and strong sustainability is reflected in the extent of substitution possibilities between K_N and K_H implied in equation (1.7). Proponents of *strong sustainability* argue that sustainability requires that the level of K_N be non-declining while proponents of *weak sustainability* argue that it only requires that the sum of K_N and K_H be non-declining. In narrow terms, 'weak sustainability" requires that the generalized production capacity of an economy is maintained intact or order to enable constant consumption per capita through time. Some authors call this criterion leading to a constant per capita consumption the *Very Weak Sustainability* (VWS) requirement. In a broader sense, *weak sustainability* requires that the welfare potential of the overall capital base remains intact (*non-decreasing social welfare*). However, this is not restricted to sustaining a material standard of living, or consumption, but also includes the values that are related to non-consumptive uses (existence and bequest values) and the public good character (amenity and recreational values) of the environment.

The concept of strong sustainability, on the other hand, implies a physical principle which is founded upon the laws of thermodynamics and processes of biological growth and requires the maintenance of the "quality of the environmental resources" (*constant environment quality*). This is a function of the stocks of biological resources, ecosystem space, and other environmental assets that are essential for the integrity of the ecosystem, and provide use and non-use values to society. A Very Strong Sustainability (VSS) criteria requires a set of stationary-state condition (defined by the rate of regeneration and the assimilative capacity of the environment and a minimum standard conservation).

Each approach, weak or strong sustainability, has its strength and weaknesses (see Box 1.4 below). As indicated, the former is the approach adopted by many environmental economists while the latter is particularly favoured by ecologists and most ecological economists. Three concepts stand out clearly in the ecologists' approach to sustainable development: **resilience** of ecosystems, **uncertainty** and **irreversibility**. The approach highlights the belief that our understanding of how natural systems function is very incomplete and our ability to predict the ecological consequences of our behaviour highly imperfect. Uncertainty pervades the behaviour of ecological systems, ensuring that we cannot know in advance whether some system is or is not resilient.[7] Thus, in thinking about how to manage them in our interests, we have to recognize

[7] Some authors have suggested that some indicators are useful as monitoring devices: they can be used to make inferences about potential changes in the degree of resilience of ecosystems in which we are interested. Schaeffer et al. 1988) propose a set of indicators, including changes in the number of native species; changes in standing crop biomass; changes in mineral micronutrient stocks; changes in the mechanisms of and capacity for damping oscillations. But suggestive as these and other indicators might be, none can ever be a completely reliable instrument in the sense that a satisfactory rating can be taken as a guarantee of resilience (see Perman et, al. 2003. p, 94).

that there is great uncertainty. Coupled with this is irreversibility: the fact that for most natural and environmental resources, it is impossible to undo a damage once it has been made.

Box 1.4. The Debate on Weak and Strong Sustainability

Historical experience tends to support the idea that physical, human and intellectual capital accumulation can offset any problems arising as stocks of natural resources are depleted. It is also true that there are many opportunities for substitution as between particulars of the general class of natural resources. But there is not yet compelling evidence that man-made capital can substitute for the life-support and amenity services that natural capital provides. As spacecraft have already demonstrated, it is possible to use K_H to provide necessary life-support services such as temperature control, breathable air, etc., but only on a small scale. It has yet to be demonstrated, or even seriously argued, that human-made capital could replace natural capital in providing life-support services for several billions of humans. In regard to amenity services, some take the view that a lack of contact with the natural environment is dehumanizing, and would argue that we should, as an ethical matter, regard the possibilities of K_H for K_N substitution as limited.

In general, the weak versus strong sustainability question is multi-faceted, and does not permit of firm precise answers, except in particular contexts. And, in some particulars, the answer is as much a matter of taste and/or ethics as it is a matter of science and technology. In terms of policy, if we are going to maintain the size of the total stock of capital (in line with the proponents of weak sustainability) or the size of the natural capital stock (in line with those that believe in strong sustainability), it is necessary to be able to measure the size of the natural capital stock (because the natural stock is itself part of total capital). Here we encounter many problems as natural capital consists of many qualitatively different components (how do we add two lakes and one forest into a single value for natural capital, for example?). For market goods, prices could serve as weights to use for aggregation but there are no market prices in relation to most environmental goods. Even if we can find prices for some environmental goods, there are many reasons why one would not be willing to accept them as correct reflections of 'true' values. (These issues are discussed in Module 5).

To make strong sustainability an operational principle, several authors have attempted to translate the constant natural capital rule into a set of ecological criteria, including the "Safe Minimum Sustainability Standards' (SMSS) or stationary state principle (defined by the rate of regeneration and the assimilative capacity of the environment) and the minimum standard conservation (SMC).

Sources: Perman et, al. chapter 4, Freeman, 1993).

If all resource-use decisions were reversible, then much of the force behind sustainability arguments would be lost. Reversibility implies that nothing would have been irretrievably lost. But many decisions about the use of environmental services cannot be reversed, particularly those that involve the extraction of resources or the development of undisturbed ecosystems.

When irreversibility is combined with imperfect knowledge of the future then optimal decision rules can change significantly. In these circumstances, there are good reasons for keeping options open and behaving in a relatively cautious manner.

Table 1.2 below summarizes the main points in the distinctions between *weak sustainability* (capital approach to sustainable development) and *strong sustainability* (the ecological approach to sustainable development).

Table 1. 2. Main Differences between Weak and Strong Sustainability.

	Strong sustainability	**Weak sustainability**
Key idea	The substitution of natural capital with other types of capital is severely limited	Natural capital and other types of capitals (manufactures are perfectly substitutable.
Consequences	Certain human actions can entail irreversible consequences	Technological innovation and monetary compensation for environmental degradation.
Sustainability issue	Conserving the irreplaceable "stocks" of critical natural capital for the sake of future generations	The total value of aggregate stock of capital should be at least maintained or ideally increased for future generation
Key concept	Critical natural capital	Optimal allocation of scarce resources
Definition of thresholds and environmental norms	Scientific knowledge as input for public deliberation (procedural rationality)	Technical/scientific approval for determining thresholds and norms (instrumental rationality)

Source: Adapted from Mancebo, 2013

1.2. 2. Factors responsible for Environmental Degradation

1.2. 2.1 Economic activities, Market Failure and Government Failure

In a broad sense, environmental degradation refers to the erosion in the quality, and hence impairment in the integrity and functions, of the environmental resource base (ERB). This can be due to

- the extraction of renewable natural resources at a faster rate than they are replaced, or
- the release of waste into the environment at a level that exceeds the assimilative capacity of the ERB: pollution

As we saw in Module 1.1, human (economic) activities are behind most cases of environmental degradation. Human activities that lead to the degradation of the environment may be driven by poverty. The sheer burden of survival may make household to extract natural resources beyond sustainable levels. This is especially serious in societies that are largely dependent on such resources, as in many communities in sub-Saharan Africa. In contrast, such activities may also be driven by the crave for affluence. Whatever is the motive, what is important to note is that these activities take place because of certain institutional failures in society. These failures can be classified into two broad groups: those due to the existing functioning of markets (market failure) and those due to public policy (government or policy failure).

Market failure occurs when the market fails to allocate resources efficiently because of the absence of certain necessary conditions. Environmental and most natural resources are not like other resources (goods and services) that are traded in the market. They have unique characteristics which makes it impossible for the market system to reckon with them in an efficient manner, if at all. Ideally, such failure of the market system would require an effective government intervention. However, in many cases, for one reason or the other, **government fails** to intervene, or does so in an inefficient or ineffective manner, sometimes making a bad situation worse. In addition, government can promote policies to achieve certain goals (like economic growth, poverty alleviation) which are good in themselves but nevertheless degrade the environment. Consider the use of subsidies on fuel (gasoline) to assist poor households, which until recently was very popular in many developing countries. While this *may* help poor households, it encourages greater fossil use and hence more pollution. Consider also the policies adopted in the extractive sectors of many resource-dependent countries which gives little or no attention to environmental factors. We will be looking at market and government failures in relation to the use of environmental and natural resource more closely in Module 2.

1.2.2.2 Environmental Impact of Economic Activities: the IPAT Identity

In a series of papers in 1970-74, Paul Ehlrich and John Holdren (Ehrlich and Holdren, 1971; Ehrlich and Ehrlich, 1990) proposed an equation aimed at estimating the overall impact of economic activity on the environment. The IPAT equation provides a useful theoretical framework to analyze the determinants of environmental impact and offers guidance on possible directions for policies to mitigate the negative effect of human activities on the ERB.

Economic activities involve extraction of resources (and service flows) from the environment and the insertion into it in the form of wastes. The amount of extraction and insertion at any point in time and for any particular resource (or waste) are determined by the size of the human population and how much *impact* each individual makes (the per capita impact). The per capita impact depends on how much each individual consumes and on the technology of production (and consumption). The consumption level of an individual is an indication of the level of affluence. The technology of production is reflected in the amount of natural/environmental resource used to produce a unit of output or the waste generated in producing that unit. Technological improvement implies a reduction in one or both measures.

If we represent the impact on the environment by I, population by P, consumption per capita by A (where A represents affluence) and the technology of production (and consumption) by T, we can define the identity

$$I \equiv P \times A \times T \qquad (1.8)$$

As a matter of practice, *I* is usually measured as mass or volume of a particular environmental resource, e.g. global carbon dioxide (CO_2) emissions, biodiversity loss, reduction in forest stock etc., and A as per capita income (GDP/population). T is measured the same way as I (what is used up in production or emitted as waste through existing technology is also a measure of the environmental impact) except that the measure is divided by the overall quantity of goods produced (consumed) (GDP) to get the amount of resource used or waste generated per unit of output.

Thus, we can express (1.8) as

$$\textit{Resource use/ Waste} \equiv \textit{Population} \times \frac{GDP}{Population} \times \frac{\textit{Resource use/ Waste}}{GDP} \qquad (1.9)$$

Equation (1.9) assists us to see why the IPAT equation is an identity. Cancelling the two population terms and the two GDP terms leaves us with Resource use/waste as an identity[8] .We now illustrate how the IPAT identity can be employed to analyses different environmental degradation scenarios. We take the case of global carbon dioxide (CO_2) emissions. Table1.3 presents the calculation of the various elements in the IPAT identity for various global CO_2 scenarios**.**

[8] Different authors have proposed various versions of the IPAT equation. For example, Dietz and Rosa (1994) proposed that IPAT accounting equation would be recast into a stochastic form in order to allow random errors in the estimation parameters. The authors consider the following formulation, I = $ap^bA^cT^de$ (where a, b, c and d are parameters and e is an error term). The authors refer to the reformulated version of IPAT as STIRPAT, an acronym of Stochastic Impact on Population, Affluence and Technology. Additionally, Schulze (2002) proposes to modify I = PAT as I = PBAT, where B are behavioral choices. The author suggests that per capita impact also depends on modifying behavior, capturing more clearly the determinants of environmental impact. Its further benefit serves to draw attention to the many behavioral choices that are immediately available to all individuals.

Table 1.3. Applying IPAT for under various global Co2 scenarios

Scenarios	Description	P (billions) 1999	A (PPP US $) 1999	T (tons per $) (derived)	I (billions of tons) (1996)
Scenario 1	Base	5.8627	6948	0.0005862	23.881952
Scenario 2	P x 1.5 (A, T remains the same)	8.8005	6948	0.0005862	35.843711
Scenario 3	P × 1.5 and A × 2 (T remains the same)	8.8005	13896	0.0005862	71.687417
Scenario 4	P × 1.5 and A × 2 (I remains same)	8.8005	13896	0.0001952	23.881952

Source: Perman et al.2003.p30)

The following can be noted from the analyses

- **Scenario one (base scenario)**: figure for T is gotten from the figures for P, A and I which were all taken from data provided by the UNDP (2001) and the World Resource Institute, WRI (2000).
- **Scenario two** uses the T figure from scenario one(base) to show the implications for I of a 50% increase in world population, given constant affluence (A) and technology (T).
- **Scenario three**: uses the T figure from Baseto show the implications for I of a combination of a 50% increase in P (world population is expected to increase by 50% between 1999 and 2100) with an 100% increase in affluence (a doubling of per capita GDP). Note also the effect on the environment.
- **Scenario four:** solves IPAT for T when I is set equal to its base level and P and A are as in scenario three**.** It shows that to keep Co2 emissions (I) at its base value given a 50% population increase and a doubling of affluence would require a reduction in the resource use or waste generated by the technology of production (T) to one third its base level.

The kind of analysis carried out in scenario four is of particular interest to climate change advocates, many of which believe that the current level of carbon dioxide emissions is dangerously high. However, climate change analyses are much more complex.

1.2.2.3. IPAT Identity and Determinants of Environmental Degradation

(a). Population growth (P)

Human population growth contributes to many environmental pressures—greater demand for food production, higher rates of resource depletion, more waste generation for a given level of production technology etc. Intensification of food production to meet the needs of a rising population imposes significant environmental costs, including land degradation, pollution from fertilizers and pesticides, and overtaxing of water supplies. There has been dramatic increase in world population since the Industrial Revolution. Global population which was approximately 1 billion in 1800, doubled to 2 billion by 1930 and reached 3 billion in 1960. Over the next 40 years (1960-2000), it doubled again, reaching 6 billion by 2000. As of 2018, it had reached 7.6 billion. The United Nations predicts that global population could reach between almost 8 to over 16 billion by 2100, depending on the average global growth, or fertility rate assumed (Figure 1.8).

Although population growth rates have declined from 2.1 percent annually in the 1960s to approximately 1 percent today, the human population is still increasing by about 80 million people per year, equivalent to the population of Germany. However, historical evidence generally suggests that the United Nations has underestimated how quickly fertility rates have fallen. The percentage rate of increase of global population is already well below its historical peak, having decreased in recent years in all regions of the world. Growth rates are currently less than 0.5% per year in developed countries and just over 2% in developing countries. Several countries (for example, Germany, Austria, Denmark and Sweden) now have falling populations, and many others are expected to move into this category in the near future. In many countries (including all industrialized countries and China), fertility rates are below the replacement rates that are required for a population size to be stationary in the long run. For these countries, population is destined to fall at some point in the future even though the momentum of population dynamics implies that population will continue to rise for some time to come.

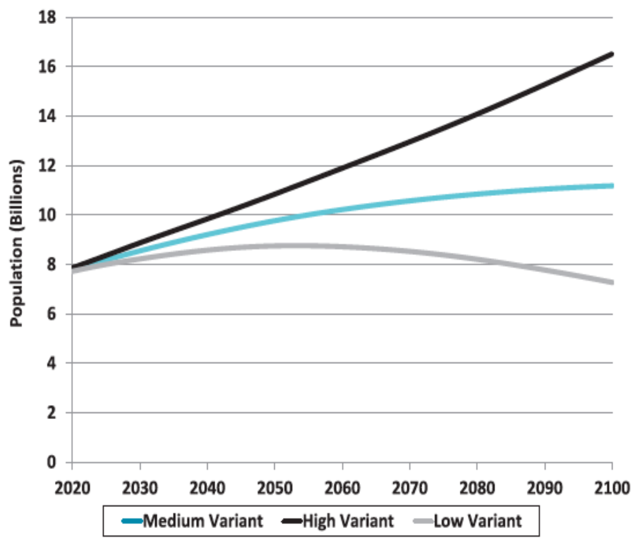


Fig. 1.8 United Nations Global Population Projections, 2020-2100. In Jonathan and Roach, 2017

The take-home point in all of these is that even slight changes in fertility rates can have profound effect on global population and on the environment. Humanity's environmental impacts will be quite different in 2100 if only 7 billion humans are on the planet as opposed to 16 billion! Thus, the potential for a sustainable future may hinge upon what happens to fertility rates around the world in the coming decades.

Policies aimed at reducing fertility rates can be designed and they offer the prospect of eliminating overpopulation as a contribution to the sustainability problem. How can (and how have) these been done and what have been the results? The microeconomics of fertility and population growth suggests that policies that raises the marginal cost (MC) of having additional children and/or reduces the marginal benefits (MB) could work. The marginal cost includes direct costs (such as the costs of childbearing, child rearing and education) and indirect costs (that is the opportunity costs of parental time in the above activities, including lost income). The marginal benefits include the psychic benefits of children, contribution of children to family income, and the extent to which old age security is enhanced by larger family size.

Policies designed to influence population growth must alter desired family size by shifting MC and or MB so that cost of having an extra child exceeds the benefit. These will include increased levels of education, particularly education of women, creating greater employment access for females, use of financial incentives (such as adjusting existing fiscal and welfare provisions, such as tax deductions, scholarships, bursaries, subsidized food, costs of access to health and educational facilities etc., in a way that they are not available beyond a certain number of children per household)[9], provision of care for and financial support of the elderly, financed by taxation on younger generation, and general economic development, including the replacement of subsistence agriculture by modern farming practices and giving farm workers the chance of

[9]Taking this too far raises question of fairness and may generate tension.

earning labour market incomes. Current declining population growth rates have been attributed to numerous factors: the widespread availability of birth control, higher costs associated with raising children, and, perhaps most importantly, a focus on educating girls.

(b). Affluence (A)

Despite the countervailing influence of population growth, the world has witnessed remarkable growth in output and consumption per man. Maddison (1995) showed that per capita GDP grew from about $1000 in 1820 to about $8500 in 1992 in a sample of 57 countries, which currently account for over 90% of world population. Over the period 1975 to 1999, world average GDP per capita grew at 1.3% per annum. At that rate of growth, over 50 years, the level of world average GDP would just about double. However, income growth has not been evenly distributed. The world average in 1999 was about twice that for China, and about 20% of that for the USA.

Economists have had an age-long fascination with economic growth. Some believe it is the panacea to all problems (the poverty of the world, the environmental degradation, and even the problem of population growth!). Two of the behavioural relationships that gained traction in economics emphasized the importance of economic growth(A) and its interaction with population growth (P) and the environment (or the technology of consumption, T). We will look at one (the demographic transition theory) here and the other (the Environmental Kuznets Curve) in our consideration of the role of T in the IPAT identity.

If affluence reduces population growth, then it could have an indirect benefit on the environment since smaller population implies lower pressure on natural and environmental resources. The **demographic transition theory** attempts to explain an observed negative correlation between income level and population growth rate (Todaro, 1989). It postulates four stages through which population dynamics progresses (Figure 1.9).

- **First stage**: high birth-rates and high death- rates. In some cases, death-rates may reflect intentions to keep populations stable, and so include infanticide, infant neglect and senilicide (Harris and Ross, 1987).
- **Second stage**: rising real incomes leading to improved nutrition, developments in public health. These lead to declines in death-rates and rapidly rising population levels.
- **Third stage**: reduced fertility rates occasioned by economic forces, such as increasing costs of childbearing and family care, reduced benefits of large family size, higher opportunity costs of employment in the home, and changes in the economic roles and status of women. Consequently, population growth begins to decline.
- **Fourth (final) stage**: relatively high income per person leads to low, and approximately equal, birth and death rates, and so stable population sizes.

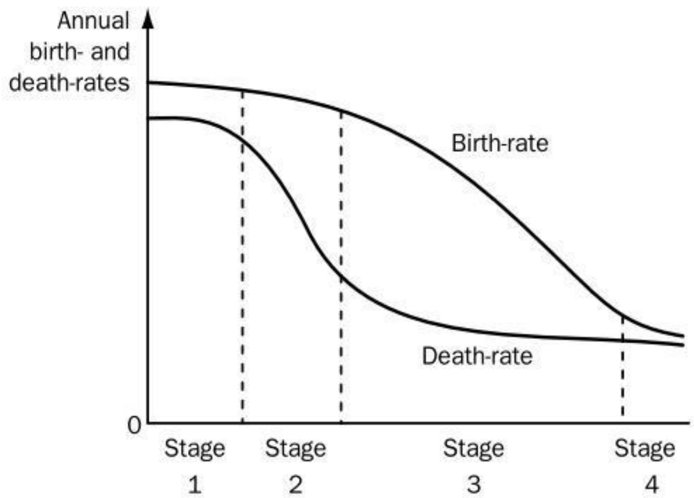


Fig. 1.9. Demographic Transition Theory. **Source:** Perman et al. 2003.p34)

The demographic transition theory appears to have fitted well into the observed population dynamics of many developed countries. However, it does not seem to have general applicability to today's developing countries. For most of these countries, the second stage was reached not as a consequence of rising real income but rather as a consequence of knowledge and technological transfer (rapid transfer of public health measures and disease control techniques from overseas led to unprecedented fall in mortality rates and drastically shortened the time it took to reach the second stage). In addition, reductions in birth rates have not been as forthcoming in many developing countries as was experienced by the developed contrives following reductions in mortality rates. For many reductions in birth rate are still lagging behind falls in mortality rate. There may also be a vicious circle of poverty in which the resources required for economic development (and so for a movement to the third stage of the demographic transition) are crowded out by rapid population expansion (Dasgupta, 1992).

(c) Technology of production and consumption (T)

Can affluence lead to a reduction in environmental by making people choose more environmentally-friendly way of doing things? The Environmental Kuznets Curve (EKC), named after Simon Kuznets (1955), suggests the existence of an inverted U relationship between income growth and environmental damage. As an economy transits from subsistence agriculture into more intensified agriculture, resource extraction and industrialization, the rates of resource depletion begin to exceed the rates of resource regeneration, and waste generation increases in quantity and toxicity. However, as income levels continue to grow, the economy transits towards information-intensive industries and services. In addition, environmental awareness increases leading to greater demand for a "clean" environment, increased enforcement of environmental regulations, use of more environmental-friendly technology and higher environmental expenditures. All of these result in levelling off and gradual decline of environmental degradation. In other words, sufficient wealth and technology allow countries to adopt clean production methods and move to a service-based economy. Environmental quality is a "normal good," meaning that people will demand more of it as they become wealthier. We will briefly examine the model and the empirical evidence.

Let per capita emissions of some pollutant into the environment (*e*) be a function of affluence (represented by per capita income, *y*).

$$e = \alpha y \qquad (1.10)$$

where α is a factor that shows how changes in *y* translates to changes in *e.*

The EKC hypothesis suggests that the coefficient α is not constant but depends on y. Thus, we can represent α as a linear function of y. as follows

$$\alpha = \beta_0 - \beta_1 y \qquad (1.11)$$

Substituting (1. 11) into (1.10) gives

$$e = \beta_0 y - \beta_1 y^2 \qquad (1.\ 12)$$

The EKC hypothesis assumes that β_1 is sufficiently small in relation to β_0 so that relationship between *e* and *y* takes the form of an inverted U: economic growth means higher emissions per capita until per capita income reaches a turning point after which further growth reduces emissions per capita.

If the evidence supports this hypothesis, it would imply that policies that foster macroeconomic growth will eventually promote a cleaner environment as well. In that case, we do not need to search for separate policies for the environment. Environmental improvement will come naturally with growth. But what does the evidence show? Assessing the validity of this conclusion involves two questions. First, are the data generally consistent with the EKC hypothesis? Second, if the EKC hypothesis holds, does the implication that growth is good for the global environment follow?

Available data seem to support the EKC relationship for some pollutants. Figure 1.10 shows the findings of a study that estimated the relationship between the average particulate matter (PM10) concentration in a country and a country's per capita income. At very low levels of income, the expected PM10 concentration tends to rise quickly as a country develops economically. PM10 concentration peaks when a country reaches an average income of around US$1,300 per person. Air pollution levels then fall steadily with further economic advancement. The World Health Organization (WHO) recommends that PM10 levels be below 20µg/m3. On average, countries achieve this standard when income per person rises above US$17,000 per person. Evidence supporting the EKC hypothesis has also been found for municipal solid waste and other air pollutants such as sulfur dioxide and carbon monoxide.

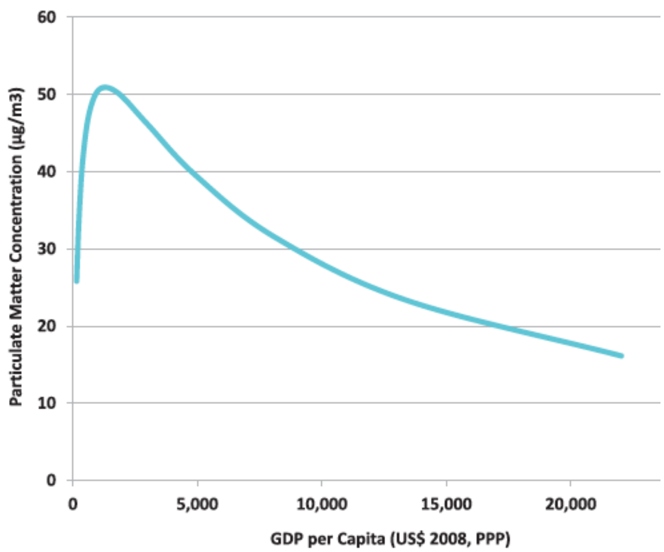


Fig. 1.10. Environmental Kuznets Curve for Particulate Matter **Source:** Mazurek, 2011).In Jonathan and Roach,

But the hypothesis does not appear to hold for some other pollutants (see Box 1.5). Other studies have consistently shown a positive relationship between CO_2 emissions and average income. In Figure 1.11, a simple statistical test to fit an inverted-U curve through the data finds that there is no turning point: per-capita CO_2 emissions continue to rise as GDP/capita increases.

Thus, promoting economic growth does not appear to be an effective means to address the issue of global CO_2 emissions. The relationship between economic growth and the environment is, in reality, more complex than implied by the EKC hypothesis. Even if the data appear to confirm that the EKC fits the experience of individual countries, it does not follow that further growth is good for the global environment. (Stern et al, Arrow et al. 1995). Economic growth is not a panacea for environmental quality; indeed, it is not even the main issue. Policies that promote gross national product growth are not substitutes for environmental policy.

Box 1.5 The EKC: Empirical Evidence

One of the earliest empirical studies on the EKC (Shafik and Bandyopadhyay, 1992) tested the validity of the relationship for ten different environmental indicators (lack of clean water, lack of urban sanitation, ambient levels of suspended particulate matter in urban areas, urban concentrations of sulphur dioxide, change in forest area between 1961 and 1986, the annual rate of deforestation between 1961 and 1986, dissolved oxygen in rivers, faecal coliforms in rivers, municipal waste per capita, and carbon dioxide emissions per capita).

Lack of clean water and lack of urban sanitation were found to decline uniformly with increasing income. The two measures of deforestation were found not to depend on income. River quality tends to worsen with increasing income. Two of the air pollutants were found to conform to the EKC hypo thesis but CO2 emissions, a major contributor to the 'greenhouse gases' do not fit the EKC hypothesis, rising continuously with income, as do municipal wastes.

The authors summarized the implications of their results by stating as follows: it is possible to 'grow out of' some environmental problems, but there is nothing automatic about doing so. Action tends to be taken where there are generalized local costs and substantial private and social benefits (**Source**: Perman et al., 2000, p

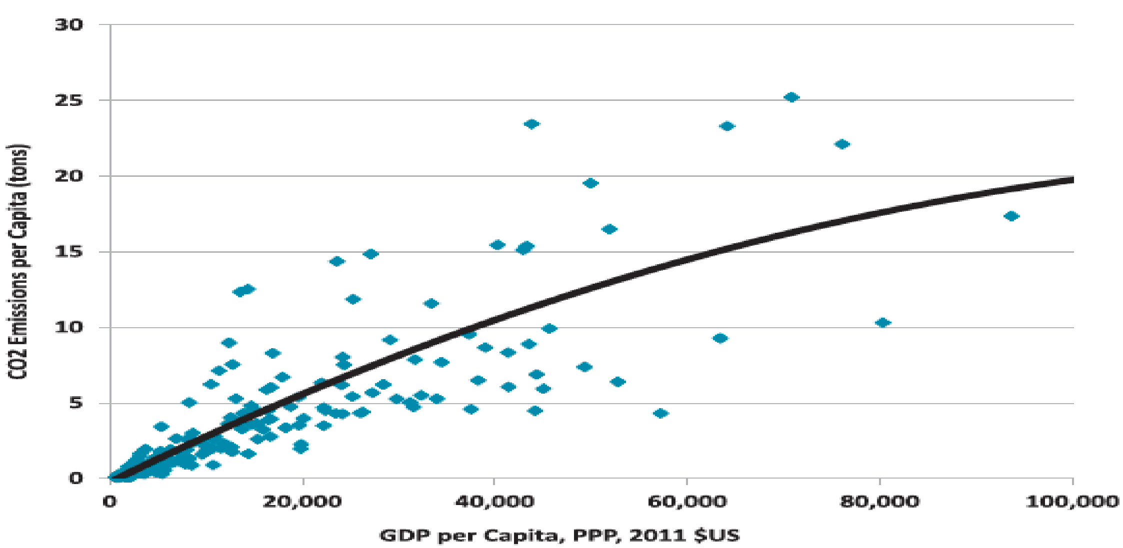


Fig.1.11. Relationship between Carbon Dioxide Emissions and GDP per Capita, 2014 (World Bank, World Development Indicators database), **Source:** Jonathan and Roach, 2017.

1.2.2.4 Economic activities and Environmental Degradation: Some Critical Notes and Observation on Recent Trends

This Module has illustrated the effect of factors, such as population growth, growth in consumption per capital, and production and consumption technologies on the environment. The IPAT identity provided a useful and simple device to show this interrelatedness. It must be strongly emphasized, however. that in reality the dynamics of environmental degradation goes far beyond the simplistic ideas represented in the IPAT equation. It is not to be expected, for example, that the other various components in the IPAT equation will remain constant when one changes (which leads to such simplistic statements as "double the population and you would double the 'environmental problems'", for example). Thus, while the Chinese "one child policy" may have contributed to limiting the country's "footprint", for example, there are many underlying complex relationships between population, growth and the environment that are not accounted for. It is also important to reiterate the fact that the various stages of progress implied in the EKC hypothesis should not be interpreted literally. As demonstrated in the preceding section, the hypothesis has not found much support in the empirical literature; many actually believe it has outlived its usefulness in terms of its ability to predict the relationship between growth and environmental quality. In any case, the idea that a country has to be first thoroughly polluted before it can start to clean up does not appeal even to common sense. In the remaining part of this subsection, we will present some observations on recent global trends on economic activities and environmental degradation.

Current trends show continual growth in population, food production, energy use, and gross world product. But population growth has slowed, and energy use has become somewhat more efficient, as shown by the energy trend line falling below that of gross world product starting in the 1980s (Figure 1.12). Global emissions of three major pollutants increased significantly from 1940 to 1980. But since then, emissions of Sulfur Dioxide (SO_2) have declined by about 20 percent, and the growth in Nitrogen Oxides (NO_x) and Volatile Organic Compounds (VOC) emission has slowed (Figure 1.13).

The most rapid increase in emissions has been in upper middle-income countries, including China—up 180 percent since 1990. Emissions have also increased significantly in lower middle-income countries, about 120 percent since 1990, primarily as a result of emissions growth in India. As a result of these trends, CO_2 emissions from high-income countries now comprise less than half of the global total (Figure 1.14). Except mitigating measures are taken, the aspirations of many developing countries to alleviate poverty through rapid growth is likely to further raise emission levels.

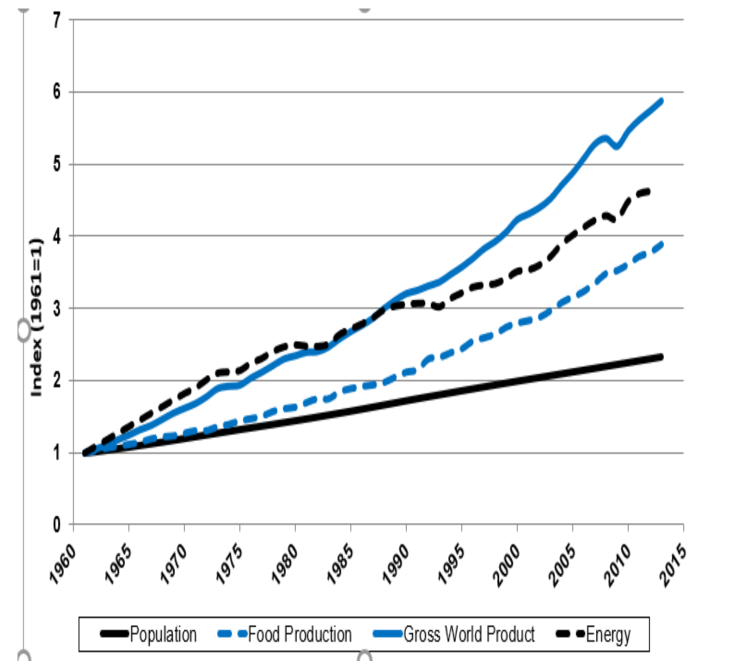


Fig. 1.12: Growth in Population, Food Production, Economic Production, and Energy Use, 1961-2013. ***Sources***: Population, food production, and gross world product from the World Bank, World Development Indicators database; energy data from the U.S. Energy Information Administration, International Energy Statistics. In Jonathan and Roach, 2017

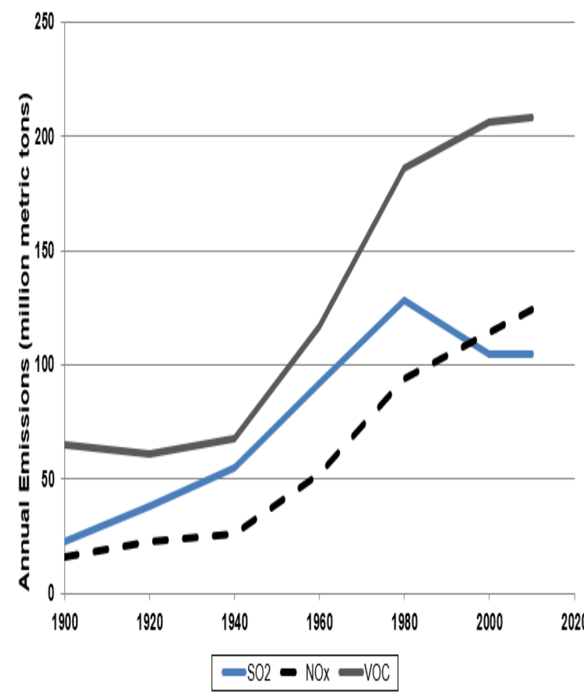


Fig. 1.13: Global Emission Trends for Sulfur Dioxide, Nitrogen Oxides, and Volatile Organic Compounds, 1900-2010. *Source*: HTAP, 2010. Also in Jonathan and Roach, 2017

The global chemicals industry has expanded significantly in recent decades, with annual production growing from less than $200 billion in 1970 to more than $4 trillion by 2010. Prior to 2000, chemicals production took place primarily in developed nations. But since then the global chemicals industry has expanded significantly in recent decades, with annual production growing from less than $200 billion in 1970 to more than $4 trillion by 2010. Prior to 2000, chemicals production took place primarily in developed nations. But since then production has shifted to developing countries, with China now being the world's largest developing countries, with China now being the world's largest producer (Figure 1.15). Increasing pressure is also being exerted on resource use. Global water use increased by more than a factor of five during the twentieth century, with further increases projected in the future. (Figure 1.16). Many countries are becoming increasingly dependent upon groundwater, which is essentially a non-renewable resource. At the same time, significant deforestation is occurring in Latin America (including the Amazon Forest in Brazil) and Africa, although recently at lower rates than during the 1990s. Reforestation in China, where an estimated 66 billion trees have been planted in recent decades, has led to a net increase in forested area in Asia (Figure 1.17).

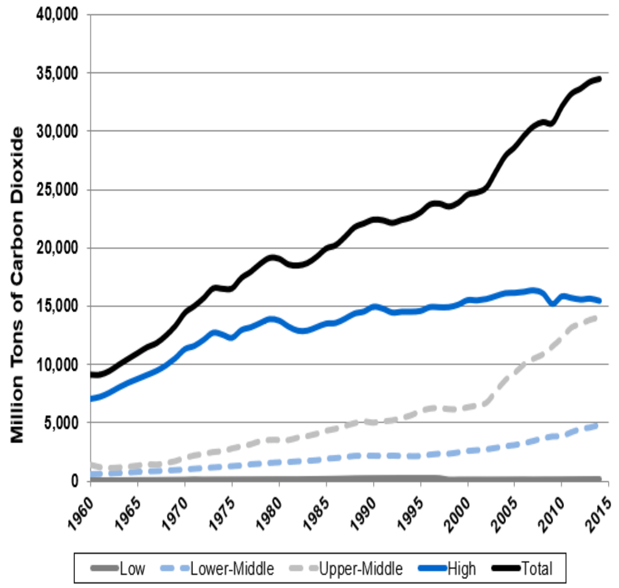


Fig. 1.14: Global Carbon Dioxide Emissions,1960-2014, by Country Income Category. *Source:* Jonathan and Roach, 2017.

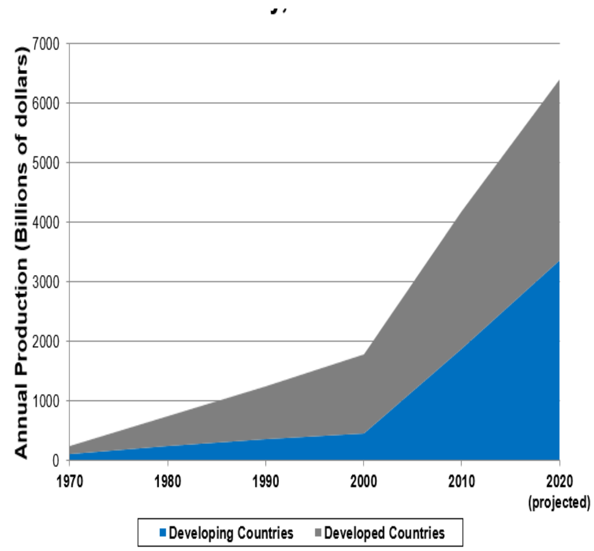


Fig. 1.15: Growth of Chemicals Production Industry, 1970-2020. ***Source***: UNEP, 2013, Also, in Jonathan and Roach, 2017.

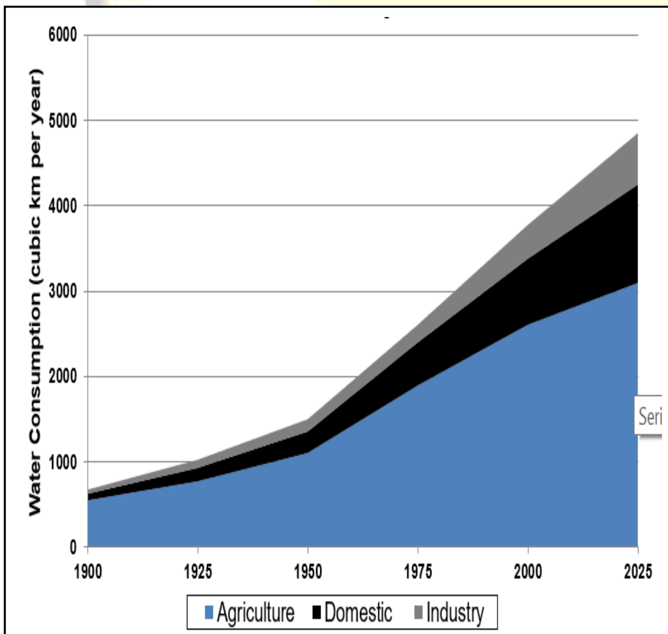


Fig.1.16 Global Water Extraction,1900-2025, by Use, ***Source:*** UNEP, 2008. Also, in Jonathan and Roach, 2017.

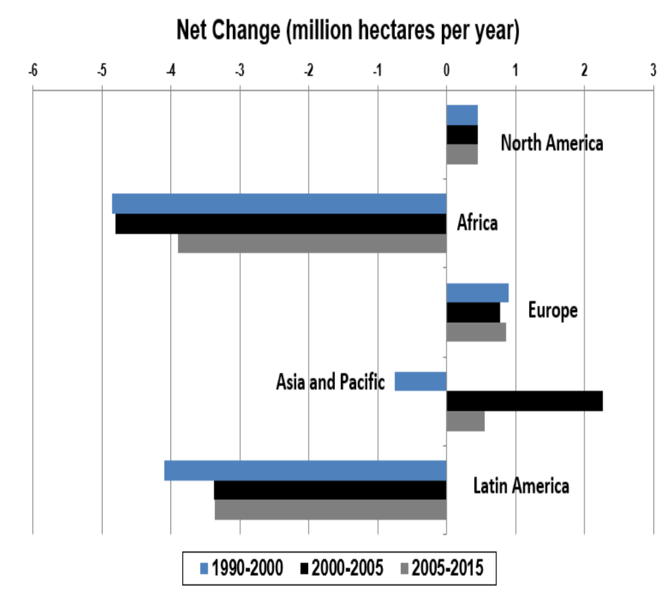


Fig.1.17 Net Change in Forest Area, by Region and Time Period. **Source:** UNEP, 2010.Jonathan and Roach, 2017.

1.2.3. Green Growth

The concept of "green growth" is used to described as a path of economic growth that relies on natural resources in a sustainable manner. The concept is now embraced globally as an alternative to the typical industrial economic growth paradigm. This path now leads to a novel concept known as a green economy.

The emergence of green growth as a concept and policy focus was against the backdrop of the 2008 global economic crisis, acknowledged as the most severe in more than 70 years. The economic downturn followed two decades of accelerated economic growth around the world, which lifted almost 1 billion people out of extreme poverty. However, the trend also left about a billion people in extreme poverty, with 1.1 billion people without electricity and 2.5 billion without access to sanitation. Thus, the growth was hardly inclusive and came at the expense of the environment, which suffered considerable degradation, characterized by pollution, biodiversity loss, and a changing climate. The global economic downturn precipitated multiple crises across sectors, from banking and finance, industry, as well as international trade and investments. The contagion unleashed credit squeeze, provoking economic slowdown across the world, with grave implications for rising poverty and inequality. The social impact of the global economic downturn was critical, prompting governments to intervene with massive bailouts in the United States and the European Union, among other regions of the world. The global economic downturn drove several economies into prolong recession, with grave difficulties for the vulnerable in many societies. It was also a further testimony to the unsustainable nature of current growth paradigm.

The concept of green growth emerged as a pre-requisite for building a green economy in the context of sustainable development and poverty reduction. The issue was raised at the inter-governmental discussions of the fifth Ministerial Conference on Environment and Development (MCED) in Asia and the Pacific, held in 2005 in Seoul, Republic of Korea. At the meeting, Green growth, or environmentally sustainable economic growth, was defined as a pathway of sustaining economic growth and job creation necessary to reduce poverty in the face of worsening resource constraints and climate crisis.

Promoters of Green growth believe it can secure a strong, stable and sustainable future for all in the face of the wide range of socio-economic and environmental challenges that are pushing the world to a tipping point. A key challenge will be meeting the energy, food and water needs of 9 billion people by 2050 and ensuring that they have clean and healthy environments. A recent modeling exercise by the OECD[10] projects large costs and potentially irreversible consequences of failing to adjust economic growth to avoid environmental risks, which will directly affect human health and wellbeing with dramatic consequences for developing countries. In contrast,

[10](OECD) (2013), Putting Green Growth at the Heart of Development. Summary for Policy Makers

Governments that put green growth at the heart of development policy and practice can achieve sustainable economic growth and social stability, safeguard the environment and conserve resources for future generations. Indeed, reconciling development with environmental protection and sustainable natural resources management is critical to avoid natural capital depletion, climate change and social insecurity.

Integrating economic and environmental policies can be particularly challenging. It requires concerted efforts to instil change, a common vision of the future and robust cooperation across government agencies. It also involves trade-offs in the short term, which need to be managed and reconciled with the long-term benefits that are to be achieved. A common argument against going green is that it is too expensive. However, this fear may actually be unfounded. Those who believe that adopting green growth would come at an additional cost assume that the current economy operates efficiently and that business as usual will continue to translate into economic benefits. However, there is growing evidence which proves that a "grow first, clean up later" approach will be costly in future. Indeed, evidence also suggests that the current economic system does not function optimally, as ecosystem goods and services are often provided free, while depletion is not adequately factored unto current decision making. As a result, when water become polluted or run out, for example, the price of this "free good" becomes extremely expensive and costing governments, businesses and the end consumer more than the cost of taking steps to manage the resource sustainably. Table 1.4reveals the benefits associated with green growth and the costs when it is not adopted.

Table 1.4: Benefits associated with green growth and costs incurred without its adoption

Embracing Green growth will likely lead to:	**Not adopting green growth will likely lead to:**
Greater economic returns	Decline in crop yields
More affordable energy access	Lower productivity
Increased employment, poverty reduction and less scarcity of public goods.	Reduced GDP
Decrease in health-related costs	Increase in malaria and premature death form indoor air pollution

Source: AfDB, 2016

In addition, available evidence suggests that the benefit from environmental regulations outweighs the cost. For example, while EPA regulations in the United States impose about 60–70 percent of all federal regulatory costs, they generate about three-quarters of the benefits of all regulations. Thus, EPA regulations result in slightly higher benefit-cost ratios, on average, than other federal regulations (Table 1.5). A United Nations Environmental Programme (UNEP) model shows a "green" economic development path has slightly negative effects on GDP in the short term, but brings increases in GDP per capita as well as substantial environmental improvements in the medium and longer terms. Investments in a green GDP path are job-intensive and bring particular benefits to the world's poorest people who often depend on natural resources for their livelihoods (Figure 1.18).

Fortunately, many countries, including developing ones, have begun to embrace green growth policies. In the mid-2000s only about one-quarter of global clean energy investments took place in developing countries. But while green energy investments in developed countries increase by a factor of 3.5 over 2004–2015, investments in developing countries increased by a factor of 17 during this period. By 2015, clean energy investments in developing countries exceeded those in developed countries (Figure 1.19).

Table 1.5 Costs and Benefits of Major Federal Regulations, 2004–2014

Agency	Number of rules	Annual benefits (billions)	Annual costs (billions)
Department of Agriculture	3	1.0–1.4	0.9–1.1
Department of Energy	20	16.4–229.0	6.3–9.0
Department of Health and Human Services	15	17.6–35.7	1.1–4.7
Department of Labour	8	9.0–26.0	2.8–6.2
Department of Transportation	25	18.2–32.4	8.1–16.1
Environmental Protection Agency	30	158.5–782.2	36.9–44.4
Joint DOT and EPA	3	33.0–60.0	8.9–16.9
Total	104	253.7–1166.7	65.0–98.4

Source: U.S. OMB, 2015. Also, in Jonathan and Roach, 2017

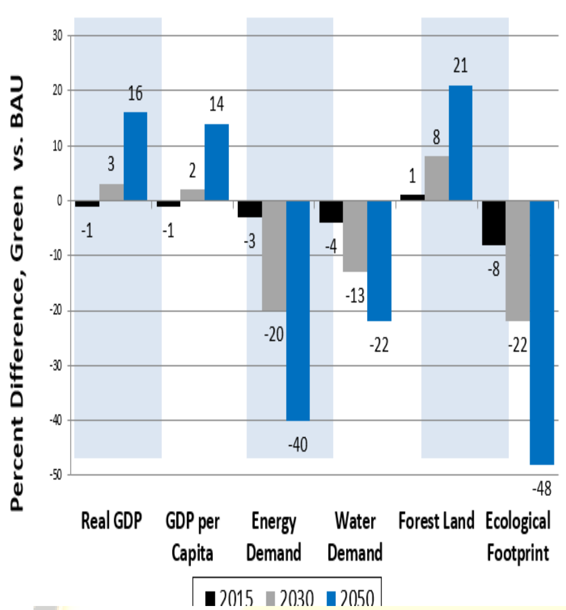


Fig.1.18 Environmental and Economic Projections, Green Economy Scenario versus Business as Usual. ***Source:*** UNEP, 2011a. Also, in Jonathan and Roach, 2017. *Note:* BAU = business as usual; GDP = gross domestic product

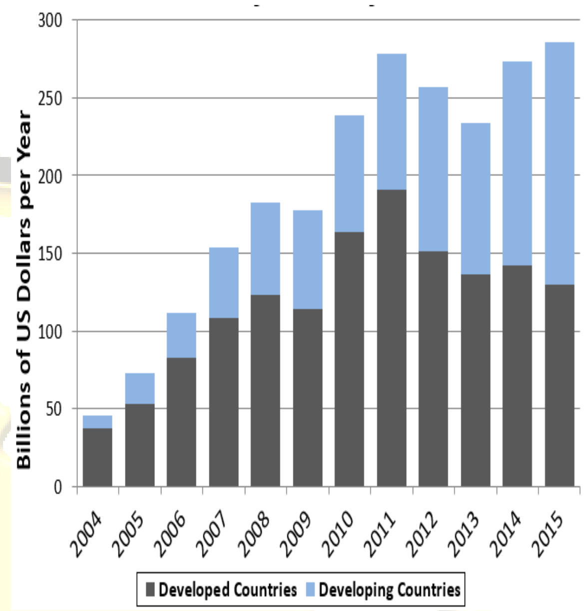


Fig. 1.19: Global Clean Energy Investments, 2004-2015, by Country Classification
Source: UNEP and Bloomberg New Energy Finance, 2016, Also, in Jonathan and Roach, 2017

Summary

- For many years, it was thought that the eradication of poverty required well-designed development programmes that were largely independent of considerations relating to the natural environment. The 1970s witnessed some shift in perspectives. While the pursuit of economic growth and development continues, it was recognized that the maintenance of growth has an important environmental dimension.

- The first major challenge to the claim to growth as a solution to the world's problems questioned the sustainability of growth itself. *The Limits to Growth* posits that environmental limits would cause the collapse of the world economic system in the middle of the twenty-first century.

- However, the paradigm shift in relation to economic growth and environmental sustainability came with the publication of the Brundtland report by the World Commission on Environment and Development (WCED. The Report provided much information about the sustainability problem, setting out the nature of economy–environment interdependence, identifying a number of potential environmental constraints on future economic growth, and arguing that current trends cannot be continued far into the future.

- The concern for sustainability derives from an ethical concern for future generations who should also enjoy the benefit of existing natural and environmental resources, together with an appreciation of the fact that such concern needs to be incorporated into current decision making.

- From the economists' perspective, a fundamental question is under what conditions is constant consumption forever possible, notwithstanding that production uses inputs of a non-renewable resource available only in finite total amount?

- Constant consumption through generations requires that the technology of production allow for some degree of substitution between natural resources and physical capital, that the production technology exhibit constant returns to scale, and that the returns to capital be greater than that of natural resources. Given these conditions, very high levels of output can be produced with very small levels of resource input, and there exists a programme of capital accumulation such that consumption can be maintained constant forever.

- The programme of capital accumulation that guarantees constant consumption over time, given the above conditions, is the Hartwick rule. This requires that at every point in time, the total rents from depletable natural resources should be saved and invested in reproducible capital.

- The issue of substitutability between capital and natural resources is fundamental to the sustainability debate. The possibility (or impossibility) or degree of substitution between environmental (and natural) resources and man-made capital in the production of goods and services leads to two different ways of looking at the concept of sustainability: 'weak sustainability' and 'strong sustainability'.

- Weak sustainability is a definition of sustainable development which allows for imperfect (and in some cases even perfect) substitution between environmental (and natural) resources and man-made capital. Strong sustainability is a definition of sustainable development which allows for no (or in some cases very little) substitution between environmental (and natural) resources and man-made capital.

- Many environmental economists adopt a weak sustainability perception of sustainable development, while ecologists and most ecological economists favour the strong sustainability approach.

- Three concepts stand out clearly in the ecologists' approach to sustainable development: resilience of ecosystems, uncertainty and irreversibility. The approach highlights the belief that our understanding of how natural systems function is very incomplete and our ability to predict the ecological consequences of our behaviour highly imperfect. In addition, damages caused to the environmental resource base are often irreversible.

- Environmental degradation refers to the erosion in the quality, and hence impairment in the integrity and functions, of the environmental resource base (ERB). This can be due to the extraction of renewable natural resources at a faster rate than they are replaced, or the release of waste into the environment at a level that exceeds the assimilative capacity of the ERB.

- Human activities that lead to the degradation of the environment may be driven by poverty. The sheer burden of survival may make household to extract natural resources beyond sustainable levels. In contrast, such activities may also be driven by the crave for affluence. However, these activities take place because of certain institutional failures in society.

- Institutional failures that lead to environmental degradation includes market failure and government or policy failure.

- Market failure occurs when the market fails to allocate resources efficiently because of the absence of certain necessary conditions. Environmental and most natural resources are not like other resources (goods and services) that are traded in the market. They have unique characteristics which makes it impossible for the market system to reckon with them in an efficient manner, if at all.

- Ideally, such failure of the market system would require an effective government intervention. However, in many cases, for one reason or the other, government fails to intervene, or does so in an inefficient or ineffective manner, sometimes making a bad situation worse.

- The IPAT identity provides a useful theoretical framework to analyze the determinants of environmental impact and offers guidance on possible directions for policies to mitigate the negative effect of human activities on the environmental resource base. It shows that policy to control population growth and to alter the technology of production and consumption activities in ways that are environmentally friendly could be helpful.

- The concept of Green growth is used to described a path of economic growth that relies on natural resources in a sustainable manner. It focuses on building a green economy in the context of sustainable development and poverty reduction.

- Promoters of Green growth believe it can secure a strong, stable and sustainable future for all in the face of the wide range of socio-economic and environmental challenges that are pushing the world to a tipping point. Governments that put green growth at the heart of development policy and practice can achieve sustainable economic growth and social stability, safeguard the environment, and conserve resources for future generations.

Discussion/Review Questions and Exercises

1. What ae the factors behind environmental degradation?

2. Identify an environmental degradation problem in your community. What factors do you think are responsible?

3. What are the prospects of using population control policies to address environmental sustainability in Sub-Saharan Africa?

4. Discuss the prospects for green growth in Sub-Saharan Africa.

5. The figure below shows six alternative time paths of consumption labelled to $\boldsymbol{C_1}$ **to** $\boldsymbol{C_6}$. In addition, the horizontal line denoted C^{MIN} represents the level of consumption which is the minimum that society deems as being socially and morally acceptable, while the line, C^{SURV}, represents the biophysical minimum consumption level.

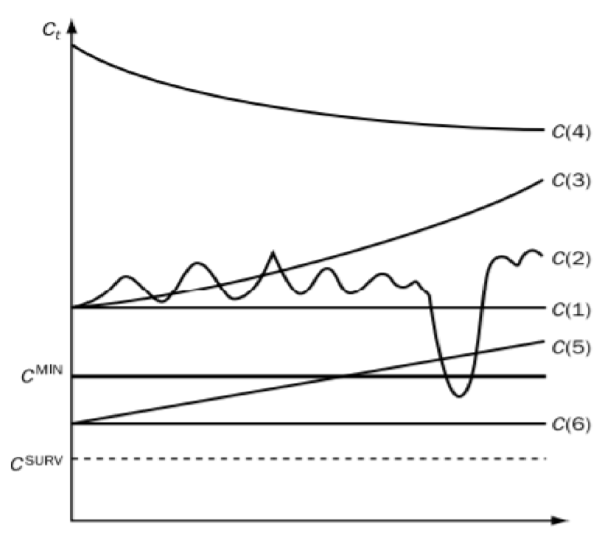


Fig. 1.7. Consumption paths over time

(a) Assume you are a social planner aiming to do the best for society over many generations and you want to rank the six alternative time paths based on three popular concepts of sustainability

- non-declining consumption
- survivable development
- minimum condition,

what will your ranking look like?

(b) What are your observations? Do you agree that an apparently sound ethical principle could in some circumstances lead to outcomes that are not obviously sensible?

(c) Suppose the ethically accepted poverty line corresponds to the horizontal line labeled C^{MIN}. How does this affect your analyses?

(d) Which paths satisfy all of the three criteria above.

(e) Which will you choose among them if your aim is to maximize a conventional utilitarian intertemporal social welfare function? How does your choice compare with C_4 ?

(f) What can you learn from the exercise?

6. Consider three different specifications of the production function represented in (1) to (3) below

$$Q_t = \alpha K_t + \beta R_t \qquad (1)$$

$$Q_t = K_t^{\alpha} R_t^{\beta} \qquad (2)$$

$$Q_t = \min(\alpha K_t, \beta R_t) \qquad (3)$$

(a) Identify each production technology, explain its characteristics and draw its isoqaunt?
(b) Which of the three specifications potentially meets the requirement for constant consumption over an infinite time horizon and why? What additional conditions must be imposed?

7. Attempt carrying out the same analyses as in Table 1.3 for CO_2emission or any other indicator of environmental impact using more recent data from the World Development Indicators (World Bank), Human Development Indicators (UNDP) and/or World Resource Institute.

8. Discuss the role of resilience, uncertainty and irreversibility in the ecological approach to sustainable development.

Materials used for this Lecture

African Development Bank (AfDB) (2016): Transitioning the African continent toward green growth. Green growth in Africa.

Freeman A.M. (1993). **The Measurement of Environmental and Resource Value**s. Washington DC, RFF

Jonathan M. Harris and Brian Roach (2017), **Environmental and Natural Resource Economics** 4th Edition, Routledge.

Pearce D. and R. Turner (2004). **Economics of Natural Resources and the Environment.** Harvester Wheatsheaf. London.

Perman, R., Ma Y., McGilvray J. and Common M. (2003). **Natural Resource and Environmental Economics,** 3rd edition**.** Edinburgh, Longman.

Tietenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics.

Module 2.1 Review of Welfare Economics and Market Outcomes (6 hours)

Learning objectives

This Module expose the reader to the basic foundation of welfare economics and its usefulness in environmental and natural resource economics. At the end of the module, the reader should

- ✓ know the basic principles behind the market system and the conditions required for an ideal system
- ✓ understand why and how the market system achieves efficiency in resource allocation
- ✓ understand the concept of 'Pareto efficiency' and the underlying conditions.
- ✓ Know how the market system relates to the issues of efficiency and fairness
- ✓ understand the concept of Pareto improvements, its role as well as its limitations, in resource allocation
- ✓ know the various types of compensation tests often use to guide policy related to allocation of resources, including natural and environmental resources.
- ✓ understand the concept of market failure and what factors are at the root of such occurrences.
- ✓ gain insights into some complexities surrounding property rights in Sub-Saharan Africa

Outline

2.1.1 Markets and Pareto Efficiency Resource Allocation

2.1.1.1The basic framework

2.1.1.2 Some Qualifications

2.1.2. Market efficiency, Fairness and Welfare theorems

2.1.3 Pareto Improvements and Compensation tests

2.1.4 Market failure

2.1.5 Focus: Property rights in Africa - A case of Land Rights and Land Ownership

Summary

Review/Discussion Questions and Exercises

Materials used for this Lecture

2.1.1 Markets and Pareto Efficiency Resource Allocation

Economists draw upon the basic results of welfare economics in addressing policy questions relating to the environment and management of natural resources. We will be looking at some of these results that are most relevant to environmental policy problems. Efficiency and optimality are the two basic concepts of welfare economics. They both relate to resource allocation. A resource allocation consists of a particular distribution of inputs into the production of a particular combination of outputs and a particular distribution of the outputs among consumers. We can examine the properties of any allocation prevailing at a particular point in time or the allocations that emerge at various points in time as to whether they meet the requirements for efficiency and optimality. Thus, efficiency and optimality can be analyzed from the static or intertemporal viewpoint. Our focus here will be on the static problem: the allocation of inputs across firms and of outputs across individuals at a point in time.

A variety of institutional arrangements might be employed to allocate resources: dictatorship, central planning, free markets. Any of these can, *in principle*, achieve an efficient allocation of resources. Economists are particularly interested in the free-market system of resource allocation. For dictatorship and central planning to achieve allocative efficiency, it is necessary that the dictator or central planner know all of the economy's production and utility functions. This is clearly infeasible, and is one of the reasons that attempts to run economies in these ways have been unsuccessful. Free markets as a way of organizing economic activity do not require any institution or agent have such knowledge; they are *decentralized* information-processing systems. Indeed, the market economy is now the dominant mode of organizing production and consumption in human societies.

Welfare economics theory points to a set of circumstances such that a system of free markets would sustain an efficient allocation of resources. These "institutional arrangements" or conditions are listed in Box 2.1. It is worth noting the importance of the above assumptions because they help us to see why market typically fails in relation to environmental and natural resource. First, if there are goods and services for which markets do not exist, then the market system cannot produce an efficient allocation, as that concept applies to all goods and services that are of interest to any agent (either as a producer or as a consumer). Second, a market in a resource or commodity can only exist where there are private property rights in that resource or commodity. A property right is the right to ownership and use of both tangible (land, buildings, etc.) and intangible (patents, trademarks, copyrights, etc.) property. To be meaningful, such rights must not be subject to revocation without clear transparent legal procedures and proper compensation for the value of the property.[11] Markets can only work where there are private property rights and a justice system to enforce and protect such rights.

[11] In technical terms, a property right is a bundle of characteristics that convey certain powers over a good to the owner of the right. These characteristics concern conditions of appropriability of returns, the ability to divide or

Box 2.1. Institutional Requirements for an Ideal Market System

- Markets exist for all goods and services produced and consumed.
- All markets are perfectly competitive.
- All transactors have perfect information.
- Private property rights are fully assigned in all resources and commodities.
- No externalities exist.
- All goods and services are private goods. That is, there are no public goods.
- All utility and production functions are "well behaved".
- All agents (actors in market system including firms and individuals) are maximizers.

In economics, utility and production functions are 'well behaved' when they have the 'right properties', that is, their technology allows for some degree substitution between goods. In technical terms, it means that the indifference curves (associated with utility functions and isoquants (associated with production functions) are continuous and have the bowed-toward-the-origin shape (they are convex).

Source: Perman et al., 2003. P 116

Third, for there to be private property rights and for them to be enforceable, all goods must be private goods (private individuals have property rights to them) and there must be no externalities in production or consumption. As we shall show in Modules 2.2 and 2.3, where also we shall be looking at the concept of private goods and externalities in details, these characteristics are basically lacking in almost all environmental and natural resources.

Agents in a market system are assumed to behave in a certain way: they always strive to do the best for themselves that they can in the circumstances that they find themselves in. Practically, firms always want to maximize profits from production and individuals want to maximize utility from consumption. In other words, all agents are maximizers. As maximizers, agents behave rationally. This means that they will not make a choice in production or consumption when there is another choice that can yield higher profit or utility. The basic tenet of welfare economics is that an efficient allocation would be the outcome in a market economy populated entirely by maximizers and where all of the institutional arrangements listed above are in place. In addition, perfect competition in markets means that all buyers and sellers are treated the same way by

---

transfer the right, the degree of exclusiveness of the right, and the duration and enforceability of the right. Where a right is exclusive to one person or corporation, a private property right is said to exist[11] (Hartwick and Olewiler, 1986).

the system of allocation; they all act as 'price-takers', believing that the prices that they face cannot be influenced by their own behaviour. No agent acts in the belief that he has any power in the market.

The above conditions which define a market system as an allocative mechanism are very stringent. They actually do not accurately describe any real market economy but what should be in an *ideal* economy. This can then be used in the welfare analysis of actual economies as a benchmark against which to assess performance in regard to efficiency criteria in such economies, and to devise policies to improve the performance. We now review the basic framework that defines the efficiency conditions in a competitive market system.

2.1.1.1 The basic framework

Given the conditions in Box 2.1, assume for simplicity, that the economy consists of two persons and two goods are produced. The production of each good uses two inputs, each of which is available in a fixed quantity. Let U denote an individual's total utility, which depends only on the quantities of the two goods that he or she consumes, so that Individual *h's* utility function can be written as

$$U^h(X_{h1},X_{h2},Z_h) \qquad (2.1)$$

where X_{hi} is h's consumption of commodity *i* and Z_h is h's supply of an input.

The individual's initial endowments of the inputs are denoted $\bar{Z}_h$. We assume non-satiation, so that the marginal utility of commodity I for individual h is always positive, i.e. U_i^h>0 (I =1.2 and h = 1,2). And the marginal utility of individual h of supplying more of labour input Z is always negative i.e. $U_z^h < 0$.

Firm I produce the total output X_i of commodity *i* according to the production function;

$$X_i = f^i(Z_{i1}, Z_{i2}), i = 1,2 \qquad (2.2)$$

Where Z_{ih} is the quantity of input type supplied by individual *h* used in the production of good *i*. The marginal product of input *Z* in producing good *i* is positive; $f_h^i = \frac{\partial f^i}{\partial Z_{ih}} > 0$.

Total consumption of good *i* by individuals does not exceed its total output so that

$$X_i \geq \sum_{h=1}^{2} X_{hi\ldots} \qquad (2.3)$$

Similarly, total use of input Z by the firms does not exceed the supply by individual h;

$$Z_h \geq \sum_{i=1}^{2} Z_{ih,\ \ h=1,2} \qquad (2.4)$$

Also, the supply of input Z by individual h, does not exceed his/her total endowment;

$$\bar{Z}_h \geq Z_h \qquad (2.5)$$

We restrict our attention to non-corner allocations in which both individuals consume both goods and supply inputs, and both firms use both types of inputs.

An allocation of resources is said to be **Pareto efficient** if it is not possible to make one or more persons better off without making at least one other person worse off. Conversely, an allocation is inefficient if it is possible to improve someone's position without worsening the position of anyone else. A gain by one or more persons without anyone else suffering is known as a **Pareto improvement**. When all such gains have been made, the resulting allocation is sometimes referred to as **Pareto optimal**, or Pareto efficient. Allocative efficiency implies a state in which there is no possibility of Pareto improvements. This requires simultaneous fulfillment of three efficiency conditions: efficiency in consumption, efficiency in input supply, efficiency in input mix, and efficiency in output or product mix.[12]

These conditions are derived from the solution to the problem:

$$Max\ U^1\left(X_{1,1}, X_{1,2}, Z_1\right) s.t. U^2\left(X_{2,1}, X_{22}, Z_2\right) \geq \bar{U}^2 \qquad (2.6)$$

Setting up the Lagrangian, we have

$$L = U^1(X_{11}, X_{12}, Z_1) + \lambda(U^2(X_{21}, X_{22}, Z_2) - \bar{U}^2) + \sum \rho_i(X_1 - \sum X_{hi}) + \sum \omega_h(Z_h - \sum Z_{ih}) + \sum \mu_i\left[f^i(Z_{i1}, Z_{12}) - X_i\right] \qquad (2.7)$$

Obtaining first order conditions with respect to:$X_{1i},, Z_1, Z_2, Z_{ih} and\ X_i$ we obtain the following;

- **Efficiency in consumption**

Obtaining partial derivatives with respect to (w.r.t.) X_{1i}, X_{2i} and rearranging, we obtain;

$$MRS_{21}^1 = \frac{U_1^1}{U_2^1} = \frac{\rho_1}{\rho_2} = \frac{U_1^2}{U_2^2} = MRS_{21}^2 \qquad (2.8)$$

Optimality/efficient allocation of total output X_1 and X_2 between two consumers, requires that consumers are given a bundle which equalizes their marginal rates of substitution (MRS) between the two goods. MRS_{21}^1 measures the rate at which individual 1 is willing to exchange/substitute commodity 1 for 2. It also means individual 1's marginal valuation of commodity 1 in terms of good 2. Equality of the MRS is a necessary condition for a Pareto efficient allocation of goods among consumers.

[12] Some authors resolve these conditions into three: **efficiency in consumption**, **efficiency in production**, and **efficiency in product-mix**.

- **Efficient input supply**

Differentiating w.r.t. $X_{1i,}X_{2i\,,}Z_1$ and Z_2 and combining the respective conditions on individual h's consumption of good *i* and supply of input Z_h yields.

$$\frac{-U_2^h}{U_1^h} = \frac{\omega_h}{\rho_i} \qquad (2.9)$$

which implies that

$$MRS_{iz}^h = -\frac{U_z^h}{U_i^h} = \frac{\omega_h}{\rho_i} = f_h^i\,, h, i, = 1{,}2 \qquad (2.10)$$

The left-hand side of equation (2.10) is individual h's MRS between his input supply and consumption of commodity *i*. It is the rate at which h must be compensated by being given more of commodity *i* if he is to increase his supply of labour input Z_h by one unit. The right-hand side is the marginal product of labour input Z_h in the production of good *i*. Efficiency requires that the additional output produced by an extra unit of Z_h is just equal to the marginal cost in terms of good *i* of Z_h to individual *h*.

- **Efficiency in input use**

Differentiating (2.7) w.r.t Z_{ih} for i,h=1,2 and rearranging, gives

$$MRTS_{21}^1 = \frac{f_1^1}{f_2^1} = \frac{\omega_1}{\omega_2} = \frac{f_1^2}{f_2^2} = MRTS_{21}^2 \qquad (2.11)$$

$MRTS_{21}^i$ is the rate at which input 1 can be substituted for input 2 without changing the output of good *i*.

- **Efficiency in output mix**

This requires the mix of outputs X_1 and X_2 produced by the firms to be efficient with the supply of inputs held constant.

Differentiating (2.7) w.r.t X_i, we get $\rho_{i=}\,\mu_i$, and using equation (2.8), we get;

$$f_h^1 = \frac{\omega_h}{\rho_1},\; f_h^2 = \frac{\omega_h}{\rho_2}, h = 1{,}2 \qquad (2.12)$$

Dividing the conditions on good 2 by conditions on good 1, gives the efficient output mix conditions;

$$MRT_{21}=\frac{f_1^2}{f_1^1}=\frac{f_2^2}{f_2^1}=\frac{\rho_1}{\rho_2}=\frac{U_1^1}{U_2^1}=\frac{U_1^2}{U_2^2}=MRS_{21} \qquad (2.13)$$

The second term in (2.13) is the ratio of marginal product of input 1 in the production of goods 2 and 1. This ratio is the marginal rate of transformation between the two goods, i.e. the rate at which the output of good 2 falls as the output of good 1 is increased as input 1 is transferred from firm 2 to firm 1 with the total input supply held constant. Optimal output mix, requires that the marginal rate of transformation MRT_{21} between the two goods is equal to the consumer's marginal rate of substitution MRS_{21} between the goods. In other words,

$$MRTS_{21}^1 = MRS_{21}^1 = MRS_{21}^2 = MRTS_{21}^2 \qquad (2.14)$$

These conditions are illustrated in Figure 2.1, To fix ideas we denote the two individuals as A and B, the two goods as X and Y, and the two inputs as Capital (K) and labour (L). Panel a illustrates the consumption efficiency condition. It shows a typical Edgeworth box. An efficient allocation of X and Y as between *A* and *B* is uniquely identified by point *b* in the box. That is the only point where equilibrium can be reached, such that there is no longer scope for a Pareto improvement. Notice that if this condition (which we now write as the equality of A and B's marginal rate of utility substitution between the two goods, $MRUS^A$ = $MRUS^B$) were not satisfied, it would be possible to rearrange the allocation as between *A* and *B* of whatever is being produced so as to make one better off without making the other worse off. In other words, at *b*, the slopes of the indifference curves of *A* and *B* are equal. Think of what happens at points where the rates are not equal!

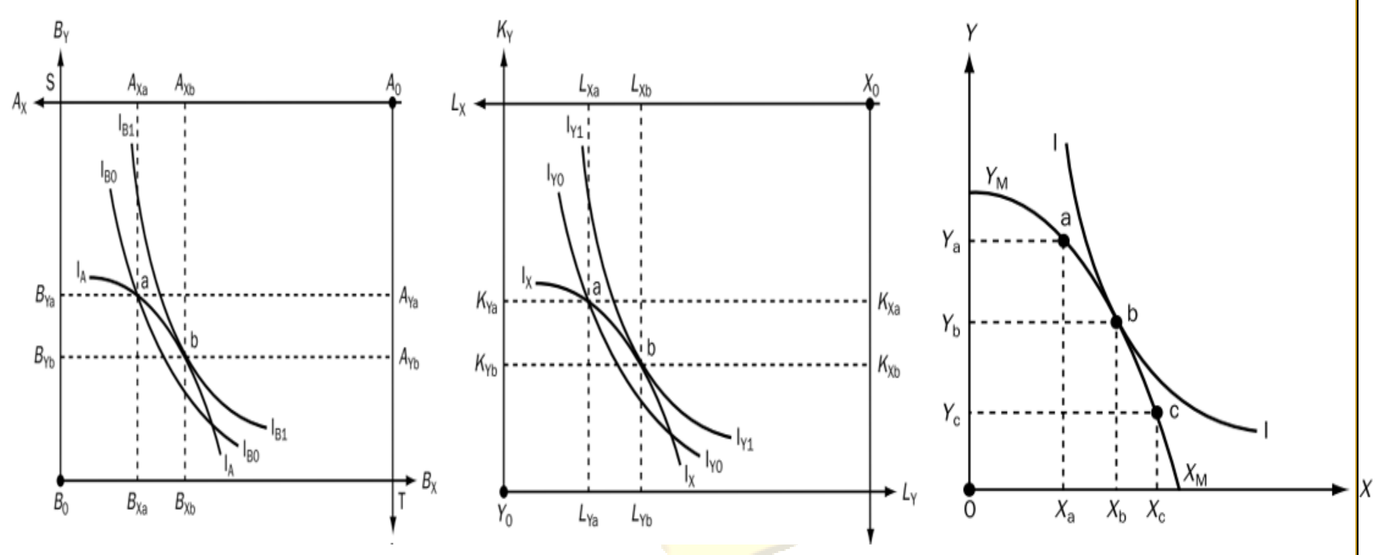


Fig. 2.1a. Efficiency in consumption

Fig. 2.1b. Efficiency in production

Fig. 2.1c. Efficiency in product-mix

In Fig 2.1b (constructed in a way similar to Figure 2.1a), **production efficiency** (which implies efficiency in input supply and input use) is achieved at point *b* where the slopes of the isoquants in each line of production are equal (the marginal rates of technical substitution are equal). If these rates are not equal, then clearly it would be possible to reallocate inputs as between the two lines of production so as to produce more of one commodity without producing any less of the other.

In Fig 2.1c, *A* and *B's* indifference curve are represented in a common indifference curve, since they have the same slopes at the point of equilibrium as implied in equation (2. 8). The production-possibilities frontier, $Y_M X_M$ shows the output combinations that the economy could produce using all of its available resources. Its slope is the MRT between X and Y. Points like *b*, where the slopes of the indifference curve and the production possibility frontier are equal, corresponds to an efficient product mix output levels for X and Y. At that point, the utility of the representative individual is maximized, given the resources available to the economy and the terms on which they can be used to produce commodities. At a combination of X and Y where (2.9) does not hold, some adjustment in the levels of X and Y is possible which would make the representative individual better off.

These results readily apply to economies with many inputs, many goods and many individuals. The only difference will be that the three efficiency conditions will have to hold for each possible pairwise comparison that one could make. In the next subsection we will make two important observations in relation to these results and with respect to an efficient allocation in a competitive market.

2.1.1.2 Some Qualifications

- **An efficient allocation of resources is not unique**

For an economy with given quantities of available resources, production functions and utility functions, there will be many efficient allocations of resources. In other words, the criterion of efficiency in allocation does not serve to identify a particular allocation. Given various possible indifference curves for both A and B and various possible initial endowment of resources, there are many possible points where the slopes of their indifference curves would be the same. Thus, for given available quantities of X and Y, there are an indefinitely large number of allocations as between A and B that satisfy $MRUS^A$ = $MRUS^B$. This point is illustrated in Figure 2.2a. The same thing applies in production. There are many possible combinations of X and Y output levels that satisfy $MRTS^X$ = $MRTS^Y$. In addition, for any particular combination of X and Y there are many allocations as between A and B that are consistent with allocative efficiency.

Fig.2.2b puts these thoughts together in one diagram. The utility-possibility frontier ($U^A U^B$) shows the various possible distribution of utility between A and B as reflected in the quantities of society's goods, X and Y, each one is able to consume. It is the locus of all possible combinations of U^A and U^B that correspond to efficiency in allocation.

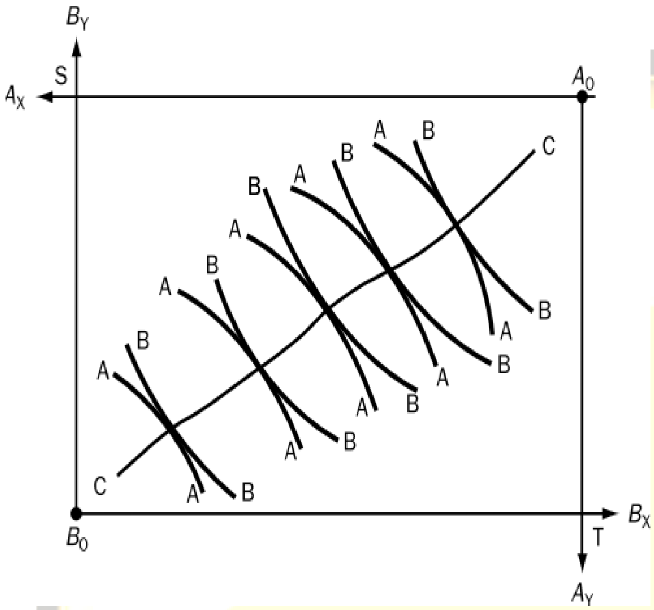


Fig. 2.2a Set of consumption-efficient allocations

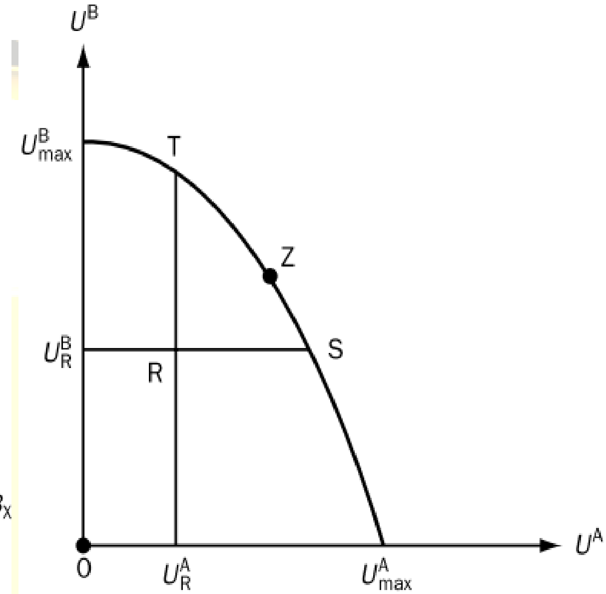


Fig. 2.2b. The utility-possibility frontier

At any point inside the utility possibility frontier, such as R, there are possible reallocations that could mean higher utility for both A and B (point Z for example). Thus, the move from R to Z would be a Pareto improvement. So would be a move from R to T, or to S, or to any point along the frontier between T and S. On the utility possibility frontier, there is no scope for a Pareto improvement.

- **An efficient allocation is not necessarily fair**

Points on the utility-possibility frontier implies different allocations of goods between A and B, and hence, different utility levels. For example, at point *a* on Figure 2.3a below, we can reasonably assume that B has more of society's good, and hence, enjoy a higher level of utility than A. The situation is different at point *c,* where A has more of society's goods, and hence, enjoy a higher level of utility than B. Yet both points are Pareto efficient. One may argue that an allocation, such as *a* or *c* is not fair, in that it gives a disproportionately larger amount of society's goods to an individual at the expense of the other.

Is it possible, using the information available, to say which of the points on the utility-possibility frontier (a, b or c) is best from the point of view of society? The answer is that it is not possible, for the simple reason that the criterion of economic efficiency does not provide any basis for making interpersonal comparisons. Efficiency does not give us a criterion for judging which allocation is best from a social point of view. Choosing a point along the utility possibility frontier (for example, moving from point *a* to *c* in Figure 2.3) is about making moves that must involve making one individual worse off in order to make the other better off, and the efficiency criteria do not cover such choices.

Making choices between efficient allocations is an ethical decision requiring the use of a social welfare function (SWF). A SWF is like a utility function for the society; it shows how the welfare of society is affected by changes in the utility of individuals. The structure depends on the weight society attach to the utility level of an individual (or group) relative to another. A SWF can be used to rank alternative allocations. Assume, for example, a SWF for the two-person economy of the general form

$$W = W(U^A, U^B) \qquad (2.15)$$

with $W_A = \partial W/\partial U^A > 0$ and $W_B = \partial W/\partial U^B > 0$

In other words, social welfare is non-decreasing in U^A and U^B. In Fig 2. 3, equilibrium at point b where the slope of the social welfare indifference curve $[\frac{W_A}{W_B}]$ is the same as the slope of the utility possibility frontier $[\frac{U^B{}_X}{U^A{}_X}$ or $\frac{U^B{}_Y}{U^A{}_Y}]$. At this point, all of the necessary conditions for efficiency (equations 2.8, 2.10, 2.11, and 2.14) are satisfied. In addition to these efficiency conditions, we have

$$\frac{W_A}{W_B} = \frac{U^B{}_X}{U^A{}_X} = \frac{U^B{}_Y}{U^A{}_Y} \qquad (2.16)$$

Equation (2.16) implies that the efficiency conditions are not only satisfied but social welfare is maximized. This is shown by the equality of the slopes of the indifference curve and the utility-possibility frontier, so that it is not possible to increase social welfare by transferring goods, and hence utility, between persons.

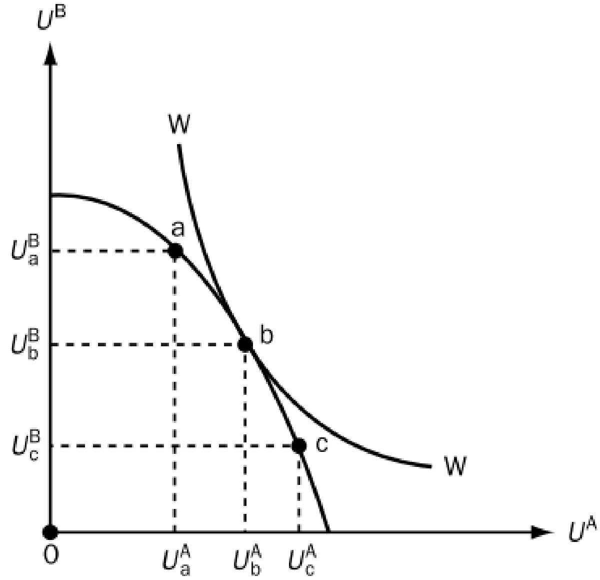


Fig.2.3a. Efficiency and social welfare

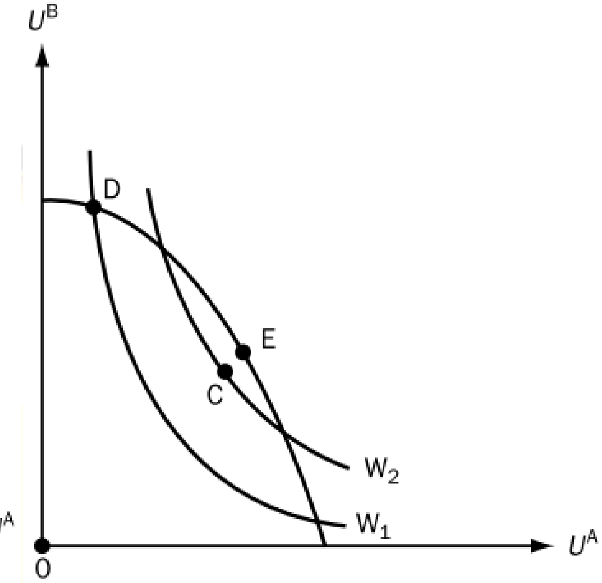


Fig.2.3b. Maximizing social welfare

Again, it is important to note that while allocative efficiency is a necessary condition for Pareto optimality, it is not generally true that moving from an allocation that is not efficient to one that is efficient must represent a welfare improvement. Such a move might result in a lower level of social welfare. For example, consider, Figure 2.3b, the allocation at *D* is not Pareto efficient, while that at D is. However, the allocation at C gives a higher level of social welfare than does the allocation at D. Whenever there is an inefficient allocation, there is always some other allocation which is both efficient and superior in welfare terms. For example, compare points C and E in Figure 3.2b. The latter is allocative efficient while C is not, and E is on a higher social welfare indifference curve. Thus, the move from C to E is a Pareto improvement in which both individuals A and B gain, and hence, involves higher social welfare. On the other hand, going from C to D replaces an inefficient allocation with an efficient one, but the change is not a Pareto improvement: B gains, but A suffers. It also involves a reduction in social welfare (**make sure you understand the logic of these arguments!).**

The specific kind of SWF employed here means that any change which is a Pareto improvement must increase social welfare (since social welfare is non-decreasing in U^A and U^B. But, it is not necessarily so for all types of SWFs. Consider a SWF that places emphasis on increasing utility for B (the poor, for example) and reducing utility of A (the rich, for example). If a change in allocation raises the utility of A without reducing the utility of B, then such a change, though Pareto-improving, will reduce, rather than increase, social welfare, because it does not satisfy the underlying ethics of the SWF.

2.1.2. Market efficiency, Fairness and Welfare theorems

A system of free markets will produce an optimal allocation according to any particular social welfare function chosen by society. In the absence of a SWF which defines society's definition of what is fair and help chose an efficient allocation that meets the criterion, it cannot be claimed that markets will achieve what might generally or reasonably be regarded as fair allocations. The attainment of efficiency is simply the exhaustion of the possibilities for voluntary mutually beneficial exchange. Clearly, if the agents involved have different initial endowment, we would not expect the voluntary trade process to lead to equal endowments after trade. Voluntary trade on the basis of self-interest is not going to equalize wealth. As the initial endowments of agents vary, so will the positions reached when all voluntary exchanges have been made.

These thoughts are summed up in the two theorems which constitute the formal foundations for modern welfare economics and its application to policy analysis in market economies (the theorems take it that all agents are maximizers, and that the ideal institutional conditions for an efficient market system holds).

- The **first welfare theorem** states that a competitive market equilibrium is an efficient allocation. In other words, competitive markets lead to an allocation that exhausts all possible gains from trade and that cannot be improved upon given agents' initial endowments. To change the allocation would imply making one or more persons better off at the expense of another or some others.

- The **second welfare theorem** states that to every efficient allocation, there is corresponding competitive market equilibrium based on a particular distribution of initial endowments. In other words, the efficient allocation realized by a competitive equilibrium is conditioned on the distribution of initial endowments.

The implication of these two welfare theorems is that any efficient allocation can be realized as a competitive market equilibrium by changing the initial distribution of endowments through appropriate set of lump-sum taxes and transfers to individual agents. If initial endowments are such that the resulting efficient allocation is considered inequitable (unfair), altering them by lump-sum taxes and transfers will produce another efficient allocation. And if the taxes/transfers redistribute from the better to the worse off, the new efficient allocation will be more equitable.

These two theorems influence the way that economists approach policy analysis in an economy mainly run by markets. They suggest that there are essentially two separable dimensions to the economic problem: the problems of efficiency and equity. Society can, via government, take a view on equity and achieve what it wants there by a system of redistributive taxes and payments, and then leave it to markets to achieve efficiency in allocation given the distribution of endowments after the tax/transfer. Notwithstanding that the conditions under which the two welfare theorems hold are not fully satisfied in any actual economy, the overwhelming majority of economists do approach practical policy analysis on the basis that the problems of efficiency and equity can be dealt with independently.

Because fairness considerations involve value judgments (they are matters of ethics), most economists prefer to focus on efficiency (the positive side of analysis) than address the equity dimension of resource allocation. This is more so because it is impossible to agree on what society's SWF should like. The relative weights to be assigned to the utilities of different individuals in a SWF are an ethical matter and economists would prefer to avoid specifying one if they can!

2.1.3 Pareto Improvements and Compensation tests

The Pareto improvement criterion helps avoid the need to refer to the SWF to decide on whether or not to recommend a particular reallocation. However, it has two limitations. First, it still does not answer the question of what the unique allocation should be (convince yourself that this is true by looking again at Figure 2.3b). Secondly, in considering policy issues there will be very few proposed reallocations that do not involve some individuals gaining and some losing, so the Pareto improvement criterion is not very useful practically. Only rarely will the economist be asked for advice about a reallocation that improves somebody's lot without damaging somebody else's. Most reallocations that require analysis involve winners and losers and are, therefore, outside of the terms of the Pareto improvement criterion.

Welfare economists have tried to devise ways of comparing allocations where there are winners and losers which do not require the use of a SWF. These are called **compensation tests**. Three of these which are largely used are the Kaldor potential compensation, Hicks potential compensation test, and the Kaldor– Hicks–Scitovsky potential compensation

- **Kaldor potential compensation test** (henceforth KCT): named after its originator, Nicholas Kaldor; it says that an allocation is superior to another if the winner from the reallocation can compensate the loser and still be better off. In considering a policy choice that changes allocations between agents, the policy move is beneficial if the winner(s) based on the outcome of the policy can potentially compensate the loser and still be better off. In other words, the overall gain more than compensate for the overall loss.

- **Hicks potential compensation test** (henceforth HKT): named after Hicks J.R.; it says that an allocation is superior to another if the loser could compensate the winner and be no worse off. In considering a policy choice that changes allocations between agents, the policy move is rejected if the loser can potentially compensate the winner for forgoing the move and be no worse off than if the move took place. In this case, the overall loss with the move is higher than the loss without the move.

- **Kaldor– Hicks–Scitovsky potential compensation** test (henceforth K-H-S CT): named after Tibor Scitovsky; it says that a reallocation is desirable if (i) the winner(s) could compensate the loser(s) and still be better off and (ii) the losers could not compensate the winners for the reallocation not occurring and still be as well off as they would have been if it did occur.

Note that these tests do not assume that the actual compensation must take place (that is a matter for the government to implement if it is judged desirable) but takes a position from the viewpoint of overall benefits (hence the use of the words, "potential"!). If the compensations do indeed take place, then we would be looking at situations associated with Pareto improvements. The example that follow illustrates the application of the three tests.

Suppose there are two allocations, denoted by 1 and 2, to be compared. The allocations are presented in Table 2.1a (we assume that both individuals have utility functions such that U= XY). Form the table, it is obvious that moving from allocation 1 to allocation 2 involves one individual gaining and the other losing. A is the winner for a move from 1 to 2 while B loses from such a move. The Kaldor compensation test suggests that allocation 2 is superior to allocation 1. However, unlike the KCT, the HCT suggests that allocation 1 is superior to allocation 2.

The values in the above example have been carefully designed to show what problem can sometimes arise with the use of the KCT and the inconsistency that may be observed from comparing the results from the KCT and the HCT. The KCT may approve a move from one allocation to the other, while at the same time approving a move from the latter allocation back to the former (You are to show that this is indeed the case for the KCT in the example in Table 2.1a).

Table 2.1a. Compensation tests may display inconsistency

	Allocation 1				**Allocation 2**		
	X	Y	U		X	Y	U
A	10	5	50		20	5	100
B	5	20	100		5	10	50

The K-H-S CT helps to overcome these issues (that is the possible ambiguity with the KCT and the inconsistency that may arise from the predictions of the KCT and HCT). It is based on the argument that for an unambiguous result from a (potential) compensation test, it is necessary to use both the Kaldor and the Hicks criteria. Applying the K-H-S CT on the case at hand, it is obvious that the move from 1 to 2 cannot be approved (why?).

It is important to note that just like the Pareto efficiency criterion, compensation tests may not necessarily lead to outcomes that are considered fair. Consider the example in Table 2.1b below.

Table 2.1b Compensation tests may not produce fairness

	Allocation 1				Allocation 2		
	X	Y	U		X	Y	U
A	10	5	50		10	4	40
B	5	20	100		15	16	240

In a move from allocation 1 to 2, A is the loser and B is the winner. According to both the KCT and the HCT, allocation 2 is superior to 1. The reallocation also passes the K-H-S CT (be sure that you can explain why). However, note that A is the poorer of the two individuals in the first instance, and that the reallocation sanctioned by the compensation tests makes A even worse off, and B better off. This raises questions of equity or fairness. In sanctioning such a reallocation, the compensation test is either saying that fairness is irrelevant or there is an implicit SWF such that the reallocation is consistent with the notion of fairness that it embodies.

Compensation tests inform much of the application of welfare economics to environmental problems because they avoid the issue of equity by treating winners and losers equally. No account is taken of the fairness of the distribution of well-being. When considering some policy intended to lead to a reallocation, the job of the economic analyst is to ascertain whether the gains exceed the losses. If they do, the policy can be recommended on efficiency grounds, since it is known that the beneficiaries could compensate the losers and still be better off. It is a separate matter, for government, to decide whether compensation should actually occur, and to arrange for it to occur if it is thought desirable.

In examining the concepts of efficiency and optimality, we have used a general equilibrium approach (this looks at all sectors of the economy simultaneously). Much (although by no means all) of the huge body of theory that makes up resource and environmental economics analysis are based on this approach. However, many practical applications use a partial equilibrium framework that enables them to focus on one good while ignoring all others. The results for efficiency are basically the same, except that it is now applied to just one good.

2.1.4 Market failure

Whenever the market system is not able to achieve allocative efficiency because of the absence of one or more of the institutional arrangements required for an ideal system of resource allocation (listed in Box 2.1), we say there is a **market failure**. It is important to note that market failure has nothing to do with the failure to achieve desirable (ethical) goals such as equity (fairness), or even sustainability. Rather, it is concerned with the inability of the market to achieve its avowed goal of allocative efficiency in resource allocation because of the absence of one or more of the conditions needed for this to be achieved.

Market failure creates inefficiencies and generate losses for society. The loss to society, often referred to as a **deadweight loss**, can be analyzed by comparing the efficiency outcome of a competitive market system with the prevailing situation under market failure. A good illustration is to compare outcomes under perfectly competitive markets with what obtains under imperfect competition, such as monopoly (a situation where participation in the supply side of the market for a resource is restricted to a single agent due to certain factors).

Figure 2.4a and 2.4b provides such an analysis. In Figure 2.4a, the perfectly competitive market long-run equilibrium is compared with that of monopoly in terms of output levels and prices: y_m

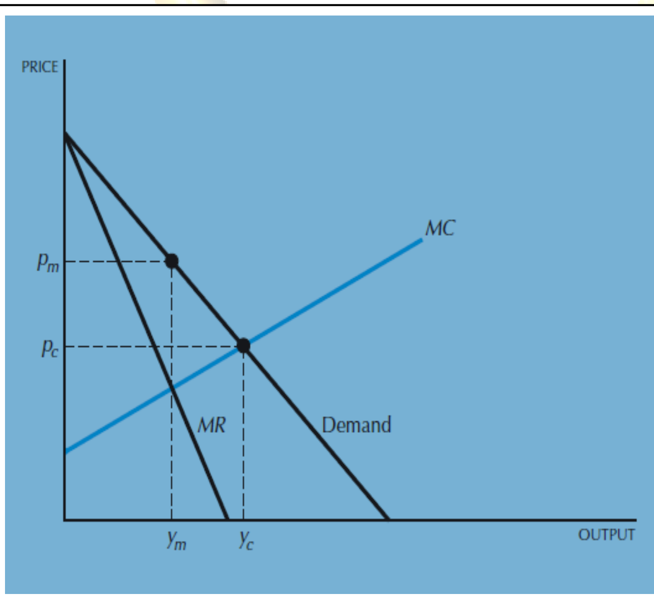


Fig. 2.4a Pareto inefficiency under monopoly. **Source:** Varian, 2010. p.446

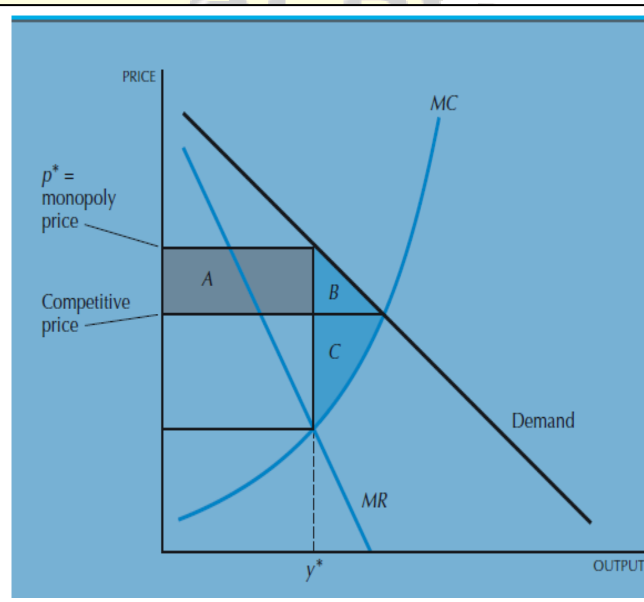


Fig. 2.4b Deadweight loss of monopoly inefficiency. **Source:** Varian, 2010. p.448

is the equilibrium output level of the monopolist and y_c is the long-run equilibrium output level of the competitive industry, P_m is the monopoly equilibrium price and P_c is the competitive price. It is easy to show that the monopoly output level, y_m is not Pareto efficient. Since the price of the product p(y) is greater than the marginal cost of production, MC(y) for all the output levels

between y_m and y_c, there is a whole range of output where people are willing to pay more for a unit of output than it costs to produce it. Pareto improvement is possible within this range of output level. Efficiency is achieved when the willingness to pay for an extra unit of output just equals the cost of producing this extra unit.

How can we measure the loss in social welfare associated with this inefficiency? We are can look of the reduction in the gains from trade. The loss to consumers from having to pay the monopoly price, P_m rather than the price obtains if the market were functioning efficiently, that is the competitive price, P_c is captured by the change in consumers' surplus. In contrast, the gain in profits to the firm from charging P_m rather than P_c is reflected in the change in producer's surplus. From another standpoint, the change in producer's surplus measures how much the owners of the firm would be willing to pay to get the higher price under monopoly, and the change in consumers' surplus measures how much the consumers would have to be paid to compensate them for the higher price. If we treat both cases as symmetric, the difference between these two numbers should give a sensible measure of the net benefit or cost of the monopoly. This is shown in Figure 2.4b, which illustrates the changes in the producer's and consumers' surplus from a movement from monopolistic to the competitive output.

At the competitive output level, price is lower but output is higher. The monopolist's surplus goes down by A due to the lower price on the units he was already selling but goes up by C due to the profits on the extra units it is now selling. The consumers' surplus goes up by A, since the consumers are now getting all the units they were buying before at a cheaper price; and it goes up by B, since they get some surplus on the extra units that are being sold. The area A is just a transfer from the monopolist to the consumer; one side of the market is made better off and one side is made worse off, but the total surplus doesn't change. The area B + C represents a true increase in surplus—this area measures the value that the consumers and the producers place on the extra output that has been produced.

The area B+C is known as the deadweight loss due to the inefficiencies associated with the presence of imperfect competition, in this case, monopoly. It provides insight into how much worse off people are paying the monopoly price than paying the competitive price. The deadweight loss due to monopoly, like the deadweight loss due to a tax, measures the value of the lost output by valuing each unit of lost output at the price that people are willing to pay for that unit.

The fundamental factors at the root of market failure are **the absence of property rights**, which makes it impossible for agents to engage in any mutually beneficial transaction over some resources, **high transactions costs**, which makes bargaining too costly and hence preclude the possibility of beneficial trades, and the lack of or **imperfection in information.**

As we shall show in the next two modules (Modules 2.2 and 2.3), most natural and environmental resources do not have the characteristics that are required for goods that are to be transacted in an ideal market system. They are often not pure private, goods; thus, private property rights are often lacking. Also, in most cases, there are externalities associated with their use. In addition, we do not have perfect information concerning them. Thus, market failure naturally occurs in the case of environmental goods. An understanding of the underlying factors that leads to market failure in relation to natural and environmental resources is necessary for effective policy intervention.

2.1.5. Focus: Property rights in Africa - A case of Land Rights and Land Ownership

The question of who owns the world's lands and natural resources is a major source of contestation around the globe, affecting prospects for rural economic development, human rights and dignity, cultural survival, political stability, conservation of the environment, and efforts to combat climate change. To inform advocacy and action on community land rights, Rights and Resource (RRI) has published *Who Owns the World's Land? A global baseline of formally recognized indigenous & community land rights* ("the global baseline"), which identifies the amount of land national governments have formally recognized as owned or controlled by Indigenous Peoples and local communities across 64 countries constituting 82 percent of global land area. The report focuses on community-based tenure regimes, which include any system where formal rights to own or manage land or terrestrial resources are held at the community level, including lands held under customary tenure regimes. The global baseline study included 19 Sub-Saharan Africa countries, and below are the key findings regarding the 19 SSA countries.

- Only 13 percent of the total land of the countries studied in Sub-Saharan Africa is owned or controlled by Indigenous Peoples and local communities, compared with 18 percent globally.

- All 19 countries studied in Sub-Saharan Africa have enacted laws to enable the recognition of community ownership or control of land; however, implementation of these laws is often weak or non-existent.

- In eight of the 19 countries, indigenous peoples and local communities own or control less than 1percent of the country's land area, including both agricultural and forested lands.

2.1.5.1 Current Reforms Geared toward Alternative Land Rights

Box 2.2 provides a highlight of land reforms in Liberia. Most land reform agendas are either driven by efficiency or equity objectives, or both. Understanding the dynamics associated with different types of land rights is crucial to any land reform efforts. *Customary land rights* offer access to land and security of tenure to many poor households. However, because they provide limited access to formal credit and input markets and to sales outside the group, opportunities for productive exchange and access to credit are limited. In a shift toward titling, *registered customary land rights* boost the possibilities for land transactions in both formal and informal markets and for access to formal credit institutions. Once advocated as the optimal solution for granting tenure security and land access, *land titles* often involve high transaction costs. While titling may benefit farmers of high-value commodities, it is usually impractical for poor resource farmers. In addition, the links between land titling and tenure security, credit availability, and investments have not been well established in Africa.

Redistributive land rights aim to reduce inequalities in landowning emanating from previous imbalances. As confiscated land initially becomes state land, redistributive rights provide limited opportunities for sale and rental. In some areas, redistributive rights have proved to fulfill both efficiency and equity objectives by providing more land access to women and younger, more productive households. Recent reforms in southern Africa encouraging *market-based land polices* were aimed at facilitating equity and efficiency while avoiding the negative effects of land confiscation. Unfortunately, there is evidence that white farmers acquired more land under these policies than disadvantaged black farmers.

Subsidized market-based reforms provide land right holders with financial support to pay for part of the cost of acquiring land. If well targeted, such programs could benefit women and poor people

Box 2.2: Land Tenure Reforms in Liberia

In 2008, Liberia established the National Land Commission and began the process of working towards land reform as part of its broader efforts to achieve lasting peace and stability. In 2013 Liberia enacted a national Land Rights Policy recognizing customary tenure. Now, in 2015, the Liberian legislature considered a draft Land Rights Act that would recognize customary tenure as a matter of law without requiring titling, using an approach similar to that of Mozambique, Tanzania, Uganda, and Zambia. This would be a significant development because experts estimate that 71 percent of Liberia's land area is held under customary tenure. A significant concern, however, is the status of concessions, as one draft provision states that existing concessions on community land will be honoured. Therefore, communities will only be able to exercise their customary rights after concessions expire.

This is particularly problematic because estimates indicate that the government has issued concessions over approximately 75 percent of Liberia's land area. Even to the extent that legislative reforms protect customary tenure, they will not eliminate tenure insecurity without robust procedural requirements and the administrative capacity to identify and respect customary land holdings. Government agencies must be willing and able to coordinate their actions in order to avoid infringements on customary tenure when issuing concessions, designating protected areas, and taking other actions that could adversely affect communities' property rights. Communities across over 30 percent of Liberia's land area have already obtained titles for their customary lands in the form of Public Land Sale Deeds and Aboriginal Land Grant Deeds. Nevertheless, a lack of technical capacity, interagency coordination, and due process has led to instances where titled community lands have been expropriated without compensation to make way for concessions or protected areas.

Source: Extracted from *Fact Sheet (2015): Who Owns the Land in Africa?* Formal recognition of community-based land rights in Sub-Saharan Africa. https://rightsandresources.org/wp-content/uploads/FactSheet_WhoOwnstheWorldsLand_web2.pdf

Getting Africa on a path of land reform that facilitates efficient, equitable, sustainable use of its land and natural resources requires understanding the intended beneficiaries of land reform programs and their environment. Examining benefits and costs of alternative land rights regimes is vital to a successful land rights reform agenda. Reforms should address all processes, including the capability of governments to undertake the necessary reforms. Some issues to take note of include

- *Reforming customary rights and local institutions:* Simple inexpensive registration programs for customary rights can help make these rights legally recognized. The aim is to improve efficiency by enhancing tenure security and land transfer, and facilitating access to credit and other inputs for rights holders.

- *Improving land rights gained from redistribution:* The majority of current land redistribution programs result in restricted land rights with rights holders denied the right to sell land. Although it is important to ensure that mass land sales do not follow such programs, it is equally important to recognize the need to allow these rights to evolve.

- *Addressing constraints in market-based reform programs:* The valuation system should be reformed by making a distinction between improved and non-improved lands. This would reduce the price of unimproved lands and make them more affordable to governments to acquire for redistribution or to poor farmers who wish to buy land.

- *Decentralized land administration:* Reforms geared toward elected authority for local land administration would increase responsiveness to local interests and needs. The government, however, must provide the broad framework and principles, rules of tenure and access, and ensure transparency and accountability of these institutions.

- *Enhancing mechanisms for land and natural resource dispute resolution:* The effectiveness of any dispute resolution mechanism depends on the ability to anticipate conflict. This calls for early warning and strategic planning. Short-term capacity-building efforts can strengthen institutions that handle refugee repatriation and integration. The ability of internally displaced persons to participate in dispute resolution should also be strengthened. Resettlement programs should be reviewed with the aim of reducing conflicts among different land uses. Programs for civic education aimed at enhancing peaceful coexistence could be useful. Institutional, legal, and policy responses to conflict should aim for comprehensive programs that work through well-established forms of redress. Improved land registration and affordable mechanisms for demarcating boundaries are essential, as are law reforms geared toward recognizing rights of communities to natural resources. Finally, trends toward improved governance are a welcome sign that inefficiency and corruption in land administration will be addressed.

Summary

- Economists draw upon the basic results of welfare economics in addressing policy questions relating to the environment and management of natural resources.

- The basic tenet of welfare economics is that the given certain institutional arrangements, the market system would lead society to an efficient allocation of its resources.

- The market system is used in the welfare analysis of actual economies as a benchmark against which to assess performance in regard to efficiency and to devise policies to improve performance.

- An allocation of resources is said to be Pareto efficient if it is not possible to make one or more persons better off without making at least one other person worse off. A gain by one or more persons without anyone else suffering is known as a Pareto improvement. When all such gains have been made, the resulting allocation is sometimes referred to as Pareto optimal, or Pareto efficient.

- Allocative efficiency implies a state in which there is no possibility of Pareto improvements. This requires simultaneous fulfillment of three efficiency conditions: efficiency in consumption, efficiency in production, and efficiency in product mix.

- For an economy with given quantities of available resources, production functions and utility functions, there will be many efficient allocations of resources. Thus, an efficient allocation is not unique.

- An efficient allocation is also not necessarily fair**.** **T**he criterion of economic efficiency does not provide any basis for making interpersonal comparisons. Making choices between efficient allocations is an ethical decision requiring the use of a social welfare function.

- A system of free markets will produce an optimal allocation according to any particular social welfare function chosen by society. In the absence of a SWF which defines society's definition of what is fair and help chose an efficient allocation that meets the criterion, it cannot be claimed that markets will achieve what might generally or reasonably be regarded as fair allocations.

- Society can achieve any ethical standard of fairness it desires by altering initial endowments of agents through taxes and lump-sum transfers and then use the market system to achieve efficiency for the given distribution of initial endowments.

- Because fairness considerations involve value judgments (they are matters of ethics), most economists prefer to focus on efficiency (the positive side of analysis) than address the equity dimension of resource allocation. This is more so because it is impossible to agree on what society's SWF should like.

- Pareto improvement is not a very useful concept in practice. In considering policy issues there will be very few proposed reallocations that do not involve some individuals gaining and some losing. Compensation tests allow us to compare allocations involving winners and losers and which do not require the use of a SWF.

- Just like the Pareto efficiency criterion, compensation tests may not necessarily lead to outcomes that are considered fair. But they inform much of the application of welfare economics to environmental problems because they avoid the issue of equity by treating winners and losers equally.

- Market failure occurs whenever the market system is not able to achieve allocative efficiency because of the absence of one or more of the institutional arrangements required for an ideal system of resource allocation. Inefficiencies associated with market failure creates losses for society.

- The fundamental factors at the root of market failure are the absence of property rights, high transactions costs, and the lack of or imperfection in information.

Discussion/Review Questions and Exercises

1. Examine the contortions necessary for the proper functioning of markets and their roles in leading markets to an efficient allocation of resources.

2. Derive and explain the conditions for pareto efficiency in a market economy.

3. State and explain the practical implications of the two welfare theorems.

4. How does welfare economics treat the issues of efficiency and fairness? How does this apply to natural and environmental resources?

5. What are the limitations of 'pareto improvement' as a guiding principle to natural resource and environmental policy? What other criterion/criteria can you recommend?

6. Consider the Table below

	Allocation 1				Allocation 2		
	X	Y	U		X	Y	U
A	10	5	50		10	4	40
B	5	20	100		15	16	240

(a) assume that the utility function is given as U = XY show that application of compensation tests may not produce a result that may be considered fair.

(b) If the SWF is given as W = 0.5UA + 0.5UB, will the application of compensation tests lead to a reallocation that can be considered fair from society's viewpoint?

(7). What do you understand by market failure? Explain the fundamental factors behind market failure.

Materials used for this Lecture

Gravelle H. and Rees R. (2004). *Microeconomics* 3rd edition Prentice Hall Edinburg Gate Harlow

Perman, R., Ma Y., McGilvray J. and Common M. (2012). *Natural Resource and Environmental Economics*, 4th edition. Edinburgh, Longman.

Prato T. (1998). *Natural Resource and Environmental Economics*, Iowa State University Press Ames.

Varian H.R. (2010). *Intermediate Microeconomics a Modern Approach*, 8th edition W.W. Norton and Company New York.

Module 2.2 Public Goods (5 hours)

Learning Outcomes

Most natural and environmental resources take the form of public goods. This Module deals with public goods and why markets fail to provide such goods, or allocate them in an efficient manner. After going through this module, the reader should

- ✓ Know what a public good is, its characteristics and how to differentiate it from a private good.
- ✓ understand why free markets cannot provide public goods efficiently.
- ✓ Know how to determine the efficient level of provision of a public good.
- ✓ Appreciate why there are inefficiencies in the use of many natural and environmental resources.
- ✓ gain insights into applications of common property resource regime in Sub-Saharan Africa.

Outline

2.2.1 Definition and characteristics of public goods and its distinction from private goods

2.2.2 Efficient Provision of Public Goods

2.2.2.1 Efficient Provision of Discrete Public Goods

2.2.2.2 Private Provision of Discrete Public Good and the free rider problem

2.2.2.3 Voting for Discrete Public Good

2.2.2.4 Efficient Provision of a Continuous Public Good

2.2.2.5 Private Provision of Continuous Public Good

2.2.3 Problems of managing the Commons

2.2.4. Focus: Management of Common Property in Africa

Summary

Review/Discussion Questions and Exercises

Material used for this Lecture

2.2.1 Definition and characteristics of public goods and its distinction from private goods

Public goods are goods that can be consumed simultaneously by a large number of people without the consumption by one imposing a cost on others. Two characteristics helps to distinguish between a private and a public good: non-rivalry in consumption (or indivisibility), and non-excludability.

Non-rivalry in consumption. This means that consumption by one consumer does not restrict consumption by others. Non-rivalry means that a product or service does not reduce in availability as people consume it. When there is rivalry in consumption, one agent's consumption (use) is at the expense of another's, in which case, the two individuals are rivals in consumption. Assume we can break a good into its discrete units (divisibility), rivalry implies that it is not possible for two or more agents to consume the same unit of the good. Only one person can consume a unit at any given time, and what is consumed by one is not available for the other to consume. In other words, the consumption of a unit of the good by one agent reduces the amount available for others to consume.

Non-excludability. Excludability is the property of a good that makes it possible for an agent to be prevented from consuming it.[13] The benefits derived from pure public goods cannot be confined solely to those who have paid for it. Indeed, non-payers can enjoy the benefits of consumption at no financial cost.

In contrast to a public good, a **private good** is a product that must be purchased to be consumed, and its consumption by one individual prevents another individual from consuming it. In other words, private goods are characterized by rivalry in consumption and excludability. A good is considered to be a private good if there is competition between individuals to obtain the good and if consuming the good prevents someone else from consuming it. With private goods, consumption ultimately depends on the ability to pay

Non-rival consumption means that public goods are efficiently allocated if provided at a zero price, something markets are seldom inclined to do. Moreover, the inability to exclude nonpayers gives rise to the free-rider problem, which further inhibits the voluntary exchange of public goods through markets. The **free-rider problem** posits that a rational agent will not contribute to the provision of a public good because he does not need to contribute to benefit. For example, if a person does not pay his taxes, he still benefits from the government's provision of national defense or public roads by free riding on the tax payments of his fellow citizens. Private goods are less likely to experience the free-rider problem because a private good has to be purchased; it is not readily available for free.

[13] The question of excludability is a matter of law and convention, as well as physical characteristics. In some cases where the law enables excludability, it is infeasible to enforce it. However, the feasibility of exclusion is a function of technology. Consider, for example, the invention of barbed wire and its use in the grazing lands of North America.

Table 2.1 below uses the two characteristics to differentiate between various types of goods, ranging from pure private goods to pure public goods.

Table 2.1: Characteristics of Private and Public goods

	Excludable	Non- excludable
Rivalrous	**Pure private good** (e.g. a bottle of soda, a cup of ice cream)	**Open-access resource** (e.g. fishing in territorial ocean waters)
Non- rivalrous	**Congestible resource** (e.g. game park, highway)	**Pure public good** (e.g. national defence)

Pure private goods exhibit both rivalry and excludability: they are 'ordinary' goods and services we see everywhere around us, ranging from a bottle of soda to a luxury car. **Pure public goods** exhibit neither rivalry nor excludability. Two often-cited examples are street lighting and the services of the national defense force. Another example may be a dam. It is non-rivalry and non-excludable in that all people within a society benefit from its use without reducing the availability of its intended function. Some goods have characteristics of both private and public goods. For example, an open access good is rivalrous (just like a private good), yet non-excludable (just like a public good). Examples are found in many natural resources, such as ocean fishery outside territorial waters. A congestible good is one that is non-rivalrous (just like a public good) yet is excludable (like a private good). Examples are wilderness area, a public park etc. For a congestible good or resource, rivalry sets in only when congestion sets in.

Table 2.1 illustrate that some type of public goods can have property of excludability while some private goods can be non-excludable. A public good is considered excludable when it has a nominal cost that creates a low barrier to consuming the good. The post office, for example, is an excludable public good because even though the service is provided for the public, there are some costs, such as stamp expenses, that prevent people who have not paid from using it. In the same manner, private goods, such as a basic FM radio show are considered non- excludable since anyone with a radio can consume them. Sometimes, the phrase, **quasi-public goods** is used to describe public goods that have characteristics of being non-rivalry but are excludable in the sense that they are subject to congestion which limits access for people, or public goods that have some form of consumption rivalry. Roads are a good example of the former. All infrastructures are built for the benefit of the public, but as more of the public uses the infrastructure, it causes traffic and congestion, lowering the value of the good. Open access fishing in territorial waters is an example of the former.

2.2.2 Efficient Provision of Public Goods

2.2.2.1 Efficient Provision of Discrete Public Goods

Assume two individuals and two goods: a private good X and a public good G. Each individual is initially endowed with wealth w_i which is used to purchase the private good x_i and also contribute g_i towards the provision of the public good so that

$$w_i = x_i + g_i \qquad (2.17)$$

and $$x_i = w_i - g_i \qquad (2.18)$$

where x_i is the amount of consumption of the private good by individual i.

Utility is strictly increasing with the consumption of the public and private good .

$$U_i(g_i, x_i) \qquad (2.19)$$

The public good is only available in discrete amount i.e. either provided in the required amount or not provided at all. If C is the cost of producing a public good, then the technology is given by

g = 1, $if\ g_1 + g_2 \geq C$

g = 0 if $g_1 + g_2 < C$. (2.20)

Providing a public good will Pareto dominate not providing if,

$$g_1 + g_2 \geq C$$

and $$U_1(1, w_1 - g_1) > U_I(0, w_1)$$

$$U_2(1, w_2 - g_2) > U_2(0, w_2) \qquad (2.21)$$

$let\ r_i$ be maximum amount of the private good that the agent i will be willing to give up, to get one unit of the public good, then r_i is called the maximum willingness to pay or reservation price. By definition r_i must satisfy the equation.

$$U_i(1, w_i - r_i) > U_i(0, w_i) \qquad (2.22)$$

Equation (2.21) will imply that

$U_i(1, w_i - g_i) > U_i(0, w_i)$ = $U_i(1, w_i - r_i)$ for i = 1,2

$$U_i(1, w_i - g_i) > \ U_i(1, w_i - r_i) \qquad (2.23)$$

Since utility is strictly increasing in private consumption then $(w_i - g_i) > \ (w_i - r_i)$ for i =1,2. This implies $r_i > g_i$, i= 1,2, and $r_2 > g_2$, and the two; $r_i + r_2 > g_i + g_2$

To provide a public good $g_i + g_2 \geq C$ = $r_i + r_2 > g_i + g_2 \geq C$.

For it to be Pareto improving to provide a public good we must have $r_i + r_2 > C$, that is total sum of willingness to pay for public good must exceed the cost of providing it. Note that for private goods the willingness to pay equals the cost.

2.2.2.2 Private Provision of Discrete Public Good and the free rider problem

As already noted, for a public good to be provided, the total cost of the good, should be less than the total willingness to pay. This would therefore require governments to know the marginal willingness to pay of each individual, which is however difficult. It is by nature that the individuals will have an incentive not to reveal their preferences or the marginal willingness to pay, because once the public good is provided, all individuals can consume equal amounts of the good without any exclusion on the account of non-payment. In other words, individuals will attempt to 'free-ride' with respect to public good provision.

Consider a situation where there are two individual who privately decide to contribute towards construction of a road that costs $150. Suppose each individual's marginal willingness to pay is $100. That is r_i = 100 for i =1, 2, the cost is C = 150, and $r_i + r_2$ > C. Each agent decides privately whether or not to buy the public good, and since the good is public once bought, there is no exclusion. We can represent the pair in a simple prisoner dilemma game each with two strategies; Buy (B) and don't buy (DB)..

P_1 strategy \ P_2 strategy	B	DB
B	-50, -50	-50, 100
DB	100, -50	0, 0

Fig. 2. 5 Two-player-two-strategy pollution abatement game

If player 1 (P_1) wants to buy the good, he has to pay 150, although his willingness to pay is 100 worth of benefits, so his net worth is -50. If P_1 buys and player 2 (P_2) does not buy, then P_2 gets 100 worth of benefits for free, in which case, we say that P_2 is free riding on P_1. The dominant strategy equilibrium (and hence, Nash equilibrium) is for both players not to contribute i.e. (DB, DB) with payoff of (0,0). The end result is that the good is not provided at all, although it would be efficient to do so. This shows that purely independent decisions may not results in an efficient amount of public good being provided. This result can be generalized to public goods in which many individuals have a stake. In Module 7.1, we show that the same kind of game characteristics apply in the case involving global public goods.

2.2.2.3 Voting for Discrete Public Good

Suppose there are three individuals who live across a stream; one has a car; another has a bike and another walks. Suppose the cost of constructing a bridge is $90. If the three individuals have to vote for the good to be provided, then each would pay $30. Supposing their willingness to pay (the reservation price) is as follows; $r_1 = 70$ (Car owner) and $r_2 = 20$ (Bike owner); $r_3 = 10$ (Foot); so that we have $r_1+r_2+r_3 > C$. In this case the good is provided. However, only the car owner will vote for the provision of the good in this case since his reservation price exceeds his share of the cost. Therefore in majority voting, the good may not be provided unless the car owner (consumer 1) is willing to compensate other consumers. Note that majority voting only measures ordinal preferences, yet efficient conditions requires comparison of willingness to pay.

2.2.2.4 Efficient Provision of a Continuous Public Good

First, we need to understand what we mean by a continuous public good. For simplicity, let us consider a marram road stretching say 100Km that needs to be tarmacked. Suppose part of the road, say 70Km are tarmacked, and the remaining 30Km are marram. Since it is a public good, all individuals have free access to both sections of the road and one can drive on the tarmac section, then on the marram or vice versa without any hindrance. This is as opposed to the bridge example we had earlier under discrete public good. In other words, you do not need to first tarmac all the 100Km before the road can be used. This is an example of a continuous good.

Now suppose a public good is to be provided in continuous amounts. Suppose two individuals, 1 and 2, use this road and they decide to contribute amounts g_1 and g_2 respectively, towards its rehabilitation. Their total contribution will be such that the amount of public good provided is;

$$G = f(g_1 + g_2) \qquad (2.24)$$

In other words, the portion of the road that shall be tarmacked depends on the total contribution by the individuals.

Now the utility derived by each individual is

$$U_i = f\{(g_1 + g_2), w_i - g_i\} \qquad (2.25)$$

The first part on the RHS of equation (2.25) is, the amount of public good and the second part is the private good. This equation therefore states that the individual derives utility from consumption of both the public good after contributing amount g_i and leaving the rest of the endowment for the purchase of private goods.

Equation (2,25) could also be written as

$$U_i\{\quad\} = U_i(G, X_i) = U_i(f(G), X_i) \qquad (2.26)$$

The first order conditions (FOC) are obtained by maximizing a weighted sum of utilities of the two consumers/individuals, such that;

$$Max_{g1,g2\geq 0} a_1 U_1(g_1 + g_2), w_1 - g_2) + a_2 U_2(g_1 + g_2), w_2 - g_2) \quad (2.27)$$

The FOC for g_1 and g_2 can be written as;

$$\frac{\partial Z}{\partial g1} = a_1 \frac{\partial U1}{\partial G}(G, X_1) - a_1 \frac{\partial U1}{\partial X1}(G, X_1) + a_2 \frac{\partial U2}{\partial G}(G, X_2) = 0 \qquad (2.28)$$

Equation (2.28) could also be written as

$$a_1 \frac{\partial U1}{\partial G}(G, X_1) + a_2 \frac{\partial U2}{\partial G}(G, X_2) = a_1 \frac{\partial U1}{\partial X1}(G, X_1) \qquad (2.29)$$

Similarly,

$$\frac{\partial Z}{\partial g_2} = a_1 \frac{\partial U1}{\partial G}(G, X_1) + a_2 \frac{\partial U2}{\partial G}(G, X_2) - a_2 \frac{\partial U2}{\partial X2}(G, X_2) \qquad (2.30)$$

Equation (2.30) could also be written as;

$$a_1 \frac{\partial U_1}{\partial G}(G, X_1) + a_2 \frac{\partial U_2}{\partial G}(G, X_2) = a_2 \frac{\partial U_2}{\partial X_2}(G, X_2) \qquad (2.31)$$

Comparing equations (2.29) and (2.31), we note that the LHS are equal, this implies that the RHS must also be equal. That is,

$$a_1 \frac{\partial U1}{\partial X1} = a_2 \frac{\partial U2}{\partial X2} \qquad (2.32)$$

Given equation (2.32), we next divide the LHS of equation (2.29) or (2.31), by the RHS and we get,

$$\frac{\frac{\partial U1}{\partial G}(G,X1)}{\frac{\partial U`(G,X1}{\partial X1}} + \frac{\frac{\partial U2}{\partial G}(G,X2)}{\frac{\partial U2}{\partial X2}(G,X2)} = 1 \qquad (2.33)$$

We note that the LHS are sum of the ratios of marginal utilities of individuals 1 and 2. Therefore, equation (2.33) can also be stated as;

$$MRS^1_{G,X} + MRS^2_{G,X} = 1 \qquad (2.32)$$

Equation (2.32) is the condition for efficiency in the case of provision of a continuous public good. It states that the sum of the marginal willingness to pay equals the marginal cost of provision. We note here that marginal cost of provision (MC) equals to unit because the public good is the sum of the contributions.

In case of a quasi linear utilities such as, $U_i\ (G) +\ X_i$, the efficiency condition becomes

$$U_1'(G) + U_2'(G) = 1 \qquad (2.33)$$

The condition in equation (2.33) gives rise to a unique level of the public good, unlike that of equation (2.32), which rise to several allocations of commodities (G, X_1 , X_2).

2.2.2.5 Private Provision of Continuous Public Good

Now let us consider two individual who privately or independently decide to contribute towards provision of a public good. Since it is a public good such as a road, once it is provided, then any one is free to use it without any restrictions. And since individuals privately decide to contribute, then there is a possibility that some individual may not contribute but would go ahead and enjoy the provided public good. Suppose the two individuals contribute amounts g_1 and g_2 respectively towards provision of the public good, individual 1's utility maximization problem can be expressed as:

$$Max_{g1>0}\, U_1\, (g_1 +\ g_2\, , W_1 - g_1\,)$$

$$s.t\ g_1 W_1\ \geq 0 \qquad (2.34)$$

The constraint in equation (2.34) , means that the amount of a public good cannot be reduced by anyone, instead it increases. To solve the above constrained problem with an inequality constraint, we use the Kuhn-Tucker approach, so the FOC is;

$$\frac{\partial U1}{\partial G}\ (G, X_1\ ) - \frac{\partial U1}{\partial X1}\ (G, X_1\ ) \leq 0 \qquad (2.35)$$

Rearranging equation (2.35), we get;

$$\frac{\partial U1}{\partial G}\ (G, X_1) / \frac{\partial U1}{\partial X1}\ (G, X_1) \leq 1 \qquad (2.36)$$

But equation (2.36) is the ratio of marginal utilities for individual 1. That is

$$MRS_{G,X}^1\ \leq 1 \qquad (2.37)$$

Equation (2.37) means that for an individual to substitute a public good for a private good (contribute towards provision of a public good), then the marginal rate of substitution between the two goods must at least be equal to the marginal cost i.e. equal to unit (MC=1). But if the marginal rate of substitution is less than the marginal cost of providing the public good, then the individual shall not contribute. This is illustrated in Figure 2.6

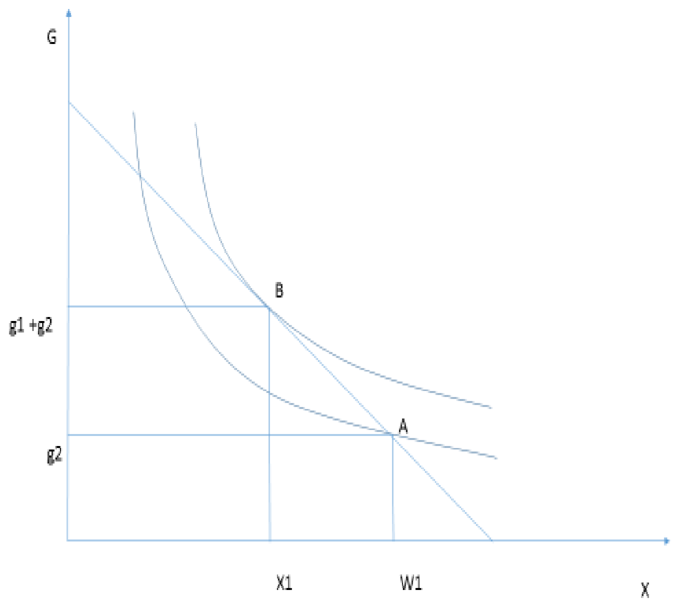


Fig 2.6 Contribute or not towards public good provision

Agent 1's endowment is (g_2, W_1), that is, he spends all his income/endowment W_1 on the private good if he contributes nothing and gets amount g_2 of public good (Point A in figure 2.5). This is because individual 2 contributes amount g_2 towards the provision of the public good and once provided, individual 1 is free to use it. We however note that because it is only individual 2 contributing towards the provision of the public good, individual 1 who decides to be a Free rider is on a lower indifference curve.

But now suppose individual 1 also decides to contribute towards the public good and he contributes an amount g_1, then he gets a total amount of public good equal to g_1+g_2 but less private good equal to X_1. However, the individual now lies on a higher indifference curve because he now has more public good. If we consider our earlier example of a continuous public good of tarmacking 100KM road. If only individual 2 contributes amount g_2, a smaller section of the road say only 30KM are tarmacked and the remaining section of the road shall be dusty. But if individual 1 also contributes amount g_1 then both shall tarmac a larger section of the road say 70KM, hence less dust and less chance of acquiring dust related diseases and thus a higher indifference curve and at equilibrium at point B.

We can express this problem in a game theoretic approach such that a Nash Equilibrium of this game is a set of contributions (g_1*, g_2*), which is such that each agent shall be contributing an optimal amount given the contributions of the other individual. So, for the two agents/individuals, if some public good is to be provided, then it should be such that

$$\frac{\frac{\partial U_1 (G^*, X_1^*)}{\partial G^*}}{\frac{\partial U_1}{\partial X_1}(G^*, X_1^*)} \leq 1, \ \ and \quad \frac{\frac{\partial U_2 (G^*, X_2^*)}{\partial G^*}}{\frac{\partial U_2}{\partial X_2}(G^*, X_2^*)} \leq 1 \qquad (2.38)$$

If some positive amount of public good G is to be provided, then at least one of these inequalities should be an equality. Note that the ratio of the marginal utilities with respect to public good G and private good X is the marginal rate of substitution between the public and the private good. So, if the MRS is equal to MC=1, then it would imply that that individual contributes towards the provision of the public good, or rather the individual is willing to exchange some of the private good for a public good. But if the MRS is less than MC (less than unit) then the individual does not contribute to the provision of the public good. And if both do not contribute, then there will be no public good. In contrast, if both contribute, then there will be much more public good provided.

An alternative way of solving for the Nash equilibrium is to solve for reaction function of agent *i*. This gives rise to the amount that agent *i* wants to contribute as a function of the other agent's contribution. We can then express individual 1's maximization problem as;

$$Max \geq_{g1,x1} U_1(g_1 + g_2, X_1) \ Such\ that\ g_1 + g_2 = W_1\ ;\ g_1 \geq 0 \qquad (2.39)$$

But since the public good G is as a result of the total contribution by both individuals, i.e. $G = g_1 + g_2$, then we can re-write the problem as;

$$MaxU_1(G, X_1)\ s.t.\ G + X_1 = W_1 + g_2\ ; G \geq g_2 \qquad (2.40)$$

Equation (2.40) presupposes that individual 1 considers the contribution of individual 2, i.e. g_2 as part of his income, so he effectively chooses the total amount of public good subject to his budget constraint and the amount that he chooses must be at least as large as the amount provided by the other individual. The budget constraint says that the total value of his consumption must equal the value of his endowment (W_1+ g_2)

Now suppose $f_1(W)$ is individual 1's demand for the public good, which is a function of his wealth. Then, the amount of public good that solves the maximization problem is given by;

$$G = Max[f_1(W_1 + g_2), g_2] \qquad (2.41)$$

Subtracting g_2 from both sides, we get;

$$g_1 = Max[f_1(W_1 + g_2) - g_2, 0] \qquad (2.42)$$

This is the reaction function for individual 1 and it shows his optimal contribution as a function of the other individual's contribution.

A Nash Equilibrium is the set of contributions (g_1^*, g_2^*) such that;

$$g_1^* \quad = Max[f_1(W_1 + g_2^*) - g_2^*, 0]$$

$$g_2^* = Max[f_2(W_2 + g_1^*) - g_1^*, 0] \qquad (2.43)$$

Assuming quasi linear Utility functions, then the Nash Equilibrium condition is;

$$U_1'(g_1^* + g_2^*) \leq 1 \qquad (2.44)$$

$$U_2'(g_1^* + g_2^*) \leq 1 \qquad (2.45)$$

Generally, at least one of the two constraint should be binding if the public good is to be provided. If individual 1 places a higher marginal value on the public good than individual 2, so that $U_1'(G) > U_2'(G)$ for all G, then individual 1 will always contribute towards the provision of the public good while individual 2 will always freeride. For instance, consider two individuals the rich-individual 1 and the poor-individual 2. The rich will always want to contribute towards repairing the road and fixing potholes, but the poor who has no vehicle but only walks will not mind using the road in the way it is. And once the road is repaired the rich cannot stop the poor from using it because it is a public good). However, both individuals will contribute if they have same marginal value for the public good.

2.2.3 Problems of managing the Commons

Several common environmental resources (both stock and flows) have the characteristics of public goods. The benefits from biological diversity, the services of wilderness resources, the climate regulation mechanisms of the earth's atmosphere, and the waste disposal and reprocessing services of environmental sinks all constitute public goods, provided the use made of them is not excessive. Much of public policy towards this class of environmental resources can be interpreted in terms of regulations or incentives designed to prevent use breaking through their threshold levels. Some other environmental resources, such as naturally renewing resource systems share public goods properties. Examples include water resource systems, open fisheries etc. In these cases, consumption by one person does potentially reduce the amount of the resource available to others (there is rivalry in consumption) so the resource could be 'scarce' in an economic sense. But this will not be relevant in practice as long as consumption rates are low relative to the system's regenerative capacity.

One of the underlying reasons why markets fail to provide public goods is because of the absence of private property right which allows a provider to restrict access or collect an appropriate price for its use. A property right that is capable of producing efficient allocations in a well-functioning market economy must possess the basic properties of exclusivity, transferability and enforceability.

Exclusivity implies that all benefits and costs accrued as a result of owning and using the resources should accrue to the owner, and only to the owner, either directly or indirectly by sale to others. **Transferability** means that the property right should be transferable from one owner to another in a voluntary exchange, while **enforceability** implies that the property right should be secure from involuntary seizure or encroachment by others. An owner of a resource with a well-defined property right (one exhibiting these three characteristics) has a powerful incentive to use that resource efficiently because a decline in the value of that resource represents a personal loss. For example, farmers who own the land have an incentive to fertilize and irrigate it because the resulting increased production raises their incomes etc.

Also, when well-defined property rights are exchanged, as in a market economy, this exchange facilitates efficiency. Because the seller has the right to prevent the consumer from consuming the product in the absence of payment, the consumer must pay to receive the product. Given a market price, the consumer decides how much to purchase by choosing the amount that maximizes his or her individual consumer surplus (where marginal benefit is just equal to marginal willingness-to- pay).

The manner in which producers and consumers use environmental resources depends on the property rights governing those resources (see Perman et al, 2003, pp128-130 for illustration in relation to biodiversity). Apart from **private property** right, there are other possible ways of defining entitlements to resource use. These include

- **state-property regimes**: where the government owns and controls the property. Parks and forests, for example, are frequently owned and managed by the government in many counties.
- **common-property regimes**: where the property is jointly owned and managed by a specified group of co-owners), and
- **res nullius or open-access regimes**: where no one owns or exercises control over the resources.

All of these create rather different incentives for resource use. For open-access regime, the resources can be exploited on a first-come, first-served basis because no individual or group has the legal power to restrict access. The situation typically leads to what has become known popularly as the "tragedy of the commons": the tendency for a resource to be exploited to the point of ruin.[14] We will be looking at this more closely in the context of renewable natural resources in Module 3.5.

[14] The phrase attributed to Hardin (1968), was originally applied to grazing land.

Common property regimes provide an alternative way of solving the problem associated with 'open-access'. The use of the term 'common property' has always been controversial. According to Berkes and Farvar (1989), this stems from philosophical differences between the socio-anthropological and resource management economists. The traditional economists' view is that property is either private or it belongs to the state. According to this view, resources that cannot be appropriated by private individuals or the state are called 'common property resources'. In contrast, the socio-anthropological perspective recognizes the possibility that a resource can be owned collectively by a defined group. According to this view, the term common property should be restricted to resources for which there exist arrangements for the exclusion of non-owners and for allocation among co-owners. The distinction between the economists' and the socio-anthropologists' views is crucially important with regard to Hardin's (1968) 'tragedy of the commons' model. This model leads to the conclusion that resources should be either privatized or controlled by government authority to ensure sustainable use.

Common-property regimes exhibit varying degrees of efficiency and sustainability, depending on the rules that emerge from collective decision-making. While some very successful examples of common-property regimes exist, unsuccessful examples are even more common. Box 2.3 provides an example of successful local common-property regime. Unfortunately, that kind of stability may be the exception rather than the rule, particularly in the face of heavy population pressure. The more common situation is illustrated by reference to another real-life situation in the Box 2.4

Box 2.3 An example of successful common-property regime

One successful example of a common-property regime involves the system of allocating grazing rights in Switzerland. Although agricultural land is normally treated as private property in Switzerland, grazing rights on the Alpine meadows have been treated as common property for centuries. Overgrazing is protected by specific rules, enacted by an association of users, which limit the amount of livestock permitted on the meadow. The families included on the membership list of the association have been stable over time as rights and responsibilities have passed from generation to generation. This stability has apparently facilitated reciprocity and trust, thereby providing a foundation for continued compliance with the rules.

Source: Tietenberg & Lewis, 2012, p28)

Box 2.4 An example of unsuccessful common-property regime

In Mawelle, a small fishing village in Sri Lanka, a complicated but effective rotating system of fishing rights was devised by villagers to assure equitable access to the best spots and best times while protecting the fish stocks. Over time, population pressure and the infusion of outsiders raised demand and undermined the collective cohesion sufficiently that the traditional rules became unenforceable, producing overexploitation of the resource and lower incomes for all the participants.

Source: Tietenberg & Lewis, 2012, p28)

We conclude this section by looking at property regimes in the management of Africa's wildlife.

2.2.4. Managing Africa's wildlife

Wildlife is widely becoming an important vehicle for rural development in Sub-Saharan Africa There are various management schemes governing wildlife preservation in the continent. Three particular forms are worthy of note. They are

- **Protected areas:** These lands are especially found in West and Central Africa but a few also exist in West and South Africa. They are often managed and funded by the country's government,

- **Private Reserves: These reserves are mostly found in southern Africa, including Namibia, South Africa, Botswana and Zimbabwe. In most of the cases, these private territories are fenced. The landowner chooses the fences depending on the type of**

animals living in the reserve. He chooses how to manage his land and can decide to build infrastructures, lodges for tourists and make them pay an entry fee. However, some guidelines are to be respected regarding wildlife protection with hunting quotas. The same kind of rules apply for private reserve only proposing tours and safaris. For example, in Namibia, a private owner may have several black rhinos on his land but – in theory – it is totally forbidden to allow hunters to kill one of them as the species is registered on the International Union for Conservation of Nature (IUCN) Red List as in critical danger.

Private reserves in southern Africa buy and sell animals on markets. This is a common practice authorized by the law but owners cannot do what they want. For instance, they need a specific state authorization to buy a species coming from other countries. Besides, the wildlife they introduce has to be adapted to the habitat. The rules on animal protection and management are decided by the government. International regulation on endangered species trade is made by the CITES (Convention on International Trade of Endangered Species).

- **Conservancies:** Conservancies embody a third management model for protected areas in Africa. Widespread in Namibia, the country accounts for more than 80 conservancies. Following the Namibian Independence, the government kept only a few national parks like Etosha and Namib-Naukluft and gave the management of other areas to the local communities. They are in charge of tourism, hunting and wildlife management. These conservancies can range from several thousands to millions of hectares. In 2006, there were about 40 conservancies in Namibia. They are much more today. They are based on the CBNRM model (Community Based Natural Resources Management). Each conservancy chooses their business model: eco-tourism, hunting and sometimes both.

Depending on the type of wildlife living on the territory, a conservancy can focus on trophy hunting to generate revenue. But the government decrees strict quotas and animal population follow-up. Villagers deal with private companies (travel agencies, safari agencies, etc.) to install camps and attract tourists. The staff (cook, cleaning lady, and guide) is recruited among the local community.

Some conservancies prefer to rely on eco-tourism rather than hunting. Here too, local villagers sign contracts with private companies in order to build infrastructures and lodges. It also implies attracting more tourists because observation generates less revenues than hunting. Indeed, hunters pay an extra fee depending on the trophy. In order to stimulate the local economy, there is a preference for local products harvested by farmers. In Namibia, the Caprivi region is wet enough to allow villagers to grow crops and sell their products to lodges. In both cases (hunting and observation), conservancies sign contracts with private businesses in charge of promotion, communication and marketing to attract tourists.

Some National Parks in Sub-Sharan Africa have lost from 60% to 70% of their fauna, sometimes even more, Threats are numerous: loss of habitats because of unsustainable farming and mining, poachers killing animals and installing traps, etc. In general, na**tional parks are failing to protect their fauna and lack funding.** According to the International Union for Conservation of Nature (IUCN), **the conservancy model based on local communities shows the best results regarding wildlife conservation in Namibia, Kenya or Tanzania. Thus, c**ommunities hold the key to expanding conservation impact in Africa [15]**.**

Many believe that communities should be given an active role in their natural capital preservation. On the other hand, the economic dimension of a project plays a major role in its success. Without economic returns for the locals, failure is bound to happen. It can be challenging to associate revenues and environment protection, but it is necessary. However, policymakers are usually not informed about the needs and wants of poor rural households and roll out programmes that are not tailor-made to suit their desires, which often results in policy failure. In addition while rural communities desire a regime that gives them full control over wildlife, poachers and those who are generally good at extracting resources from the environment tend to oppose change (Ntuli, Muchapondwa and Okumu, 2020).

Activity- and actor-led land and natural resource conflicts in Sub-Saharan Africa, are a cause for concern. Disputes mostly center on the demarcation, ownership, and inheritance of land; or from the weakening of customarily held rights of pastoralists. The causes include unsuitable land legislation, especially where there is no comprehensive land policy or ambiguous laws do not address overlapping rights and claims to land. Dysfunctional and inaccessible land administration also contribute to disputes, as do land grabbing and land invasions. In addition, disputes are fueled by the pressure of increasing population.

The next subsection focus on wildlife management regime in Zimbabwe, believed by some to have been relatively successful.

2.2.4.1 Focus: Managing Africa's wildlife - The Case of Zimbabwe[16]

In most of the world, the ownership of wildlife lies in the hands of governments. State agencies closely regulate the use of wildlife, including the amount permitted, when, and where hunting

[15] https://www.iucn.org/crossroads-blog/201810/communities-hold-key-expanding-conservation-impact.

[16] This section is due to **Kay Muir-Leresche and Robert H. Nelson (2001),** Managing Africa's Wildlife, Property and Environment Research Center, PERP, Volume 19, No.3, Fall 2001.

can occur. However, in the southern African nations of Zimbabwe, South Africa, Namibia and Botswana, an important experiment has been taking place over the past forty years (G. Child 1995). To a considerable degree, these nations have legalized and privatized the use of wildlife, encouraging hunting, tourism, and the sale of meat, hides, and horns. The wildlife remains state-owned without a formal owner (*res nullius*), but certain conditions grant private landowners the full rights to control the use of wildlife on their land.

The experience of Zimbabwe illustrates this development. Unfortunately, the violence of recent months in that country, caused by the government's refusal to protect property rights, may well undermine this extraordinary progress in wildlife management. In Zimbabwe, before the mid-1960s, farmers were not allowed to hunt, cull, or sell venison. They relied heavily on the Department of National Parks and Wildlife Management for control of problem animals and had no incentive to encourage healthy wildlife populations. Many game animals were killed in order to control the tsetse fly, which wild animals were believed to host (Murindagomo 1997, 433). Yet, cattle ranching has always been economically marginal in much of Zimbabwe's semi-arid rangelands. The ranching industry long depended on government subsidies, encouraging excessive cattle stocking and degraded rangeland.

One proposed early solution was to use the rangelands for producing wild game for meat. Some observers argued that wild animals were better adapted to Zimbabwe habitat and that introducing game could result in more meat (and higher profits) than cattle operations were showing (Dasmann and Mossman 1961). So, a 1961 Conservation Act allowed large-scale commercial farmers to obtain permits to harvest wildlife for meat. However, capturing and killing widely dispersed wildlife populations proved expensive. This and other problems, such as the absence of established channels of distribution, made wild-game meat production unprofitable (Muir 1989).

Luckily, a more promising approach emerged: **safari hunting**. Safari hunting was concentrated primarily in Kenya and of east Africa for much of the twentieth century, but a safari industry began growing rapidly in Zimbabwe in the 1960s. By 1974, one study found 17 ranchers actively managing their ranches for safari hunting, with another 150 ranchers showing interest (B. Child 1988, 178). Safari hunting does not depend on large numbers of animals, but more on the presence of trophy animals. It thus offered opportunity for larger revenues while lowering burdens on both the land and the wildlife, creating less harmful environmental impacts on semi-arid rangelands. At the same time, transferring responsibility for wildlife on private land from the state to the landholder allowed the Department of Natural Parks and Wildlife Management to concentrate on wildlife research and its protected areas.

In 1975, the Parks and Wildlife Act delegated management control of safari hunting, harvesting, and other wildlife activities to large-scale commercial farmers. Commercial farmers were designated as the "appropriate authority" for deciding the wildlife use of the land. By either controlling safari operations or leasing hunting rights to an independent safari operator, the owner could now make important decisions regarding their wildlife, including: the time and place of hunting, the number of animals to be hunted, the age and sex of the animals, minimum acceptable trophy sizes, and other hunting conditions. Also, the Department of National Parks and Wildlife Management established detailed training programs and tightened licensing requirements for hunting guides. Today, the guides of Zimbabwe are often said to be the best qualified in Africa.

The principal author of the 1975 act, then the director of the Department of National Parks and Wildlife Management, Graham Child, stated that with the exception of specially protected animals, “land holders are better placed than anyone else to conserve their wildlife” (quoted in B. Child 1988, 179Ð80). The government abolished most license fees and allowed landowners to charge for hunting and fishing as well.

In some areas of Zimbabwe, commercial wildlife operations became more profitable than cattle-raising by the 1980s. Ranchers located near the Matetsi Safari area in northwest Zimbabwe especially benefitted from tourism because of nearby Victoria Falls and Hwange National Park. Most ranchers who converted to wildlife operations experienced significant financial gains.

One ranch in the northwest, the Rosslyn Ranch, was originally managed for cattle beginning in 1948. Poor financial returns led the owners to convert to wildlife ranching in 1967. Wildlife populations expanded in four years by as much as 50 percent. In terms of meat production alone, the ranch proved viable, but the growing revenues from safari operations gave the greatest boost to profitability. (After 1972, the ranch area was taken over by the Rhodesian government and incorporated into the Matetsi Safari area, where it is now regularly leased as state-owned land to safari operators) (B. Child, 1988, 305Ð10).

In 1986, Brian Child surveyed ranchers in the south-eastern area (the “lowveld”) of Zimbabwe, asking them to name the most profitable use of their land. None named cattle ranching alone. Thirty percent said “mostly cattle, some wildlife.” Forty percent said “mostly wildlife, some cattle,” and 30 percent said “wildlife only.” By 1987, 10 percent of all private farm and ranch land owners were registered as wildlife producers (Muir 1989, 311).

Kreuter and Workman (1994) analyzed 15 large cattle-only operations, 7 wildlife, and 13 mixed ranches in the Midlands Province, where the climate is less arid and the land more productive. Among these, ranches devoted exclusively to wildlife were less financially successful than those with cattle only. However, the mixed ranches had the highest profitability.

By 1995, 18 percent of all Zimbabwe farmers were registered as being in the wildlife business (at least in part). A 1995 survey (achieving 50 percent coverage) of Wildlife Producers Association members showed that their lands held 250,000 wild plains animals, including 10,000 sable, 10,000 zebra and more than 2,000 giraffe (Muir 1998, 9). Private land has begun to make a major contribution to species diversity in Zimbabwe. By 1994, ninety-four percent of the eland in Zimbabwe were on privately owned commercial farm and ranch lands. Additionally, 64 percent of the kudu, 63 percent of the giraffe, 56 percent of the cheetah, and 53 percent of both sable and impala found refuge on private lands. In fact, private lands contained a majority of every plains game species found in Zimbabwe except zebra, 46 percent of which were found on private lands (Hill 1994). Tsessebe were once threatened throughout Zimbabwe but were able to recover on private ranches, subsequently allowing their restoration to many other private and public lands in Zimbabwe (Zimbabwe Trust 1992, 42). All in all, except for large and dangerous game, Zimbabwe wildlife is better preserved through private management for financial gain than by government protection. This happened because of the property rights innovations of 1961 and 1975, which have promoted species diversity and sustainable development.

Summary

- Public goods are goods that can be consumed simultaneously by a large number of people without the consumption by one imposing a cost on others. They are usually characterized by non-rivalry in consumption and non-excludability.

- Non-rival consumption means that public goods are efficiently allocated if provided at a zero price, something markets are seldom inclined to do. Moreover, the inability to exclude nonpayers gives rise to the free-rider problem, which further inhibits the voluntary exchange of public goods through markets

- For it to be Pareto improving to provide a discrete public good, the total sum of willingness to pay must exceed the cost of providing it. Individuals have an incentive to not to reveal their preferences or marginal willingness to pay, because once the public good is provided, all individuals can consume equal amounts of the good without any exclusion on the account of non-payment.

- Efficiency in the case of provision of a continuous public good requires that the sum of the marginal willingness to pay equals the marginal cost of provision.

- For an individual to substitute a public good for a private good (contribute towards provision of a public good), the marginal rate of substitution between the two goods must at least be equal to the marginal cost

- An allocation of a public good using a price system is called a Lindahl allocation.

- One of the underlying reasons why markets fail to provide public goods is because of the absence of private property right which allows a provider to restrict access or collect an appropriate price for its use. A property right that is capable of producing efficient allocations in a well-functioning market economy must possess the basic properties of exclusivity, transferability and enforceability.

- Several common environmental resources have the characteristics of public goods. The manner in which producers and consumers use environmental resources depends on the property rights governing them.

- Apart from private property right, there are other possible ways of defining entitlements to resource use, including state-property regimes, common-property regimes and open-access regimes. All of these create different incentives for resource use.
- For open-access regime, the resources can be exploited on a first-come, first-served basis because no individual or group has the legal power to restrict access. The situation typically leads to the 'tragedy of the commons': unlimited access destroys the incentive to conserve.

Discussion/Review Questions and Exercises

1. Derive and explain the efficiency conditions for the provision of a discrete public good. What are the implications for market provision?

2. What problems are associated with managing the commons?

Materials used for this Lecture

Gravelle H. and Rees R. (2004). **Microeconomics** 3rd edition Prentice Hall Edinburg Gate Harlow

Hara, Mafaniso, Turner, Stephen, Haller, Tobias and Matose, Frank (2009)'Governance of the commons in southern Africa: knowledge, political economy and power', *Development Southern Africa*,26:4,521 — 537 DOI: 10.1080/03768350903181324
URL: http://dx.doi.org/10.1080/03768350903181324

Perman, R., Ma Y., McGilvray J. and Common M. (2012). **Natural Resource and Environmental Economics**, 4th edition. Edinburgh, Longman.

Prato T. (1998). **Natural Resource and Environmental Economics**, Iowa State University Press Ames.

Tietenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics

UNDP, Land Rights for African Development from Knowledge to Action

Varian H.R. (2010**). Intermediate Microeconomics a Modern Approach,** 8th edition W.W. Norton and Company New York.

Module 2.3: Environmental Externalities (5 hours)

Learning Outcomes

Most natural and environmental resource problems often take the form of externalities. This Module exposes the reader to the subject matter of environmental externalities and what options can be used to address them. After going through this module, the reader should

- ✓ understand the concept of externalities and the various possible types of externalities.
- ✓ know why the presence of an externalities lead to market failure.
- ✓ Know what options can be used to correct for externalities.
- ✓ understand the Coase theory and how Coase bargaining can be used to correct for externalities.
- ✓ understand the limitations of private bargaining as a means of addressing the externality problem
- ✓ understand the concept of Pigouvian taxes, how it can be used to correct for negative externalities, and the problems associated with its use.
- ✓ understand why externalities are often associated with the use of many natural and environmental resources.

Outline

2.3.1 Definition and Classifications of Externalities

2.3.2. Externalities and Resource Allocation

2.3.3. Solutions to Externality Problems

2.3.3.1 The Private Approach

2.3.3.2. Coase Theorem: A formalization

2.3.3.3 The Public Approach: Pigouvian taxes

Summary

Review/Discussion Questions and Exercises

2.3.1 Definition and Classifications of Externalities

An externality exists when the consumption or production choices of one person or firm negatively or positively affect the utility or production of another entity in *an unintended manner.* When someone's behaviour increases or decreases another's utility or profit, without compensation, then we say that the latter is exposed to an externality. Externalities could be positive (when it increases the welfare of the agent impacted) or negative (when it reduces his or her welfare).

Negative Externalities occur when a consumer /producer's activities negatively impact another consumer utility /producer's production function and is not compensated for by one who caused it. Examples are when a factory upstream is polluting a river, and hence, leading to less fish catch by a fishery downstream; cigarette smoke from an individual , which is affecting others in the neighbourhood, loud music coming from a housing unit affecting residents in the neighbourhood who want to sleep or need some quietness.

Positive Externalities occur when a consumer /producer's activities positively influence another consumer utility /producer's production function, e.g. getting pleasure from observing a neighbour's flower garden; crops being pollinated by bees from a neighbouring bee farm.

Note that the phrase "unintended way" in the definition of an externality above implies that the agent producing the effect did not take it into consideration in his/her transaction. Thus, he/she did not pay (in the case of a negative externality) or secure (in the case of a positive externality) any compensation for the effect being generated on the other party (parties). In other words, an externality exists because there are things that people care about that are not priced, that the market system does not take account of. Positive externality is not as much an issue in economics as when externalities are negative. After all, it is easier to "reward" an agent for a positive externality!

Some authors identify a different form of externalities: **pecuniary externalities** (that is externalities which arise when the external effect is transmitted through altered prices). Suppose that a new firm moves into an area and drives up the rental price of land. That increase creates a negative effect on all those paying rent and, therefore, is an external diseconomy. However Pecuniary externalities do not cause a market failure because the associated price changes reflect changes in scarcity value: without pecuniary externalities, the price signals would fail to sustain an efficient allocation. Hence, pecuniary externalities are not usually the focus when we talk about externalities and relate it to market failure. Box 2.5 presents an example of externalities arising from shrimp farming in a region of Thailand.

Box 2.5 Shrimp farming and externalities in Thailand

In the Po village on the coast of Surat Thani Province in Thailand, more than half of the 1,100 hectares of mangrove swamps have been cleared for commercial shrimp farms. Although harvesting shrimp is a lucrative undertaking, mangroves serve as nurseries for fish and as barriers for storms and soil erosion. Following the destruction of the local mangroves, Tha Po villagers experienced a decline in fish catch and suffered storm damage and water pollution. Can market forces be trusted to strike the efficient balance between preservation and development for the remaining mangroves? Calculations by economists Sathirathai and Barbier (2001) demonstrated that the value of the ecological services that would be lost from further destruction of the mangrove swamps exceeded the value of the shrimp farms that would take their place. Preservation of the remaining mangrove swamps would be the efficient choice. Would a potential shrimp-farming entrepreneur make the efficient choice? Unfortunately, the answer is no. This study estimated the economic value of mangroves in terms of local use of forest resources, offshore fishery linkages, and coastal protection to be in the range of $27,264–$35,921 per hectare. In contrast, the economic returns to shrimp farming, once they are corrected for input subsidies and for the costs of water pollution, are only $194–$209 per hectare. However, as shrimp farmers are heavily subsidized and do not have to take into account the external costs of pollution, their financial returns are typically $7,706.95–$8,336.47 per hectare. In the absence of some sort of external control imposed by collective action, development would be the normal, if inefficient, result. The externalities associated with the ecological services provided by the mangroves support a biased decision that results in fewer social net benefits, but greater private net benefits.

Source: Suthawan Sathirathai and Edward B. Barbier. "Valuing Mangrove Conservation in Southern Thailand" *Contemporary Economic Policy*, Vol. 19, No. 2 (April 2001), pp. 109–122.; Titenberg and Lewis, 2012. p27)

External effects can be classified according to what sort of economic activity they originate in and what sort of economic activity they impact on. An economic situation involves a **consumption externality** if one consumer cares directly about another agent's consumption or production. For example, the amount of pollution produced by nearby automobiles are examples of negative consumption externalities. Similarly, a **production externality** arises when production possibilities of one firm are influenced by the choices of another firm or consumer. For example, an apple orchard located next to a beekeeper. In this case, there are mutual positive production externalities – each firm's production positively affects the production possibilities of the other firm. On the other hand, a fisheries cares about the amount of pollutants dumped into a fishing area, since this will negatively influence the amount of fish catch.

Given the above, and using a two-person, two-(private)-commodity framework, we will have a six-fold classification of externalities as illustrated in Table 2.2.

Table 2.2. Classification of externalities

	Arising from	Affecting	Utility/production function
1	Consumption	Consumption	$U^A(X^A, Y^A, X^B)$
2	Consumption	Production	$X(K^X, L^X, Y^A)$
3	Consumption	Consumption and production	$U^A(X^A, Y^A, X^B)$ and $Y(K^Y, L^Y, X^B)$
4	Production	Consumption	$U^A(X^A, Y^A, X)$
5	Production	Production	$X(K^X, L^X, Y)$
6	Production	Consumption and production	$U^A(X^A, Y^A, Y)$ and $X(K^X, L^X, Y)$

2.3.2. Externalities and Resource Allocation

Externalities may or may not have the characteristics of a public good (non-rivalry and non-excludability) but those that are most relevant for policy analysis exhibit non-rivalry and non-excludability. This is especially the case with external effects that are associated with environmental problems such as pollution. As in the case of public goods, the presence of externalities leads to market failure. Given that all of the other institutional conditions for a pure market system to realize an efficient allocation hold, if there is a beneficial externality associated with a good, the market will produce too little of it in relation to the requirements of allocative efficiency. In the case of a harmful externality, the market will produce more of it than efficiency requires. Since our concern in this course is application of welfare economics to environmental problems, we will focus on negative externalities, and show that in the absence of corrective policy, the market system will 'over-supply'.

Externalities affect resource allocation because the market fails to fully set the price for the external effects generated by some economic activities. This is because market prices reflect the private costs and benefits that agents derive from goods and services (which are based on the personal utility derived), while ignoring the costs/benefits imposed on third parties. Thus, the pricing mechanism fails to reflect the true or social costs or benefits of economic activity. With externalities, private and social costs/benefits are not the same, so that resources allocated on the basis of private consumption and/or production decisions will not be efficient. We will illustrate with cases of consumption as well as production externalities.

Scenario I: Negative externality in a consumer-to-consumer case

Suppose that A and B live in adjacent flats (apartments). A is a saxophone player, who enjoys practicing a lot. B does not like music, and can hear A practicing. The utility functions are

$$U^A = U^A(M^A, S^A) \qquad 2.50$$

$$U^B = U^B(M^B, S^A) \qquad 2.51$$

where M represents wealth and S^A is the hours that A plays the saxophone each week, with $\partial U^A/\partial M^A > 0$, $\partial U^B/\partial M^B > 0$, $\partial U^A/\partial S^A > 0$ and $\partial U^B/\partial S^A < 0$. Let

MB = marginal benefit of playing to A (see this as the amount that A would pay, if it were necessary, to play a little more, or the amount of compensation that would be required to leave A as well off given a small reduction in playing).

MEC = marginal external cost (the marginal cost) of playing to B (see this as the amount that B would be willing to pay for a little less playing or the amount of compensation that would be required to leave B as well off given a small increase in hours of A's saxophone playing).

Given that A does not in fact have to pay anything to play her saxophone in her flat, she will increase her hours of playing up to the level, where MB is equal to zero (the point M_0) in Figure 2.10. At that level, A's total benefit from playing is given by the sum of the areas of the triangles a, b and d, and B's total suffering is measured in money terms by the sum of the areas b, d and c. This is not a Pareto efficient outcome, because at M_0, MEC > MB. The efficient outcome is at M* where MEC = MB. At any M to the left of M*, MB > MEC, so that for a small increase in M, A would be willing to pay more than would compensate B for that increase. At any M to the right of M*, MEC > MB so that for a small decrease in M, B would be willing to pay more for a small decrease in M than would be required to compensate A for that decrease.

.

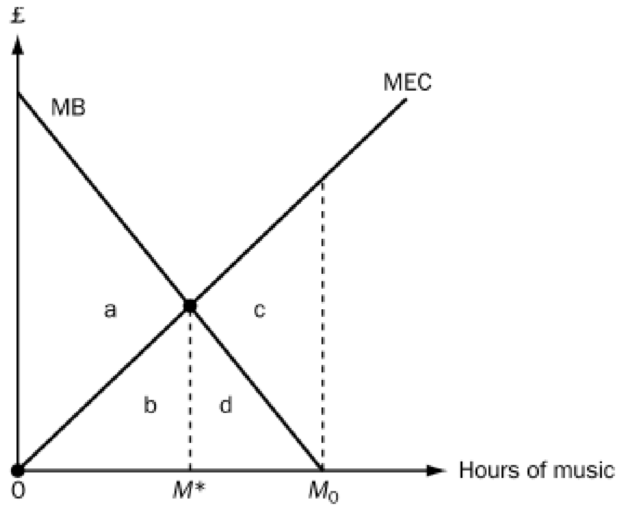


Fig. 2.8 Market allocation with a negative externality in a consumer-to-consumer case. **Source:** Perman et al., 2003. p137.

Note that the inefficient level of saxophone playing at M_0 comes about because there are no payments in respect of variations in M (no market in M), so that the effect on B is "unintentional" on the part of A (it is not taken into account or internalized by A). In other words, A does not compensate B, because B does not have any legal right to such compensation. B does not have a property right in a domestic environment unpolluted by saxophone music.

Scenario II: negative externality in a producer-to-producer case

Consider two firms with production functions

$$X(K^X, L^X, S) \quad (2.52)$$

$$Y(K^Y, L^Y, S) \quad (2.53)$$

where S stands for pollutant emissions arising in the production of Y, which emissions affect the output of X for given levels of K and L input. As an example, suppose two firms are located by a river. The first produces steel (Y), while the second, somewhat downstream, operates a resort hotel producing hospitality services (X). Both use the river, although in different ways: the steel firm uses it as a receptacle for its waste; the hotel uses it to attract customers seeking water recreation. We assume for simplicity, that $\partial Y/\partial S > 0$, so that for given levels of K^Y and L^Y, lower S (emissions) means lower Y output, and that $\partial X/\partial S<0$, so that for given levels of K^X and L^X, higher S means lower X.

For the same reasons adduced in the consumer-to-consumer case, if these two facilities have different owners, an efficient use of the water is not likely to result (make sure you can explain why). Figure 4.10 also applies. We can reinterpret and relabel the horizontal axis so that it refers to S, with S_0 replacing M_0 and S* replacing M*. For profits in the production of X we have

$$\pi^X = P_X X(K^X, L^X, S) - P_K K^X - P_L L^X \quad (2.54)$$

where ∂πX/∂S < 0. This is the marginal external cost, MEC, the impact of a small increase in S on profits in the production of X.

For profits in the production of Y we have

$$\pi^Y = P_Y Y(K^Y, L^Y, S) - P_K K^Y - P_L L^Y \qquad (2.55)$$

where ∂πY/∂S > 0. This is the marginal benefit, MB, the impact of a small increase in S on profits in the production of Y. In the absence of a well-defined property right S will be too large for efficiency.

Scenario III: negative externality in a producer-to-consumer case

This kind of externality usually impact on many agents, and with respect to them is non-rival and non-excludable. Hence it is the kind of pollution problems seen as most relevant to policy determination. As it is traditional, we assume for the sake of simplifying the analysis that there are just two agents whose consumption are impacted by the activities of a given firm. In the language of our two-person, two-commodity economy we have

$$U^A = U^A(X^A, Y^A, S) \qquad (2.56)$$

with $\partial U^A/\partial S < 0$

$$U^B = U^B(X^B, Y^B, S) \qquad (2.57)$$

with $\partial U^B/\partial S < 0$

$$X = X(K^X, L^X) \qquad (2.58)$$

$$Y = Y(K^Y, L^Y, S) \qquad (2.59)$$

with ∂Y/∂S > 0

Emissions arise in the production of Y and adversely affect the utilities of A and B. The pollution experienced by A and B is non-rival and non-excludable. An example of the above problem is a fossil-fuel-burning electricity plant located in an urban area. Its emissions pollute the urban air shed, and, to a first approximation, all who live within the affected area experience the same level of atmospheric pollution.

Our task in the next subsection is to see what types of solutions are available to solve the externalities problem in each of the three scenarios examined above.

2.3.2. Solutions to Externality Problems

2.3.2.1 The Private Approach

The private approach to the externality problem occurs when affected individuals are able to find a means of solving the problems of external damage or benefit. For example, if one neighbour is inflicting external damage on other neighbours, then these neighbours can get together and agree upon how to solve this problem (i.e. without government intervention). This approach is an application of the **Coase theorem** (Coase, 1960**),** which states that *if property rights are well defined and transactions costs are low, private parties can internalize an externality.* That is, conflicting parties can solve their externality problem by bargaining to a common point prior to assigning the property rights. According to the Coase theorem, two private parties will be able to bargain with each other and find an efficient solution to an externality problem. Before we go into a formal consideration of Coase theorem, we will apply the proposition to the three scenarios above.

In scenario I, we saw that A does not compensate B because B does not have any legal right to such compensation. B does not have a property right in a domestic environment unpolluted by saxophone music. A possible solution will be to establish such a property right. For example, we can give B the legal right to a domestic environment that is not noise-polluted. We can show that such legal arrangements would support bargaining which would lead to an efficient outcome (achieve M* as the level of M in Figure 2.8, which is reproduced below again for convenience). To see this, consider that to the left of M*, with MB > MEC, because B has property right over a noise-free atmosphere, A will be willing to pay more in compensation for a small increase in M than B requires, and so, will pay and play more. But A will not increase M beyond M* because the compensation that would be necessary to pay B would be greater than the worth to A of the small increase thereby attained. In other words, A will pay B up to the point her willingness to pay is just equal to what B is willing to accept in compensation (at point M*).

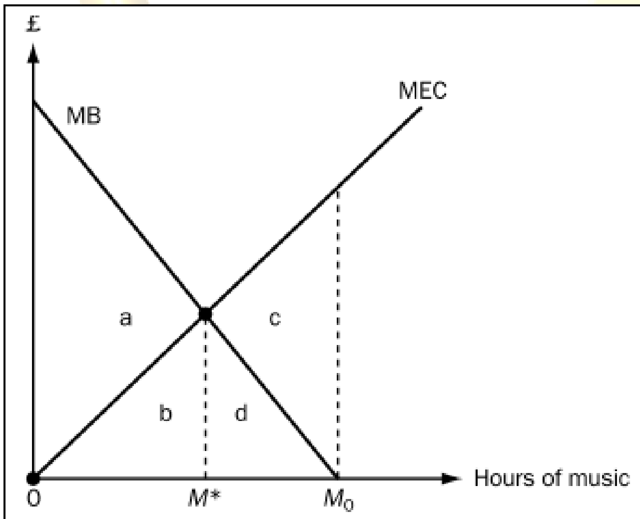


Fig. 2.8 Market allocation with a negative externality in a consumer-to-consumer case. **Source:** Perman et al., 2003. p137.
.

The reduction in M is (M_0 – M*) and A pays B an amount equal to the area of triangle b, which is the money value of B's suffering at the efficient outcome M*. With the assignment of right to B, A would have gotten zero benefits. But by bargaining she is able to secure a net benefit of a (the total benefit from pollution less the amount paid to B). B gets the same zero loss she would have gotten if there was no pollution because she is paid exactly the amount that compensates for the damages suffered. Thus, the transaction is Pareto improving.

The Coase theorem suggests that it really does not matter on efficiency grounds to whom property right is assigned in an externality problem (polluter or victim). Either way an efficient outcome can be achieved through bargaining. If polluter has the legal right to pollute, it would be in the interests of the victim to pay the polluter not to exercise the right to the full if the benefit that would be gained exceeds the payment that has to be made. The only effect the decision to whom to property right is assigned has on the outcome is that it changes the distribution of surplus among the affected parties.

For example, assume that in the present case, property right is assigned to polluter, A. This could happen for example, if a law is passed saying that all saxophone players have an absolute right to practice up to the limits of their physical endurance. This gives A legal right to play as much as she wants. But a legal right can be traded. So, the opportunity now exists for A and B to bargain to a contract specifying the amount that A will actually play. That amount will be M* in Figure 2.8. To see this, note that to the right of M*, MEC > MB, so B's willingness to pay for a small reduction in the hours of saxophone played is greater than the compensation that A requires for that small reduction. Starting at M_0 and considering successive small reductions, B will be offering more than A requires until M* is reached, where B's offer will exactly match the least that A would accept. A and B would not be able to agree on a level of M to the left of M*, since, in that region, B's willingness to pay is less than A requires by way of compensation. In this case, M is reduced by (M_0 − M*) and B pays A an amount equal to the area of triangle d, the money value of A's loss as compared with the no-property-rights situation. A gets the same overall benefit as before while B gets a net benefit of c. With the right assigned to A, the latter would have polluted to the level M_0 and B would have suffered damages to the sum of b, d and c. But with bargaining, B is able to reduce his loss to the sum of b (damage from pollution) and d (the amount paid to A). Thus, the transaction is Pareto improving.

Note that even though the outcomes are the same from the efficiency standpoint, they are not the same from the view point of wealth distribution. The way the property right is assigned affects the wealth of A or B. To be granted a new property right is to have one's potential monetary wealth increased. It is also possible that increases in wealth affect the receiving individual's tastes. This will lead to shifts in the MB and MEC curves and an outcome that is still efficient but different from M* in Fig. 2.8. In addition, the way property right is allocated determines how the cost of obtaining the efficient level of output is shared between the parties. When the property right is assigned to the polluter, the cost is borne by the victim (part of the cost is the damage and part is the payment or bribe to reduce the level of damage). When the property right is assigned to the victim, the cost is borne by the polluter who now must compensate for all damages. We cannot say categorically that either way of assigning property rights necessarily promotes equity because it not always the case that externality sufferers are relatively poor and generators relatively rich, or vice versa.

Again, for the same reasons as outlined above, the situation in scenario II where we had a negative externality of a producer-to-producer sort can be solved by assigning property rights to the downstream sufferer or to the upstream generator. Bargaining could then, in either case, produce an efficient outcome. However, an alternative way of internalizing the externality in that situation would be to have the firms collude so as to maximize their joint profits. In other words, the two firms will be operated as if they were a single firm producing two commodities. Here, the ideal conditions will apply because there will be no impact on the firm's activities the level of which is unintentionally set by others.

The private bargaining situation can be applied to various similar examples. Consider a farm and a ranch next to each other. The rancher's cows occasionally wander over to the farm and damage the farmer's crops. The farmer has an incentive to bargain with the rancher to find a more efficient solution. If it is more efficient to prevent cattle trampling a farmer's field by fencing in the farm, rather than fencing in the cattle, the outcome of the bargaining will be to fence around the farm. Consider another example. Mr. Wafula plants pear trees on his property which is adjacent to that of Mr. Kofi. Mr. Kofi gets an external benefit from Wafula's pear trees because he picks up the pears that fall on the ground on her side of the property line. This is an externality because Mr. Kofi does not pay Mr. Wafula for the utility received from gathering fallen pears. As a result, in response, Mr Wafula can put up a net that will prevent pears from falling on Mr. Kofi's side of the property line, eliminating the externality. Alternatively, Wafula could impose a cost on Kofi if he wants to continue to enjoy the pears from the pear trees. Both parties will be better off if they can agree to the second scenario, as Kofi will continue to enjoy pears and Wafula can increase the production of pears.

Well-defined property rights can solve public goods problems in other environmental areas, such as land use and species preservation. The buffalo neared extinction and the cow did not because cows could be privately owned and reared for profit. Today, private property rights in elephants, whales, and other species could solve the tragedy of their near extinction. In Africa, for instance, elephant populations are growing in Zimbabwe, Malawi, Namibia, and Botswana, all of which allow commercial harvesting of elephants. Since 1979 Zimbabwe's elephant population in these countries has been growing. On the other hand, in countries that ban elephant hunting—Kenya, Tanzania, and Uganda, for example—there is little incentive to breed elephants but great incentive to poach them. In those countries, elephants are disappearing.

Limitations to private solution

Given the simple and compelling logic of the arguments of the Coase theorem, the question arises as to why uncorrected externalities are still a problem. If they exist by virtue of poorly defined property rights and can be solved by the assignment of clearly defined rights, why have legislatures not acted to deal with externality problems? Here are some plausible explanations. First, the case for property rights solutions is entirely an efficiency case. Legislators do not give efficiency criteria the weight that economists do – they are interested in all sorts of other criteria. Secondly, bargaining is costly even in the presence of clearly defined property rights. In addition, the costs increase with the number of participants. In particular, the presence of a larger number of victims complicates the process of negotiation (bargaining). It is the case that while expositions of the Coase theorem deal with small numbers of generators and sufferers (typically one of each), externality problems that are matters for serious policy concern generally involve many generators and/or many sufferers, and are often such that it is difficult and expensive to relate one particular agent's suffering to another particular agent's action. This makes bargaining expensive, even if the necessary property rights exist in law. The costs of bargaining (more generally called 'transactions costs') may be so high as to make bargaining infeasible. Thirdly, aside the large numbers problem, in many cases of interest, the externality has public bad characteristics which preclude bargaining as a solution.

To illustrate the limitations of Coase bargaining in resolving the inefficiency created by the presence of externalities, we now consider the third scenario of a negative externality in a producer-to-consumer case. Here, private bargaining based on some assignment of property rights will not deal with the externality problem. Joint profit maximization solution is also not relevant. Correcting the market failure in this case requires some kind of ongoing intervention in the workings of the market by some government agency. We will look at this later, after a formal consideration of the Coase theorem.

2.3.2.2. Coase Theorem: A formalization

Consider two individuals A and B, with contrasting preferences which are such that A likes to smoke and B likes clean air. Both A and B have the same initial endowment W_0, thus their consumption possibilities (indifference curves) are as shown in Figure 2.9 below: Individual A is better off when he smokes more but B is better off when he gets clean air. If A and B have the same money endowment, their initial location lies on the vertical line EE', so points E and E' are two possible initial endowments. However, the exact point of these two depends on the ownership of property rights. For instance, suppose individual A has the property right to smoke, this means that individual B's initial location is at point E' where he has zero clean air but with his money endowment. But this is not a Pareto efficient point. At point E' individual A has maximum smoke and B has minimum air, B will have to pay (bribe A) to reduce the amount of smoke. Trade between A and B will lead to a reallocation of their endowments to point X' which is Pareto efficient, however, recall that Pareto inefficient allocation is one where one individual is made better off and the other unaffected or both are made better off with a reallocation. In this case both A and B are made better off; A now has less smoke but with more money income and B has more clean air although with less income.

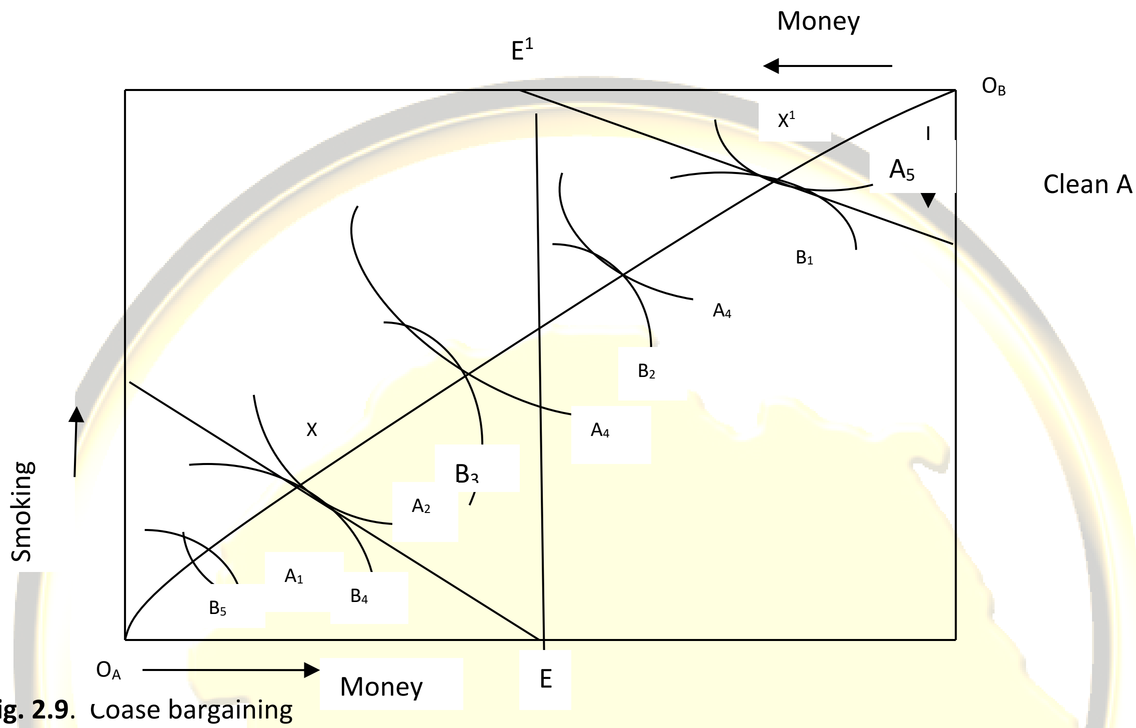


Fig. 2.9. Coase bargaining

The case where B has the property rights, the initial endowment would be at point E (maximum clean air and minimum smoke). So, B has the right to trade some of the clean air for more money, trading will give rise to point X which is Pareto optimal. In this case individual A is better off with, more smoke although less money and B is also better off; more money but less clean air. The exact position on the contract curve depends on the property right ownership: If there are well defined property rights, the good involving an externality no matter who holds the property right, the agent can trade from their initial endowment to a Pareto efficient allocation, hence the "Coarse Theorem."

The same analyses can be illustrated in mathematical form. Suppose B is assigned the property right. In this case, A is unable to engage in the externality producing activity without permission of B. Suppose the bargaining is such that B demands the payment of T in return for permission to generate externality (X). A will agree if he would be at least as well off as he would be by rejecting it (producing nothing), that is,

$$\emptyset_A(X) - T \geq \emptyset_A(0) \qquad (2.60)$$

B will choose her offer to solve the problem

$$\text{Max}\emptyset_B(X) + T, \ X>0$$

$$\text{s.t } \emptyset_A(X) - T \geq \emptyset_A(0) \qquad (2.61)$$

But since the constraint is binding this will imply that $T = \emptyset_A(X) - \emptyset_A(0)$, then by substitution A's optimal offer involves the level of X that solves

$$\text{Max } \emptyset_B(X) + (\emptyset_A(X) - \emptyset_A(0)) \quad (2.62)$$

with F.O.C. $\emptyset_B'(X^*) + \emptyset_A'(X^*) = 0$

implying that $\emptyset_B'(X^*) = (-)\emptyset_A'(X^*)$ (2.63)

This is precisely the socially optimal level.

2.3.2.2 The Public Approach: Pigouvian taxes

This occurs as government intervenes in particular situations where there are external effects. There is a range of means of intervention that the government environmental protection agency (EPA) could use in this regard (see Module 4). Here we consider just taxation as one of the instruments (Pigouvian taxes). We will first apply it to the situation portrayed in Scenario III, where we had a negative externality involving a producer and many consumers, after which we will formalize using a partial equilibrium analytical framework.

We found that in the case of negative externality involving a producer and many consumers, Coase bargaining may not be possible. Government intervention can take the form of a tax on the firm generating the externality which forces the latter to internalize the cost of its actions. The question is how can this be done in a way to achieve allocative efficiency?

Let PMC be the **private marginal co**st to the firm. These are the input costs that the firm bears and actually takes account of in determining the profit-maximizing output level of its product, Y. Thus,

$$PMC = \partial C/\partial Y \quad (2.64)$$

where C = $P_K K^Y$ + $P_L L^Y$ = C(Y), P_K and P_L are price of inputs K and L respectively. Recall that the firm's production function is given as Y = $Y(K^Y$, L^Y, S), with ∂Y/∂S > 0, where S is the pollution that is associated with the production of Y.

Since consumption of the externalities by the two representative consumers, A and B, is non-rival and non-excludable (that is the pollution is a '**public bad'**), the **marginal external cost** (MEC) is the sum of the willingness to pay of each of them (recall the case of public good). Let SMC be the **social marginal cost**. This is the marginal cost of the firm's activities to society (how much the firm's activities cost society at the margin). It consists of the private marginal cost incurred by the firm directly and the marginal external cost imposed on individuals through pollution. Thus

$$SMC = PMC + MEC \quad (2.65)$$

In Figure 2.10, the PMC curve is shown as increasing with the firm's output (Y) in the usual way. The SMC curve has a steeper slope than the PMC curve, so that MEC is increasing with Y: as the production of Y increases, pollution level, S also increases. To maximize profit, the firm produces at Y_0, where PMC is equal to the output price P_Y. But this is not the Y output that goes with efficiency, as in balancing costs and benefits at the margin: it ignores the external costs borne by the representative consumers (A and B). Efficiency is achieved at Y* where SMC equals P_Y.

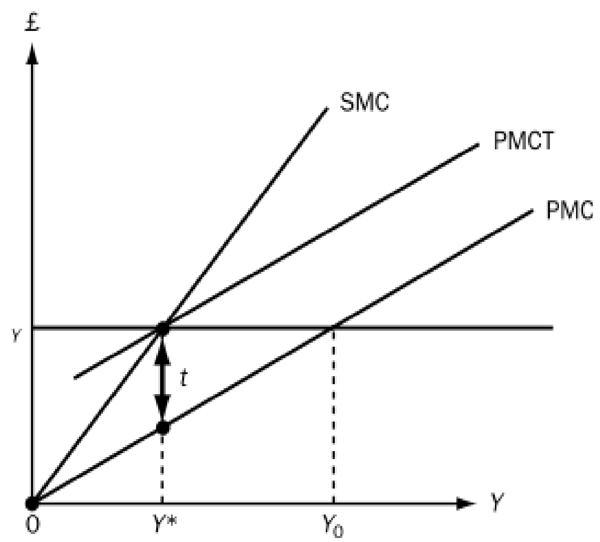


Fig. 2.10 Using taxes to correct for an externality

To correct this market failure, the EPA can tax S at a suitable rate. Let PMCT = private marginal cost to the firm plus the tax imposed. The appropriate tax rate, which will achieve efficiency in this case is

$$t = \text{SMC}^* - \text{PMC}^* = \text{MEC}^* \qquad (2.66)$$

The tax must be equal to marginal external cost at the efficient levels of Y and S. In other words, the optimal tax rate is equal to the MEC at the optimal level of pollution (that is MEC*). Thus, in order to be able to impose taxation of emissions at the required rate, the EPA would need to be able to identify Y*.

Given that prior to EPA intervention what is actually happening is Y_0, identification of Y* and calculation of the corresponding MEC* would require that the EPA knew how MEC varies with S. This would mean that the EPA have to know the utility functions of all consumers represented in A and B. However, given the nature of the case, this information is not revealed in markets. We encounter again the usual problem of preference revelation in regard to public goods/bads. What does this imply for feasible policy in respect of pollution control by taxation? We will examine the question in more details and provide answers in Module 4, where we consider environmental policies. Our analysis here has focused on one output. We can easily extend it to the case where the emissions arise in more than one production activity, say two firms, for example. In such situation, efficiency will require that emissions from both sources be taxed at the same rate, t = MEC*.

As we conclude this Module, note that negative environmental externalities are a major issue in the world today. For example, both U.S. and European estimates of the external costs of automobile use are around 10 cents per mile (see Table 2.3). Converted to damages per gallon of gasoline, the damages are $2.10 per gallon. These estimates suggest that externalities from automobile use in the United States amount to about 3 percent of GDP. Also, a study of global externalities found that in 2009 primary production and processing industries generated $7.3 trillion in unpriced externality damages, equivalent to 13 percent of world economic output (Table 2.4).

Table 2.3. External Costs of Automobile Use, U.S. Cents per Mile, United States and Europe

Cost Category	United States Estimate	Europe Estimate
Climate Change	0.3	3.3
Local Pollution (air and noise)	2.0	0.8
Accidents	3.0	3.7
Oil Dependency	0.6	Not estimated
Traffic Congestion	5.0	Not estimated
Other External Costs	Not estimated	1.2
Total	10.9	9.0

Source: Jonathan and Roach, 2017

Table 2.4: Global Environmental Externalities

Impact Category	Damages
Land use	$1.8 trillion
Water consumption	$1.9 trillion
Greenhouse gases	$2.7 trillion
Air pollution	$0.5 trillion
Land and water pollution	$0.3 trillion
Waste generation	$0.05 trillion

Source: Trucost, 2013. *Note:* Original European estimates were in euros per kilometer. Conversion to cents per mile based on 2016 currency conversion rates. *Sources:* Parry et al., 2007; Becker, et al., 2012. Also, in Jonathan and Roach, 2017.

Summary

- An externality exists when the consumption or production choices of one person or firm negatively or positively affect the utility or production of another entity in an unintended manner. Externalities could be positive (when it increases the welfare of the agent impacted) or negative (when it reduces his or her welfare).

- External effects can be classified according to what sort of economic activity they originate in and what sort of economic activity they impact on. An economic situation involves a consumption externality if one consumer cares directly about another agent's consumption or production. A production externality arises when production possibilities of one firm are influenced by the choices of another firm or consumer.

- Externalities may or may not have the characteristics of a public good, but those that are most relevant for policy analysis exhibit non-rivalry and non-excludability. This is especially the case with external effects that are associated with environmental problems such as pollution.

- As in the case of public goods, the presence of externalities leads to market failure. Given that all of the other institutional conditions for a pure market system to realize an efficient allocation hold, if there is a beneficial externality associated with a good, the market will produce too little of it in relation to the requirements of allocative efficiency. In the case of a harmful externality, the market will produce more of it than efficiency requires.

- Externalities affect resource allocation because the market fails to fully set the price for the external effects generated by some economic activities. This is because market prices reflect the private costs and benefits that agents derive from goods and services while ignoring the costs/benefits imposed on third parties.

- The private approach to solving the externality problem involves affected individuals finding a means of solving the problems of external damage or benefit. According to the Coase theorem, two private parties will be able to bargain with each other and find an efficient solution to an externality problem.

- The Coase theorem suggests that it really does not matter on efficiency grounds to whom property right is assigned in an externality problem (polluter or victim). Either way an efficient outcome can be achieved through bargaining.

- But bargaining is costly even in the presence of clearly defined property rights. The costs also increase with the number of participants. In particular, the presence of a larger number of victims complicates the process of negotiation (bargaining) so that private solution to an externality problem is no longer possible. Most environmental externalities are of this form.

- Public approach to solving externalities involves among others, the use of taxes and subsidies. Pigouvian taxes can be used to correct a negative externality. Such approach requires that the tac rate so that marginal social cost is equal to marginal social benefit. This requires the regulating agency to have knowledge of firms' production functions and consumers utility functions and how these are affected by the externality. But such information is often not available to the authorities.

Review/Discussion Questions and Exercises

1. Provide a mathematical description and explanation of the following problems
   (a) Consumption-to-consumption externality
   (b) Consumption-to-production externality
   (c) Consumption-to-consumption and production externality
   (d) Production -to-consumption externality
   (e) Production -to-production externality
   (f) Production-to-consumption and production externality

2. Establish a possible linkage between natural and environmental resources, public goods/bads and externalities

3. Coase theorem suggest that assignment of property rights can solve externalities problem. Why don't we find government often times establishing legal rights to individuals in cases of externalities?

4. How does transaction cost affect the prospect of finding a private solution to the externality problem? Can you give local examples in your country of externalities involving large number of parties and potentially high transaction costs?

5. Can you suggest what can be done to correct a positive externality?

6. Illustrate how Pigouvian taxes can be used to correct a negative externality. Are there limitations to the use of this instrument in cases of environmental externalities? Why or why not?

Materials used for this Lecture

Gravelle H. and Rees R. (2004). Microeconomics 3rd edition Prentice Hall Edinburg Gate Harlow

Perman, R., Ma Y., McGilvray J. and Common M. (2012). **Natural Resource and Environmental Economics**, 4th edition. Edinburgh, Longman.

Prato T. (1998). **Natural Resource and Environmental Economics**, Iowa State University Press Ames.

Varian H.R. (2010). **Intermediate Microeconomics a Modern Approach**, 8th edition W.W. Norton and Company New York.

Module 2.4 The Second-best Problem, Imperfect Information, and Government Failure (2 hours)

Learning Outcomes

This Module examines the second-best theory and the use of second-best policies in the face of multiple market failures, the role of imperfect information, and the case of government failure in resource allocation, with emphasis on natural and environmental resources. After going through the module, the reader is expected to

- ✓ understand the concept of second-best policy, the rationale, and its relevance.
- ✓ appreciate why second-best policy analyses may be particularly relevant in the case of natural and environmental resources and in the context of Sub-Saharan Africa.
- ✓ understand the concept of public or government failure and how to relate this with natural and environmental resource management.
- ✓ Know the factors behind government failure and the relevance in Sub-Saharan Africa.
- ✓ understand the role of imperfect information in both market and government failure.

Outline

2.4.1 The Second-best Problem

2.4.2 The Role of Imperfect Information

2.4.3 Government failure

Summary

Discussion/Review Questions and Exercises

Materials used for the Lecture

2.4.1 The Second-best Problem

In most real word markets, many of the conditions required for an ideal market system are often lacking. Thus, correction of just one failure, such as an externality, will not lead to efficiency. Consider, for example, where the firms generating external effects sell their outputs in imperfectly competitive markets (that is, they are not price-takers, a reality with many firms operating in the real world). Attempting to correct the externality as if it were the only issue will not necessarily improve matters in efficiency terms. It may even make things worse! What is required is an analysis that takes account of multiple sources of market failure, and of the fact that not all of them can be corrected at the same time or even corrected at all: a “second-best policy”.

A ‘**second-best policy’** is a package of government interventions that do the best that can be done given that not all sources of market failure can be corrected. Consider a case where the firm generating a negative externality (pollution) is a monopolist. The situation that emerges is illustrated Figure 2.11. The profit-maximizing output level of the monopolist is Y_0 (where private marginal cost, PMC, equals marginal revenue, MR_Y) with price P_{Y0}.

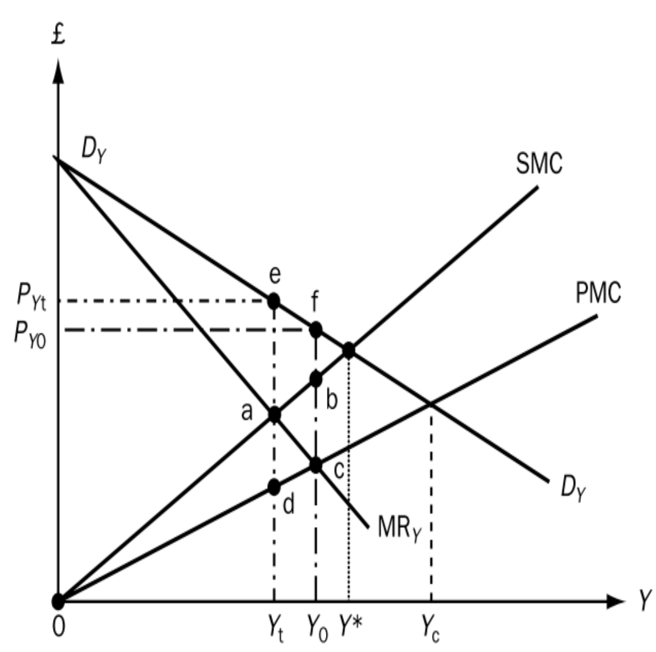


Fig. 2.11. The polluting monopolist, **Source:** Perman et al,. 2003, p 142.

From the point of view of efficiency, there are two problems about the output level Y_0.

- It is too low: because the monopolist sets marginal cost equal to marginal revenue (rather than price as would a perfectly competitive firm). Y_c is the output level that would have been produced under perfect competition.

- It is too high: because the monopolist ignores the external costs generated in the production of Y. If the monopolist had done this it would have produced at Y_t where SMC = MR_Y. The output corresponding with efficiency in this context is Y* where SMC = P_Y.

Now suppose that there is an EPA empowered to tax firms' emissions and that it does this so that for this monopolist producer of Y, SMC becomes the marginal cost on which it bases its decisions. As a result of the EPA action, Y output will go from Y_0 down to Y_t, with the price of Y increasing from P_{Y0} to P_{Yt}. The imposition of the tax gives rise to gains and losses. As intended, there is a gain in so far as pollution damage is reduced – the monetary value of this reduction is given by the area *abcd* in Figure 2.11. However, as a result of the price increase, there is a loss of consumers' surplus, given by the area P_{Yt}efP_{Y0}. It cannot be presumed generally that the gain will be larger than the loss. The outcome depends on the slopes and positions of PMC, SMC and $D_Y D_Y$. In any particular case the EPA would have to have all that information in order to figure out whether imposing the tax would involve a net gain or a net loss.

The analysis has implications for environmental policy to correct market failures caused by factors such as externalities, as in this example. Fundamentally, when dealing with polluting firms that face downward-sloping demand functions, in order to secure efficiency in allocation, the EPA needs two instruments, one to internalize the externality, and another to correct under-production due to imperfect competition (the firms' setting MC = MR rather than MC = P). However, EPAs are not given the kinds of powers that this would require. For example, they can tax emissions, but they cannot regulate monopoly.

Given **complete information** on the cost and demand functions, and on how damages vary with the firm's behaviour, the EPA could figure out a "**second-best tax rate**" to be levied on emissions: one that guarantees that the gains from its imposition will exceed the losses. This will not move the firm to the efficient point, Y* in Figure 2.11 but will guarantee that the equivalent to *abcd* that it induces (the monetary gain from pollution reduction) will be larger than the corresponding equivalent to P_{Yt}efP_{Y0} (the loss from consumer surplus).

The level of the second-best tax rate depends on three main factors. These are the damage done by the pollutant, the firm's costs, and the elasticity of demand for its output. In the case where we have many polluting monopolies to deal with, the EPA would be looking at imposing different tax rates on each (even where all produce the same emissions) based on the different elasticities of demand that they face in their output markets. But charging different firms different rates of tax on emissions of the same stuff is unlikely to be politically feasible, even if the EPA had the information required to calculate the different rates.

It may be interesting to ask why does government not correct the monopoly situation and thus eliminate the associated market failure before, or alongside correcting the negative externality. Well, it is possible to argue that a monopoly situation on its own need not lead to market failure if in the first instance, it was possible for all agents (the monopolist and the consumers) to bargain (remember Coase theorem). But given the very large number of agents involved, bargaining is certainly impossible in this case (the problem of free-riding and prohibitive transaction cost). Thus, the root of the market failure is the impossibility of bargaining due to high transaction cost. In general, complete absence of property right (or defective property right systems) and high transaction cost which makes bargaining infeasible, are major issues at the root of most cases of market failure. The logic of the second-best theory stems from the fact that in practice, it may be impossible to correct some cases of market failure in a way that can achieve allocative efficiency. For example, it may be difficult to dislodge a monopolist or break the monopolistic powers of some firms. Also, accurate information may not be available to secure the right policy decision in some cases. In some other cases, granting subsidy on a good or taxing a product may be considered necessary even though they lead to market failure on their own, etc. In many cases, political-economy factors play a significant role. Some good/right things to do may not be expedient for political-economy reasons!

2.4.2 The Role of Imperfect Information

Efficiency of the market systems require that all agents have complete and perfect information. It turns out that this condition is also required to effectively correct for market failure, including that due to externalities or the public good nature of a resource. Consider, one of the various examples cited in Module 2.3, a consumption-to-consumption externality that occurs where two individuals share a flat and one (A) is a smoker but the other (B) is not. Suppose that B does not find cigarette smoke unpleasant, and is unaware of the dangers of passive smoking. Then, notwithstanding that the government has legislated for property rights in domestic air unpolluted with cigarette smoke, B will not seek to reduce A's smoking. Given B's ignorance, the fact that bargaining is possible is irrelevant. The level of smoke that B endures will be higher than it would be if B were not ignorant. In this case, B's ignorance is the source of an uncorrected externality!

The simple example above is illustrative. There is a lot of ignorance in relation to the existence and value of many environmental and natural resources services. The consequence is that people are not willing to pay for them (you won't pay for what you don't value, lest what you don't even know about!).

The nature of the corrective policy in the case of imperfect information is clear – the provision of information. In many cases, the information involved will have the characteristics of a public good, and there is a role for government in the provision of accurate information. In some cases, the government cannot fulfill this role because it also does not have accurate and unambiguous information! Particularly where it is the future consequences of current actions that are at issue, as for example, in the case of global warming, it may be simply impossible for anybody to have complete and accurate information. But future consequences of current actions become particularly important in circumstances where those actions have irreversible consequences, as it is with the use of many environmental resources. For example, once developed, a natural wilderness area cannot be returned to its natural state).

2.4.3 Government failure

Government intervention offers the possibility of realizing efficiency gains, by eliminating or mitigating situations of market failure. Efficiency gains may be obtained if government can create and maintain appropriate institutional arrangements for establishing and supporting property rights as the basis for bargaining. However, the scope of this kind of government action to correct market failure is limited to cases where rivalry and excludability are present. Many environmental problems do involve non-rivalry and non- excludability. In such cases, possible government interventions to correct market failure are often classified into two groups: command- and-control instruments, which are rules and regulations prohibiting, limiting or requiring certain forms of behaviour, and fiscal instruments, including tax and subsidy systems, and marketable permits designed to create appropriate patterns of incentives on private behaviour. In Module 4, we discuss these interventions under the heading of Environmental Policies. Government intervention can also take the form of information provision or funding research activity that can increase the stock of knowledge.

But actual government intervention does not always or necessarily realize the efficiency gains required and may entail losses. Not all government intervention in the functioning of a market economy is either desirable or effective. As we saw in the case of second-best choices, the removal of one cause of market failure does not necessarily result in a more efficient allocation of resources if there remain other sources of market failure. In addition, government intervention may itself induce economic inefficiency. Poorly designed tax and subsidy schemes, for example, may distort the allocation of resources in unintended ways, or may simply fail to achieve desired outcomes. For example, attempt by the Greek government to reduce car usage, and hence, congestion and pollution in Athens by prohibiting entry into the city on particular days by cars with particular letters on their license plates served to promote the purchase of additional cars by households wishing to maintain freedom of mobility in the city!, Similarly, the use of quantity controls in fisheries policy (such as determining minimum mesh sizes for nets, maximum number of days of permitted fishing, required days in port for vessels, etc.) intended to address the free-access problem of overexploitation, have met with very little success. Fishermen responded by making behavioural adjustments to minimize their impact.[17]

Of equal importance is the realization that actual government interventions may not be always motivated by efficiency, or even equity, considerations. Special interest groups use the political process to engage in what has become known as rent seeking (the use of resources in lobbying and other activities directed at securing protective legislation). Successful rent-seeking activity will increase the net benefits going to the special interest group, but it will also frequently lower the surplus to society as a whole (a classic case of the aggressive pursuit of a larger slice of the pie leading to a smaller pie). Political economists believe that in a democratic setting, voters,

[17] Perman et al., 2003, p.144-145.

elected members of the legislature, workers in the bureaucracy, and pressure groups are all driven by a form of self-interest. Voters are assumed to vote for candidates they believe will serve their own interests; legislators are assumed to maximize their chances of re-election; bureaucrats are assumed to seek to enlarge the size of the bureaucracy, so improving their own career prospects, while pressure groups push special interests with politicians and bureaucrats. Given these motivations and circumstances, the outcome of government intervention is not likely to be a set of enacted policies that promote either efficiency or equity. The concepts of "**regulatory capture"** and "**government capture**" are used to describe situations where the government or its regulatory agency has been hijacked to serve some vested interest other than that of society.

Consider the example in Module 2.3 where we had a producer-to-producer externality involving a steel producing firm and a hospitality firm. The steel producing firm generates negative externality for the hospitability firm. But assume that the government, for reasons of national security, decides to subsidize the production of steel. The outcome is illustrated in Figure 2.12. The private marginal cost curve (PMC) to the steel-producing firm shifts down and to the right causing a further increase in production, lower prices, and even more pollution. Thus, the subsidy moves us even further away from where surplus is maximized at Q*. The shaded triangle A shows the deadweight loss (inefficiency) without the subsidy. With the subsidy, the dead weight loss grows to areas A + B + C. The social policy has the side effect of increasing an environmental inefficiency.

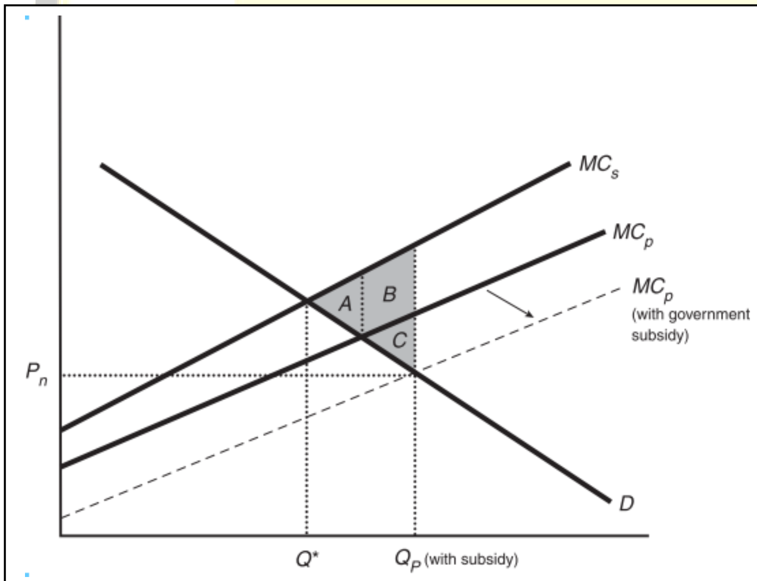


Fig. 2.12. Effect of subsidy on a polluting product. **Source**

The example is typical of many other cases, including subsidies on petroleum products, which is often used in many countries. Political-economy factors and lack of (or imperfect) information can help explain the prevalence of government failure in many context

Most economic arguments for government intervention are based on the idea that the market place cannot provide public goods or handle externalities. But markets often solve public goods and externalities problems in a variety of ways. Businesses frequently solve free-rider problems by developing means of excluding nonpayers from enjoying the benefits of a good or service. Cable television services, for instance, scramble their transmissions so that nonsubscribers cannot receive broadcasts. Both throughout history and today, private roads have financed themselves by charging at tolls to road users. Other supposed public goods, such as protection and fire services, are frequently sold through the private sector on a fee basis.

Public goods can also be provided by being tied to purchases of private goods. Shopping malls, for instance, provide shoppers with a variety of services that are traditionally considered public goods: lighting, protection services, benches, and rest-rooms, for example. Charging directly for each of these services would be impractical. Therefore, the shopping mall finances the services through receipts from the sale of private goods in the mall. The public and private goods are "tied" together.

The imperfections of market solutions to public goods and externalities problems must be weighed against the imperfections of government solutions. As noted above, governments rely on bureaucracy and have weak incentives to serve consumers. Furthermore, politicians may supply public "goods" in a manner to serve their own interests, rather than the interests of the public. Examples of wasteful government spending and pork-barrel projects are legion. Government often creates a problem of "forced riders" by compelling persons to support projects they do not desire. Private solutions to public goods problems, when possible, are usually more efficient than governmental solutions.

Again, the GERD Project discussed in Box 2.7 in Module 2.3 may help illustrate multiple problems involving the public provision of a public good. It should also be noted that governments and their agencies in Sub-Saharan Africa are often handicapped by information limitations and gaps in technical capacities. Weak governance, including limited accountability and ineffective rule of law, may make governments unable to take action against some agents where there is the need to correct for a market failure or pursue a social goal. All of these factors may hinder the effectiveness of governments in achieving even second-best policies.

Summary

- A ‘second-best policy’ is a package of government interventions that do the best that can be done given that not all sources of market failure can be corrected. The logic of the second-best theory stems from the fact that in practice, it may be impossible to correct some cases of market failure in a way that can achieve allocative efficiency.

- Efficiency of the market systems require that all agents have complete and perfect information. It turns out that this condition is also required to effectively correct for market failure, including that due to externalities or the public good nature of a resource.

- In some cases of market failures, it is possible for government intervention through the creation of property rights to provide an effective solution. In most cases of market failure involving environmental externalities of the nature of a public good/bad, government intervention can take the form of rules and regulations prohibiting, limiting or requiring certain forms of behaviour, and/or fiscal instruments (including tax and subsidy systems, and marketable permits) designed to create appropriate patterns of incentives on private behaviour.

- But actual government intervention does not always or necessarily realize the efficiency gains required and may entail losses. In many cases government may not have the relevant information to effectively correct for market failures, especially those involving natural and environmental resources. In addition, actual government interventions may not always be motivated by efficiency, or even equity, considerations.

- Political economists believe that in a democratic setting, voters, elected members of the legislature, workers in the bureaucracy, and pressure groups are all driven by a form of self-interest Given these motivations and circumstances, the outcome of government intervention is not likely to be a set of enacted policies that promote either efficiency or equity. Political-economy factors and lack of (or imperfect) information can help explain the prevalence of government failure in many context

- The imperfections of market solutions to public goods and externalities problems must be weighed against the imperfections of government solutions.

Discussion/Review Questions and Exercises

1. Provide a discussion of the second-best theory using an example drawn from your community and involving a natural or environmental resource problem and employing relevant graphical illustrations.

2. Assess the relevance of the second-best theory in the context of Sub-Saharan African countries.

3. A firm producing good X and located in an urban city is one of two firms operating in the industry. Production activities are associated with emission of dangerous chemical into the atmosphere. What are the options and constraints available to the Environmental Protections Agency in addressing this problem?

4. What possible options are available for government to intervene in the case of market failure involving environmental externality that takes the form of
   (a) private bad
   (b) a public bad?

5. Explain the role of imperfect information in market and government failure with an application to natural and environmental resources

6. Do you agree that market failure is most likely to be a more serious problem in Sub-Saharan Africa? Why or why not?

Materials used for this lecture

Gravelle H. and Rees R. (2004). **Microeconomics** 3rd edition Prentice Hall Edinburg Gate Harlow

Perman, R., Ma Y., McGilvray J. and Common M. (2012). **Natural Resource and Environmental Economics,** 4th edition**.** Edinburgh, Longman.

Prato T. (1998). **Natural Resource and Environmental Economics**, Iowa State University Press Ames.

Module 3.1. Optimal Extraction of Non-Renewable Resources: The Basic Model (6 hours)

Learning objectives

This Module introduces the reader to the basic model of optimal extraction of non-renewable resources, such as minerals. After going through the module, you should be able to

- ✓ better understand the concept of non-renewable resources and appreciate the distinctions between alternative measures of resource stock, such current reserves potential reserves.
- ✓ understand and apply the concept of dynamic efficiency in relation to non-renewable resources.
- ✓ Know the concept of 'user cost', or 'scarcity value' and its importance in the optimal management of non-renewable resources.
- ✓ Know the implications of dynamic efficiency conditions for the extraction and price profile of a non-renewable resource under perfectly competitive conditions and be able to compare these with outcomes under monopoly.
- ✓ understand Hotelling Rule, its intuition and how it is applied to non-renewable resources under perfectly competitive and monopoly situations.
- ✓ know the difference between dynamic efficiency and intertemporal fairness or sustainability, and how the latter can be reconciled with the former.
- ✓ understand the role of Hartwick Rule in achieving both efficiency and sustainability.
- ✓ Know why observed global trend in the extraction and price profile of non-renewable resources may differ from the predictions of the basic model of dynamic efficiency.

Outline

3.1.1 Introduction: Some useful concepts and categorization of non-renewable resources
3.1.2 Optimal extraction of non-renewable resources with constant extraction costs
    3.1.2.1 The Simple Mathematics of Dynamic Efficiency
    3.1.2.2 Optimal extraction in a simple two-period model
    3.1.2.3 A Graphical solution to the dynamic efficiency problem
    3.1.2.4 Dynamic efficiency and the Hotelling Rule
    3.1.2.5 Dynamic efficiency and Hotelling rule under monopoly conditions
3.1.3 Dynamic efficiency, intertemporal fairness, and Hartwick rule
3.1.4. The Empirical Evidence

Summary

Review/Discussion Questions and Exercises

Materials used for the lecture

3.1.1. Introduction: Some useful concepts and categorization of non-renewable resources

Our focus in this Module is on the optimal management of non-renewable resources and hence depletable or exhaustible by nature. A good example in this case is mineral resources. Three separate concepts are used to classify the stock of such resources.

- **Current reserves**: these are defined as known resources that can profitably be extracted at current prices. The magnitude can be expressed as a number.
- **Potential reserves**: these are the reserves or amount of the resource that is potentially available. Because they depend on a number of factors, potential reserves are most accurately defined as functions rather than numbers.
- **Resource endowment**: represents the natural occurrence of the resource in the earth's crust.

Data on current reserves of a non-renewable resource should not be confused with the maximum potential reserves. The potential reserve depends, among others, on the price people are willing to pay for it—the higher the price, the larger the potential reserves. For example, the amount of additional oil that could be recovered from existing oil fields by enhanced recovery techniques, shows that as price is increased, the amount of economically recoverable oil also increases. While the potential reserves of a non-renewable resource depend on the price, the resource endowment does not. The latter is more of a geological, rather than an economic, construct. It represents an upper limit on the availability of the resource. We should also not assume that the entire resource endowment can be made available as potential reserves at a price people would be willing to pay. If an infinite price were possible, the entire resource endowment could be exploited. However, an infinite price is not likely. The currently demonstrated economic reserves may be only a small portion of the ultimately recoverable reserves. Box 3.1 illustrates the various concepts and categorization of a non-renewable resource. The implication of this standard resource classification diagram is that current reserves are not fixed but variable, and can be expanded in two ways: by new discovery or by new technology and changing market conditions. They can also be depleted through extraction.

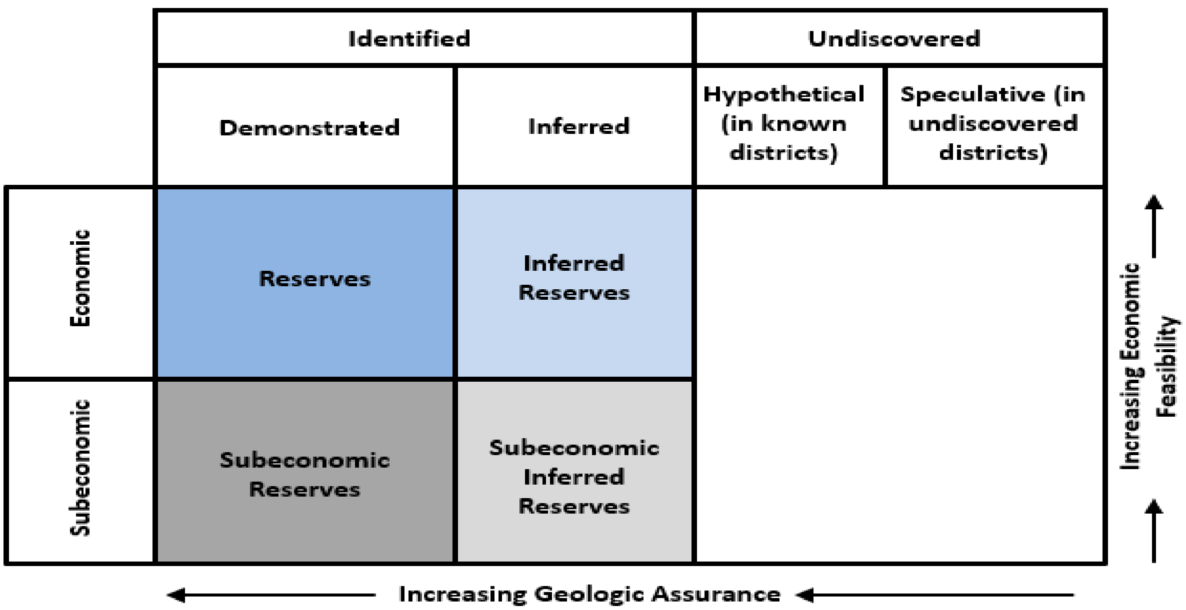
3.1.2 Optimal extraction of non-renewable resources with constant extraction costs[18]

If we are to judge the adequacy of market allocations, we must define what is meant by efficiency in relation to the management of a non-renewable resource. Since, the critical issue is the allocation of the resource over time, **dynamic efficiency** becomes the core concept to apply. This requires going beyond the condition required for **static efficiency** learnt in Module 2. The dynamic efficiency criterion assumes that society's objective is to maximize the present value of

[18] The assumption of a constant extraction cost includes also the possibility of xero cost.

net benefits coming from the resource. For a depletable resource, this requires a balancing of the current and subsequent uses of the resource.

Box 3.1 Classification of Nonrenewable Resources



Some Key terms

Identified resources: specific bodies of mineral-bearing material whose location, quality, and quantity are known from geological evidence, supported by engineering measurements.

Inferred resources: material in unexplored extensions of demonstrated resources based on geological projections.

Undiscovered resources: unspecified bodies of mineral-bearing material surmised to exist on the basis of broad geological knowledge and theory.

Hypothetical resources: undiscovered materials reasonably expected to exist in a known mining district under known geological conditions.

Speculative resources: undiscovered materials that may occur in either known types of deposits in favorable geological settings where no discoveries have been made, or in yet unknown types of deposits that remain to be recognized.

Sources: Rocky Mountain Institute, http://www.rmi.org/RFGraph-McKelvey_diagram_for_coal_gas_resources. Also, in Jonathan and Roach, 2017; U.S. Bureau of Mines and the U.S. Geological Survey, "Principle of the Mineral Resource Classification System of the U.S. Bureau of Mines and the U.S. Geological Survey." GEOLOGICAL SURVEY BULLETIN, 1976, pp. 1450-A. Also, in Jonathan and Roach, 2017.

We present below the basic model of dynamic efficiency in the extraction of a non-renewable resource. We shall also show that the static efficiency is a special case of dynamic efficiency.

3.1.2.1 The Simple Mathematics of Dynamic Efficiency

Assume, for simplicity, that the demand curve for the non-renewable resource is linear and stable over time. Thus, the inverse demand curve at any given time (say year *t*) can be written as

$$P_t = a - bq_t \qquad (3.1)$$

where P_t is the price of the resource at time t, q_t is the quantity of the resource extracted (and offered for sale) at time *t* and a and b are positive constants.

The total benefits from extracting an amount q_t in year *t* are then the integral of this function (the area under the inverse demand curve), and can be written as

$$TB_t = \int_0^{q_t} (a - bq)dq \qquad (3.2)$$

$$= aq_t - \frac{b}{2}{q_t}^2 \qquad (3.3)$$

Let us assume further that the marginal cost of extracting the resource is a constant, *c,* so that the total cost of extracting any amount q_t in year *t* can be written as

$$TC_t = cq_t \qquad (3.4)$$

If the total available amount of this resource is $\bar{Q}$, then the dynamic allocation of the resource over *n* years is the one that satisfies the maximization problem:

$$\text{Max} \sum_{i=1}^{n} \frac{aq_i - b{q_i}^2/2 - cq_i}{(1+r)^{i-1}} + \lambda \left[\bar{Q} - \sum_{i=1}^{n} q_i\right] \qquad (3.5)$$

where *r* is the interest or discount rate.

Assuming that $\bar{Q}$ is less than would normally be demanded; in other words, **the non-renewable resource is scarce**, the dynamic efficient allocation must satisfy

$$\frac{(a-bq_i) - c}{(1+r)^{i-1}} - \lambda = 0, \; i = 1 \ldots n \qquad (3.6)$$

and

$$\bar{Q} = \sum_{i=1}^{n} q_i \qquad (3.7)$$

An implication of (3.6) is that (P_t – c) increases over time at a rate equals to *r* (the interest or discount rate). The difference between the price and marginal cost at any given time is given by the value of λ and is known as the **marginal user cost** (MUC).

The MUC plays a key role in our thinking about allocating depletable resources over time. In practical terms equation 3.6 says that the present value of the marginal net benefit (PVMNB) from the last unit extracted of the resource in period i-1 must be equal to the PVMNB in period *i*. These points are demonstrated more clearly below using a simple two-period model.

3.1.2.2 Optimal extraction in a simple two-period model

Assume that the fixed quantity of the resource can be allocated over two periods, and that demand is constant in the two periods, and given as

$$P = 8 - 0.4q \qquad (3.8)$$

Marginal cost is assumed constant at $2. To illustrate the points made in the preceding section, we consider two scenarios: the first in which there is no scarcity, and the second in which the resource is scarce in the economic sense.

Scenario I

Assume total supply, $\bar{Q}$, is ≥30 and period two is the terminal period. In this case an efficient allocation would produce 15 units in each period (q_1 = q_2 = 15). This follows from the simple application of the condition required for static efficiency learnt in Module 2, that is, equating marginal revenue from the resource in each period to the marginal cost (MR=MC). The result is illustrated in figure 3.1. Let us note the characteristics of this result. First, the result is the same as what will obtain under static efficiency (if the two periods are not related). In other words, time does not really matter in this case (discounting is irrelevant). The reason for this is because **there is no scarcity**. Thirty units is sufficient to cover the demand in both periods under efficiency conditions; so that consumption in period 1 does not impose any constraint on consumption in period 2; the allocations are not interdependent.

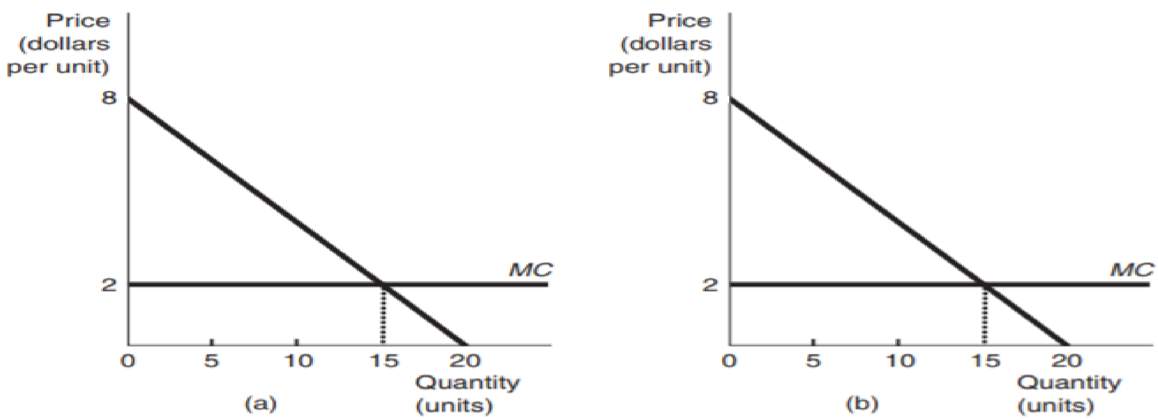


Fig. 3.1 The optimal allocation of an abundant depletable resource: (a) Period 1 and (b) Period 2. **Source:** Tietenberg and Lewis, 2012. p105)

Secondly, because demand and costs are identical in both periods and there is no scarcity, the efficient allocation over the two periods requires that $q_1 = q_2$. In other words, we have a constant extraction profile over the two periods.

Scenario II

Now assume that $\bar{Q} < 30$ (e.g. $\bar{Q} = 20$) but all other conditions remain the same. According to the dynamic efficiency criterion, and as implied in equation (3.5), the efficient allocation is the one that maximizes the present value of net benefits (PVNB), that is the sum of the present values in each of the two periods. Since $\bar{Q}$ is less than what would normally be demanded under static efficiency conditions, equations (3.6) and (3.7) are binding. In other words, the solution to (3.5) must satisfy the two equations

If we assume that *r* = 0.10 or 10%, using the information provided, the solution to the maximization problem yields

$$\frac{8-0.4q_1-2}{(i+0.1)^0} - \lambda = 0;\ 6 - 0.4q_1 = \lambda \qquad (3.9)$$

for the first period (*t* =1), and

$$\frac{8-0.4q_2-2}{(1+0.1)^1} - \lambda = 0;\ 5.465-0.364q_2 = \lambda \qquad (3.10)$$

for period *t* = 2.

Equations (3.9) and (3.10) satisfy the requirement in (3.6). To fulfill the requirement in (3.7), we must have

$$q_1 + q_2 = 20 \qquad (3.11)$$

Equating (3.9) and (3.10) gives

$$0.4q_1 - 0.364q_2 = 0.545 \qquad (3.12)$$

Solving (3.11) and (3.12) yields

$$q_1 = 10.242 \qquad (3.13)$$

$$q_2 = 9.756 \qquad (3.14)$$

Equations (3.13) and (3.14) gives us the dynamic efficient allocation of the non-renewable resource over the two periods. We can use the information to solve for the equilibrium values (dynamically efficient levels) of p_1 and p_2 as follows

$$p_1 = 8 - 0.4\,(10.242) = 3.9 \qquad (3.15)$$

$$p_2 = 8 - 0.4\,(9.758) = 4.097 \qquad (3.16)$$

We can also derive the marginal user cost (MUC = λ) in each period. It is much easier to do this for period 1 because the discount factor for the period is 1. Substituting (3.13) into (3.9) gives us

$$\lambda_1 = 1.903 \qquad (3.17)$$

Thereafter, we can find period 2's λ as the one-period future value of period 1's λ at the given interest rate. In other words,

$$\lambda_2 = \lambda_1(1 + 0.1)^{i-1} \qquad \text{(where i =2)}$$

$$\lambda_2 = 1.903\ (1.1) = 2.093 \qquad (3.18)$$

These results are quite intuitive. We note the following points

- Whenever a resource is scarce (remember in this case, $\bar{Q}$ = 20 which is less than what would normally be demanded in both periods under static conditions), the marginal user cost (MUC) will be positive (λ > 0). We can, in fact, interpreter the situation in scenario I where there was no scarcity to be a case where λ= 0 so that in each period

$$\frac{8-0.4q_1-2}{(i+0.1)^0} - \lambda = \frac{8-0.4q_1-2}{(i+0.1)^0} - 0 \qquad (3.19)$$

  In this case, efficiency just required setting $8 - 0.4q_1$ = 2.

- Thus, the MUC in each period (λ_t) is a reflection of the **scarcity value** of the resource in that period. Note that this is just the difference between price and marginal cost for the period. For example, in the second-case scenario, we have

$$\lambda_1 = p_1 - c = 3.903 - 2 = 1.903 \qquad (3.20)$$

$$\lambda_2 = p_2 - c = 4.097 - 2 = 2.097 \qquad (3.21)$$

  Thus, we can interpret the MUC in terms of the **scarcity rent** of the resource for each period. In the absence of scarcity, the scarcity rent will be zero.

Our analyses show that in the absence of scarcity, and given the assumptions of constant demand and constant marginal cost, dynamic efficiency in the use of the non-renewable resource requires that the resource be extracted at a fixed rate and that the price remains constant. However, with scarcity,

- the price of the resource increases over time ($p_2 > p_1$): the price profile rises over time,
- the extraction of the resource decreases over time ($q_2 < q_1$): the extraction profile decreases over time,
- the MUC (scarcity value of the resource or scarcity rents) rises over time at the rate of interest ($\lambda_2 > \lambda_1$), so that
- the present value of marginal user cost (PVMUC) remains constant over time.

- the higher the interest rate, the greater will be the increases in the MUC (scarcity value) of the resource over time, the more will extraction be tilted towards the present rather than the future, and the higher will be the future price.

Scarcity of resources is a fact of real life, not less for non-renewable resources. Intertemporal scarcity imposes an opportunity cost on using a depletable resource (the MUC). When resources are scarce, greater current use diminishes future opportunities. The MUC is the present value of these forgone opportunities at the margin. Specifically, uses of those resources, which would have been appropriate in the absence of scarcity, may no longer be appropriate once scarcity is present. Thus, an efficient market would have to consider not only the marginal cost of extraction for this resource but also the marginal user cost (MUC), the opportunity cost of extracting an extra unit of the resource today (in this period) rather than tomorrow (in the next period). Whereas in the absence of scarcity, price is equal to marginal cost (P = MC), with scarcity, price is equal to the sum of marginal cost and the marginal user cost (P = MC + MUC).

3.1.2.3 A Graphical solution to the two-period dynamic efficiency problem

We can also use a graphical approach to solve the dynamic efficiency problem in the example above where the non-renewable resource is scarce. The solution is illustrated in Figure 3.2. The two lines depict the PVMNB for each of the two periods. The PVMNB curve for Period 1 ($PVMNB_1$) is to be read from left to right while that for period 2 ($PVMNB_2$) goes from right to left. (The zero axis for the period 2 net benefits is on the right, rather than the left, side.). All points along the horizontal axis yields a total of 20 units (the sum of the two allocations is always 20). The left-hand axis represents an allocation of all 20 units to Period 2, and the right-hand axis represents an allocation entirely to Period 1. Any point on the axis picks a unique allocation between the two periods.

The present value of the marginal user cost (that is, the additional value created by scarcity) is graphically represented by the vertical distance between the quantity axis and the intersection of the two curves. It is identical to the present value of the marginal net benefit in each of the periods. Note that this value is higher for period 2 because of the positive discount factor. Both the size of the marginal user cost and the allocation of the resource between the two periods is affected by the discount rate. Because of discounting, the efficient allocation allocates somewhat more to Period 1 than to Period 2. A discount rate larger than 0.10 would mean rotating the Period 2's curve an appropriate amount toward the right-hand axis, holding fixed the point at which it intersects the horizontal axis. (Can you explain why?). The larger the discount rate, the greater the amount of rotation required. Thus, the amount allocated to the second period would be necessarily smaller with larger discount rates.

Fig. 3.3 illustrates the equilibrium that emerged from the dynamic efficiency criteria in this case of resource scarcity (compare this with equation 3.1).

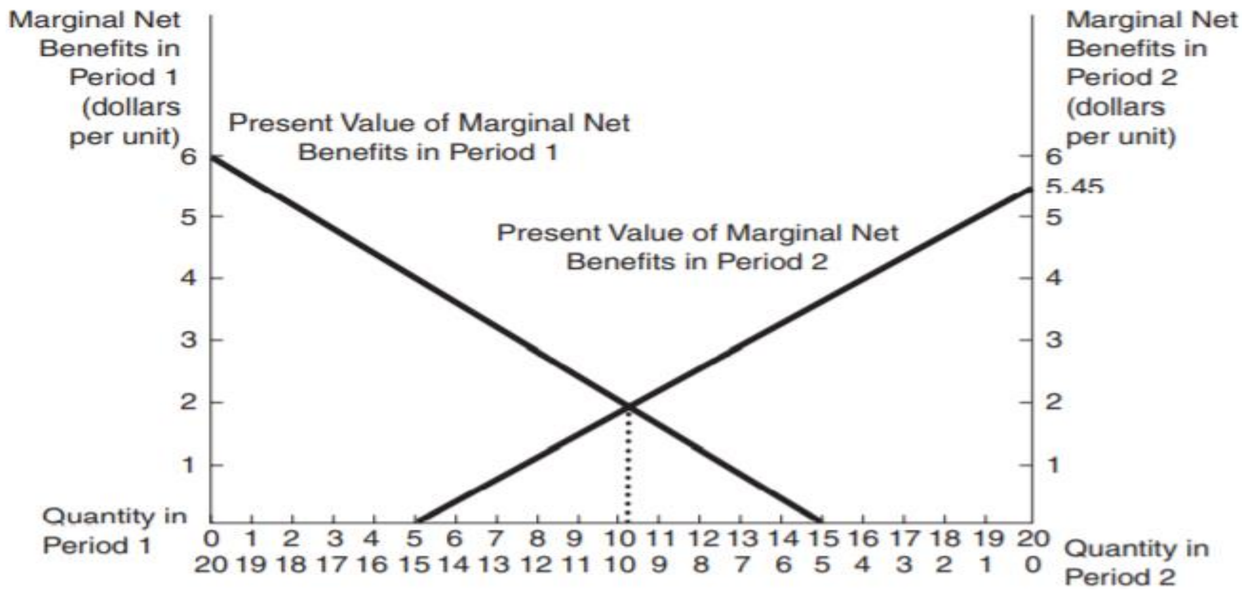


Fig. 3.2 Graphical solution to the dynamic efficient allocation in a two-period model with resource scarcity. **Source:** Tietenberg and Lewis, 2012. p105).

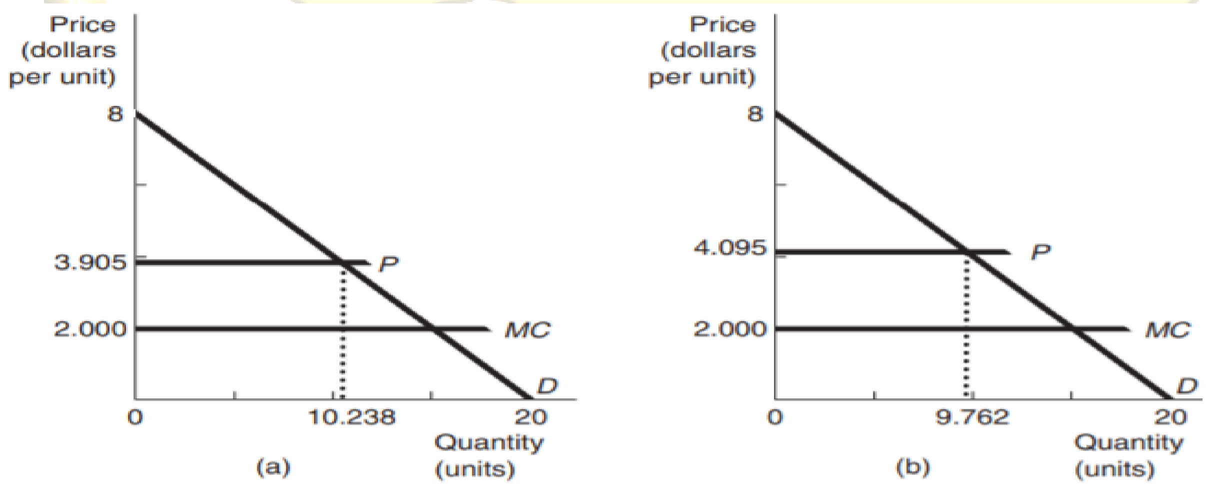


Fig. 3.3 The Efficient Market Allocation of a Depletable Resource with Constant-Marginal Cost: (a) Period 1 and (b) Period 2

Note that the dynamic efficiency criterion is independent of an institutional context. It is equally appropriate for evaluating resource allocations generated by markets, government rationing, or even the whims of a dictator. However, while any efficient allocation method must take scarcity into account, the details of precisely how that is done depend on the context.

3.1.2.4 Dynamic efficiency and the Hotelling Rule

The dynamic efficiency conditions maximize the sum of the present value of marginal net benefits (PVMNB). In our two-period model, this can be written as

$$\text{PVMNB}_T = \text{MNB}_1 + \left(\frac{1}{(1+r)}\right)\text{MNB}_2 \qquad (3.22)$$

Where PVMNB_T = present value of total marginal net benefit and $\left(\frac{1}{(1+r)}\right)$ is the discount factor for period 2.

This implies a trade-off between q_1 and q_2. The optimal output is chosen such that changes in the two periods' marginal net benefit s are equal,

$$\Delta\text{MNB}_1 = \left(\frac{1}{(1+r)}\right)\Delta\text{MNB}_2 \qquad (3.23)$$

Ignoring the Δ component, equation (3,23) can also be expressed as

$$\text{P}_1\text{- MC}_1 = \left(\frac{1}{(1+r)}\right)\text{P}_2\text{-MC}_2 \qquad (3.24)$$

As already noted, the implication of the dynamic efficiency conditions for the time profiles of production and price is that the latter must be rising while the former must be decreasing. From (3.24), it is obvious that since $r > 0$ (positive interest rate), we must have (P_1- MC_1) < (P_2- MC_2) and if $MC_1 = MC_2 = c$ as assumed, then $P_1 < P_2$ and $q_1 > q_2$.

An equivalent statement of (3.24) is

$$\text{Rent}_1 = \left(\frac{1}{(1+r)}\right)\text{Rent}_2 \qquad (3.25)$$

In other words, the optimal extraction path is one that yields rents that rises at the same rate as the discount rate. In this perspective, the depletable resource is viewed as a capital asset. An efficient extraction profile is the one that causes the value of the asset to appreciate at the same rate as other forms of capital (man-made and financial capital) that owners may own in their portfolio. Finding the efficient extraction rate is the same thing as maximizing the portfolio of capital assets. These results generalize into what is traditionally called the Hotelling Rule. Hotelling (1931) shows that under perfectly competitive conditions, profit is maximized when the portfolio price of a resource rises at the same rate as the interest (or discount) rate. In other words, Hotelling Rule requires that

$$\frac{\partial P/\partial t}{P(t)} = \frac{\dot{P}}{P(t)} = r \qquad (3.26)$$

- If $\frac{\dot{P}}{P(t)}$ > *r*, profit is not being maximized; the extraction rate should be increased.
- *If* $\frac{\dot{P}}{P(t)}$ < *r*, profit is not being maximized; the extraction rate should be reduced.

The efficient extraction rate that maximizes profit (the sum of marginal net benefits) is the one that fulfills (3.26).

A second way to interpret the Hotelling Rule is to rewrite (3.26) as

$$P_t^* = \left(\frac{1}{(1+r)^t}\right) P_t \qquad (3.27a)$$

where P_t^* is the discounted value of P_t. In continuous time, (3.27a) can be written as

$$P_t = P_0 e^{rt} \qquad (3.27b)$$

Equations (3.27a) and (3.27b) state that the discounted price of the natural resource is constant along an efficient resource extraction path. In other words, Hotelling Rule requires that the discounted value of the resource should be the same at all dates. The result is merely a special case of a general asset–efficiency condition; the discounted (or present value) price of any efficiently-managed asset will remain constant over time. This way of interpreting Hotelling Rule shows that there is nothing special about a non-renewable natural resources *per se* when it comes to thinking about efficiency. The resource is just like any other asset, and all efficiently-managed assets will satisfy the condition that their discounted prices should be equal at all points in time. Firms that produce non-renewable resources will view it just like any other viable asset and manage it in a way that maximizes returns over time. However, unlike ordinary competitive firms, they will not produce where price is equal to marginal cost (P = MC), but will take account of the scarcity rent or user cost and balance the resource rent available from producing today with the expected rent available from producing in the future. In so doing, high-quality resources, with high recoverable rents, will be exploited first, while lower-grade resources may be held in reserve.

In theory, the non-renewable resources should face eventual exhaustion (given its depletability). The price of the resource should rise over time (in accordance with Hotelling's rule), while the quantity extracted should gradually decline, reaching zero when the "**choke price**", or highest possible market price for the resource, is reached. However, the complete life cycle of a non-renewable resource includes an early period of declining prices and increasing consumption, followed eventually by rising prices and decreasing consumption (Figure 3.4). In Table 3.1 we show the global production level, the global reserves, and the expected lifetime for some non-renewable (mineral) resources in 2015.

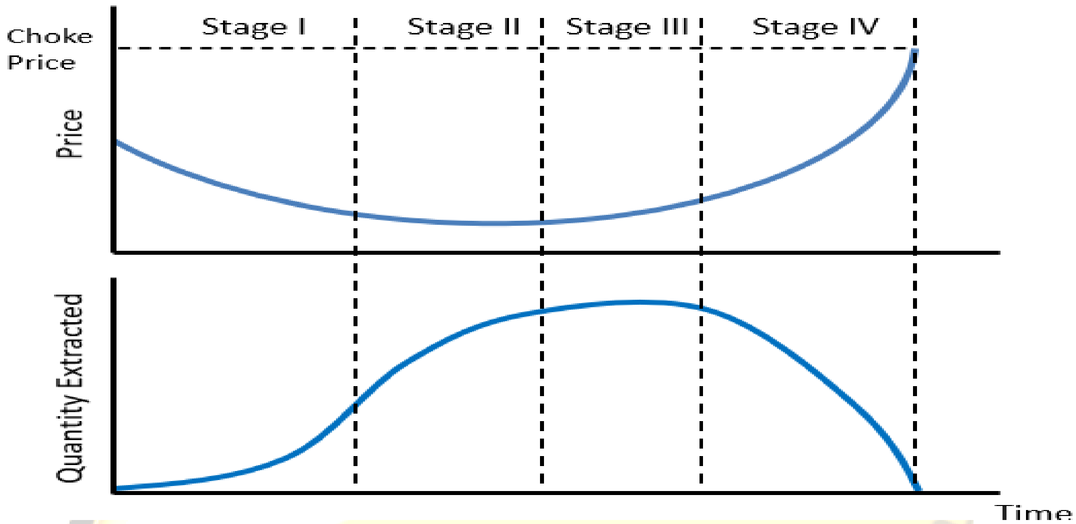


Fig. 3.4 Hypothetical non-renewable resource use profile. Source: Adapted from Hartwick and Olewiler, 1998. Also, in Jonathan and Roach, 2017

Table 3.1. Expected Resource Lifetimes, Selected Minerals

Mineral	2015 global production (thousand metric tons)	Global reserves (thousand metric tons)	Expected resource lifetime, years (static reserve index)
Aluminium	274,000	28,000,000	102
Cobalt	124	7,100	57
Copper	18,700	720,000	39
Iron ore	3,320,000	85,000,000	26
Lead	4,710	89,000	19
Lithium	32.5	14,000	431
Mercury	2.34	600	256
Nickel	2,530	79,000	31
Tin	294	4,800	16
Tungsten	87	3,300	38
Zinc	13,400	200,000	15

Source: U.S. Geological Survey, Minerals Commodities Summaries, 2016. Also, in Jonathan and Roach 2017. Note: Aluminium data for bauxite ore, the primary source of aluminium.

3.1.2.5. Dynamic efficiency and Hotelling Rule under monopoly conditions

The dynamic efficiency criteria we have examined are based on a model of perfect competition. If we have a non-renewable resource with a reasonably competitive extraction sector, we might expect to see a standard result depicting diminishing extraction rate and increasing market price overtime. As in the static case, the conditions maximize the sum of producer and consumer surpluses (social welfare from the efficiency viewpoint). They are the same conditions that will be chosen by a social planner that wishes to maximize society's benefits if markets are indeed perfectly competitive, and the market interest rate is equal to the **social consumption discount rate** (the discount rate that society uses to value future consumption).

In reality, non-renewable resource industries are not perfectly competitive. As we noted in Module 2, perfect competition is more of an ideal market state. Most markets in the real word, including the markets for non-renewable resources, are imperfect. To show the effect of imperfect competition on the dynamic efficiency conditions, we illustrate with the extreme case of monopoly.

Under perfect competition, the market price is exogenous to (fixed for) each firm. Thus, we are able to obtain the result that in competitive markets, marginal revenue equals price (P=MR). However, in a monopolistic market, price is not fixed, but will depend on the firm's output choice. With a downward-sloping demand curve, marginal revenue will be less than price in this case (P ≠ MR, P < MR). Since MR is what is equated to MC to determine output, in the case of the monopolist, MR (not price) is what will be required to rise at the rate of interest in order to maximize the present value of marginal net benefits. Thus, the necessary condition for profit maximization in a monopolistic market requires that the marginal profit (and not the net price or royalty) should increase at the rate of interest in order to maximize the discounted profits over time. In contrast to a perfectly competitive firm, a monopolist will extract the non-renewable resource at a rate such that the proportional MR increases at the rate of interest (or the discount rate). In other words, for a monopolist, the dynamic efficiency condition requires that

$$MR_1 - MC_1 = \left(\frac{1}{(1+r)}\right) MR_2 - MC_2 \qquad (3.28)$$

and the Hotelling Rule in the case of a monopolist may be written as

$$\frac{\partial MR/\partial t}{MR(t)} = \frac{\dot{MR}}{MR(t)} = r \qquad (3.29)$$

The monopolist schedules extraction so that its marginal revenue rises at the interest rate; equivalently, so that its discounted marginal revenue is constant through time.

How do the extraction paths differ between a monopolistic and competitive firm? If the discounted monopoly price is decreasing overtime,

- the current price must be increasing at a rate less than the interest rate and therefore less the rate at which it would increase under competitive extraction. Thus. $\frac{\dot{MR}}{MR(t)} = r$ implies that $\frac{\dot{P}}{P(t)} < r$ and $\frac{P(\dot{t})^m}{P(t)^m} < \frac{P(\dot{t})^c}{P(t)^c}$, where $\frac{P(\dot{t})^m}{P(t)^m}$ is the rate of increase of current price under monopoly and $\frac{P(\dot{t})^c}{P(t)^c}$ is the rate of increase of current price under perfect competition.

- the initial monopoly price must be higher than the initial competitive price. $P_0^m > P_0^c$ where P_0^m is the initial monopoly price and P_0^c is the initial competitive price, and

- since for the given market demand curve, higher prices mean lower quantities and vice-versa; it must be that the monopolist begins by extracting smaller quantities than the competitive firm, $q_0^m < q_0^c$, where q_0^m is initial quantity extracted under monopoly and q_0^c is initial output under perfect competition.

In addition, we can also conclude that the rate at which extraction declines is slower under monopoly, so that the monopolist's extraction plan must extend over a longer period than that of the competitive firm. By implication, the monopolist is then extracting greater quantities than the competitive firm towards the end of its planning period.

To summarize, a monopolistic firm will take a longer time to fully deplete the non-renewable resource than a perfectly competitive market in our model. Extraction of the resource will be slower at first in monopolistic markets, but faster towards the end of the depletion horizon. In addition, the initial net price will be higher in monopolistic markets, and the rate of price increase will be slower. In other words, the monopolist will restrict output and raise prices initially, relative to the case of perfect competition. The rate of price increase, however, will be slower than under perfect competition. Eventually, an effect of monopolistic markets is to increase the time horizon over which the resource is extracted. However, this does not mean that monopoly extraction is "better" in terms of social welfare, since we have already seen that the competitive extraction path maximizes social welfare under some assumed conditions, implying that the monopoly outcome must be sub-optimal in a social-welfare-maximizing sense.

3.1.3. Dynamic efficiency, intertemporal fairness, and Hartwick Rule

In Module 2, we learnt that the economists' idea of sustainable development implies that society maintains a constant consumption level over time. The idea of bequeathing to future generations the same level of consumption possibilities that the present generation has, is itself derived from a given notion of fairness or intergenerational equity[19]. If we accept this notion of

[19] John Rawls (1971), *A Theory of Justice*. Harvard University Press.

fairness or conception of sustainable development, the critical question then is: do the dynamic efficiency conditions satisfy the requirement for fairness? In other words, are efficient allocations of a non-renewable resource fair?

In the numerical example we have constructed, it certainly does not appear that the efficient allocation satisfies the sustainable criterion. In the two-period example, more resources are allocated to the first period than to the second. Therefore, net benefits in the second period are lower than in the first. Sustainability does not allow earlier generations to profit at the expense of later generations, and this example certainly appears to be a case where that is happening. Yet choosing this particular extraction path does not prevent those in the first period from saving some of the net benefits for those in the second period. If the allocation is dynamically efficient, it will always be possible to set aside sufficient net benefits accrued in the first period for those in the second period, so that those in the second period will be at least as well off as they would have been with any other extraction profile.

To illustrate, we take a numerical example that compares a dynamic efficient allocation with sharing to an allocation where resources are committed equally to each generation. We continue to assume that the quantity of the non-renewable resource that is available is 20 and that all other conditions remain the same. Suppose, for example, you believe that setting aside half (10 units) for each period would be a better allocation than the dynamic efficient allocation. The net benefits to each period from this alternative scheme would be $40.[20] We can compare this with an allocation of net benefits that could be achieved with the dynamic efficient allocation. If the dynamic efficient allocation is to satisfy the sustainability criterion, we must be able to show that it can produce an outcome such that each generation would be at least as well off as it would be with the equal allocation. Applying the same technique, we can show that in the dynamic efficient allocation, the net benefits to the first period were 40.466, while those for the second period were 39.512. Clearly, if no sharing between the periods took place, this example would violate the sustainability criterion; the second generation is worse off.

But suppose the first generation was willing to share some of the net benefits from the extracted resources with the second generation. If the first generation keeps net benefits of $40 (thereby making it just as well off as if equal amounts were extracted in each period) and saves the extra $0.466 (the $40.466 net benefits earned during the first period in the dynamic efficient allocation minus the $40 reserved for itself) at 10 percent interest for those in the next period, this savings would grow to $0.513 by the second period [0.466(1.10)]. Add this to the net benefits received directly from the dynamic efficient allocation ($39.512), and the second generation would receive $40.025. Those in the second period would be better off by accepting the dynamic

[20] This is gotten by applying the formula: sum of PVMNB is $(P_1 - MC_1)q_1 + (P_2 - MC_2)q_1$ for period 1, and $(P_2 - MC_2)q_2 + (P_1 - MC_1)q_2$ for period 2.

efficient allocation with sharing than they would if they demanded that resources be allocated equally between the two periods.

This example demonstrates that although dynamic efficient allocations do not automatically satisfy sustainability criteria, they could be compatible with sustainability, even in an economy relying heavily on depletable resources. The possibility that the second period can be better off is not a guarantee; the required degree of sharing must take place. Box 3.2 illustrates with an example of one of the most popular sharing schemes for non-renewable resource.

Box 3.2 The Alaska permanent Fund

One interesting example of an intergenerational sharing mechanism currently exists in the State of Alaska. Extraction from Alaska's oil fields generates significant income, but it also depreciates one of the state's main environmental assets. To protect the interests of future generations as the Alaskan pipeline construction neared completion in 1976, Alaska voters approved a constitutional amendment that authorized the establishment of a dedicated fund: the Alaska Permanent Fund. This fund was designed to capture a portion of the rents received from the sale of the state's oil to share with future generations. The amendment requires: At least 25 percent of all mineral lease rentals, royalties, royalty sales proceeds, federal mineral revenue-sharing payments and bonuses received by the state be placed in a permanent fund, the principal of which may only be used for income-producing investments. The principal of this fund cannot be used to cover current expenses without a majority vote of Alaskans. The fund is fully invested in capital markets and diversified among various asset classes. It generates income from interest on bonds, stock dividends, real estate rents, and capital gains from the sale of assets. To date, the legislature has used some of these annual earnings to provide dividends to every eligible Alaska resident, while using the rest to increase the size of the principal, thereby assuring that it is not eroded by inflation. The 2010 dividend was $1,281.

Although this fund does preserve some of the revenue for future generations, two characteristics are worth noting. First, the principal could be used for current expenditures if a majority of current voters agreed. To date, that has not happened, but it has been discussed. Second, only 25 percent of the oil revenue is placed in the fund; assuming that revenue reflects scarcity rent, full sustainability would require dedicating 100 percent of it to the fund. Because the current generation not only gets its share of the income from the permanent fund, but also receives 75 percent of the proceeds from current oil sales, this sharing arrangement falls short of that prescribed by the Hartwick Rule. Source: The Alaska Permanent Fund Web site: http://www.pfd.state.ak.us/, Also in Tietinberg and Lewis, p110

The idea of sharing resource wealth with future generations is closely linked to the Hartwick Rule. The **Hartwick Rule** (discussed in Module 2) provide us a more practical and enforceable way of implementing the sustainability criterion. It requires society to invest all the rents accruing from a non-renewable resource in other forms of capital in order to maintain the portfolio of capital assets, and thus, provide a necessary condition for maintaining consumption over time. To understand the intuition behind this Rule, suppose a grandparent left you an inheritance of $10,000, and you put it in a bank where it earns 10 percent interest. What are the choices for allocating that money over time and what are the implications of those choices? If you spent exactly $1,000 per year, the amount in the bank would remain $10,000 and the income would last forever; you would be spending only the interest, leaving the principal intact. If you spend more than $1,000 per year, the principal would necessarily decline over time and eventually the balance in the account would go to zero. In the context of this discussion, spending $1,000 per year or less would satisfy the sustainability criterion, while spending more would violate it. Thus, the Hartwick Rule affords us a way of judging the sustainability of an allocation by examining whether or not the value of the total capital stock is non-declining. We will pick up this thought again in Module 6 where we will be examining how to account for the use of environmental and natural resources in our income accounting systems.

3.1.4. The Empirical Evidence

The dynamic efficiency conditions suggest a falling quantity profile and a rising price profile for a non-renewable resource. What does the empirical evidence reveal? Figure 3.5 shows the price profile for some mineral resources for the 25-year period between 1990 and 2015, while figure 3.6 display the global economic reserves for two of the minerals over the period 1996-2016. The evidence shows that after a long period of stable or declining prices, prices for key minerals rose rapidly in 2006-7. Since then they have fallen back somewhat, though in most cases they remain above previous levels. Also, even though production has risen steadily since 1950, reserves for most minerals have increased, not decreased. The two pieces of information appear to suggest that until recently, we have been in Stage I or II of Figure 3.4, although we may now be moving into Phase III with rising prices for many resources. However, this kind of statement must be taken with reservations. This is because the market for most non-renewable resources (commodity markets in particular) are susceptible to various shocks, some of which could be permanent. In general, the empirical evidence has not **perfectly** followed the predictions of the dynamic efficiency model. This is more pronounced in the case of the production profile.[21] In addition, reserve base estimates indicate that shortages are not imminent, although higher-quality reserves may run short in the relatively near future.

Observed differences between the predictions of the theoretical model and the empirical evidence may be due to many factors. Some important considerations are

- Effects of **population growth** and **growth in real income**. These act to shift the demand curve for non-renewable resources so that demand is no longer constant over time.

[21] An accurate assessment will require examination of the production and price profiles for each firm.

- Falling cost of exploration, discovery and development (EDD) due to **technological advancement**. This invalidates our assumption of constant marginal cost.
- **New discoveries** and emergence of **substitutes** for non-renewable resources, a factor ignored in the basic model.
- The dual effect of **changes in interest (discount) rate**. Apart from playing a direct role in determining the time path of price, interest rate affects some part of costs, which translate to decisions on extraction rate. Thus, the overall effect of interest rate on extraction and price profiles may be ambiguous.

In Module 3.2, we will examine how the dynamic-efficiency conditions changes in response to some of these factors.

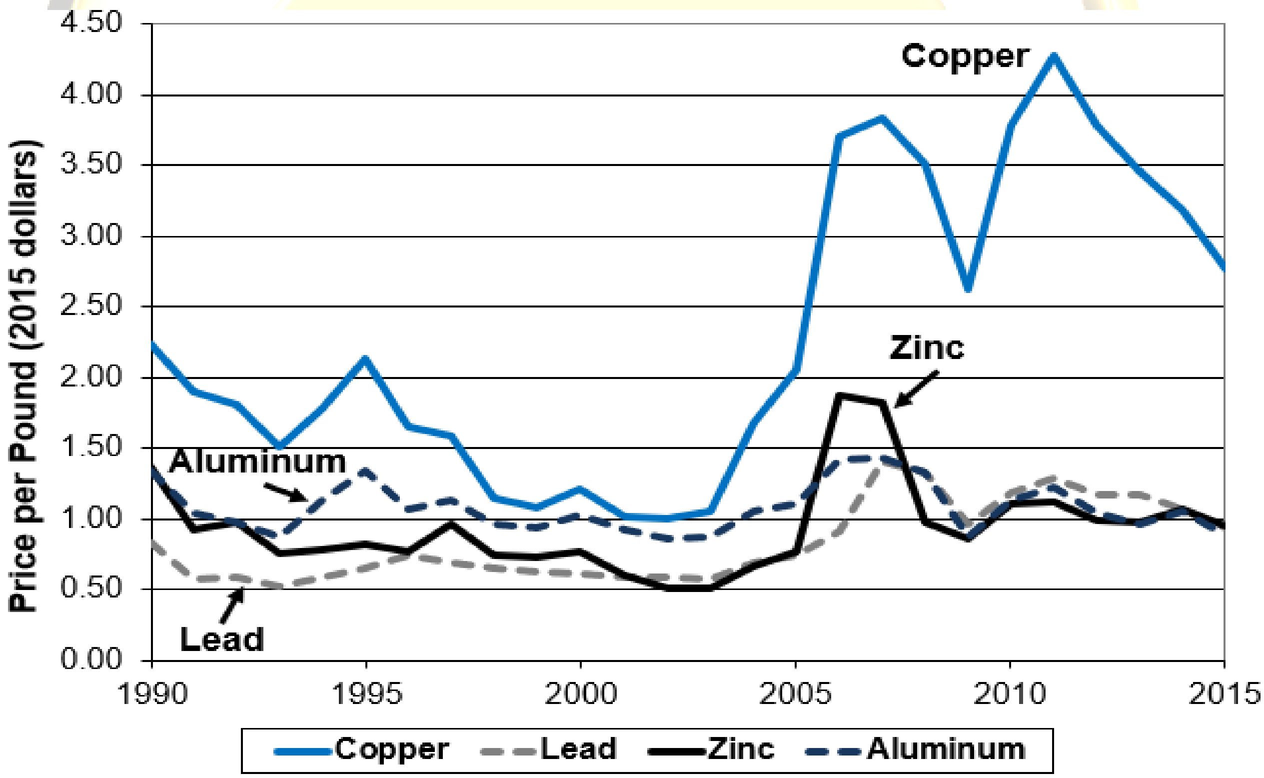


Fig. 3.5. Prices for Selected Minerals, 1990-2015. **Source:** U.S. Geological Survey, *Minerals Commodities Summaries*, various years. Also, in Jonathan and Roach 2015.

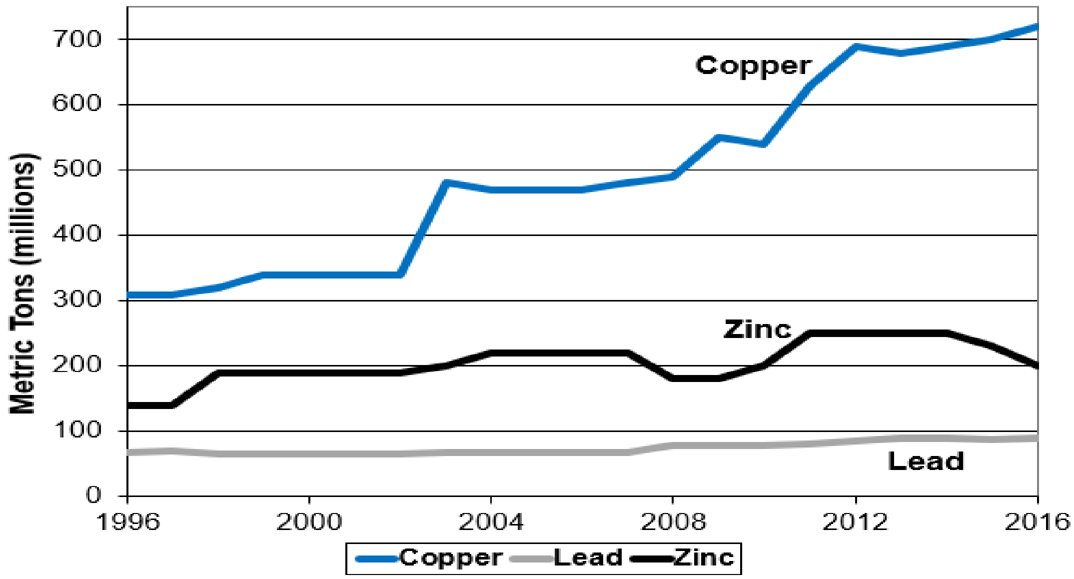


Figure 3.6. Global Economic Reserves for Selected Minerals, 1996-2016. Source*:* U.S. Geological Survey, Minerals Commodities Summaries, various years

Summary

- Three separate concepts are often used to classify the stock of non-renewable (depletable) resources, such as minerals. They include current reserve (the known resources that can profitably be extracted at current prices), potential reserves (the reserves or amount of the resource that is potentially available) and resource endowment (the natural occurrence of the resource in the earth's crust). While the potential reserves of a non-renewable resource depend on the price, the resource endowment is more of a geological, rather than an economic, construct and represents an upper limit on the availability of the resource. Current reserves can be expanded in two ways by new discovery or by new technology and changing market conditions. They can also be depleted through extraction.

- Intertemporal scarcity imposes an opportunity cost on using a depletable resource, which is the marginal user cost. When resources are scarce, greater current use diminishes future opportunities. The MUC is the present value of these forgone opportunities at the margin. An efficient market would have to consider not only the marginal cost of extraction for this resource but also the marginal user cost, the opportunity cost of extracting an extra unit of the resource today rather than tomorrow. Whereas in the absence of scarcity, price is equal to marginal cost, with scarcity, price is equal to the sum of marginal cost and the marginal user cost.

- The dynamic efficiency conditions maximize the sum of the present value of marginal net benefits. in the absence of scarcity, and given the assumptions of constant demand and constant marginal cost, dynamic efficiency in the use of the non-renewable resource requires that the resource be extracted at a fixed rate and that the price remains constant.

- However, with scarcity, the price of the resource increases over time, the extraction of the resource decreases over time, and the marginal user cost (which reflect the scarcity value of the resource or scarcity rents) rises over time at the rate of interest. The higher the interest rate, the greater will be the increases in the marginal user cost, the more will extraction be tilted towards the present rather than the future, and the higher will be the future price.

- The optimal extraction path for a non-renewable resource is one that yields rents that rises at the same rate as the discount rate. A depletable resource is viewed as a capital asset. An efficient extraction profile is the one that causes the value of the asset to appreciate at the same rate as other forms of capital (man-made and financial capital) that owners may own in their portfolio. Finding the efficient extraction rate is the same thing as maximizing the portfolio of capital assets. By the Hotelling Rule, under perfectly competitive conditions, profit is maximized when the portfolio price of a resource rises at the same rate as the interest (or discount) rate, or, put differently, when the discounted price of the natural resource is constant along the extraction path.

- As in the static case, the dynamic-efficiency conditions maximize the sum of producer and consumer surpluses (social welfare from the efficiency viewpoint). They are the same conditions that will be chosen by a social planner that wishes to maximize society's benefits if markets are indeed perfectly competitive, and the market interest rate is equal to the social consumption discount rate.

- Thus, the dynamic efficiency conditions for profit-maximization in a monopolistic market requires that the proportional marginal revenue (not the net price or royalty) should increase at the rate of interest. This implies that extraction of the resource will be slower at first in monopolistic markets, but faster towards the end of the depletion horizon. In addition, the initial net price will be higher in monopolistic markets, and the rate of price increase will be slower. Thus, a monopolistic firm will take a longer time to fully deplete the non-renewable resource than a perfectly competitive market all else being equal. However, this does not mean that monopoly extraction is "better" in terms of social welfare.

- Observed differences between the predictions of the dynamic efficiency model and the empirical evidence may be due to many factors, including the effects of population growth and growth in real income, falling cost of exploration, discovery and development (EDD) due to technological advancement, new discoveries and emergence of substitutes for non-renewable resources, and the effect of changes in interest (discount) rate.

- A dynamically efficient allocation can satisfy a certain requirement for fairness or sustainability if society follows the Hardwick Rule.

Review/Discussion Questions and Exercises

1. Draw a diagram showing the typical non-renewable resource use profile over time. Include the path for resource prices and consumption. Identify the four stages and briefly discuss what is happening to price and consumption in each stage.

2. Identify a major non-renewable resource that is produced in your country and examine (a) the price profile, and (b) your country's production profile, over the last 20 years or since commencement of production, depending on which is shorter. Comment on your results in the light of the dynamic efficiency conditions for non-renewable resources.

3. Explain why we will likely never "run out" of a non-renewable resource such as oil. Does this also imply that we will always be able to extract all the oil we need? Explain.

4. Explain and demonstrate the importance of the marginal user cost in the efficient management of a non-renewable resource.

5. Explain the implication of Hotelling Rule for (a) the price profile of a depletable resource (b) the quantity profile of a depletable resource (c) maximization of social welfare under (i) perfectly competitive conditions, (ii)monopoly conditions.

6. Explain (a) Hotelling Rule, (b) Hartwick Rule, and show their relevance to the optimal and sustainable management of a non-renewable resource.

7. What has been the general trend in non-renewable resource prices in the past several decades? What change has been observed in recent years? Are these trends consistent with Hotelling Rule? Discuss.

8. In the numerical example given in the module the inverse demand function for the depletable resource is P = 8 – 0.4q and the marginal cost of supplying it is $2.
   (a) If 20 units are to be allocated between two periods, in a dynamic efficient allocation how much would be allocated to the first period and how much to the second period when the discount rate is zero?

   (b) Given this discount rate what would be the efficient price in the two periods?

   (c) What would be the marginal user cost in each period?

9. Assume the same demand conditions as stated in Question 8, but let the discount rate be 0.10 and the marginal cost of extraction be $4. How much would be produced in each period in an efficient allocation? What would be the marginal user cost in each period? Would the static and dynamic efficiency criteria yield the same answers for this problem? Why?

9. One current practice is to calculate the years remaining for a depletable resource by taking the prevailing estimates of current reserves and dividing it by current annual consumption. How useful is that calculation? Why?

Materials used for the Lecture Notes

Jonathan M. Harris and Brian Roach (2017), **Environmental and Natural Resource Economics** 4th Edition, Routledge.

Tietenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics

Module 3.2. Optimal Extraction of Non-Renewable Resources: Extensions to the Basic Model (6.5 hours)

Learning outcome

This Module introduces some extensions to the basic model of optimal extraction of non-renewable resources. After going through this chapter, you should be able to

- ✓ construct and solve simple models of optimal resource depletion.
- ✓ carry out simple comparative dynamic analysis in the context of resource depletion models, and thereby determine the consequences of changes in interest rates, known stock size, demand, price of backstop technology, and resource extraction costs.
- ✓ understand the effect of uncertainty on optimal extraction of non-renewable resources.
- ✓ understand the meaning of a socially optimal depletion programme, and why this may differ from privately optimal programmes.
- ✓ understand the role of resource substitution possibilities and the ideas of a backstop technology and a resource choke price.
- ✓ gain greater understanding of the effect of recycling on non-renewable resource utilization.
- ✓ understand how environmental externalities associated with extraction of no-renewable resources affect the optimal extraction and price paths and how failure to take such factors into account lead to inefficiencies.

Outline

3.2.1. The N-period constant-cost case: optimal extraction of a non-renewable resource over an unlimited period

In Module 3.1, we saw that due to the fixed and finite supplies of depletable resources, production of a unit today precludes production of that unit tomorrow. Therefore, production decisions today must take forgone future net benefits into account. The marginal user cost (MUC) is the opportunity cost measure that allows intertemporal balancing to take place. In the two-period model, the marginal cost of extraction is assumed to be constant, but the value of the marginal user cost rises over time. When the demand curve is stable over time and the marginal cost of extraction is constant, the rate of increase in the current value of the marginal user cost is equal to *r*, the discount rate. Thus, in Period 2, the marginal user cost would be 1 + r times as large as it was in Period 1.[22] MUC rises at rate *r* in an efficient allocation in order to preserve the balance between present versus future production.

So far, our examination of the optimal extraction of the non-renewable resource assumes only two periods and that demand and marginal cost remain constant over the periods. We also assumed that the resource has no substitute. In addition, we have up to this point ignored the role of population and income growth on demand, or of the exploration for new resources or technological progress. These are historically significant factors in the determination of actual consumption paths of non-renewable resources. In this Module, we are going to extend the basic model by relaxing these assumptions. We begin this generalization by retaining the constant-marginal-extraction-cost assumption while extending the time horizon within which the resource is allocated.

Assume the same inverse demand function for the non-renewable resource as in the two-period case considered in Module 3.1, that is $P= 8 - 0.4q$. We assume that marginal cost remains constant at c = 2 and the discount rate remains at 10%. But now instead of two periods, we spread the allocation over a larger (unlimited) number of years and increase the total recoverable supply from 20 to 40. In this case, how long the resource will last is no longer predetermined; thus, the time of exhaustion must be derived as well as the extraction path prior to exhaustion of the resource. The equations describing the allocation that maximizes the present value of net benefits derived in Module 3.1 are

$$\frac{(a-bq_t)-c}{(1+r)^{t-1}} - \lambda = 0, \quad t = 1 \ldots n \qquad (3.30)$$

and

$$\sum_{t=1}^{T} q_t = \bar{Q} \qquad (3.31)$$

[22] The condition that marginal user cost rises at rate r is true only when the marginal cost of extraction is constant

The allocation and value of T that satisfies these conditions for the given parameter values are illustrated in the Table 3.2 below. The optimality of this allocation can be verified by substituting these values into the above equations. (Due to rounding, these add to 39.999, rather than 40.000.).

Table 3.2. Dynamic efficient allocation for a non-renewable resource in an N-period case

q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	T
8.004	7.305	6.535	5.689	4.758	3.733	2.607	1.368	0.000	9

Figure 3.7a demonstrates how the efficient quantity extracted varies over time, while Figure 3.7b shows the behaviour of the marginal user cost (MUC) and the marginal cost of (MC) extraction. Total marginal cost refers to the sum of the two. The marginal cost of extraction is represented by the lower line, and the marginal user cost is depicted as the vertical distance between the marginal cost of extraction and the total marginal cost. To avoid confusion, note that the horizontal axis is defined in terms of time, not the more conventional designation—quantity.

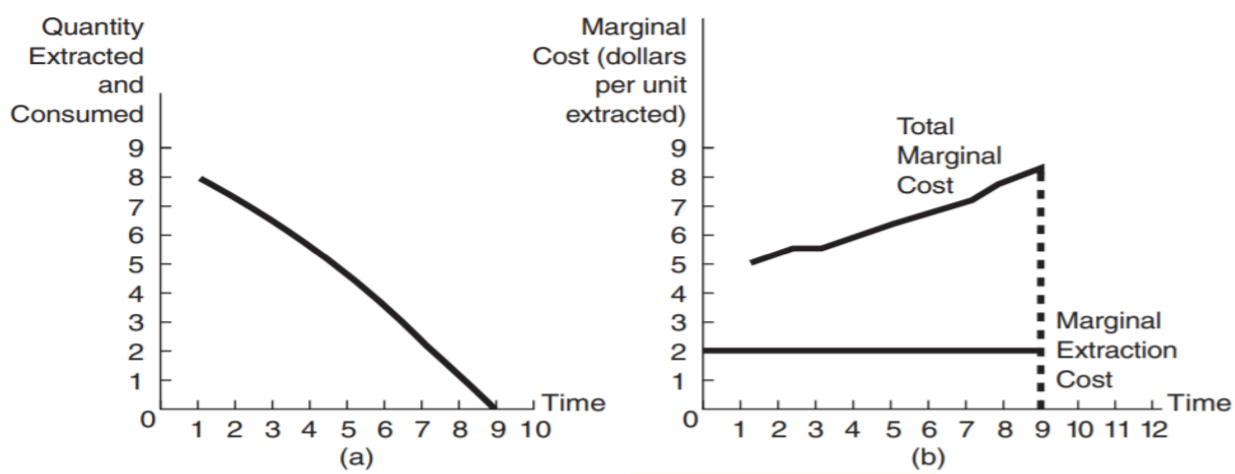


Fig. 3.7 (a) Quantity Profile for a non-renewable resource with constant marginal extraction cost and no substitute resource (b) Marginal cost profile for a non-renewable resource with constant marginal extraction cost and no substitute resource. **Source**: Tietinberg and Lewis, 2012. P.124

Several trends are worth noting. First of all, in this case, as in the two-period case, the efficient MUC rises steadily in spite of the fact that the marginal cost of extraction remains constant. This reflects increasing scarcity and the accompanying rise in the opportunity cost of current consumption as the remaining stock dwindles. In response to these rising costs over time, the extracted quantity falls over time until it finally goes to zero, which occurs precisely at the moment when the total marginal cost becomes $8. At this point, total marginal cost is equal to the highest price anyone is willing to pay, so demand and supply simultaneously equal zero. Thus, even in this challenging case involving no increase in the cost of extraction, an efficient allocation envisions a smooth transition to the exhaustion of a resource. The resource does not "suddenly" run out, although in this case it does run out. The highest possible price for which demand for the non-renewable resource is zero, is called the **choke price**. At that price, a substitute for the resource, if available, becomes economically more attractive. In this example, in the absence of a substitute, the choke price is $8.

We will now examine the effect on the optimal extraction path of a change in some of the parameters of the basic (N-period) model.

3.2.1.1. Effect of changes in the interest or discount rate

We have earlier mentioned in passing (in Module 3.1) the effect of a change in the discount rate on the present value of marginal net benefit of a non-renewable resource. We will provide a further discussion here in the case where the extraction period is unlimited. With an increase in the discount rate, the optimal price path will be lower at the initial point but will grow more quickly, and will reach its final (**choke**) price earlier in time (before t = T). This result can be explained by the following observations. First, the choke price itself (K in Figure 3.8), is not altered by the interest rate change. Second, as we have already observed, the new price path must rise more steeply with a higher interest rate. Third, we can deduce from using the resource exhaustion constraint, that it must begin from a lower initial price level. The change in interest rate does not alter the quantity that is to be extracted; the same total stock is extracted whatever the interest rate might be. If the price path began from the same initial value (P_0) then it would follow a path such as that shown by the curve labelled *A* in Figure 3.8 and would reach its choke price before *t* = *T*. But then the price would always be higher than along the original price path, but for a shorter period of time. Hence the resource stock will not be fully extracted along path A and that path could not be optimal.

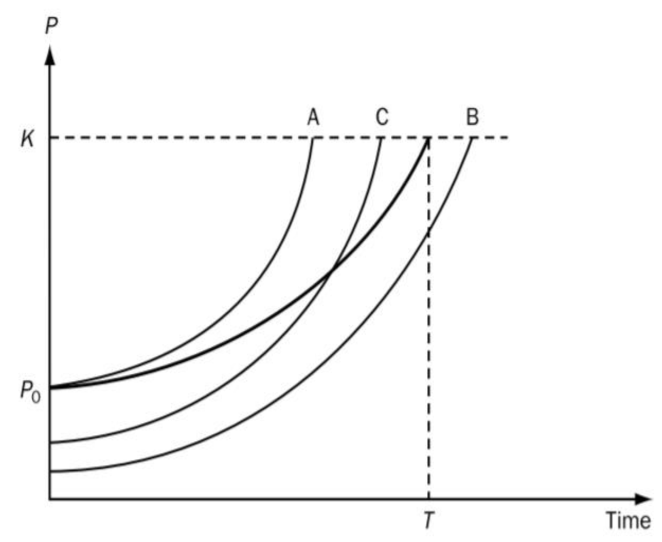


Fig. 3.8. The effect of an increase in the interest rate on the optimal price of the non-renewable resource. **Source:** Perman et al., 2003. p.520

A path such as *B* is not feasible. Here the price is always lower (and so the quantity extracted is higher) than on the original optimal path, and for a longer time. But that would imply that more resources are extracted over the life of the resource than were initially available. This is not feasible. The only feasible and optimal path is one such as *C*. Here the price is lower than on the original optimal path for some time (and so the quantity extracted is greater); then the new price path crosses over the original one and the price is higher thereafter (and so the quantity extracted is lower).

Note that because the new path must intersect the original path from below, the optimal depletion time will be shorter for a higher interest rate. This is intuitively reasonable. Higher interest rate means greater impatience. More is extracted early on, less later; and total time to full exhaustion is quicker.

3.2.1.2 Effect of changes in demand

Suppose that there is an increase in demand for the resource, possibly as a result of population growth or rising real incomes. This will cause the demand curve to shift outwards. Given this change, the old royalty or net price path would result in higher extraction levels, which will exhaust the resource before the net price has reached *K*, the choke price. Hence the net price must increase to dampen down quantities demanded; the time until the resource stock is fully exhausted will also be shortened.

3.2.1.3 Effect of an increase in the size of the known resource stock

In practice, estimates of the size of reserves of non-renewable resources, such as minerals are under constant revision. As prices rise, what were previously uneconomic stocks become economically recoverable. Exploration and discovery can lead to increases in the available reserves of a non-renewable resources.

Consider the case of a single new discovery of a fossil fuel stock. Other things being unchanged, if the royalty path were such that its initial level remained unchanged at P_0, then given the fact that the rate of royalty increase is unchanged, some proportion of the reserve would remain unutilized by the time the choke price, *K*, is reached. This is clearly neither efficient nor optimal. It follows that the initial royalty must be lower and the time to exhaustion is extended.

3.2.1.4 Effect of a change in the constant resource extraction cost

Consider the case of an increase in the constant extraction costs, possibly because labour charges rise in the extraction industry. A rise in extraction costs will raise the initial price, P_0, slow down the rate at which price increases, and lengthen the time to complete exhaustion of the stock. What about a fall in extraction costs? This can happen as a consequence of **technological progress** decreasing the costs of extracting the resource from its reserves. A fall in extraction costs will lower the initial price, increase the rate at which price increases, and shorten the time to complete exhaustion of the stock. However, if the changes in extraction cost were very large, then our conclusions may need to be amended. For example, if a cost increase were very large, then it is possible that the new price in period 0, P_0', will be above the choke price. It is then not economically viable to deplete the remaining reserve – an example of an economic exhaustion of a resource, even though, in physical terms, the resource stock has not become completely exhausted.

We can illustrate each of the four scenarios discussed above in the context of the mathematical model developed earlier by taking each at a time, assuming others remain constant. The dynamic efficiency results from using the new value of the parameter can then be compared with the initial results.

3.2.2. Effects of Uncertainty

The Hotelling Rule (examined in Module 3.1) is an abstract analytical tool; its operation in actual market economies is dependent upon the existence of a set of particular institutional circumstances. In many real situations these institutional arrangements do not exist and so the Rule lies at a considerable distance from the operation of actual market mechanisms. In addition to the **discount rate equivalence** (that is the quality between the discount rate used by a competitive firm in deciding the present value of benefit flows from extracting the non-renewable resources - the private discount rate- and the social discount rate employed by a social planner), two assumptions are required to ensure a social optimal extraction of a non-renewable resource in the case of perfect competition. First, the resource must be owned by competitive agents. Second, each agent must know at each point in time all current and future prices. In other words, agents have perfect foresight. In the absence of perfect foresight, knowledge of these prices requires the existence of both **spot markets** and a complete set of **forward markets** for the resource in question. But no resource actually possesses a complete set of forward markets, and in these circumstances, there is no guarantee that agents can or will make rational supply decisions.

In addition, uncertainty is prevalent in decision-making regarding non-renewable resource extraction and use. There is uncertainty, for example, about stock sizes, extraction costs, how successful research and development will be in the discovery of substitutes for non-renewable resources (thereby affecting the cost and expected date of arrival of a backstop technology), pay-offs from exploration for new stock, and the action of rivals.

What do optimal extraction programmes look like when there is uncertainty, and how do they compare with programmes developed under conditions of certainty? Let us assume an owner of a natural resource (such as a mine) wishes to maximize the net present value of utility over two periods. (We are limiting the analyses to two periods to make it more tractable). Thus, the optimization problem can be stated as

$$\text{Max } (u_o + \frac{u_1}{1+\rho}) \qquad (3.32)$$

If there is a probability (π) of a disaster (for example, the market might be lost) associated with the second period of the extraction programme, then the owner will try to maximize the expected net present value of the utility (if he or she is risk-neutral):

$$\text{Max } \left[u_o + \pi.0 + (1-\pi)\left(\frac{u_1}{1+\rho}\right)\right]$$

$$= \text{Max } \left[u_o + (1-\pi)\left(\frac{u_1}{1+\rho}\right)\right]$$

$$= \text{Max } \left[u_o + \left(\frac{u_1}{1+\rho^*}\right)\right] \qquad (3.33)$$

$$\text{where } \frac{1}{1+\rho^*} = \frac{1-\pi}{1+\rho}$$

Note that $(1+\rho^*)(1-\pi) = 1+\rho$

$$\Rightarrow \rho^* - \rho = \pi(1+\rho^*) > 0 \qquad (\text{if } 1 \geq \pi > 0)$$

$$\Rightarrow \rho^* > \rho$$

Therefore, in this example, the existence of risk is equivalent to an increase in the discount rate for the owner, which implies, as we have shown before, that the price of the resource must rise more rapidly and the depletion is accelerated.

3.2.3 Transition to a substitute

So far, we have discussed the allocation of a depletable resource when no substitute is available to take its place. In reality many non-renewable resources have substitutes, which may be man-made, or even a renewable natural resource. For example, oil and natural gas can be replaced with solar to supply energy, exhaustible groundwater can be replaced with a surface-water substitute. Indeed, renewables, including solar, wind, hydro, bio fuels and others, lie at the heart of the transition to a less carbon-intensive and more sustainable energy system. In this section we shall examine the optimal management of a non-renewable resource in the presence of a perfect substitute. We will begin with the case of a substitute renewable resource and thereafter, examine the case where the substitute is also a depletable resource.

3.2.3.1 Transition to a renewable substitute

How do we define an efficient allocation for a non-renewable resource in the presence of a renewable substitute, which is available for example, at constant marginal cost? As in the case already examined in section 3.2.1 above, the depletable (non-renewable) resource would be exhausted in this case also, but that will be less of a problem, since we will merely switch to the renewable one at the appropriate time.

For the purpose of our numerical example, assume the existence of a perfect substitute for the depletable resource that is infinitely available at a cost of $6 per unit. The transition from the depletable resource to this renewable resource would ultimately transpire because the renewable resource marginal cost ($6) is less than the maximum willingness to pay ($8). The total marginal cost for the depletable resource in the presence of a $6 perfect substitute would never exceed $6, because society could always use the renewable resource instead, whenever it was cheaper. Thus, while the maximum willingness to pay (the choke price) sets the upper limit on total marginal cost when no substitute is available, the marginal cost of extraction of the substitute sets the upper limit when a perfect substitute is available at a marginal cost lower than the choke price. This is an important observation and the example is carefully structured to bring out the point. We now present a formal model to derive the dynamically efficient allocations of both the depletable resource and its substitute.

Let q_t be the amount of the non-renewable resource extracted in year t and q_{st} the amount used of the renewable perfect substitute. We use $d to depict the marginal cost of the substitute. Given these conditions, the total benefit and cost formulas become

$$\sum_{t=1}^{T} a(q_t + q_{st}) - \frac{b}{2}(q_t + q_{st})^2 \qquad (3.34)$$

$$\text{Total cost} = \sum_{t=1}^{T}(cq_t + dq_{st}) \qquad (3.35)$$

The objective function is thus

$$\text{PVNB} = \sum_{t=1}^{T} \frac{a(q_t + q_{st}) - \frac{b}{2}(q_t{}^2 - q_{st}{}^2 + 2q_t q_{st}) - cq_t - dq_{st}}{(1+r)^{t-1}} \qquad (3.36)$$

subject to the constraint on the total availability of the depletable resource

$$\bar{Q} - \sum_{t=1}^{T} q_t \geq 0 \qquad (3.37)$$

Necessary and sufficient conditions for an allocation maximizing this function are expressed in Equations (3.38), (3.39), and (3.40):

$$\frac{a - b(q_t + q_{st}) - \text{c}}{(1+r)^{t-1}} - \lambda \leq 0, \quad \text{t} = 1, \ldots, \text{T} \qquad (3.38)$$

Any member of equation set (3.38) will hold as an equality when q_t > 0 and will be negative when

$$a - b(q_t + q_{st}) - d \leq 0, \quad t = 1, \ldots, T \qquad (3.39)$$

Any member of equation set (3.39) will hold as an equality when q_{st} > 0 and will be negative when q_{st} = 0

$$\bar{Q} - \sum_{t=1}^{T} q_t \geq 0 \qquad (3.40)$$

For the parameter values, a = \$8, b = 0.4, c = \$2, d = \$6, $\bar{Q}$ = 40, and r = 0.10 assumed in the previous example, it can be readily verified that the optimal conditions are satisfied by the allocation and value of *t* and λ shown in Table 3.3.

Table 3.3. Dynamic efficient allocation for a non-renewable resource with a renewable substitute.

q_1	q_2	q_3	q_4	q_5	q_6	q_{s6}	q_{st} for *t > 6*	q_{st} for *t < 6*	λ
8.798	8.177	7.495	6.744	5.919	2.863	2.137	5.000	0.000	2.481

The depletable resource is used up before the end of the sixth period and the switch is made to the substitute resource at that time. From equation set (3.39), in competitive markets, the switch occurs precisely at the moment when the resource price rises to meet the marginal cost of the substitute. The transition point (called the **switch point**) in this example is earlier than in the previous example. Since all characteristics of the problem except for the availability of the substitute are the same as in the previous example, the difference can be attributed to the availability of the renewable substitute. The efficient path for this situation is given in figures 3.9a and 3.9b. In this efficient allocation, the transition is once again smooth. Quantity extracted per unit of time is gradually reduced as the marginal user cost rises until the switch is made to the substitute. No abrupt change is evident in either marginal cost or quantity profiles.

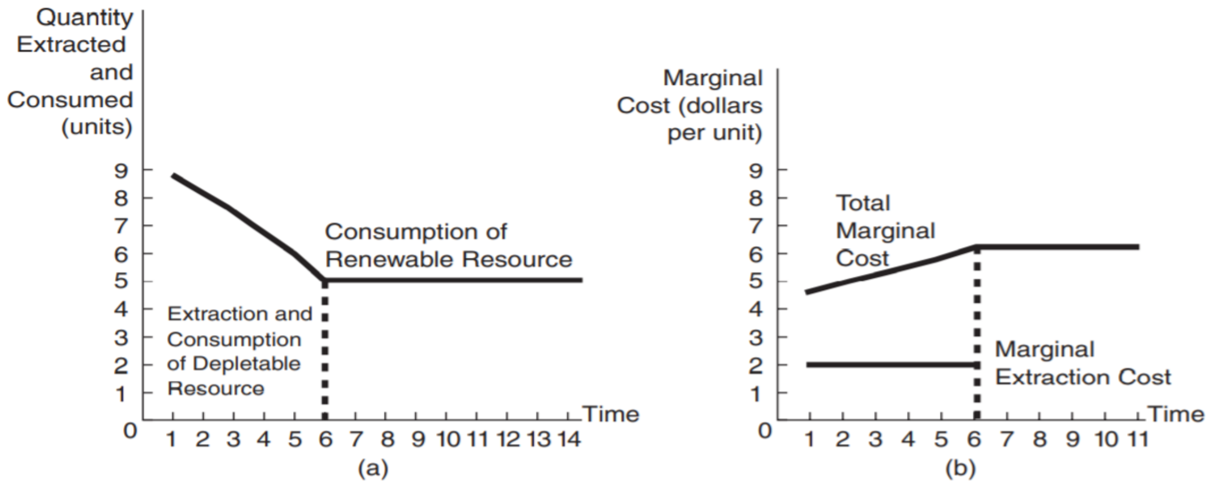


Fig. 3.9 (a) Constant Marginal Extraction Cost with Substitute Resource: Quantity Profile (b) Constant Marginal Extraction Cost with Substitute Resource: Marginal Cost Profile, **Source**: Tietinberg and Lewis, 2012. P.126.

Because the renewable resource is available, more of the depletable resource would be extracted in the earlier periods than was the case in our previous numerical example without a renewable resource. As a result, the depletable resource would be exhausted sooner than it would have been without the renewable resource substitute. In this example, the switch is made during the sixth period, whereas in the last example the last units were exhausted at the end of the eighth period. That seems consistent with common sense. When a substitute is available, the need to save some of the depletable resource for the future is less pressing (in other words, the opportunity cost of extraction is now lower). At the transition point, called the **switch point**, consumption of the renewable resource begins. Prior to the switch point, only the depletable resource is consumed, while after the switch point only the renewable resource is consumed.

This sequencing of consumption pattern results from the cost patterns. Prior to the switch point, the depletable resource is cheaper. At the switch point, the marginal cost of the depletable resource (including marginal user cost) rises to meet the marginal cost of the substitute, and the transition occurs. In our numerical example, due to the availability of the substitute resource, after the switch point consumption never drops below five units in any time period. This level is maintained because five is the amount that maximizes the net benefit when the marginal cost equals $6 (the price of the substitute). (Convince yourself of the validity of this statement by substituting $6 into the willingness-to-pay function and solving for the quantity demanded).

3.2.3.2 Transition to a non-renewable substitute with higher marginal cost

Will the situation be different if the substitute is another depletable (non-renewable resource) with a higher marginal cost (consider the transition from coal to gas for example)? It is not difficult to see how an efficient allocation would be defined when the transition is from one constant marginal-cost depletable resource to another depletable resource with a constant, but higher, marginal cost (see Figure 3.10). The total marginal cost of the first resource would rise over time until it equalled that of the second resource at the time of transition (T*). In the period of time prior to transition, only the cheapest resource would be consumed; all of it would have been consumed by T*.

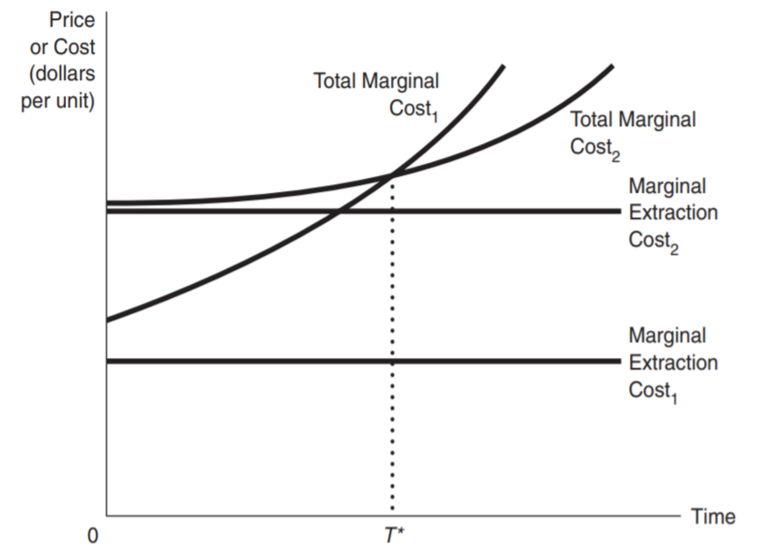


Fig. 3.10. The transition from one constant-cost depletable resource to another. **Source**: Tietinberg and Lewis, 2012. P.127

A close examination of the total-marginal-cost path reveals two interesting characteristics worthy of our attention. First, even in this case, the transition is a smooth one; total marginal cost never jumps to the higher level. Second, the slope of the total marginal cost curve over time is flatter after the time of transition. The first characteristic is easy to explain. The total marginal costs of the two resources have to be equal at the time of transition. If they weren't equal, the net benefit could be increased by switching to the lower-cost resource from the more expensive resource. Total marginal costs are not equal in the other periods. In the period before transition, the first resource is cheaper and therefore used exclusively, whereas after transition the first resource is exhausted, leaving only the second resource.

The slope of the marginal cost curve over time is flatter after transition simply because the component of total marginal cost that is growing (the marginal user cost, MUC) represents a smaller portion of the total marginal cost of the second resource than of the first. The total marginal cost of each resource is determined by the marginal extraction cost plus the marginal user cost. In both cases the marginal user cost is increasing at rate *r*, and the marginal cost of extraction is constant. As seen in Figure 3.10, the marginal cost of extraction, which is constant, constitutes a much larger proportion of total marginal cost for the second resource than for the first. Hence, total marginal cost rises more slowly for the second resource, at least initially.

3.2.3.3 A fall in the price of backstop technology

Our examination of the transition to a renewable substitute affords us some insight into what will be the effects of a fall in the price of **backstop technology** on the optimal price and extraction paths of a non-renewable resource. Recall, we assumed the existence of a choke price, *K*, for the non-renewable resource, such that If the price were to rise above K, the economy will cease consumption of the non-renewable resource and switch to an alternative source – the **backstop source**. We saw that when the marginal cost of the renewable resource (and hence, the price) is less than the maximum willingness to pay for the non-renewable resource, the lower marginal cost of the renewable substitute becomes the choke price. The transition to the substitute will occur at that price.

Now, suppose that technological progress occurs, increasing the efficiency of a backstop technology. This will tend to reduce the price of the backstop, so that the choke price will fall further. Given this scenario, the initial value of the resource price on the original optimal price path, P_0, can no longer be optimal. In fact, it is now too high since the net price would reach the new choke price before T, leaving some of the economically useful resource unexploited. So, the initial price of the non-renewable resource, P_0, must fall to a lower level, $P_0{}'$, to encourage an increase in demand so that a shorter time horizon is required until complete exhaustion of the non-renewable resource reserve. When the resource price reaches the new, reduced choke price, demand for the non-renewable resource falls to zero.

3.2.4. Optimal extraction in the case of variable (rising) marginal cost

We have expanded our examination of the efficient allocation of depletable resources to include longer time horizons and the availability of other depletable or renewable resources that could serve as perfect substitutes. As we move closer to realism, we now consider a situation in which the marginal cost of extracting the depletable resource rises with the cumulative amount extracted, that is a situation where we have a marginal cost that is variable (a function instead of a constant). This is commonly the case, for example, with minerals, where the higher-grade ores are extracted first, followed by an increasing reliance on lower-grade ones. We continue with the assumption that there is a renewable substitute, a backstop source.

Analytically, this case is handled in the same manner as the previous case in section 3.2.3.1, except that the function describing the marginal cost of extraction is slightly more complicated. Let us assume, for example that the marginal cost of extraction is represented in the function

$$MC_t = \$2 + 0.1Q_t \qquad (3.41)$$

where Q_t is cumulative extraction to date; thus, implying that marginal cost increases with the cumulative amount extracted.

The dynamic efficient allocation of this resource is found by maximizing the present value of the net benefits, using this modified cost of extraction function. The results of that maximization are portrayed in Figures 3.11a and 3.11b.

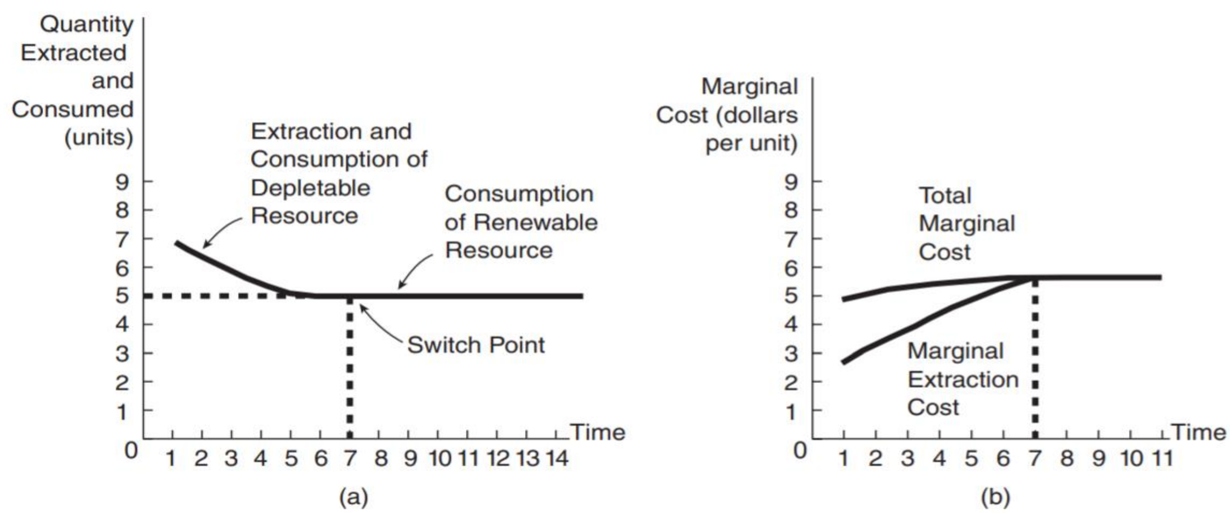


Fig. 3.11. (a) Increasing Marginal Extraction Cost with Substitute Resource: Quantity Profile (b) Increasing Marginal Extraction Cost with Substitute Resource: Marginal Cost Profile. **Source**: Tietinberg and Lewis, 2012. p.128.

The most significant difference between this case and the other lies in the behaviour of marginal user cost. In the previous case, we noted that marginal user cost (MUC) rose over time at rate *r*. However, in this case where the marginal cost of extraction increases with the cumulative amount extracted, MUC declines over time until, at the time of transition to the renewable resource, it goes to zero. Why? Remember that marginal user cost is an opportunity cost reflecting forgone future marginal net benefits. In contrast to the constant marginal-cost case, in the increasing marginal-cost case every unit extracted now raises the cost of future extraction. Therefore, as the current marginal cost rises over time, the sacrifice made by future generations (as an additional unit is consumed earlier) diminishes; the net benefit that would be received by a future generation, if a unit of the resource were saved for them, gets smaller and smaller as the marginal extraction cost of that resource gets larger and larger. By the last period, the marginal extraction cost is so high that earlier consumption of one more unit imposes virtually no sacrifice at all. At the switch point, the opportunity cost of current extraction drops to zero, and total marginal cost equals the marginal extraction cost.

Note that total marginal cost cannot be greater than the marginal cost of the substitute. In the increasing marginal extraction cost case, at the time of transition, the marginal extraction cost also must equal the marginal cost of the substitute. If that weren't true, it would imply that some of the resource that was available at a marginal cost lower than the substitute would not be used. This would clearly be inefficient, since net benefits could be increased by simply using less of the more expensive substitute. Hence, at the **switch point**, in the rising marginal-cost case, the marginal extraction cost has to equal total marginal cost, implying a zero marginal user cost.

The increasing-cost case differs from the constant-cost case in another important way as well. In the constant-cost case, the depletable resource reserve is ultimately completely exhausted. In the increasing-cost case, however, the reserve is not exhausted; some is left in the ground because it is more expensive to use than the substitute. In sum, the complication of increasing marginal cost changes the time profile of the marginal user cost, but it does not alter the basic finding of declining consumption of depletable resources coupled with rising total marginal cost.

3.2.5. The Effect of Recycling

Recycling can be described as transforming materials from their original state into another form after being used. In many instances recycling is referred to as processing waste into a new product. Thus, what would otherwise be regarded as waste can either be converted to raw materials for making new product or get transformed directly into another product entirely. Originally, recycling was aimed at reducing the need to extract more resources from the environment. Today, recycling has emerged as an instrument not only to mitigate resources depletion, but also to keep the environment clean for both people and wildlife. Recycling has also emerged as an important facet in technology development, as new systems are devised to help recycle a variety of resources, including metals, plastic, and water. Indeed, recycling has assumed a growing dimension in the international development community over the past few decades, particularly against the backdrop of rapid depletion of non-renewable resources. The 3R (reduce, reuse, recycle) concept aims at fostering a sound material-cycle society, which does not waste valuable goods or materials and limits the impact of economic development on extraction of non-renewable resources.

Figure 3.12 shows the effect of recycling on the extraction path of a renewable resource. As with the case of the substitution of a backstop (ideally renewable) resource, the expansion of resource recycling can stretch out the lifetime of a non-renewable resource. The equilibrium levels of recycling can be identified based on marginal costs of virgin and recycled materials. Producers will tend to operate where these marginal costs are equalized. When the environmental cost of extracting non-renewable resources is added to the private marginal cost, the marginal cost of using virgin materials be higher so that equilibrium can be achieved at a higher level of recycling. This is illustrated in figure 3.13.

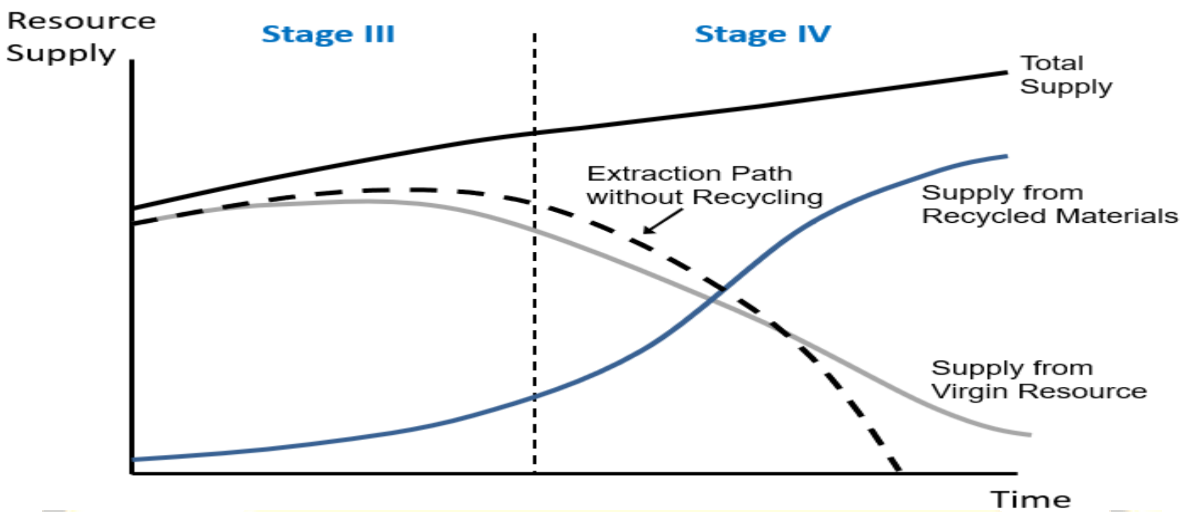


Fig. 3.12. Impact of Recycling on Virgin Resource Extraction Path. **Source:** Jonathan and Roach, 2017

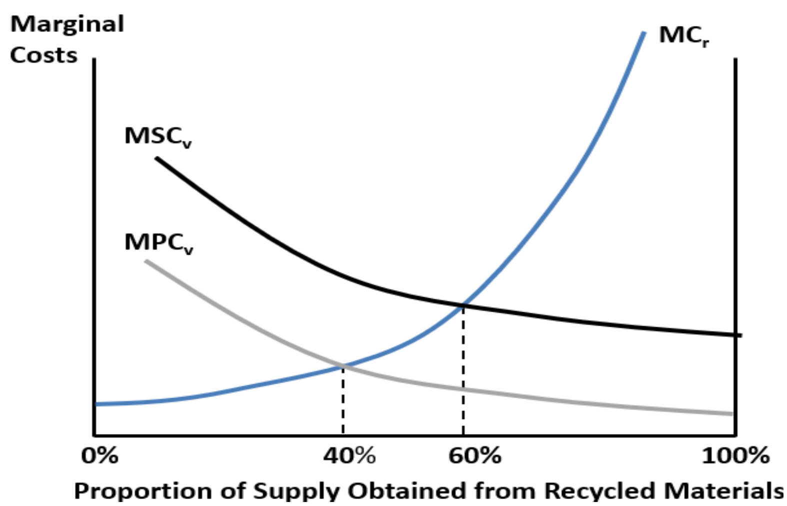


Fig. 3.13. Marginal Costs of Recycling. **Source:** Jonathan and Roach, 2017

We will demonstrate in the next section the effect of internalizing the environmental cost of extraction on the extraction and price profiles of a non-renewable resource.

Figure 3.14 shows recycling rates for some metals in the United States. The proportion recycled is highest for lead, a durable metal with high environmental impacts. About half of the country's iron, steel, and aluminium are also recycled.

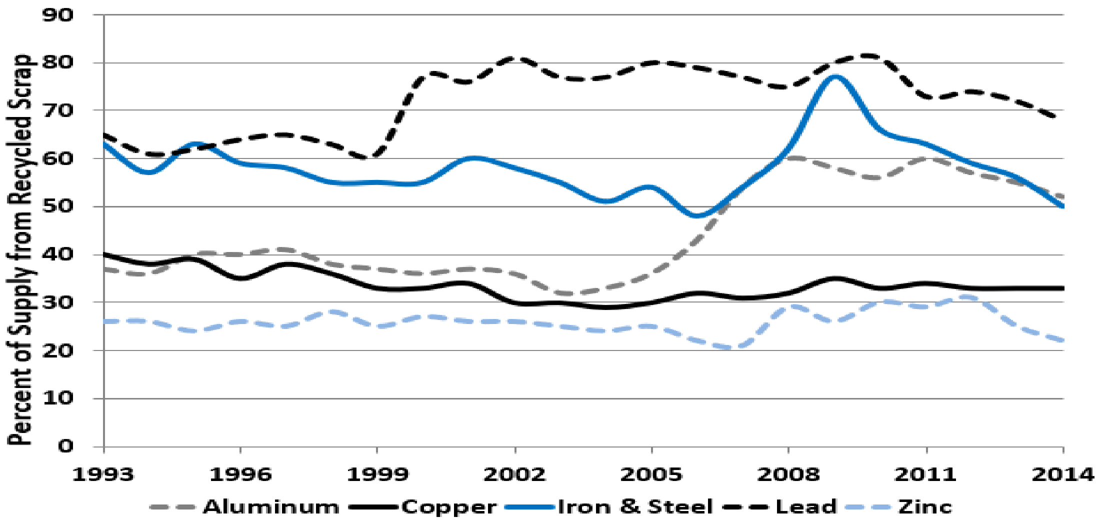


Fig. 3.14. Scrap Metal as a Percentage of U.S. Supply, 1993-2014. **Source:** Various editions of *Minerals Yearbook*, USGS. Also, in Jonathan and Roach, 2017.

3.2.6 Market Allocations of depletable resources: some considerations

In this section, we examine the question of whether actual markets can be expected to produce an efficient allocation of a non-renewable resource. Can the private market, involving millions of consumers and producers each reacting to his or her own unique preferences, ever result in a dynamically efficient allocation? Is profit-maximization compatible with dynamic efficiency?

3.2.6.1 Appropriate property right structures

The most common misconception of those who believe that even a perfect market could never achieve an efficient allocation of non-renewable (depletable) resources is based on the idea that producers want to extract and sell the resources as fast as possible, since that is how they derive the value from the resource. This misconception makes people see markets as myopic and unconcerned about the future. However, as long as the property rights governing natural resources have the characteristics of exclusivity, transferability, and enforceability (see Module 2), as it is in the case of some non-renewable resources, such as minerals, where the state, a group, or even individuals, exercises legally enforceable property right, the markets in which those resources are bought and sold will not **necessarily** lead to myopic choices. When bearing the marginal user cost, the producer acts in an efficient manner. A resource in the ground has two potential sources of value to its owner (1) a use value when it is sold, and (2) an asset value when it remains in the ground. As long as the price of a resource continues to rise, the resource in the ground is becoming more valuable. The owner of this resource accrues this capital gain, however, only if the resource is conserved. A producer who sells all resources in the earlier periods loses the chance to take advantage of higher prices in the future.

A profit-maximizing producer attempts to balance present and future production in order to maximize the value of the resource. Since higher prices in the future provide an incentive to conserve, a producer who ignores this incentive would not be maximizing the value of the resource. We would expect resources owned by a myopic producer to be bought by someone willing to conserve and prepared to maximize its value. As long as social and private discount rates coincide (discount rates equivalence), property rights structures are well defined, and reliable information about future prices is available, a producer who pursues maximum profits simultaneously provides the maximum present value of net benefits for society. The implication of this analysis is that, in competitive resource markets, the price of the resource equals the total marginal cost of extracting and using the resource. Thus, an efficient allocation is the same allocation that will be produced by an efficient market. When used to describe an efficient market, the total marginal cost curve associated with the efficient allocations considered in Figures 3.9 through 3.11 above describes the time path that prices could be expected to follow.

3.2.6.2 Environmental Costs of non-renewable resource extraction

One of the most important situations in which property rights structures may not be well defined is that in which the extraction of a natural resource imposes an environmental cost on society which is not internalized by the producers. The aesthetic costs of strip mining, the health risks associated with uranium tailings, and the acids leached into streams from mine operations are all examples of associated environmental costs. Table 3.5 present a compilation of some potential negative externalities associated with non-renewable resource extraction.

Table 3.4 Potential Environmental Impacts of Mining

Activity	Potential Impacts
Excavation and ore removal	Destruction of plant and animal habitat, human settlements, and other features (surface mining) Land subsidence (underground mining) Increased erosion; silting of lakes and streams Waste generation Acid drainage and metal contamination of lakes, streams, and groundwater
Ore concentration	Waste generation (tailings) Organic chemical contamination Acid drainage and metal contamination

Activity	Potential Impacts
Smelting/refining	Air pollution (including sulfur dioxide, arsenic, lead, cadmium, and other toxics) Waste generation (slag) Impacts of producing energy (most energy used for mineral production goes into smelting and refining)

Source: Young, 1992.

Not only is the presence of environmental costs empirically important, but also it is conceptually important, since it forms one of the bridges between the traditionally separate fields of environmental economics and natural resource economics.

Suppose, for example, that the extraction of the depletable resource caused some damage to the environment that was not adequately reflected in the costs faced by the extracting firms. As we saw in Module 2, this is an external cost. The cost of getting the resource out of the ground, as well as processing and shipping it, is borne by the resource owner and considered in the calculation of how much of the resource to extract. The environmental damage, however, is not automatically borne by the owner and, in the absence of any outside attempt to internalize that cost, will not be part of the extraction decision. How would the market allocation, based on only the former cost, differ from the efficient allocation, which is based on both costs?

We can examine this issue by modifying the numerical example used earlier in section 3.2.3. Assume the environmental damage can be represented by increasing the marginal cost by $1. Thus, the new marginal cost function can be written as

$$MC_t = \$3 + 0.1Q_t \qquad (3.42)$$

The additional dollar reflects the cost of the environmental damage caused by producing another unit of the resource. What effect do you think this would have on the efficient time profile for quantities extracted? The answers are given in Figures 3.15a and 3.15b. The result of including environmental cost on the timing of the switch point is interesting because it involves two different effects that work in opposite directions. On the **demand side**, the inclusion of environmental costs results in higher prices, which tend to dampen demand. This lowers the rate of consumption of the resource, which, all other things being equal, would make it last longer. All other things are not equal, however as there is a **supply-side effec**t. The higher marginal cost also means that a smaller cumulative amount of the depletable resource would be extracted in an efficient allocation. (Remember the effect of an increase in marginal cost). In our illustrative example, the efficient cumulative amount extracted would be 30 units instead of the 40 units extracted in the case where environmental costs were not included. This supply-side effect tends to hasten the time when a switch to the renewable resource is made, all other things being equal.

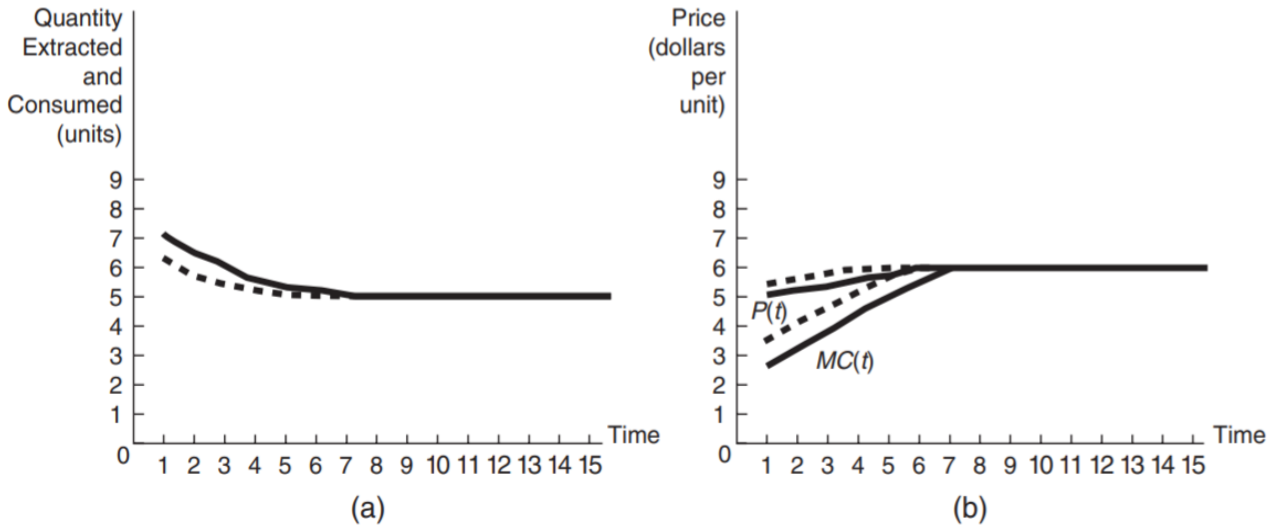


Fig. 3.15 Increasing marginal extraction cost with Substitute resource in the presence of environmental costs (a) Quantity Profile (b) Price Profile (Solid line—without environmental costs; dashed line—with environmental costs)

Which effect dominates—the rate of consumption effect or the supply effect? In our numerical example, the supply-side effect dominates and, as a result, the time of transition for an efficient allocation is sooner than for the market allocation. But in general, the answer depends on the shape of the marginal-extraction-cost function. With constant marginal cost, for example, there would be no supply-side effect and the market would transition later. Also, if the environmental costs were associated with the cost of the renewable resource, rather than the depletable resource, the time of transition for the efficient allocation would have been later than the market allocation. (Can you explain why?).

Our analyses show that Ignoring external costs associated with the extraction of a non-renewable resource leaves the price of the resource too low, so that too much of the resource would be extracted from an efficiency viewpoint. This, once again, shows the interdependencies among the various decisions we have to make about the future. Environmental and natural resource decisions are intimately and inextricably linked.

Summary

- For a non-renewable resource extracted at constant marginal cost, in a dynamic efficiency setting, the marginal user cost rises steadily in spite of the fact that the marginal cost of extraction remains constant. This reflects increasing scarcity and the accompanying rise in the opportunity cost of current consumption as the remaining stock dwindles. In response to these rising costs, the extracted quantity falls over time until it finally goes to zero. This occurs when the total marginal cost is equal to the highest price anyone is willing to pay, so demand and supply simultaneously equal zero.

- The highest possible price for which demand for the non-renewable resource is zero, is called the choke price. At that price, a substitute for the resource, if available, becomes economically more attractive.

- An increase in the discount rate lowers the optimal price path at its initial point but causes it to rise more quickly so that the choke price is arrived at earlier in time. Higher interest rate means greater impatience. More is extracted early on, less later; and total time to full 'exhaustion' is quicker.

- An increase in demand causes the demand curve to expand, leads to an increase in price and a reduction in the time it takes for the resource stock to be fully exhausted.

- Additions to the available reserves of a non-renewable resource will lead to a reduction in price (a lower initial price) and an extension of the time to exhaustion.

- An increase in the constant extraction costs will raise initial price, slow down the rate at which price increases, and lengthen the time to complete exhaustion of the stock, while a fall in extraction costs will lower the initial price, increase the rate at which price increases, and shorten the time to complete exhaustion of the stock. However, other scenarios may unfold if the changes in extraction cost are very large.

- Agents that operate in the market for non-renewable resources do not have perfect foresight. There are also no complete set of forward markets. In addition, uncertainty is prevalent in decision-making regarding non-renewable resource extraction and use. The presence of uncertainties (the existence of risk) acts like an increase in the discount rate for the owner of a non-renewable resource, which implies that the price of the resource must rise more rapidly and the depletion is accelerated.

- While the maximum willingness to pay (the choke price) sets the upper limit on total marginal cost when no substitute is available, the marginal cost of extraction of the substitute sets the upper limit when a perfect substitute is available at a marginal cost lower than the choke price.

- When a renewable substitute to a non-renewable (depletable) resource is available, more of the depletable resource would be extracted in the earlier periods than in the case where no renewable substitute is available. As a result, the depletable resource would be exhausted sooner than it would have been without the renewable resource substitute.

- Where marginal cost of extraction is not constant but increases with the cumulative amount extracted, the marginal user cost declines (rather than increases) over time until it becomes zero at the time of transition to the renewable resource. At that point, the available reserves of the non-renewable resource is not exhausted; some is left in the ground because it is more expensive to use than the substitute.

- As with the case of the substitution of a backstop resource, the expansion of resource recycling can stretch out the lifetime of a non-renewable resource.

- As long as social and private discount rates coincide, property rights structures are well defined, and reliable information about future prices is available, the price of a non-renewable resource in competitive markets, equals the total marginal cost of extracting and using the resource. Thus, an efficient allocation is the same allocation that will be produced by an efficient market.

- One of the most important situations in which property rights structures may not be well defined is that in which the extraction of a natural resource imposes an environmental cost on society which is not internalized by the producers. On the demand side, the inclusion of environmental costs results in higher prices, which tend to dampen demand. This lowers the rate of consumption of the resource, which, all other things being equal, would make it last longer. On the supply-side. higher cost tends to hasten the time when a switch to the renewable resource is made, all other things being equal.

- Ignoring external costs associated with the extraction of a non-renewable resource leaves the price of the resource too low, so that too much of the resource would be extracted from an efficiency viewpoint.

Review/Discussion Questions and Exercises

1. Given the demand function for a non-renewable resource, Q_t = 100 - Q_t (Q is in tons, P is in $/ton). Assume that the marginal cost of extraction is constant at c = 10, the interest rate, r is 10% and the fixed available reserve is R = 153 tons.

(a) Derive the dynamically efficient price and extraction profiles assuming that exactly 1 unit is sold in period T

(b) During how many periods does extraction take place?

2. Explain the effect of an increase in the interest rate on the dynamically-efficient production and price path of a non-renewable resource and the time it takes for the resource to be exhausted.

3. Explain and illustrate the effect of an increase in the following on the initial price, price path and exhaustion period of a non-renewable resource
(a) increase in the demand
(b) increase in the constant marginal cost
(c) increase in known reserves
(d) an increase in uncertainty and risk

4. Define and explain the importance of the marginal user cost in optimal management of a non-renewable resource. Explain and compare the intertemporal behaviour of marginal user cost in the case of
   (a) a non-renewable resource with constant or zero marginal cost
   (b) a non-renewable resource with variable (increasing) marginal cost.
   (c) Does the introduction of a variable rather than a constant (or zero) marginal cost affect the basic conclusions of the dynamic efficiency model as it relates to price and extraction profiles? Why or why not.

5. What effects is the presence of a renewable substitute likely to have on the price and extraction profile of a depletable resources is owners are profit-maximizing?

6. Assume an extractive firm is compelled to internalize the environmental cost of extracting a renewable resource, what are the likely effects on
(a) The amount of the resource that is recycled, assuming recycling is possible
(b) The extraction profile of the resource
(c) The price profile of the resource
(d) The duration it will take for the resource to be 'exhausted' or to make a switch to a renewable substitute.

Materials used for the Lecture Notes

Jonathan M. Harris and Brian Roach (2017), **Environmental and Natural Resource Economics** 4th Edition, Routledge.

Perman, R., Ma Y., McGilvray J. and Common M. (2003). **Natural Resource and Environmental Economics,** 3rd edition**.** Edinburgh, Longman.

Tietenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics.

3.3. Optimal Extraction of Non-renewable Resources: Application to Sub-Saharan Africa (1.5 hours)

Africa has significant natural resource wealth. The continent holds more than half of the world's rare minerals About 30 per cent of all global mineral reserves are found in Africa. The continent's proven oil reserves constitute 8 per cent of the world's stock, while those of natural gas amount to 7 per cent. Minerals account for an average of 70 per cent of total African exports and about 28 per cent of the gross domestic product (GDP). Earnings from recent oil, gas and mineral discoveries could lead to an increase in government revenues of between 9 per cent and 31 per cent in the first 10 years of production in countries like Mozambique, Tanzania and Uganda (AfDB/BMGF 2015: 6).

Table 3. 5 reveals the high dependence on natural resources by African economies, with three quarters of the countries listed in the Table falling within the classification of countries heavily dependent on non-renewable natural resources.

Table 3.5 Primary Commodity Exports in Selected African countries (% of GDP) (2000 – 2015)

Congo	72	Chad	31
Equatorial Guinea	67	Zambia	26
Angola	57	Namibia	25
Libya	51	Nigeria	23
Gabon	50	Somalia	22
Botswana	38	Zimbabwe	22
Algeria	36	Swaziland	21
Mauritania	33	Ghana	21
Seychelles	33	Mozambique	20
Cote d' Ivoire	32	Togo	20

Source: UNCTAD 2016 Database

Africa's abundant natural resources provide a unique opportunity to foster human and economic development. However, the continent suffers from the paradox of plenty, meaning that abundant endowments of natural resources do not lead to equivalent levels of prosperity, broad-based development and resource-based industrialization. One key obstacle preventing African countries from realizing this potential is poor governance. Poor governance refers to the lack of strong institutions and weak policies, aimed at short-term gains rather than at long-term development objectives. In addition, easy access to and capture of revenues are factors that make governments less accountable and more likely to preserve the interests of the minority governing elite with limited benefits for the population (United Nations Conference on Trade and Development 2012). The level of poverty in most resource-rich countries in Africa has not been reduced, as predictions based on their economic growth performance would suggest. Overall, the wealth of natural resources has not yielded positive benefits in terms of inclusion and poverty reduction in Africa.

While many sub-Saharan African countries have made significant legal reforms in oil, gas, and mining over the past decades, figure 3. 16 shows that, in almost all countries, there remains an implementation gap between what laws say and how resource governance works in practice. This gap prevents countries from fully capitalizing on their resources and on the investments, they have made in legal reforms. The largest implementation gaps exist in two areas: transparency of environmental and social impacts and sharing of revenues with local government. Sub-Saharan Africa also lags behind the rest of the world in implementation of laws related to transparency and oversight of key institutions (including state-owned enterprises) and sovereign wealth funds, as well as with compliance with fiscal rules.

A further problem that impedes development benefits from natural resource wealth is the status and structure of Africa's extractive industries. Most countries in the continent remain exporters of unprocessed or lightly processed commodities. According to the UN Economic Commission for Africa (UNECA 2013), Sub-Saharan Africa's dependence on primary products is high and there is low value addition to commodities before export. Due to this dependence, the region is exposed to high commodity volatility and limited linkage of the commodity sector to the local economy. Furthermore, according to the Africa Progress Panel, 'a study by the Southern African Development Community [SADC] on the value chain for a range of minerals in Africa shows that the value of processed products was typically 400 times greater than the equivalent unit value (by weight) of the raw material' (2013: 45). Thus, Africa needs to climb the value-added chain of mineral processing and manufacturing to unlock the full economic potential of its natural resources. It needs to foster local resource-based industrialization and value addition and embrace it as a legitimate aspirational goal. The goal is to use natural resource endowments to develop a competitive local supply industry that, through employment creation, value addition, technology and knowledge transfer, fosters broad-based sustainable development.

The fact that the potential benefits of natural resources have failed to materialize has been recognized by decision-makers in the continent. Many African governments have expressed a commitment to invest revenues from natural resources in order to enhance development outcomes, including better health, better education and access to quality social services (**follow Hartwick rule?**). At the continental level, a new African normative framework on natural resource governance—the Africa Mining Vision (AMV)—was adopted by heads of state in 2009. The AMV provides a compelling thrust towards 'transparent, equitable and optimal exploitation of mineral resources to underpin broad-based sustainable growth and socio-economic development'. It further calls for greater value addition of African mineral resources. The African Mineral Governance Framework, designed to facilitate the realization of the AMV, has also been drafted.

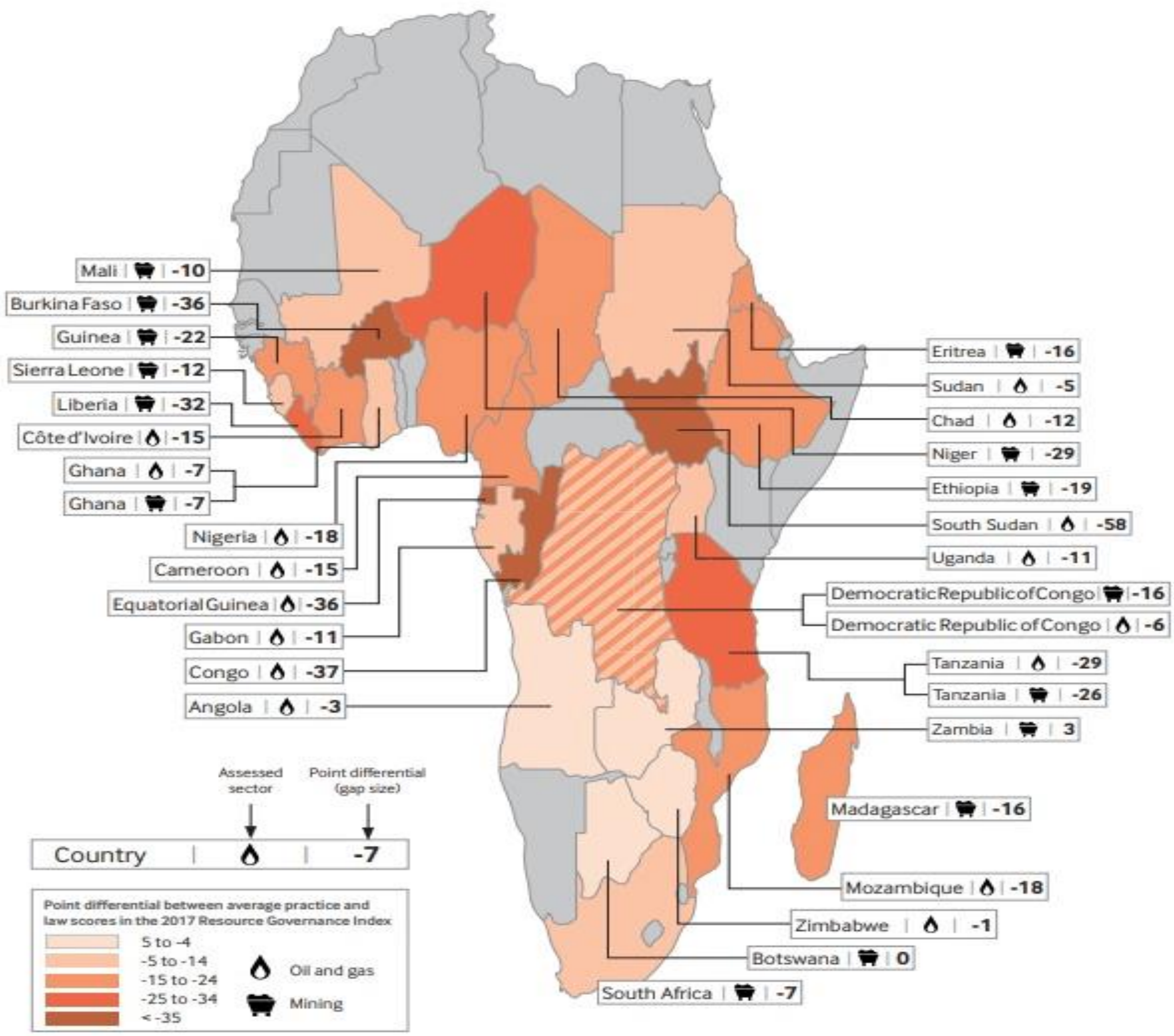


Fig. 3.16. The natural resources implementation gap**. Source:** Resource Governance Institute, Media Briefing February 2018. Also available at https://www.brookings.edu/blog/africa-in-focus/2019/04/24/figure-of-the-week-natural-resource-governance-in-africa/

More recently, in 2015 the African Union (AU) adopted the first 10-year Implementation Plan for Agenda 2063, which places natural resources at the center of the construction of a developmental state and of the socio-economic transformation of African countries, and which is also a milestone of the 2030 Agenda for Sustainable Development. Regional institutions such as the AfDB Natural Resource Center (ANRC) and the African Minerals Development Center (AMDC/ECA) are also supporting African states in their efforts to improve their natural resource governance. International frameworks

At the level of international institutions and partners—including non-state actors—the focus is generally on the governance of natural resources. Efforts have concentrated on the fight against corruption, as well as tax mobilization, transparency and accountability, as reflected in the latest Resource Governance Index released by the Natural Resource Governance Institute (NRGI 2017). Other initiatives reflecting this trend include the Organization for Economic Co-operation and Development (OECD) Due Diligence Guidance for Responsible Supply Chains of Minerals from Conflict-Affected and High-Risk Areas; the Kimberley Process; the Dodd-Frank Act in the United States; the European Union's transparency and accounting directives; the Extractive Industries Transparency Initiative; and Publish What You Pay (European Centre for Development Policy Management 2017: 3).

Despite these efforts, natural resources are not yet a vehicle for the transformative development of African states. In fact, mineral resources continue to constitute one of the main sources of violent conflict, and represent a threat to the democratization process in most African states (Tana Forum 2017). Some advocate a need for a paradigm shift in the reflection and action on natural resources governance to use this sector as a key leverage for inclusive, participatory and owner-based socio-economic transformation of these countries.

Some of the factors that students/readers may want to consider in reflecting on the optimal management of non-renewable resources in Sub-Sahara Africa in addition to those highlighted above include

- The role of **high discount rates** and **uncertainties.** Economic and political uncertainties may make all actors (including state actors) to discount future benefits heavily, thus, leading to an inefficient and unsustainable exploitation of mineral resources.

- The **failure to account for the environmental cost of extractive activities**. This is due to combination of factors, including a lack of political will, heavy reliance on mineral resource wealth, questionable affinity between extractive firms and those controlling political power, regulatory capture, and inadequate technical capacity on the part of regulatory agencies.

- The potential for greater recycling.

- The political economy of resource funds: why dos some countries have resource funds that save some of non-renewable resource wealth for future generations while others do not?

Students/ readers may want to explore one or more of these issues using their own countries as case studies.

Additional Material used for this Lecture

Enhancing natural resource governance in Africa, Information Brief, Africa and west Asia Programme, July 2017. Available at https://www.idea.int/sites/default/files/publications/enhancing-natural-resource-governance-in-africa.pdf

.

Module 3.4. Optimal Management of Renewable Resources: The Basic Model (4.5 hours)

Learning objectives

This Module introduces the reader to the basic model of optimal management of non-renewable resources. After going through the notes, the reader should

- ✓ Know the broad classes of renewable resources and the concept of stock and flow with respect to renewable resources.
- ✓ understand the biological growth function of a renewable resource and associated concepts, such as density-dependent and property of compensation.
- ✓ be able to interpret the simple logistic growth model and distinguish between a stable and unstable equilibrium.
- ✓ understand the idea of a sustainable yield and the maximum sustainable yield.
- ✓ understand the concept of steady-state harvesting and alternative steady states.
- ✓ be able to distinguish between steady-state outcomes and dynamic adjustment processes that may (or may not) lead to a steady-state outcome.
- ✓ understand the concept of static and dynamic efficiency in relation to renewable resources and why efficient harvest levels differ from the maximum sustainable yield.
- ✓ understand the kind of equilibria that can prevail in a model of renewable resource harvesting and the implications for property right structure?

Outline

3.4.1 Introduction: Natural growth, regeneration and cyclical resources

3.4.2 Biological growth processes

3.4.3 Steady-state harvesting, Maximum Sustainable Yield, and Efficiency

 3.4.3.1. Steady-state harvesting

 3.4.3.2 Efficiency: Static and Dynamic

3.4.4. Equilibria in Renewable resource Harvesting Model

Summary

Discussion/Review Questions and Exercises

Materials used for the Lecture

3.4.1 Introduction: Natural growth, regeneration and cyclical resources

As we learnt in Module 1.1 environmental resources are described as renewable when they have a capacity for reproduction and growth. The class of **renewable resources** is diverse. It includes populations of biological organisms such as fisheries and forests which have a natural capacity for growth, and water and atmospheric systems which are reproduced by physical or chemical processes. While the latter do not possess biological growth capacity, they do have some ability to assimilate pollution inputs (thereby maintaining their quality) and, at least in the case of water resources, can self-replenish as stocks are run down (thereby maintaining their quantity). It is also conventional to classify arable and grazing lands as renewable resources. In these cases, reproduction and growth take place by a combination of biological processes (such as the recycling of organic nutrients) and physical processes (irrigation, exposure to wind, etc.). Fertility levels can regenerate naturally so long as the demands made on the soil are not excessive. We may also consider more broadly defined environmental systems (such as wilderness areas or tropical moist forests) as being sets of interrelated renewable resources. The categories just described are renewable stock resources.

A broad concept of renewables would also include flow resources such as solar, wave, wind and geothermal energy. These share with biological stock resources the property that current harnessing of the flow does not mean that the total magnitude of the future flow will necessarily be smaller. Indeed, many forms of energy-flow resources are, for all practical purposes, non-depletable. Most of the literature on the economics of renewable resources focus on the harvesting of animal species ('hunting and fishing') and the economics of forestry.[23] In order to narrow down our coverage, we will in this and the next module be considering animal species using fishery as an illustration.

We can distinguish between the stock and flow of a non-renewable resource. The **stock** is a measure of the quantity of the resource existing at a point in time, measured either as the aggregate mass of the biological material (the biomass) in question (such as the total weight of fish of particular age classes or the cubic metres of standing timber), or in terms of population numbers. The **flow** is the change in the stock over an interval of time, where the change results either from biological factors, such as 'recruitment' of new fish into the population through birth or 'exit' due to natural death, or from harvesting activity. One similarity between renewable and non-renewable resources is that both are capable of being fully exhausted (that is, the stock being driven to zero) if excessive and prolonged harvesting or extraction activity is carried out. In the case of non-renewable resources, exhaustibility is a consequence of the finiteness of the stock. For renewable resources, although the stock can grow, it can also be driven to zero if

[23] Agriculture could also be thought of as a branch of renewable resource harvesting. But agriculture – particularly in its more developed forms – differs fundamentally from other forms of renewable resource exploitation in that the environmental medium in which it takes place is designed and controlled.

conditions interfere with the reproductive capability of the renewable resource, or if rates of harvesting continually exceed net natural growth.

How the stock of a common-pool resource, such as fisheries, forestry and other commercially-viable species changes depends on whether those who are best positioned to protect them have sufficient incentives to do so. A commercially valuable species is like a double-edged sword: On the one side, the value of the species to humans provides a strong, current reason for human concern about its future. On the other hand, its value may promote excessive harvest. It is evident that enforceable private property rights do not exist for many forms of renewable resource. In the absence of regulation or collective control over harvesting behaviour, the resource stocks are subject to open access. We will show in Module 3.5 that open-access resources tend to be overexploited in both a biological and an economic sense, and that the likelihood of the resource being harvested to the point of exhaustion is higher than where private property rights are established and access to harvesting can be restricted.

3.4.2 The Biological growth processes

In order to investigate the economics of a renewable resource, it is first necessary to describe the pattern of biological (or other) growth of the resource. As noted above, biological populations belong to a class of renewable resources we call **interactive resources**: the size of the resource stock (population) is determined jointly by biological considerations and by actions taken by society. The postharvest size of the population, in turn, determines the availability of resources for the future. Thus, humanity's actions affect the flow of these resources over time: The rate of harvest has intertemporal effects: Tomorrow's harvesting choices are affected by today's harvesting behaviour.

To fix ideas, we consider the growth function for a population of some species of fish. This is conventionally called a **fishery**. We suppose that this fishery has an **intrinsic** (or potential) growth rate denoted by *g*. This is the proportional rate at which the fish stock would grow when its size is small relative to the carrying capacity of the fishery, and so the fish face no significant environmental constraints on their reproduction and survival. The intrinsic growth rate *g* may be thought of as the difference between the population's birth and natural mortality rate (again, where the population size is small relative to carrying capacity). Suppose that the population stock is *S* and it grows at a fixed rate *g*. Then in the absence of human predation the rate of change of the population over time is given by

$$\frac{dS}{dt} = \dot{S} = g(S(t)) \qquad (3.42)$$

By integrating this equation, we obtain an expression for the stock level at any point in time:

$$S_t = S_0 e^{gt} \qquad (3.43)$$

in which S_0 is the initial stock level. In other words, for a positive value of g, the population grows exponentially over time at the rate g and without bounds. This is only plausible over a short span of time. Any population of fish exists in a particular environmental milieu, with a finite carrying capacity, which sets bounds on the population's growth possibilities. A simple way of representing this effect is by making the actual (as opposed to the potential) growth rate depend on the stock size. Then we have what is called **density-dependent growth**. Using the symbol χ to denote the actual growth rate, the growth function can then be written as

$$\dot{S} = \chi(S)S \qquad (3.44)$$

where *χ(S)* states that χ is a function of S, and shows the dependence of the actual growth rate on the stock size. If this function has the property that the proportionate growth rate of the stock ($\dot{S}$/S) declines as the stock size rises, then the function is said to have the **property of compensation.** Now let us suppose that under a given set of environmental conditions there is a finite upper bound on the size to which the population can grow (its carrying capacity). We will denote this as S_{MAX}. A commonly used functional form for χ(S) which has the properties of compensation and a maximum stock size is the simple **logistic function**:

$$\chi(S) = g\,(1 - \frac{S}{S_{MAX}}) \qquad (3.45)$$

in which the constant parameter g > 0 is the intrinsic or potential growth rate of the population. Where the logistic function determines the actual population growth rate, we may therefore write the biological growth function as

$$\dot{S} = \frac{dS}{dt} = g\,(1 - \frac{S}{S_{MAX}})S \qquad (3.46)$$

where $\dot{S}$ and dS/dt refer to the net effect of natural changes and human predation.

In the absence of human predation, the logistic biological growth function can be written as

$$G(S) = g\,(1 - \frac{S}{S_{MAX}})S \qquad (3.47)$$

If we assume that there is some positive population threshold level, S_{MIN}, such that the population would inevitably and permanently decline to zero if the actual population were ever to fall below that threshold (as it is the case with many biological resources), we can rewrite (3.47) as.

$$G(S) = g\,(S - S_{MIN})(1 - \frac{S}{S_{MAX}}) \qquad (3.48)$$

Note that if S_{MIN} = 0, equation 3.48 collapses to the special case of equation.3.47.

The logistic form is a good approximation to the natural growth processes of many fish, animal and bird populations (and indeed to some physical systems such as the quantity of fresh water in an underground reservoir). It was first applied to fisheries by Schaefer (1957).[24] The generalization given by (3.48) is illustrated in Figure 3.17.a. The graph reflects population biology, not economics. It shows the typical pattern of growth of a specie in a supporting environment, at first exponentially increasing, then slowing to a maximum equilibrium level. This pattern is known as a logistic curve. Species that exceed the maximum will decline back to it due to limitations such as shortage of food supply. Species that fall below a critical minimum will decline to extinction. The relevance of this curve for resource economics is that all non-renewable resource exploitation must necessarily be subject to these underlying biological laws.

Transforming the figure into a stock/annual growth format shows the same logic in terms of stock growth. Figure 3. 17b shows the relationship between the stock size and the associated rate of change of the population due to biological growth (the amount of growth, G, is a quadratic function of the resource stock size, S). It abstracts from such influences as water temperature and the age structure of the population which is assumed to counterbalance each other over the long term. An increasing population grows at first faster, then more slowly, until the natural equilibrium is reached (unless it falls below the critical minimum and negative growth leads to extinction).

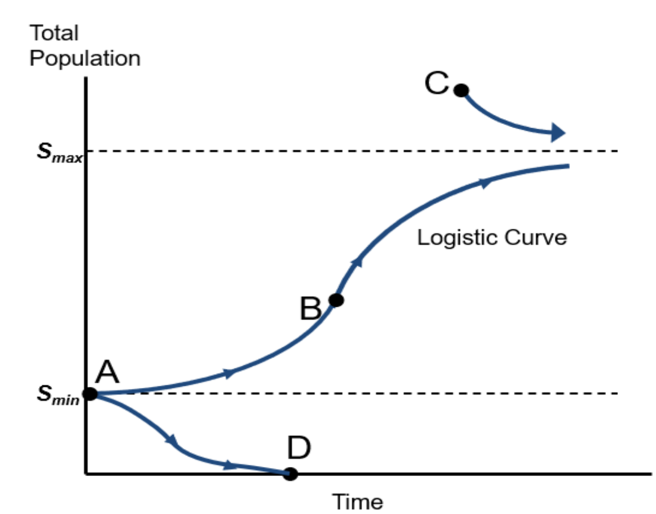


Fig 3.17a Species' population growth over time.
Source: Adapted from Jonathan and Roach, 2017

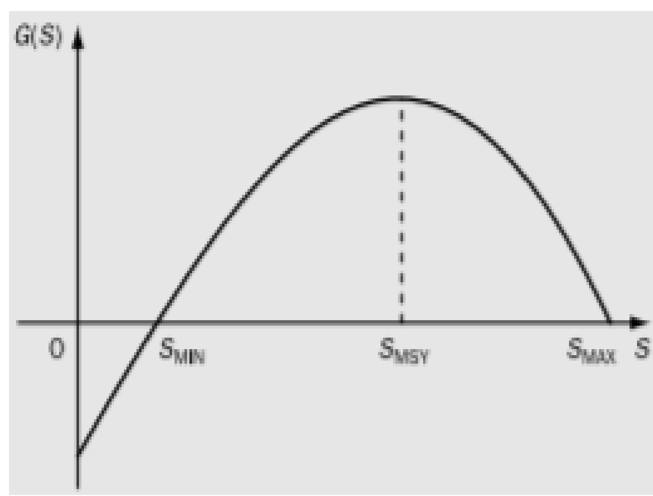


Fig 3.17b Biological growth model showing the relationship between the fish population stock and growth. Perman et al., 2003.p558.

[24] Several other generalizations of the logistic growth model exist apart from the one represented in equation 3.48. Another commonly used equation for biological growth, other than the logistic equation is the Gompertz function. For more on this, see Perman et al., 2003 pp557-560.

We note the following in relation to Figure. 3.17b.

- S_{MAX} is the maximum stock size that can be supported in the environmental milieu. This value is conditional on the particular environment circumstances that prevails, and would change if any of those circumstances (such as ocean temperature or stocks of nutrients) change.

- S_{MAX} is a **stable** or **natural equilibrium**; it is the population size that will persist in the absence of outside influence. Movements away from the population level set forces in motion to restore it. In this equilibrium, growth in the stock is zero: reductions in the stock due to mortality or out-migration would be exactly offset by increases in the stock due to births, growth of the fish in the remaining stock, and in-migration.

- S_{MIN} is the minimum viable population, the level of population below which growth in population is negative (deaths and out-migration exceed births and in-migration).

- S_{MIN} is an **unstable equilibrium**: population sizes to the right lead to positive growth and a movement along the curve to S_{MAX} and away from S_{MIN}. Population sizes to the left lead to declines and ultimately extinction.

- S_{MSY} is the **maximum sustainable yield** (MSY) population, the population that yields maximum growth rate, and hence maximum sustainable harvest. We will be looking at this concept more closely in the next subsection. For the simple logistic function, the maximum amount of growth, and hence the S_{MSY} will occur when the stock size is equal to half of S_{MAX}.

- There is a range of population sizes (S_{MIN} to S_{MSY} in Fig. 3.17b) where population growth increases as the population increases and a range (S_{MSY} to S_{MAX}) where initial increases in population lead to eventual declines in growth.

3.4.3 Steady-state harvesting, Maximum Sustainable Yield, and Efficiency

A sustainable yield of a non-renewable resource is a catch (or harvest) level that equals the growth rate of the population and can be maintained forever. As long as the population size remains constant, the growth rate (and hence the catch) will remain constant as well. There are various sustainable yield levels of a non-renewable resource, each corresponding with a given positive growth rate. However, there is a unique sustainable yield that gives that gives the highest level of sustainable harvest. This is called the **maximum sustainable yield** (MSY) (the largest catch that can be perpetually sustained) and it is the harvest level that corresponds to the growth rate of the specie at the MSY population (S_{MSY}).

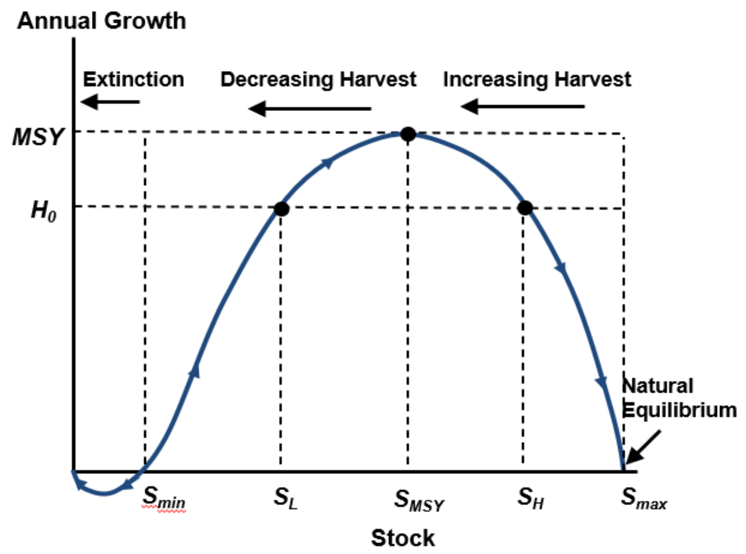


Figure 3.18a. Species population, Harvesting and annual Growth. **Source**: adapted from Jonathan and Roach, 2017

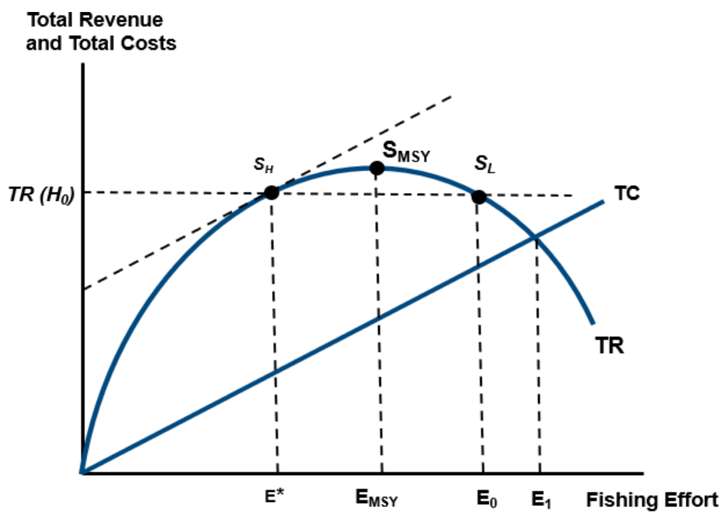


Fig. 3. 18b Total revenues, total costs and harvest effort in a fishery. **Source**: adapted from Jonathan and Roach, 2017

The economic analysis of fisheries, or similar renewable resources, derives from the previous (biological) curve. In Figure 3.18a and figure 3.18b, we introduce economic activities (harvesting) into the biological growth model. In Figure 3.18a, economic activity, reflected in harvest of the specie, is represented as a movement from right to left. At first, harvesting effort actually increases stock growth rates, but after the MSY population is reached, growth rates and annual harvest decline. The illustration shows that extreme over-harvesting can lead to species collapse and extinction. Larger catches would be possible in the short run, but would lead to reduced population sizes and may lead to extinction.

Figure 3.18b shows the relationship between total revenue/cost of harvest and **harvesting effort.** Benefits (revenues) and costs are portrayed as a function of fishing effort. In this case, effort is reflected in a movement from left to right. (Take note of the change in the positions of S_L and S_H in the two diagrams).

The shape of the revenue function is dictated by the shape of the biological function in Figure 3.18a. Multiplying the quantity of output by the market price of fish (assumed constant in this case) transforms the curve into Total Revenue terms. Total Costs is shown on the same graph as a straight line reflecting an assumption of constant marginal costs). The net benefit is the difference (vertical distance) between benefits (prices times the quantity caught) and costs (the constant marginal cost of effort times the units of effort expended).

As sustained levels of effort are increased, eventually a point is reached (E_{MSY}) at which further effort reduces the sustainable catch and revenue for all years. That point corresponds to the MSY population S_{MSY}) in Figure 3.18a, meaning that both points reflect the same population and growth levels. In fact, every effort level portrayed in Figure 3.18b corresponds to a specific population level in Figure 3.18a.

3.4.3.1. Steady-state harvesting

Consider a period of time in which the amount of the stock being harvested (H) is equal to the amount of net natural growth of the resource (G). Suppose also that these magnitudes remain constant over a sequence of consecutive periods. We call this **steady-state harvesting**, and refer to the (constant) amount being harvested as a sustainable yield. Defining F as the actual rate of change of the renewable resource stock, with F = G − H, it follows that in steady-state harvesting F = 0 and so the resource stock remains constant over time. What kinds of steady states are feasible?

The harvest level corresponding to the maximum sustainable yield population is a steady state harvest because at that harvest level the quantity of net natural growth (corresponding to the point MSY on the vertical line in Figure 3.18a) is at a maximum. If at a stock of S_{MSY} harvest is set at the constant rate MSY, we obtain a maximum sustainable yield (MSY) steady state. A resource management programme could be devised which takes this MSY in perpetuity. It is sometimes thought to be self-evident that a fishery, forest or other renewable resource should be managed so as to produce its maximum sustainable yield. We shall see later that economic theory does not, in general, support this proposition. The MSY is not the only possible steady-state harvest. Indeed, Figure 3.18b shows that any harvest level between zero and MSY is a feasible steady-state harvest, and that any stock between zero and S_{MSY} can support steady-state harvesting. For example, H_g is a feasible steady-state harvest if the stock size is maintained at either S_L or S_H. Which of these two stock sizes would be more appropriate for attaining a harvest level of H_g is also a matter we shall investigate soon.

Before moving on, it is important to understand that the concept of a steady state is a heuristic device: useful as a way of organizing ideas and structuring analysis. But, like all heuristic devices, a steady state is a mental construct and using it uncritically can be inappropriate or misleading. Fisheries and other resource stocks are rarely, if ever, in steady states. Conditions are constantly changing, and the 'real world' is likely to be characterized by a more-or-less permanent state of disequilibrium. For some problems of renewable resource exploitation, the analysis of transition processes is more important or insightful than information about steady states. We shall examine some of these ‘dynamic’ matters in Module 3.5. Nevertheless, we will assume that looking at steady states is useful, and investigate their properties under various institutional circumstances.

3.4.3.2 Efficiency: Static and Dynamic

It is easy to locate the static efficient level of harvest in this basic model. The **static efficient level of effort** is shown to be at point is E^* in Figure 3.18b. At this level of effort, the vertical distance between benefits and costs is maximized. Note that this is where the marginal benefit (the slope of the total benefit curve) is equal to marginal cost (the constant slope of the total cost curve) as required by statice efficiency conditions. Levels of effort higher than E^* are inefficient because the additional cost associated with them exceeds the value of the fish obtained.

From the diagram, it is obvious that the level of effort associated with MSY population, E_{MSY}, is not efficient. Diminishing returns take place as the MSY is approached. The marginal revenue obtained by expanding catch from the economic optimum to the maximum sustained yield falls below the marginal cost. The MSY would be efficient only if the marginal cost of additional effort were zero. (Can you explain why?). The level of effort necessary to harvest the maximum sustainable yield is higher than the static-efficient level of effort. Thus, the static-efficient level of effort leads to a larger fish population, but a lower annual catch than the maximum sustainable yield level of effort.

It will be interesting to consider what would happen to the static efficient sustainable yield if a technological change were to occur, for example, the use of sonar detection, lowering the marginal cost of fishing. A lower marginal cost would result in a rotation of the total cost curve to the right. With this new cost structure, the old level of effort would no longer be efficient. The marginal cost of fishing (slope of the total cost curve) would now be lower than the marginal benefit (slope of the total benefit curve) at E^*. Since the marginal cost is constant, the equality of marginal cost and marginal benefit can result only from a decline in marginal benefits. This implies an increase in effort. The new static efficient sustainable yield equilibrium implies more annual effort, a lower population level, a larger annual catch, and a higher net benefit for the fishery.

As we saw in the case of non-renewable resources, dynamic efficiency requires optimization over time. As in the former case, temporally interdependent allocations over time give rise to a marginal user cost measuring the opportunity cost of increasing current effort. This reflects the forgone future net benefits when more resources are extracted in the present. For efficient interdependent allocations, the marginal willingness to pay is equal to the marginal user cost plus the marginal cost of extraction. Again, as in the case of non-renewable resources, the interest or discount rate plays a crucial role. The static efficient sustainable yield can be interpreted as the special case of the dynamic efficient sustained yield where the discount rate is zero.

The effect of a positive discount rate for the management of a fishery is similar to its influence on the allocation of depletable resources—the higher the discount rate, the higher the cost (in terms of forgone current income) to the resource owner of maintaining any given resource stock. Thus, when positive discount rates are introduced, the efficient level of effort would be increased beyond that suggested by the static efficient sustained yield with a corresponding decrease in the equilibrium population level.

The increase in the yearly effort beyond the efficient sustained yield level would initially result in an increased net benefit from the increased catch. (Convince yourself that this is true by looking again at Figure 3.18b). However, since this catch exceeds the sustained yield for that population size, the population of fish would be reduced and future population and catch levels would be lower (Fig. 3.18a). Eventually, as that level of effort is maintained, a new, lower equilibrium level would be attained when the size of the catch once again equals the growth of the population. As the discount rate is increased, the dynamic efficient level of effort is increased until, with an infinite discount rate, it would become equal to E_1 in Fig. 3.18b, the point at which net benefits go to zero (Colin Clark, 1976). Note that with an infinite discount rate, the marginal user cost (MUC) is zero, because no value is received from future allocations. This implies that the marginal cost of extraction equals the constant price, and total benefits equal total costs.

While the static efficient sustained yield implies a larger fish population than the MSY. Once discounting is introduced, it is inevitable that the dynamic efficient sustained yield would imply a smaller fish population than the static efficient sustained yield and it is possible, though not inevitable, that the sustained catch would be smaller. The likelihood of the population being reduced below the MSY population depends on the discount rate. In general, the lower the extraction costs, and the higher the discount rate, the more likely it is that the dynamic efficient level of effort will exceed the level of effort associated with the MSY. When the marginal extraction cost is zero, the static efficient sustainable yield and the MSY are equal (why?), But with zero marginal extraction costs and a positive discount rate, the dynamic efficient level of effort necessarily exceeds not only the static efficient level of effort, but also the level of effort associated with the MSY. In contrast, higher extraction cost reduces the static efficient sustainable yield but not the MSY. (MSY is a biological, not an economic, concept.). Thus, by reducing efficient effort levels, higher extraction costs reduce the likelihood that discounting would cause the population to be drawn below the MSY.

In the model considered here and under the circumstances described, a dynamically efficient management does not lead to extinction of the fishery. As Figure 3.18b shows, E_1 is the highest possible dynamically efficient level in the model (why?), and that level falls well short of the level needed to drive the population to extinction. However, in more complex models, extinction certainly can be an outcome. For extinction to occur under a dynamic efficient management scheme, the benefit from extracting the very last unit would have to exceed the cost of extracting that unit (including the costs on future generations). As long as the population growth rate exceeds the discount rate, this will not be the case. If, however, the growth rate is lower than the discount rate, extinction can occur even in an efficient management scheme if the costs of extracting the last unit are sufficiently low.

Why does the biomass rate of growth have anything to do with whether or not an efficient catch profile leads to extinction? The reason is because rates of growth determine the productivity of conservation efforts. With high rates of growth, future generations can be easily satisfied. When the rate of growth is very low, it takes a large sacrifice by current generations to produce more fish for future generations. In the limiting case, where the rate of growth is zero, we have a resource with fixed supply, and is therefore no different from an exhaustible resource. Total depletion would occur whenever the price commanded by the resource is high enough to cover the marginal cost of extracting the last unit.

Actual fisheries differ from the standard model examined here in two key ways. First, harvesting marginal costs are typically not constant but rather increase as the remaining stock size diminishes. Second, prices may not be constant; the size of the harvest can affect prices (larger harvests can depress prices). Both realities suggest additional incentives for conserving the stock. The question then is, are these incentives important in practice? In the next Module, we will examine actual market allocation and contrast it with the dynamically efficient allocation. This will enable us to entertain the possibility of various public policy corrective measures that can be applied in cases where actual market outcomes differ from that suggested by the dynamic efficiency criterion. However, we will conclude this module by looking again at the three possible equilibria that may occur in the basic non-renewable resource (in this case, fishing) model.

3.4.4. Equilibria in Renewable resource Harvesting Model

When the transform the economic figures illustrated in the total revenue and total cost curves in Figure 3.18b into marginal and average revenue terms, we find three possible equilibria in our fishing model.

- The first, which we call the **open access equilibrium** (E_1 in Figure 3.19.) results if there are no limits on entry into the fishery.

- The second, the **maximum sustainable yield** E_{MSY} is not a market equilibrium, but as we showed under the consideration of steady-state harvests, it could be achieved by policies restricting output to MSY. Such policies (including quotas or fishing licenses) are discussed in Module 4.

- The third, the **economic optimum**, E^*can be achieved by private ownership of the fishery (as in an inland pond) or where this is not feasible (as in an ocean fishery) by regulation or by social convention (as in some traditional fisheries).

As evident, these equilibria are related to underlying property right regimes. In the next module we will be taking a detailed examination of property right regimes and their implications for the optimal management of non-renewable resources.

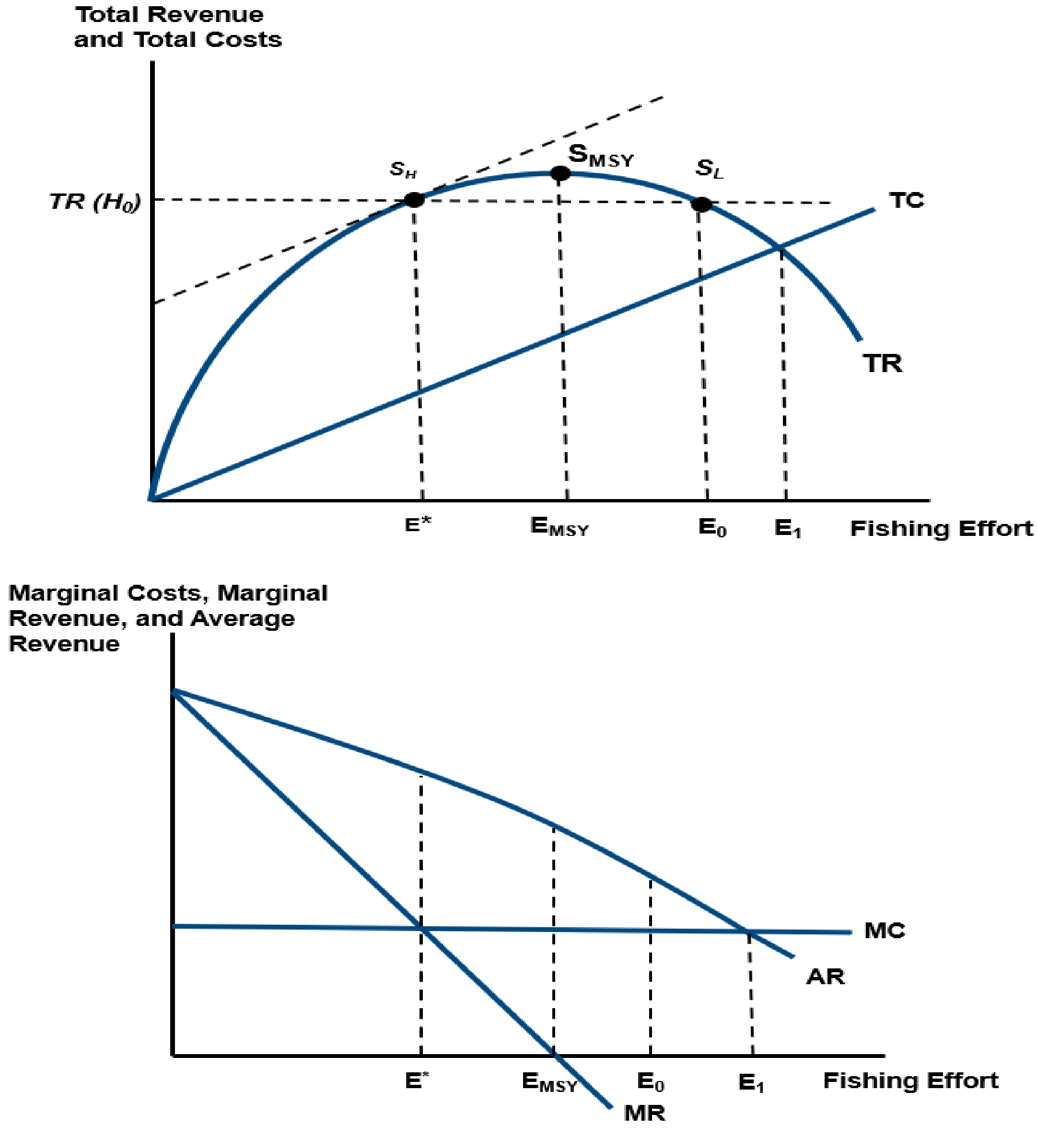


Fig.3.19. Property right regimes and Renewable resource harvesting. **Source:** Jonathan and Roach, 2017.

Summary

- The class of renewable resources is diverse. It includes populations of biological organisms such as fisheries and forests which have a natural capacity for growth, and water and atmospheric systems which are reproduced by physical or chemical processes. Most of the literature on the economics of renewable resources focus on the harvesting of animal species ('hunting and fishing') and the economics of forestry.

- We can distinguish between the stock and flow of a non-renewable resource. The **stock** is a measure of the quantity of the resource existing at a point in time; the **flow** is the change in the stock over an interval of time. The logistic growth function is a good approximation to the natural growth processes of many biological species

- Like non-renewable resources, the stock of a renewable resource can be driven to zero if conditions interfere with its reproductive capability, or if rates of harvesting continually exceed growth rate. How the stock of a common-pool resource, such as fisheries, forestry and other commercially-viable species changes depends on whether those who are best positioned to protect them have sufficient incentives to do so. A commercially valuable species is like a double-edged sword: on the one side, the value of the species to humans provides a strong, current reason for human concern about its future. On the other hand, its value may promote excessive harvest.

- A sustainable yield of a non-renewable resource is a catch (or harvest) level that equals the growth rate of the population and can be maintained forever. There are various sustainable yield levels of a non-renewable resource, each corresponding with a given positive growth rate. However, there is a unique sustainable yield that gives the highest level of sustainable harvest. This is called the Maximum sustainable yield (MSY).

- A steady-state harvesting refers to a situation where a sustainable yield of a non-renewable resource is maintained over consecutive period of time so that the resource stock remains constant over time. There are many steady-state harvesting scenarios that are possible with a non-renewable resource, including that involving the MSY. However, fisheries and other resource stocks are rarely, if ever, in steady states. Conditions are constantly changing, and the 'real world' is likely to be characterized by a more-or-less permanent state of disequilibrium. For some problems of renewable resource exploitation, the analysis of transition processes is more important or insightful than information about steady states.

- For a non-renewable resource, the static efficiency level of harvest is where the marginal benefit of harvest is equal to the marginal cost. The harvest level associated with the MSY is not efficient. The static-efficient level of harvesting effort leads to a larger fish population, but a lower annual catch than the maximum sustainable yield level of effort.

- The static efficient sustainable yield can be interpreted as the special case of the dynamic efficient sustained yield where the discount rate is zero. When positive discount rates are introduced, the efficient level of effort would be increased beyond that suggested in the static case with a corresponding decrease in the equilibrium population level. As the discount rate is increased, the dynamic efficient level of effort is increased until, with an infinite discount rate, it occurs at the point at which net benefits go to zero.

- There are three possible equilibria in the basic model of renewable resource harvesting. The first is an open-access equilibrium, the second (which can only be achieved as an equilibrium through policy intervention) is associated with the maximum sustainable yield. The third is the economic optimum. Which of the three will prevail is determined by the property right regime.

Discussion/Review Questions and Exercises

1. Explain the following concepts in relation to renewable resources

(a) density-dependent growth.

(b) property of compensation.

(c) State and explain a function depicting the property of compensation

2. Provide a simple model and illustration to express the growth dynamics of a biological (renewable) resource.

3. With the aid of necessary diagrammatic illustration show and explain why the efficient level of harvest of a non-renewable resource is not necessarily the one that is consistent with the maximum sustainable yield.

4. Explain the importance of the biomass rate of growth of a renewable resource in deterring whether or not a dynamically efficient catch profile leads to extinction?

5. Using relevant diagrams, explain the kind of equilibria that can prevail in a model of renewable resource harvesting. What are the implications for property right structure?

Materials used for the Lecture notes

Jonathan M. Harris and Brian Roach (2017), **Environmental and Natural Resource Economics** *4th Edition*, Routledge.

Perman, R., Ma Y., McGilvray J. and Common M. (203). **Natural Resource and Environmental Economics**, *3rd* Edition, Edinburgh, Longman.

Tietenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics.

Module 3.5 Renewable Resources: Optimal Harvesting under Different Property Right Regimes (7 hours)

Learning Outcomes

This Module examines optimal harvesting under different property right regimes and also compares them with the socially efficient resource harvesting. After going through the module, the reader should

- ✓ be able to distinguish between an open-access fishery and a private-property fishery, and explain the characteristics and forms of private property fishery.
- ✓ be able to distinguish between a biological and economic equilibrium in a fishery model.
- ✓ understand the kinds of equilibria that are likely to prevail in a model of renewable resource harvesting and why.
- ✓ know the implications of these for environmental and resource management policy.
- ✓ know how externality affect the benefit and cost functions for private operators in a renewable resource (such as fishery) and how benefit and cost externalities cause the outcome of present-value (private-property)-maximizing fisheries to differ from the socially efficient outcome.
- ✓ understand how a private property fishery regime may help internalize a cost externality associated with the fishery.
- ✓ know the factors that determine whether or not the population of a non-renewable resource is likely to be driven to extinction by harvesters/hunters.

Outline

3.5.1. Open-access Harvesting Model
- 3.5.1.1 The Model
- 3.5.1.2 Equilibrium
- 3.5.1.3 Dynamics of renewable resource harvesting
- 3.5.1.4 Some concluding thoughts on open-access fishery

3.5.2. The Private-property fishery
- 3.5.2.1 The static model
- 3.5.2.2 The present-value-maximizing fishery model

3.5.3. Socially efficient resource harvesting
- 3.5.3.1 Externalities in the benefits function
- 3.5.3.2 Externalities in the fishery production function

3.5.3.3 Monopolistic fisheries
3.5.4. A Safe Minimum Standard (SMS) of conservation
3.5.5. Some Empirical Evidence
Summary
Discussion/Review Questions and Exercises
Materials used for the Lecture

It will appear, I hope, that most of the problems associated with the words 'conservation' or 'depletion' or 'overexploitation' in the fishery are, in reality, manifestations of the fact that the natural resources of the sea yield no economic rent. Fishery resources are unusual in the fact of their common-property nature; but they are not unique, and similar problems are encountered in other cases of common-property resource industries, such as petroleum production, hunting and trapping, etc. (**Source:** Gordon, 1954, in Perman et al, 2004, p. 555.

In an overpopulated (or overexploited) world, a system of the commons leads to ruin....Even if an individual fully perceives the ultimate consequences of his actions he is most unlikely to act in any other way, for he cannot count on the restraint his conscience might dictate being matched by a similar restraint on the part of all others (**Source:** Garrett Hardin, Carrying Capacity as an Ethical Concept (1967), in Tietinberg and Lewis, 2012,p.320.

3.5. 1 Open-access Harvesting Model

In Module 3.4 we saw that there are three possible equilibrium in a renewable resource harvesting model and that the prevailing outcome is determined by the property right structure. We examine this thought in details in this Module. The first model of renewable resource exploitation we will examine is an open-access fishery model. It is important to be clear about what an open access fishery is taken to mean in the environmental economics literature. The open-access fishery model shares two of the characteristics of the standard perfect competition model. First, if the fishery is commercially exploited, it is assumed that this is done by a large number of independent fishing 'firms'. Therefore, each firm takes the market price of landed fish as given. Second, there are no impediments to entry into and exit from the fishery.

But the free entry assumption has an additional implication in the open-access fishery, one which is not present in the standard perfect competition model. In a conventional perfect competition model, each firm has enforceable property rights to its resources and to the fruits of its production and investment choices. However, in an open-access fishery, while owners have individual property rights to their fishing capital and to any fish that they have actually caught, they have no enforceable property rights to the *in situ* fishery resources, including the fish in the water. The lack of de facto enforceability may derive from the fact that the fish are spatially mobile, or from the fact that boats are spatially mobile (or both). On the contrary, any vessel is entitled or is able (or both) to fish wherever its owner likes. Moreover, if any boat operator chooses to leave some fish in the water in order that future stocks will grow, that owner has no enforceable rights to the fruits of that investment. This reality has implications for the kind of equilibrium that prevail.

3.5.1.1 The Model

The open-access model has two components: (i) a biological sub-model, describing the natural growth process of the fishery; and (2). an economic sub-model, describing the economic behaviour of the fishing boat owners. We shall be looking for two kinds of 'solutions' to the open-access model. The first is its equilibrium (or steady-state) solution. This consists of a set of circumstances in which the resource stock size is unchanging over time (a **biological equilibrium**) and the fishing fleet is constant with no net inflow or outflow of vessels (an **economic equilibrium**). Because the steady-state equilibrium is a joint biological–economic equilibrium, it is often referred to as **bioeconomic equilibrium**. The second kind of solution we shall be looking for is the adjustment path towards the equilibrium, or from one equilibrium to another as conditions change. In other words, our interest also lies in the dynamics of renewable resource harvesting. As we saw in Module 3.4, this has important implications for whether a fish population may be driven to exhaustion, and indeed whether the resource itself could become extinct. We first establish the biological and economic sub-models.

In the absence of harvesting and other human interference, the rate of change of the stock depends on the prevailing stock size

$$dS/dt = G(S) \qquad (3.49)$$

We assume that the particular form taken by this growth function is the simple logistic growth model examined in Module 3.4.

To derive the **harvest function** (or fishery production function), we must know the factors that determine the size of harvest H, in any given period. We consider two factors: the **amount of resources devoted to fishing** and the **size of the resource stock**. In the case of marine fishing, the amount of resources devoted to fishing could include the number of boats deployed and their efficiency, the number of days when fishing is undertaken, etc. For simplicity, we assume that all the different dimensions of harvesting activity can be aggregated into one magnitude called effort, *E*. In relation to resource stock, other things being equal, the larger the stock the greater the harvest for any given level of effort. Hence, abstracting from other determinants of harvest size, including random influences, we may take harvest to depend upon the effort applied and the stock size. That is

$$H = H(E, S) \qquad (3.50)$$

This relationship can take a variety of particular forms. One very simple form, which appears to be a good approximation to actual relationships (see Schaefer, 1954 and Munro, 1981, 1982), is given by

$$H = eES \qquad (3.51)$$

where e is a constant number, often called the **catch coefficient**.[25]

Dividing each side of (3.51) by E, we have

$$\frac{H}{E} = eS \qquad (3.52)$$

which says that the quantity harvested per unit effort is equal to some multiple (e) of the stock size.

We have already defined the fish-stock growth function with human predation as the biological growth function less the quantity harvested. That is,

$$\dot{S} = G(S) - H \qquad (3.53)$$

The total cost of harvesting, C, depends on the amount of effort being expended

$$C = C(E) \qquad (3.54)$$

For simplicity, harvesting costs are taken to be a linear function of effort,

$$C = wE \qquad (3.55)$$

where w is the cost per unit of harvesting effort, taken to be a constant.[26]

Let B denote the gross benefit from harvesting some quantity of fish. The gross benefit will depend on the quantity harvested, so we have $B = B(H)$. In a commercial fishery, the appropriate measure of gross benefits is the total revenue that accrues to firms. Assuming that fish are sold in a competitive market, each firm takes the market price P as given and so the revenue obtained from a harvest H is given by

$$B = PH \qquad (3.56)$$

Fishing profit is given by

$$NB = B - C \qquad (3.57)$$

How are fishing efforts determined under conditions of open access? A crucial role is played here by the level of economic profit prevailing in the fishery. Given that there is no method of excluding incomers into the industry, nor is there any way in which existing firms can be prevented from changing their level of harvesting effort, effort applied will continue to increase

---

[25]. The use of a constant catch coefficient parameter is a simplification that may be unreasonable, and is often dropped in more richly specified models. Note also that equation 3.51 can be regarded as a special case of the more general form H = e$E^{\alpha}S^{\beta}$ in which the exponents need not be equal to unity. In empirical modelling exercises, this more general form may be more appropriate. Another form of the harvest equation sometimes used is the exponential model H = S(1 − exp(−eE)) (see Perman et al. 2003. P564-565).

[26] The assumption that harvesting costs are linearly related to fishing effort may not be valid in many cases.

as long as it is possible to earn positive economic profit.[27] Conversely, individuals or firms will leave the fishery if revenues are insufficient to cover the costs of fishing. A simple way of representing this algebraically is by means of the equation

$$dE/dt = \delta \cdot NB \qquad (3.58)$$

where δ is a positive parameter indicating the responsiveness of industry size to industry profitability.

When economic profit (NB) is positive, firms will enter the industry; and when it is negative, they will leave. The magnitude of that response, for any given level of profit or loss, will be determined by δ. Although the true nature of the relationship is unlikely to be of the simple, linear form in equation 3.58, this suffices to capture what is essential.

3.5.1.2 Equilibrium

The biological equilibrium occurs where the resource stock is constant through time (that is, it is in a steady state). This requires that the amount being harvested equals the amount of net natural growth:

$$G = H \qquad (3.59)$$

In contrast, economic equilibrium requires that the amount of fishing effort be constant through time. Such an equilibrium is only possible in open-access fisheries when rents have been driven to zero, so that there is no longer an incentive for entry into or exit from the industry, nor for the fishing effort on the part of existing fishermen to change. We express this by the equation

$$NB = B - C = 0 \qquad (3.60)$$

which implies (under our assumptions) that PH = wE.

Notice that when this condition is satisfied, $dE/dt = 0$ and so effort is constant at its equilibrium (or steady-state) level $E = E^*$.

We can envisage an open-access fishery steady-state equilibrium by means of the **fishery's yield–effort relationship** examined in Module 3.4. To obtain this, first note that in a biological equilibrium $H = G$. Then, by substituting the assumed functions for H and G from equations 3.51 and 3.47 (see Module 3.4) respectively we obtain:

[27] The terms rent, economic rent, royalties and net price may be used as alternatives to economic profit. They all refer to a surplus of revenue over total costs, where costs include a proper allowance for the opportunity of capital tied up in the fishing fleet.

$$gS\,(1-\frac{S}{S_{MAX}}) = eES \qquad (3.61)$$

which can be rearranged to give

$$S = S_{MAX}\,(1-\frac{e}{g}E) \qquad (3.62)$$

Equation (3.62) is one equation in two endogenous variables, *E* and *S* (with parameters *g*, *e* and S_{MAX}). It implies a unique equilibrium stock at each level of effort. Substituting equation (3.62) into equation (3.51) (*H = eES*), gives

$$H = eES_{MAX}\,(1-\frac{e}{g}E) \qquad (3.63)$$

In an open-access economic equilibrium, profit is zero, so

$$PH = wE \qquad (3.64)$$

Equations (3.63) and (3.64) constitute two equations in two unknowns (*H* and *E*); these can be solved for the equilibrium values of the two unknowns as functions of the parameters alone. The steady-state stock solution can then be obtained by substituting the expressions for *H* and *E* into equation (3.51). Thus, the steady-state solutions for *E*, *H* and *S* can be derived for any particular set of assumptions about numerical values of the parameters *g*, S_{MAX}, *e, δ, P* and *w*. In general, the analytical expressions for *E*, S* and H** (where an asterisk denotes the equilibrium value of the variable in question) as functions of the model parameters alone are:

$$E^* = \frac{g}{e}\,(1-\frac{W}{PeS_{MAX}}) \qquad (3.65)$$

$$S^* = \frac{W}{Pe} \qquad (3.66)$$

$$H^* = \frac{gW}{Pe}\,(1-\frac{W}{PeS_{MAX}}) \qquad (3.67)$$

As illustrated in Module 3.4, the solution method can also be represented graphically, as shown again in Figures 3.20a and 3.20b. In Figure 3.20a, the three rays emanating from the origin portray the harvest– stock relationships (from the function *H = eES*) for three different levels of effort. If effort were at the constant level E_1, then the unique intersection of the harvest–stock relationship and biological growth function determines a steady-state harvest level H_1 at stock S_1. The lower effort level E_2 determines a second steady-state equilibrium (the pair {H_2, S_2}).

The various points of intersection satisfy equation (3.62), being equilibrium values of *S* for particular levels of *E*. Clearly there is an infinite quantity of possible equilibria, depending on what constant level of fishing effort is being applied. The points of intersection in Figure 3.20b not only satisfy equation (3.62) but they also satisfy equation (3.63). Put another way, the equilibrium {*E, S*} combinations also map into equilibrium {*E, H*} combinations. The result of this mapping from

{*E, S*} space into {*E, H*} space is shown in the fishery's yield–effort relationship (Figure 3.20b). This portrays the steady-state harvests that correspond to each possible effort level.

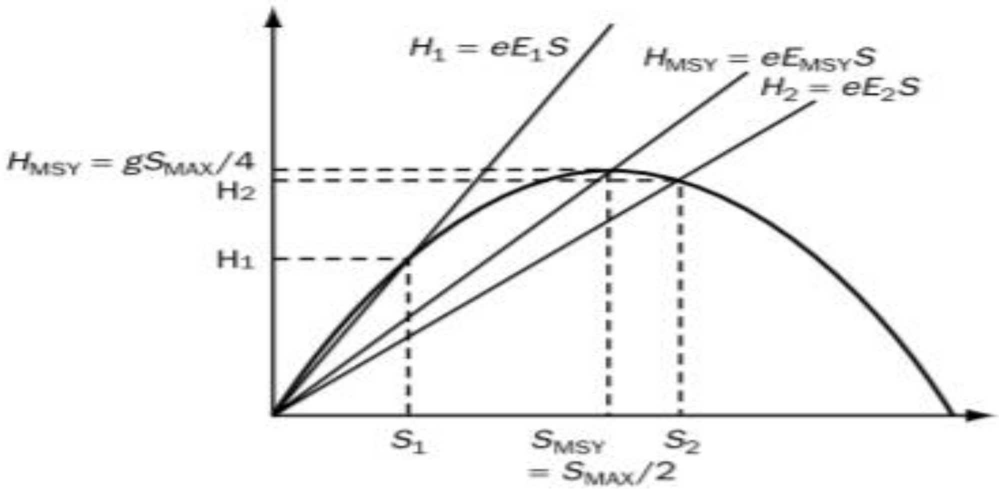


Fig 3.20a Steady-state equilibrium fish harvests and stocks at various effort levels

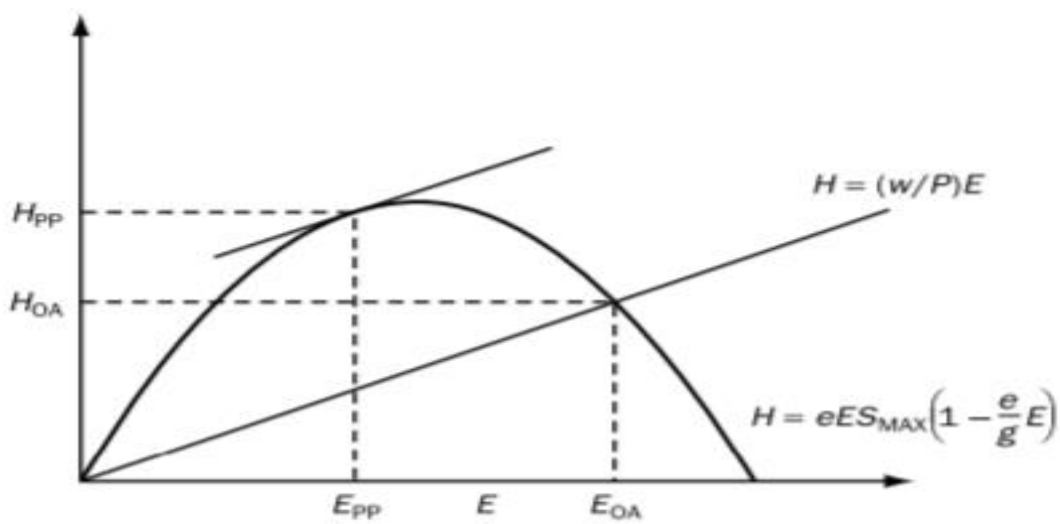


Fig. 3.20b. Steady-state equilibrium yield–effort relationship.

Source: Perman et al., 2003. P.565.

Mathematically, it is a plot of equation (3.63). The particular point on this yield–effort curve that corresponds to an open-access equilibrium will be the one that generates zero economic profit. The zero economic profit equilibrium condition *PH = wE* can be written as *H = (w/P)E*. For given values of P and w, this plots as a ray from the origin with slope w/P in Figure 3.20b. The intersection of this ray with the yield–effort curve locates the unique open-access equilibrium outcome. Alternatively, multiplying both functions in Figure 3.20b by the market price of fish, P, we find that the intersection point corresponds to *PH = wE*. This is, of course, the zero-profit condition, and confirms that $\{E_{OA}, H_{OA}\}$ is the open-access effort–yield equilibrium.

3.5.1.3 Dynamics of renewable resource harvesting

Our discussion so far has been exclusively on steady states: equilibrium outcomes which, once achieved, would remain unchanged provided that relevant economic or biological conditions remain constant. However, we may also be interested in the dynamics of resource harvesting. This would consider questions such as how a system would get to a steady state if it were not already in one, or whether getting to a steady state is even possible. In other words, dynamics is about transition or adjustment paths. Dynamic analysis might also give us information about how a fishery would respond, through time, to various kinds of shocks and disturbances.

A complete description of fishery dynamics is beyond the scope of this Module but we will give some important insights. We will focus once again, on the dynamic analysis for the open-access model. Suppose a mature fishery exists that has not previously been commercially exploited. The stock size is, therefore, at its **carrying capacity**. The fishery now becomes available for unregulated, open-access commercial exploitation. If the market price of fish, *P*, is reasonably high and fishing cost (per unit of effort), *w*, is reasonably low, the fact that stocks are high (and so easy to catch) implies that the fishery will be, at least initially, profitable for those who enter it. Have in mind equations (3.51), (3.55), (3.56) and (3.57) when thinking this through. If a typical fishing boat can make positive economic profit then further entry will take place.

How quickly new capacity is built up depends on the magnitude of the parameter δ in equation (3.58). In this early phase, effort is rising over time as new boats are attracted in, and stocks are falling. Stocks fall because harvesting is taking place while new recruitment to stocks is low (recall that the logistic growth function has the property that biological growth is near zero when the stock is near its maximum carrying capacity). This process of increasing *E* and decreasing *S* will persist for some time, but it cannot last indefinitely. As stocks become lower, fish become harder to catch and so the cost per fish caught rises. Profits are squeezed from two directions: harvesting cost per fish rises, and fewer fish are caught. Eventually, this profit squeeze will mean that a typical boat makes a loss rather than a profit, and so the process we have just described goes into reverse, with stocks rising and effort falling. For the model we are examining, the processes of adjustment are subtle in the sense that the changes do not occur as discrete switches but instead are continuous and gradual. We also find that stocks and effort (and also harvest levels) have oscillatory cycles with the stock cycles slightly leading the effort cycles. In some circumstances, these oscillations dampen down as time passes, and the system eventually settles to a steady-state outcome such as that described in the previous section.

Figure 3.21 illustrate the kind of oscillations around a steady state equilibrium that might be encountered in open-access fishery. The oscillations shown in this diagram are particularly acute and have massive amplitude; for other combinations of parameter values, the cycles may be far less pronounced. In this case, the steady-state outcome obtained by solving equations (3.65, 3.66 and 3.67) is eventually achieved (albeit very slowly). But that is not necessarily true; if the oscillations were even larger, the population may be driven down to a level from which it cannot recover, and the fishery is driven out of existence, irrespective of whether or not the equilibrium equations imply that a steady state exists with positive stock and effort.

Other things being equal, the probability that effort oscillations might cause a population to collapse before it can attain a steady-state solution is increased when

- there is a critical minimum threshold, that is the growth function has $S_{MIN} > 0$, and
- the growth function or environmental conditions are stochastic.

The first of these should be self-evident. A positive minimum population size – found when the growth function exhibits **critical depensation**, for example – implies that low populations that would otherwise recover through natural growth may collapse irretrievably to zero. The second is also relatively straightforward. If in practice there is a stochastic component to growth then the achieved growth in any period may be smaller or larger than its underlying mean value as determined by the biological growth equation. In some cases, the stochastic component may be such that growth is actually negative. If the population were at a very low level already, then a random change inducing lower than normal or negative net growth could push the resource stock below the point from which biological recovery is possible. In addition, our analysis so far has presumed that environmental conditions remain unchanged over time. However, in practice these conditions do change, usually unpredictably, and sometimes very rapidly. When open access to a resource is accompanied by worsening environmental conditions, and when these changes are either rapid or unforeseen or both, harvesting outcomes can be catastrophic.

3.5.1.4 Some concluding thoughts on open-access fishery

In an open-access fishery, overfishing does not usually happen because of ignorance; it tends to result from the forces of competition because access is poorly regulated. Self-regulation on the part of fishermen may do little to overcome the consequences of open access. As we saw in Module 2, under conditions of ‘common property’ where private property rights are vested in communities. relationships of trust and social norms *may* be sufficiently well developed and entrenched to create patterns of resource exploitation that are both sustainable and rational for the group as a whole. In contrast to **common property regimes**, open access implies the absence of any such binding norms with a tendency for exploitation to take place under conditions of 'free-for-all' individualistic competition.

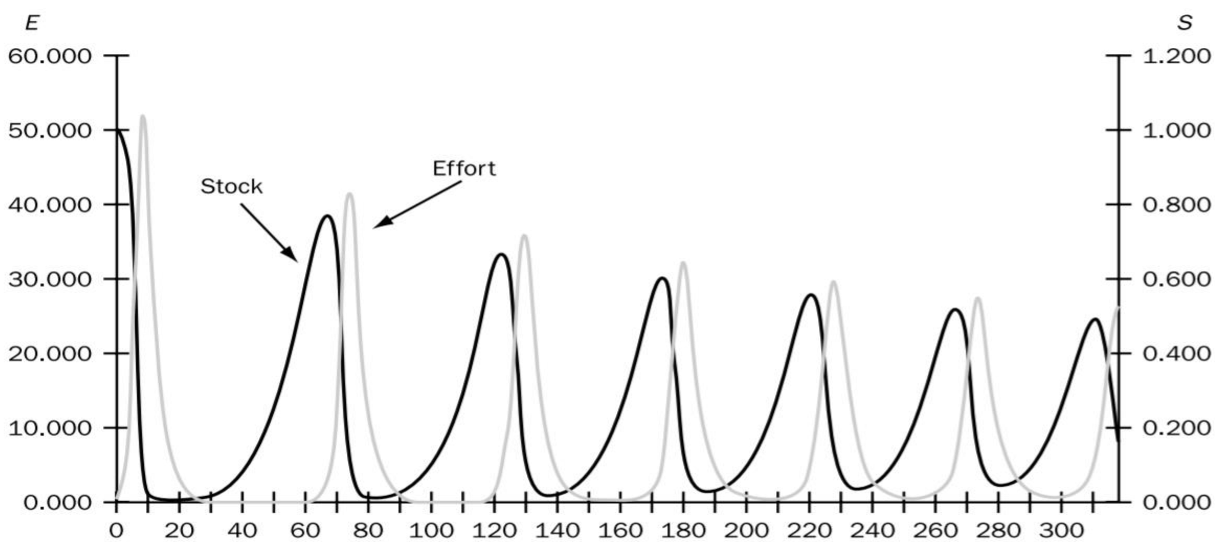


Fig. 3.21 Stock and effort dynamic paths

It is possible that any resource stock could be harvested to exhaustion, or a specie driven to extinction, under open access. While it is also true that this is possible under almost any regime, including those with enforceable private property rights, it remains true that open-access conditions increase the probabilities of those outcomes occurring. The main reason for this is that in these circumstances there is no collectively rational management of harvesting taking place. Even where what should be done is evident, an institutional mechanism to bring this about is missing. Another way of thinking about this is in terms of **stock externality effects** and economic inefficiency. Open-access harvesting programmes are inefficient because the resource harvesters are unable to appropriate the benefits of investment in the resource. If a single fisher were to defer the harvesting of some fish until a later period, all fishers would benefit from this activity. It would be in the interests of all if a bargain were made to reduce fishing effort by the industry. However, the conditions under which such a bargain could be made and not reneged upon are very unlikely to exist. Each potential bargainer has an incentive to free-ride once a bargain has been struck, by increasing his or her harvest while others reduce theirs. Moreover, even if all existing parties were to agree among themselves, the open-access conditions imply that others could enter the market as soon as rents became positive. Open-access resources thus have one of the properties of a public good – non-excludability – and this alone is sufficient to make it likely that markets will fail to reach efficient outcomes.

In the section that follows, we shall show that harvest rates are typically higher under conditions of open access compared to private property regimes. Other things being equal, the greater are harvesting rates, the higher is the likelihood of extinction.

3.5.2. The private-property (fishery) Model

In an open-access fishery, firms exploit available stocks as long as positive profit is available. While this condition persists, each fishing vessel has an incentive to maximize its catch. But there is a dilemma here, both for the individual fishermen and for society as a whole. From the perspective of the fishermen, the fishery is perceived as being overfished. Despite each boat owner pursuing maximum profit, the collective efforts of all drive profits down to zero. From a social perspective, the fishery will be economically 'overfished' and the stock level may (but will not necessarily) be driven down to biologically dangerous levels. The underlying cause of this state of affairs is the public-good nature of the open-access fishery. Although reducing the total catch today may be in the collective interest of all (by allowing fish stocks to recover and grow), it is not rational for any fisherman to individually restrict fishing effort. There can be no guarantee that he or she will receive any of the rewards that this may generate in terms of higher catches later. Indeed, there may not be any stock available in the future. In such circumstances, each firm will exploit the fishery today to its maximum potential, subject only to the constraint that its revenues must at least cover its costs.

The institutional arrangements of a **private-property fishery** may help overcome this dilemma. The private-property fishery has the following three characteristics:

- There is a large number of fishing firms, each behaving as a price-taker and so regarding price as being equal to marginal revenue. Thus, the industry can be described as being competitive.
- Each firm is profit- (or wealth-) maximizing.
- There is a particular structure of well-defined and enforceable property rights to the fishery, such that owners can control access to the fishery and appropriate any rents that it is capable of delivering.

There are several particular structure of property rights that are possible and will be consistent with the definition of private property in this context. One of two common views, regards the fishery as an aggregate of a large number of smaller individual fisheries. Each of these sub-fisheries is privately owned by one firm that has property rights to the fish which are there currently and at all points in time in the future. The owners of any fishing firm may, of course, lease or sell their property rights to another set of individuals. All harvested fish, however, sell in one aggregate market at a single market price. A second view regards the fishery as being managed by a single entity which controls access to the fishery and coordinates the activity of individual operators to maximize total fishery profits (or wealth). Nevertheless, harvesting and pricing behaviour are competitive rather than monopolistic.

Neither of these views is satisfactory as a statement of what actually does exist, nor what might realistically exist. The first faces problems in deciding how to specify ownership rights to migratory fish. Moreover, it could only be descriptively accurate if the fishery in question is huge, highly spatially aggregated, fishery. A typical researcher would not want to study at this level of aggregation. The second concept – the **coordinated fishery** – seems problematic in that we rarely, if ever, find examples of such internally coordinated fisheries (except in the case of fish farming etc.). And even if one were to find examples, it is difficult to imagine that they would operate as competitive fisheries rather than as monopolies or cartels. In fact, **monopoly fishery** and the **cartel fishery** are two other variants of the private-property fishery sometimes discussed in the literature. However, they are not common in practice.

Despite the operational difficulties associated with the two views mentioned above they help in developing public policy towards fishery regulation and management. If we are confident that a particular property rights structure would bring about socially efficient (or otherwise desirable) outcomes, then policy instruments can be designed to mimic that structure.

3.5.2.1 The static model

The analysis in thus section supposes that biological and economic conditions remain constant over some span of time, and then investigates what aggregate level of effort, stock and harvest would result if each individual owner (with enforceable property rights) managed affairs so as to maximize profits in any arbitrarily chosen period of time. This analytical approach only generates wealth-maximizing outcomes if fishermen do not discount future cash flows. More specifically, the static private-property fishery turns out to be a special case of a multi-period fishery model in which owners use a zero-discount rate.

The biological and economic equations of the static private-property fishery model are identical to those of the open-access fishery in all respects but one: the open-access entry rule (equation 3.58: $dE/dt = \delta \cdot NB$), which in turn implies a zero-profit economic equilibrium, no longer applies. Instead, owners choose effort to maximize economic profit from the fishery. This can be visualized with the help of Figure 3.21b. As we did earlier, multiply both functions by the market price of fish. The inverted U-shape yield–effort equation then becomes a **revenue–effort equation**. And the ray emerging from the origin now becomes $PH = wE$, with the right-hand side thereby denoting fishing costs. Profit is maximized at the effort level which maximizes the surplus of revenue over costs. Diagrammatically, this occurs where the slopes of the total cost and total revenue curves are equal. This in indicated in Figure 3.20b by the tangent to the yield–effort function at E_{PP} being parallel to the slope of the H = (w/P)E line.

An algebraic derivation of the steady-state solution to this problem – showing stock, effort and harvest as functions of the parameters – is as follows. The derivation initially follows exactly that given in Section 3.5.1.2, with equations (3.61), (3.62) and (3.63) remaining valid here. However, the zero-profit condition (equation 3.64) is no longer valid, being replaced by the profit-maximization condition:

$$\text{Maximize } NB = PH - wE \qquad (3.68)$$

Remembering that *H = eES,* and treating *E* as the instrument variable, this yields the necessary first-order condition,

$$\partial(PeES)/\partial E = \partial(wE)/\partial E \qquad (3.69)$$

Substituting equation (3.62) into (3.69) we have

$$\partial\left\{(PeS_{MAX}\,(1 - \frac{e}{g}\,\mathrm{E})\right\}/\partial \mathrm{E} = \partial(\mathrm{wE})/\partial \mathrm{E} \qquad (3.69)$$

from which we obtain after differentiation

$$PeS_{MAX} + 2\,PeS_{MAX}(\frac{e^2}{g}) = \mathrm{w} \qquad (3.70)$$

That is, the marginal revenue of effort is equal to the marginal cost of effort. This can be solved for E^*_{PP} (the subscript denoting 'private property') to give

$$E^*_{PP} = \frac{1}{2}\frac{g}{e}(1 - \frac{w}{PeS_{MAX}}) \qquad (3.71)$$

Substitution of E^*_{PP} into (3.62) gives

$$S^*_{PP} = \frac{1}{2}\frac{PeS_{MAX}+w}{Pe} \qquad (3.72)$$

and then using *H = eES* we obtain

$$H^*_{PP} = \frac{1}{4}g(S_{MAX} - \frac{w^2}{P^2e^2S_{MAX}}) \qquad (3.73)$$

Table 3.6 compares the values of the steady-state equilibrium stock, harvest and effort for both open-access and static private-property fishery for specific assumed values of the underlying parameters. Under the assumptions we have made about functional forms, the static private-property equilibrium will always lead to a higher resource stock level and a lower effort level than that which prevails under open access. This is confirmed for our particular parameter assumptions, with the private-property stock being three times higher and effort only half as large as in open access.

The steady-state harvest may be higher, lower or identical. This is evident from inspection of Figure 3.20b. For the particular set of parameter values used in our illustrative example, private-property harvest is larger than open-access harvest, as shown in the diagram. But it will not always be true that private property harvests exceed those under open access. The source of this indeterminacy follows from the inverted U shape of the yield–effort relationship. Although stocks will be higher under private property than open access, the quadratic form of the stock–harvest relationship implies that harvests will not necessarily be higher with higher stocks.

Table 3.6 Comparing steady-state solutions under open-access and private- property fishery

Parameter value assumptions for the illustrative numerical example		**Steady-state solutions under parameter value assumptions**		
Parameter	Assumed numerical value		Open access	Static private property
G	0.15			
S_{MAX}	1			
E	0.015	Stock	0.200	0.600
δ	0.4	Effort	8.000	4.000
P	2.00	Harvest	0.024	0.036
W	0.6			

3.5.2.2 The present-value-maximizing fishery model

The present-value-maximizing fishery model generalizes the model of the static private-property fishery. The essence of the model is that a rational private-property fishery will organize its harvesting activity so as to maximize the present value (PV) of the fishery. The individual components of the model are very similar to those of the static private fishery model except that it takes account of time, and hence, is set within the context of intertemporal optimization. We will limit our investigation in this Module to the general results and its implications.

As in the case of non-renewable resources, the key to understanding profit-maximizing behaviour when access to the resource can be regulated lies in capital theory. A renewable resource is a capital asset (think about a fishery with a single owner who can control access to the resource and appropriate all returns from it). We wish to consider the owner's decision about whether to marginally change the amount of fish harvesting currently being undertaken. Choosing not to harvest some fish is equivalent to a capital investment. The uncaught fish will be there next period; moreover, biological growth will mean that there is an additional increment to the stock next period (over and above the quantity of fish left unharvested). This amounts to saying that the asset – in this case the fishery – is productive.

A decision about whether to defer some harvesting until the next period is made by comparing the marginal costs and benefits of adding additional units to the resource stock. By choosing not to harvest an incremental unit, the fisher incurs an opportunity cost in holding a stock of unharvested fish. Holding these units sacrifices an available return. The marginal cost of the investment is *ip,* where *i* = prevailing rate of return on capital, *p* is the forgone revenue or net price of the resource and is equal to the difference between the market price and the marginal cost ($P - C_H$). Thus, *ip,* is the present value of the sacrificed return. The owner compares this marginal cost with the marginal benefit obtained by the resource investment (not harvesting the incremental unit). There are two categories of benefit:

- As a consequence of an additional unit of stock being added, total harvesting costs will be reduced by the quantity $CS = \partial C/\partial S$ (note that $\partial C/\partial S < 0$).
- The additional unit of stock will grow by the amount *dG/dS*. The value of this additional growth is the amount of growth valued at the net price of the resource.

A present-value-maximizing owner will add units of resource to the stock only if the marginal cost of doing so is less than the marginal benefit. That is:

$$ip < p\frac{dG}{dS} - \frac{dC}{dS} \qquad (3.74)$$

This states that a unit will be added to stock provided its 'holding cost' (*ip*) is less than the sum of its harvesting cost reduction and value-of-growth benefits. Conversely, a present-value-maximining owner will harvest additional units of the stock if marginal costs exceed marginal benefits:

$$ip > p\frac{dG}{dS} - \frac{dC}{dS} \qquad (3.75)$$

These imply the asset equilibrium condition

$$ip = p\frac{dG}{dS} - \frac{dC}{dS} \qquad (3.76)$$

When this is satisfied, the rate of return the resource owner obtains from the fishery is equal to *i*, which is the rate of return that could be obtained by investment elsewhere in the economy. To confirm that this equality exists, divide both sides of equation (3.76) by the net price p to give

$$i = \frac{dG}{dS} - \frac{\left(\frac{dC}{dS}\right)}{p} \qquad (3.77)$$

Equation (3.77) is one (steady-state) version of what is sometimes called the '**fundamental equation' of renewable resources**. The left-hand side is the rate of return that can be obtained by investing in assets elsewhere in the economy. The right-hand side is the rate of return that is obtained from the renewable resource. This is made up of two elements:

- the natural rate of growth in the stock from a marginal change in the stock size (dG/dS).
- the value of the reduction in harvesting costs that arises from a marginal increase in the resource stock $\left\{\frac{\left(\frac{dC}{dS}\right)}{p}\right\}$

Equation (3.77) is an implicit equation for the unknown PV-maximining equilibrium value of S. Solving the equation for S gives an analytical solution for the equilibrium stock. Given that, expressions for the PV optimal solutions for H and E can be obtained using the biological growth function and the fishery production function, we can derive the equilibrium values of *S*, *E* and *H* for different interest rate values. We will not go through these solutions here, but Table 3.7 provides an illustration of these values for rates of interest between zero and 100 per cent (and two higher values), conditional on the assumed functional forms and other baseline parameter values given in Table 3.6. The Table also shows (for the purpose of comparison) the equilibrium values for a static private-property fishery and an open-access fishery. Note that for these last two models, equilibrium values do not vary with the interest rate.

Three important results are evident from an inspection of the data in Table 3.7.

- In the special case where the interest rate is zero, the steady-state equilibria of a static private-property fishery and a PV-maximining fishery are identical. For all other interest rates, they differ. Notice that when *i= 0*, equation (3.77) collapses to

$$\frac{dG(S)}{dS}\, p = \frac{dC(H,S)}{dS} \qquad (3.78)$$

  The left-hand side of this expression is the marginal revenue (with respect to stock changes) and the right-hand side is the marginal cost (with respect to stock changes). Profit-maximizing equilibrium (in the static private-property fishery model without discounting) requires that these be equal. This is, of course, the standard result for any static profit-maximizing model.

- For non-zero interest rates, the steady-state fish stock (fishing effort) in the PV profit-maximining fishery is lower (higher) than that in the static private fishery, and becomes increasingly lower (higher) the higher is the interest rate.

- As the interest rate becomes arbitrarily large, the PV-maximining outcome converges to that of an open-access fishery. Thus, an open-access fishery can be thought of as one in which the absence of enforceable property rights means that fishing boat owners have an infinitely high discount rate. Put differently, the interest rate of boat owners in an open-access fishery is, in effect, infinity, irrespective of what level the prevailing interest rate in the rest of the economy happens to be.

Table 3.7. Steady-state equilibrium in (a) a static private-property fishery, (b) open-access fishery and (c) a PV-maximining fishery with various interest rates

	Static private fishery*			PV-maximising fishery			Open-access fishery*		
i	S	E	H	S	E	H	S	E	H
0.0	0.6	4.0	0.036	0.6000	4.0000	0.0360	0.20	8.0	0.024
0.1	0.6	4.0	0.036	0.4239	5.7607	0.0366	0.20	8.0	0.024
0.2	0.6	4.0	0.036	0.3333	6.6667	0.0333	0.20	8.0	0.024
0.3	0.6	4.0	0.036	0.2899	7.1010	0.0309	0.20	8.0	0.024
0.4	0.6	4.0	0.036	0.2677	7.3333	0.0293	0.20	8.0	0.024
0.5	0.6	4.0	0.036	0.2527	7.4734	0.0283	0.20	8.0	0.024
0.6	0.6	4.0	0.036	0.2434	7.5660	0.0276	0.20	8.0	0.024
0.7	0.6	4.0	0.036	0.2369	7.6314	0.0271	0.20	8.0	0.024
0.8	0.6	4.0	0.036	0.2320	7.6798	0.0267	0.20	8.0	0.024
0.9	0.6	4.0	0.036	0.2283	7.7171	0.0264	0.20	8.0	0.024
1.0	0.6	4.0	0.036	0.2253	7.7467	0.0260	0.20	8.0	0.024
10.0	0.6	4.0	0.036	0.2024	7.9759	0.0242	0.20	8.0	0.024
100.0	0.6	4.0	0.036	0.2002	7.9976	0.0240	0.20	8.0	0.024

* Static private fishery and open-access fishery steady-state equilibrium values do not vary with the interest rate. These equilibrium values have been copied into every row for purposes of comparison.
Source: Perman et al, 2003.p577.

As we conclude our examination of a renewable resource system under open-access and probate-property regimes, it is important to reiterate that there is nothing inevitable about population collapses in open-access conditions (as is clearly illustrated in Table 3.7). Much depends on economic factors that may be favourable or unfavourable to large population sizes. Nevertheless, the possibility that a population may be driven to zero is greater

- when the resource is harvested under conditions of open access than where enforceable property rights prevail;
- the higher is the market price of the harvested resource;
- the lower is the cost of harvesting a given quantity of the resource;
- when prices are endogenous, the more that market price rises as catch costs rise or as harvest quantities fall;
- the lower the natural growth rate of the stock;
- the lower the extent to which marginal extraction costs rise as the stock size diminishes;
- the higher is the discount rate.

Even under private property conditions, an optimal harvesting programme may drive a fish stock to zero. This is most likely where the prey is simple to catch even when the stock approaches a critical minimum threshold level, and where the harvested resource is very valuable. In this case, the optimal harvest level could exceed biological growth rates at all levels of stock. In the next section, we will examine the alternative resource management scenarios discussed above in the light of socially-efficient resource harvesting.

3.5.3. Socially-Efficient Resource Harvesting

As with all resource allocation decisions, there can be no guarantee that privately maximining decisions will be socially efficient (let alone socially optimal). For the PV-maximining fishery to be socially optimal, the following additional conditions must be satisfied.

- The market price, P, correctly reflects all social benefits
- There are no fishing externalities on the cost side and
- The private and social consumption discount rates are identical, $i = r$.

If one or more of those conditions is not satisfied, private fishing will not be socially efficient. We now investigate some reasons why such a divergence might arise.

3.5.3.1 Externalities in the benefits function

The first case is that in which social benefits depend not only on the size of the resource harvest but also on the level of the resource stock. For many biological species that are harvested, intentionally or unintentionally (as by-catch), it is evident that society does place a value on the existence of these species and is concerned about the number of them that do exist. This is clearly true for many large animals such as big cats, whales and apes, and surely extends much more widely than that. In this case, the solution will in general be one in which the socially optimal stock level is higher than that generated from PV-maximining fishery, reflecting the positive utility that the resource stock generates.

More generally, society is likely to have multiple objectives which are not well represented by the private harvester's own objective function (which will tend to be dominated by catch quantity considerations). This is very important for many terrestrial resources, particularly woodlands and forests but it also applies to marine resources. For example, society may have an interest in the maintenance of population diversity or genetic diversity; it may be willing to pay a larger risk premium to ensure high resistance to disease among marine organisms; or it may prefer to maintain stock levels much higher than would private harvesters – at a safe minimum standard – in response to uncertainty and the threats of catastrophic change. All these could be thought of as additional arguments that would appear in the social objective function (but which would not usually enter private profit functions).

3.5.3.2 Externalities in the fishery production function

A second source of social inefficiency arises from externalities operating through the fishery production function. There are two important types of harvesting externality. First, it often happens that resource harvesting inadvertently destroys other species. Beam trawling, for example, (in which a net is weighed down to the sea bed by heavy beams and is trawled along the sea bed to catch bottom-feeding fish such as cod), causes immense damage to other sea-bed creatures, and can cause those populations to collapse. This is a classic externality problem and it is clear that outcomes are most unlikely to be efficient in such cases. Some form of regulation of fishing practices would be appropriate here.

A second kind of externality in the production function is often known as a '**crowding' diseconomy**. Suppose that each boat's harvest depends on its effort and on the effort of others. Then each boat's catch imposes a contemporaneous external cost on every other boat. In effect, boats are getting in each other's way. When any boat is fishing, the costs of harvesting a given quantity of fish become higher for all other boats. This externality drives the average costs of fishing for the fleet as a whole above the marginal costs of an individual fisher. If a crowding effect of this kind exists, the function, *C(H, S)* from the point of view of an individual boat operator may differ from the function *CS(H, S)* from the social point of view. Whether it does or not depends on what institutional conditions apply.

If the private (or PV-maximining) fishery is one in which effort is somehow or other coordinated in the common interest, then it would be sensible to assume that the size of the fishing fleet as a whole (and the spatial patterns of fishing) would be optimally chosen. The optimal size of fleet would balance the additional benefits of extra boats against the additional external costs of extra boats. The crowding diseconomies become internalized in this way, leading to efficient outcomes. Alternatively, if the fishery were in private-property ownership but were carefully and effectively regulated, then it is conceivable that such regulation might also internalize the externality. Under conditions of open access, there is virtually no possibility that crowding effects would be internalized by the actions of fishermen alone. The kind of coordination we referred to above cannot happen in the competitive struggle to grab fish. Almost certainly, the unregulated industry would consist of more harvesters and more harvesting capital than is economically efficient.

It is important to note that the crowding diseconomy we have just referred to is entirely different from the '**stock externality' effect** which arises in open access. Crowding externalities are contemporaneous; stock externalities are intertemporal. The latter exist when the taking of fish today imposes additional costs in the future by virtue of the reduced future stock size. We have already remarked that this kind of externality fundamentally distinguishes the open-access and private (and PV-maximizing) fishery cases. Realistically or not, modellers typically assume that the stock externality will not be internalized in open-access conditions but will be in a private fishery. The first part of this assumption is surely true. Indeed, there is ample evidence that fishing effort is massively excessive and inefficient in many open-access fisheries throughout the world. (We illustrate this in section 3.5.5). It is of interest to note that the Schaefer (1957) form of fishery production function we have used in this module ($H = eES$) rules out the existence of crowding externalities. For a given stock size, H will change in proportion to changes in E. In other words, the marginal product of effort is constant, precluding crowding effects.

3.5.3.3 Monopolistic fisheries

The existence of monopoly ownership of a fishery may also generate inefficient outcomes. A resource market is monopolistic if there is one single price-making harvester. It is known from standard microeconomic theory that marginal revenue exceeds marginal cost at a monopolistic market equilibrium. A monopoly owner would tend to harvest less each period, and sell the resource at a higher market price, than is socially efficient. Therefore, if a renewable resource were harvested under monopolistic rather than competitive conditions, an economically inefficient harvesting level **may** result. However, as we saw in Module 2, in a world where there are other market failures pushing harvest rates to excessive levels, monopoly harvests may be closer to the second-best efficient allocation than those from 'competitive' PV-maximizing fisheries.

3.5. 4. Safe Minimum Standard (SMS) of Conservation

Our discussion of the 'best' level of renewable resource harvesting has focused almost exclusively on the criterion of economic efficiency. However, as we observed in Module 2, if harvesting rates pose threats to the sustainability of some renewable resource (such as North Atlantic fisheries or primary forests) or jeopardize an environmental system itself (such as a wildlife reserve containing extensive biodiversity) then the criterion of efficiency may be insufficient or inappropriate. Even in a deterministic world – in which population growth rates are known with certainty – the pursuit of an efficiency criterion is not sufficient to guarantee the survival of a renewable resource stock or an environmental system in perpetuity, particularly when resource prices are high, harvesting costs are low, or discount rates are high. Where biological systems are stochastic, or where uncertainty is pervasive, threats to sustainability are even more pronounced. Many writers – some economists, but particularly non-economists – argue that correcting market failure and eliminating efficiency losses should be given secondary importance to the pursuit of sustainability. This would suggest that policy be targeted to the prevention of species extinction or the loss of biological diversity whenever that is reasonably practical.

Efficiency objectives can be pursued within this general constraint. Such considerations bring us back to the principle that policy be oriented around the criterion of a safe minimum standard (SMS) of conservation.

In relation to renewable resource policy, a strict version of SMS would involve imposing constraints on resource harvesting and use so that all risks to the survival of a renewable resource are eliminated. This is unlikely to be of much practical relevance. Virtually all human behaviour entails some risks to species survival, and so a strict SMS would prohibit virtually all economic activity. In order to make the concept usable, it is necessary to impose weaker constraints, so that the adoption of an SMS approach will entail that, under reasonable allowances for uncertainty, threats to survival of valuable resource systems are eliminated, provided that this does not entail excessive cost. For decisions to be made that are consistent with that weaker criterion, judgements will be necessary particularly about what constitutes 'reasonable uncertainty' and 'excessive cost', and which resources are deemed 'sufficiently valuable' for application of the SMS criterion. These issues could be controversial.

3.5.5. Some Empirical Evidence

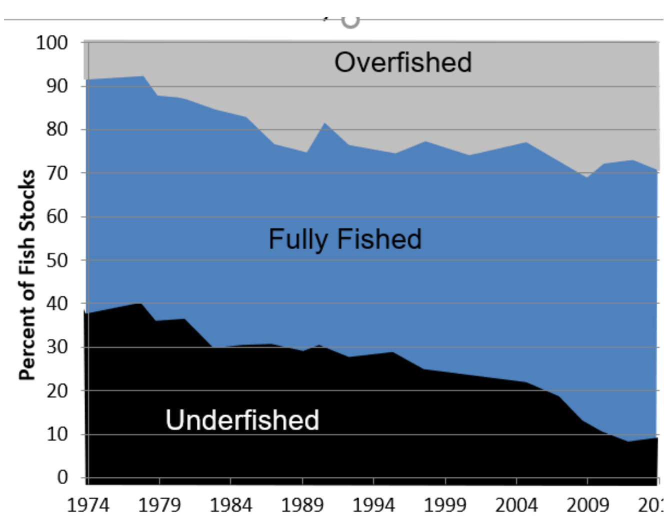


Fig. 3.22 Global Trends in the Status of Fish Stocks, 1974-2013. **Source:** FAO, 2016. Also, in Jonathan and Roach, 2017.

Available empirical evidence suggests that the world is experiencing extensive losses of many renewable resource-population stocks, and unprecedented rates of species extinction. Figure 3.22 suggest that many of the world's major fisheries have been exploited beyond their maximum sustainable yield point, with the result that they have declined significantly from peak yields. The proportion of overexploited fisheries has risen since the 1970s, while the proportion of less-than-fully-exploited fisheries has declined.

Figure 3.23 shows that the world output of wild fish has essentially reached maximum yield. There has been no overall increase in wild catch since the late 1980s, and possibly a slight decline since 1995. Rapidly increasing production through aquaculture has kept overall output rising, but large-scale aquaculture often has seriously negative environmental effects.

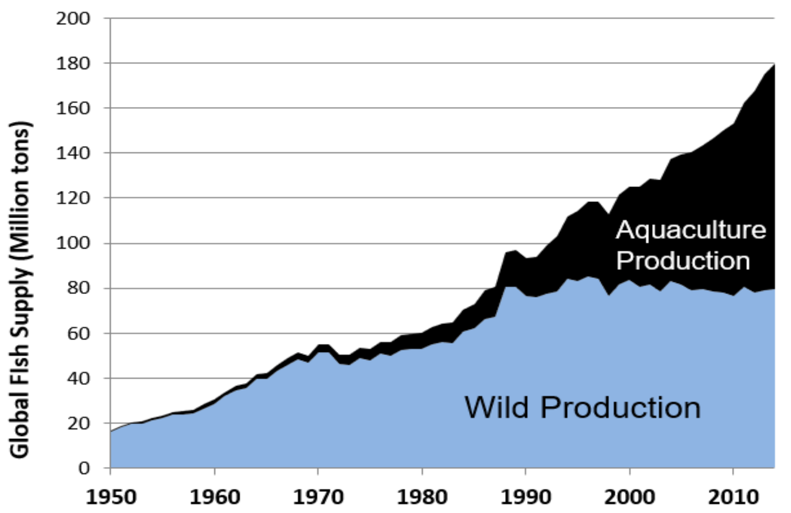


Fig. 3.23. Global Fish Harvest, Wild Catch and Aquaculture, 1950-2013. **Sources:** Food and Agriculture Organization of the United Nations (FAO), Global Capture Production Statistics; Global Aquaculture Production Statistics, accessed September 2016. Also, in Jonathan and Roach, 2017.

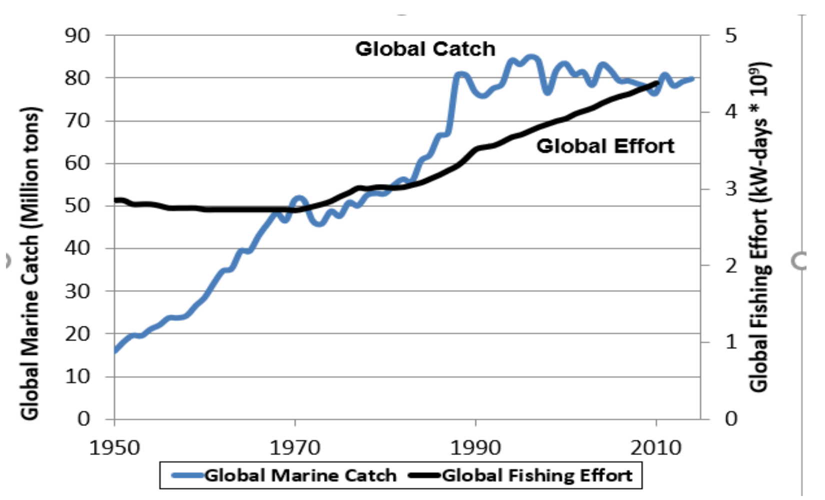


Fig. 3.24. Global Marine Fishing Harvest (1950-2014) and Global Fishing Effort (1950-2010). **Sources:** Food and Agriculture Organization of the United Nations (FAO), Global Capture Production Statistics; Anticamara, et al., 2011. *Note:* Marine catch includes marine fishes, crustaceans, molluscs, and marine mammals. The measure of kW-days*10^9 on the right axis is based on the power of fishing vessels and the number of days each year they spend fishing. Also, in Jonathan and Roach, 2017.

Figure 3.24 illustrates what we might call 'the paradox of fisheries': as global fleet capacities have steadily risen, total fish catch has not increased. This means that catch rates (measured in tons of fish landed per unit of global fishing effort) have steadily fallen. Clearly this represents extreme economic inefficiency, with ever-increasing capital investment for ever-declining returns.

A significant reduction in marine fishing would likely imply increasing land-based food production. The advantages of marine fishing include more land for wildlife habitat, reduced soil erosion, and reduced chemicals use. However, the open oceans are a surprisingly unproductive ecological habitat as compared to most land-based ecosystems, based on the amount of biomass supported by a given area. The net primary productivity of open oceans is less than half that of most agricultural land. The sustainable yield of wild fish species may thus represent one real carrying capacity limit for humans. Aquaculture is seen by some as a way of escaping the ecological limits of wild fish harvesting. But ecological problems associated with aquaculture are also significant, and reminiscent of the problems of intensive land-based agriculture: excessive resource consumption, excessive wastes, and damage to natural ecosystems. Traditional small-scale pond aquaculture is generally more sustainable than large-scale ocean aquaculture.

Summary

- An open-access fishery shares two of the characteristics of the standard perfect competition model. Each firm takes the market price of landed fish as given; and there are no impediments to entry into and exit from the fishery. However, free entry in the case of an open-access fishery means that firms have no enforceable property rights to the in situ fishery resources, including the fish in the water.

- The level of economic profit prevailing in the fishery plays a critical role in determining fishing efforts. Effort applied will continue to increase as long as it is possible to earn positive economic profit from the fishery.

- The biological equilibrium occurs where the resource stock is constant through time (that is, it is in a steady state). This requires that the amount being harvested equals the amount of net natural growth: economic equilibrium requires that the amount of fishing effort be constant through time. Such an equilibrium is only possible in open-access fisheries when rents have been driven to zero, so that there is no longer an incentive for entry into or exit from the industry, nor for the fishing effort on the part of existing fishermen to change. The unique equilibrium level of heaviest in an open-access renewable resource model is the effort level that equates total revenue and total cost (implying zero profit).

- In an open access fishery the dynamics of adjustment over time may cause the system to revert to the unique equilibrium or lead to collapse. Other things being equal, the probability that effort oscillations might cause a population to collapse before it can attain a steady-state solution is increased when there is a critical minimum threshold, and/or when the growth function or environmental conditions are stochastic. When open access to a resource is accompanied by worsening environmental conditions, and when these changes are either rapid or unforeseen or both, harvesting outcomes can be catastrophic.

- It is possible that any resource stock could be harvested to exhaustion, or a specie driven to extinction, under open access. While it is also true that this is possible under almost any regime, including those with enforceable private property rights, it remains true that open-access conditions increase the probabilities of those outcomes occurring. The main reason for this is that in these circumstances there is no collectively rational management of harvesting taking place. Even where what should be done is evident, an institutional mechanism to bring this about is missing.

- The institutional arrangements of a private-property fishery may help overcome this dilemma. There are several particular structure of property rights that are possible and will be consistent with the definition of private property. The static private-property equilibrium will always lead to a higher resource stock level and a lower effort level than that which prevails under open access. The steady-state harvest may be higher, lower or identical. Although stocks will be higher under private property than open access, the quadratic form of the stock–harvest relationship implies that harvests will not necessarily be higher with higher stocks.

- In a dynamic framework, a present-value-maximizing owner will harvest up to the point where the rate of return obtained from the fishery is equal the rate of return that could be obtained by investment elsewhere in the economy. For non-zero interest rates, this imply a lower steady-state fish stock and higher fishing effort) than in the static case. As the interest rate becomes arbitrarily large, the PV-maximining outcome converges to that of an open-access fishery.

- For any property right regime governing renewable resource management, whether the population will be driven to extinction generally depends on economic factors that may be favourable or unfavourable to large population sizes.

- As with all resource allocation decisions, there can be no guarantee that privately maximining decisions will be socially efficient (let alone socially optimal). For the present value-maximining (private-property) fishery to be socially optimal, the market price of fish must correctly reflect all social benefits, there must be no fishing externalities on the cost side, and the private and social consumption discount rates must be equal.

- Externalities in a fishery could arise from both benefit and cost sides Benefit externalities occur when social benefits depend not only on the size of the resource harvest but also on the level of the resource stock. In the case of benefit externalities, the socially-optimal stock level is higher than that generated from present value-maximining fishery, reflecting the positive utility that the resource stock generates. Cost externalities operate through the fishery production function when it imposes higher cost on operators. Under this condition, a private property regime may help internalize the externalities but not an open-access.

- The existence of monopoly ownership of a fishery may also generate inefficient outcomes. A monopoly owner would tend to harvest less each period, and sell the resource at a higher market price, than is socially efficient. However, in a world where there are other market failures pushing harvest rates to excessive levels, monopoly harvests may be closer to the second-best efficient allocation than those from ‘competitive’ PV-maximizing fisheries.

- Many, particularly in the field of ecological economists argue that correcting market failure and eliminating efficiency losses in the management of non-renewable resources should be given secondary importance to the pursuit of sustainability. This would suggest that policy be targeted to the prevention of species extinction or the loss of biological diversity whenever that is reasonably practical. A strict version of the safe minimum standard (SMS) approach would involve imposing constraints on resource harvesting and use so that all risks to the survival of a renewable resource are eliminated. However, this is unlikely to be of much practical relevance.

- Aquaculture is seen by some as a way of escaping the ecological limits of wild fish harvesting. But ecological problems associated with aquaculture are also significant, and reminiscent of the problems of intensive land-based agriculture: excessive resource consumption, excessive wastes, and damage to natural ecosystems. Traditional small-scale pond aquaculture is generally more sustainable than large-scale ocean aquaculture

Discussion/Review Questions and Exercises

1. Distinguish between a biological and economic equilibrium in an open-access fishery model.
2. What are the likely outcomes of dynamic adjustments in an open-access renewable resource model and what factors are likely to determine the outcome that prevails?
3. How will you define and explain a private-property fishery?
4. Show and analyze the relationship between a static open-access fishery and a present-value private-property fishery with a focus on the role of the discount rate.
5. Explain the factors that determine whether or not the population of a non-renewable resource is likely to be driven to extinction by harvesters/hunters.
6. What conditions are necessary to achieve socially optimal outcome in a present value private-property fishery?
7. How does externality affect the benefit and cost functions for private operators in a fishery?
8. Explain and different between 'stock externality effect' and 'crowding diseconomy' in a renewable resource model.
9. Why may the pursuit of efficiency be inadequate when it comes to the management of non-renewable resources?

Materials used for the Lecture notes

Jonathan M. Harris and Brian Roach (2017), **Environmental and Natural Resource Economics** *4th Edition*, Routledge.

Perman, R., Ma Y., McGilvray J. and Common M. (203). **Natural Resource and Environmental Economics**, *3rd* Edition, Edinburgh, Longman.

Tietenberg, T. & Lewis, L. (2012). **Environmental & Natural Resource Economics** 9th Edition, The Pearson Series in Economics.

Module 3.6. Optimal Management of Renewable Resources: Application to Sub-Saharan Africa (1.5 hours)

As noted in Module 3.3, Africa has significant natural resource wealth. The continent is rich in both renewable and non-renewable natural resources. It has the world's largest arable landmass, the second largest and longest rivers (the Congo and Nile, respectively), and the second largest tropical forests (African Capacity Building Foundation 2013). According to the African Development Bank (AfDB), 'the total value added of its fisheries and aquaculture sector alone is estimated at USD 24 billion' (AfDB 2016: 3).

There is a rich and growing literature on the management of non-renewable resources in Sub-Saharan Africa highlighting various aspects of the problem, what works under certain contexts, and what does not, and suggesting some general guidelines for reforms. (A list is provided in the supplementary readings on Topic 3 in the Course outline). In general, the evidence does not suggest that the continent has managed its renewable natural resources in an efficient or sustainable manner. For example, Honlonkou and Hassan (2014) show that natural reserves in Benin are under serious threats from legal and illegal loggers, poachers, hunters, farmers, herders, and wood traders. According to the authors, high population growth and technological change have led to increased exploitation of and excessive pressure on these resources. Available statistics show that one thousand square kilometres of forests are destroyed every year against only ten square kilometres of reforestation.

In Ghana, increased demand for resources has led to the overexploitation of wildlife and extensive modification of wildlife habitats. About 70% of the country's original 8.22 million ha of closed forest has been destroyed, and the deforestation rate is put at 220 square kilometres (sq. km) per annum. Most wild animal species are believed to be seriously depleted and at least 18 of the 222 mammalian species recorded in the country are under threat.[28] By the late 1980s, reductions in Kenya's elephant and rhino populations had reached 85% and 97% respectively. Some of the factors attributed included the country's high population rates, with some of the fastest rates of population growth in areas around parks and reserves, and the pattern of land tenure. Much of Namibia's land is arid or semiarid and faces severe environmental pressures due to drought and overgrazing by livestock (Bojo, 1996).

We consider below some specific cases of managing renewable resources (commons) in Sub-Saharan Africa. We will limit our consideration here to Southern Africa[29]. Students/readers are encouraged to study available materials on other subregions in the Continent. In southern Africa, colonialism largely led to appropriation of resources from the indigenous populations. In South Africa, Zimbabwe and Namibia, most of the people were actually evicted from their land under policies such as apartheid, leaving the local population in severely crowded communal reserves. Colonial rule and these policies eroded or discouraged traditional management practices that had been in place before the indigenous people were overrun by white settlers. Most (black majority) independent states adopted similar positivist and centralized management approaches that had been practiced under colonialism. Among the most common state property regimes in southern Africa are protected wildlife areas based on conserving biodiversity and using it for economic gain in the form of tourism. These are usually created from the residential and cultural landscapes of local populations who are evicted and moved elsewhere, with little or no benefits accruing to them. The challenges to commons governance in southern Africa are analyzed with reference to floodplains, grasslands, forests, and fisheries.

Floodplains

The floodplains cases present important lessons about institutional changes related to knowledge, politics, economics and power. One is that common pool resources are increasingly managed by agencies located far from the floodplains, which leads to a decline in local empowerment and real management options, despite governments' advocacy of participatory approaches. This again highlights the issue of the power of different actors in this context and links economic problems arising from globalization and national trends to the local level.

In Mvula and Haller's Chilwa floodplain case, one of the most important features is the fragmentation of resource management across different departments and districts since colonial times. This divided the ecosystem linked to Lake Chilwa into different levels of governance. Knowledge of its management was lost and access to resources was nationalized. Under pressure from a national economic crisis, fishermen from the overused Lake Malombe moved to the Chilwa floodplain, increasing pressure on resources and causing conflicts. There are also

[29] This section is extracted from Extracted from, Hara, Mafaniso, Turner, Stephen, Haller, Tobias and Matose, Frank (2009)'Governance of the commons in southern Africa: knowledge, political economy and power', *Development Southern Africa*,26:4,521 — 537 DOI: 10.1080/03768350903181324.
URL: http://dx.doi.org/10.1080/03768350903181324

considerable problems of legal pluralism and uncertainty about the role of traditional authorities, and confusion about which authorities the resource users are accountable to.

Similarly, the Kafue Flats and their surrounding areas have been broken up into administrative units. Haller and Chabwela describe how common pool resources, previously managed by the interrelated institutions of local groups, were placed under fragmented jurisdictions. Former common property regimes regulating the coordinated use of common pool resources have been dismantled and are now under de jure state control.

There have also been controversial incentives regarding control of land for pasture or for irrigation. This is linked to the financial crisis in Zambia since 1975, which increased the value of common pool resources and made the area very attractive. Outside users now have the right, as citizens, to state-regulated access to these resources. While negotiating this access, outsiders' bargaining power is reinforced ideologically by their claim to citizenship of a state that is legally present but institutionally absent.

Another case is the Okavango Delta where two villages are both involved in Community based natural resource management (CBNRM) and close to Wildlife Management Areas. In Ikoga, CBNRM has been a challenge due to lack of knowledge about how to implement it and difficult relations with a local Trust and a non-governmental organization – all leading to frustration with the stagnation of the process. Seronga, on the other hand, has two CBNRM programmes operating, from which villagers are generating income. But in both cases, all the local people can do is contract out an area to tourist operators. Legislation and restrictions related to tourism prevent direct resource management. Hunting is not allowed, fishing is restricted because of recreational fisheries, and cattle husbandry is limited. The people also face problems with the increase of dangerous wildlife in the area, although this wildlife increases the area's attractiveness for tourists.

Grasslands and forests

Fragmentation and alienation are also leading issues in the governance of the commons on the southern shores of Lake Kariba. A study by Nyikahadzoi et al.(2017), describes how previously integrated indigenous systems for natural resource management and use were disrupted and fragmented by the flooding of the Zambezi valley. The dam created a new kind of national hydropower and fisheries resource, but submerged local economic interests. These trends continued on the Zimbabwean side through processes of land use planning that placed protected conservation and forestry areas under separate jurisdictions, alienating their governance from the local people who depended on them. These people were marginalized by broader economic interests and broader definitions of sound environmental management.

A study by Lefatshe Magole (2009), shows how the resource governance experience of the aboriginal San people in the Mababe and Phuduhudu settlements in Ngamiland, to the east of the Okavango Delta, typifies the marginalization of this group by processes of modernization and land use planning driven by national authorities. As has so often happened across southern Africa, the San – the least powerful local actors – were largely excluded from new resource governance and use dispensations that were supposedly driven by conservation imperatives. So far, CBNRM has not proved an effective way for these disempowered groups to rebuild sustainable livelihoods. Their natural resource base is still effectively controlled by more powerful external actors – the state and the safari companies.

In the Lake Ngami area of north-western Botswana, trends in range management are shrinking the commons. Here, the Ovambandero people's flexible and environmentally adaptive resource use systems have been constrained by the award of individual title to extensive areas of grazing land through the national Tribal Grazing Land Policy. With their access to grazing and water resources now restricted, the Ovambandero fear that their herds will no longer be able to withstand the droughts that their indigenous systems had been able to cope with. In Botswana, the challenge to CBNRM is to move beyond nature conservation and ecotourism and offer viable strategies for sustainable use of the grazing commons.

A study by Matose (2016) shows that, in the forest patches of the Dwesa-Cwebe area of the Eastern Cape coast, the theme of conflicting knowledge systems recurs. Indigenous environmental knowledge and value systems were subordinated to state-imposed ones based on western science, and local people were banned from using forest resources for many decades. Linked to the South African 'betterment' programmes of land use planning in communal areas, this was the environmental dimension of rural people's subjection to the colonial state – so far only partially relieved by the post-apartheid dispensation, despite a successful community land claim over the forest reserve and the institution of co-management arrangements.

Fisheries

Fisheries provide two examples of the complexities of resource governance in southern Africa: one with regard to designing and implementing transformation policies in capital intensive natural resource economic sectors where barriers to entry are high for new entrants, and the other with regard to elite capture of co-management arrangements that may disadvantage the very people that devolution of management was supposed to benefit. The review of the kapenta fishery on Lake Kariba by Nyikahadzoi (2009) shows that cooperative management arrangements between the state and resource users are difficult to institute despite the existence of a supporting legal framework. The process of redistributing rights has largely been influenced by external political and macroeconomic factors that have undermined such attempts by the state and eroded its capacity to act as a neutral economic regulator. Global agendas have neutralized the government's efforts to achieve a more equitable access to natural resources in Zimbabwe.

Economics and power plays have been blatantly prominent in the dynamics of policy evolution in the South African fisheries. The democratic government launched a process of transformation and governance reform intended to open small pelagic fisheries to smaller enterprises owned by previously disadvantaged people. In an intricate series of manoeuvres, stakeholders responded by organizing themselves into communities in terms of economic and social interests, political ideologies and strategic options. The result has been less beneficial to the poor and disadvantaged than the reformers had hoped.

In Modules 3.4 and 3.5, we observed that even in a deterministic world – in which population growth rates are known with certainty – the pursuit of an efficiency criterion is not sufficient to guarantee the survival of a renewable resource stock or an environmental system in perpetuity, particularly when resource prices are high, harvesting costs are low, or discount rates are high. As we noticed in Module 3.3 where we focused on non-renewable resources, high discount rate on the part of actors may also be a significant factor driving inefficiency in the management of renewable resources in Sub-Saharan Africa. There are many reasons why actors (both private and public, including individuals, households, extractive firms and state actors) may place a heavy premium on the present rather than the future (that is, heavily discount the future). They include high level of poverty and economic uncertainties and insecure property rights. The effect could be heavy when this combines with low cost of harvesting (arising from low-income level and poor governance). Birungi and Hassan (2010) argue that more efficient government efforts to reduce poverty would enhance conservation generally.

A critical hypothesis in many studies is that the property rights structure is a key factor in determining the choice between wildlife and livestock utilization. Unlike many countries, Zimbabwe follows a policy of sustainable utilization and views all its mammals as renewable natural resources to be managed (Bojo, 1996). It will be interesting to explore recent developments around land ownership and wildlife conservation in the country. In a study on compliance with regulations among the artisanal fishers in Sudan, Abusin and Hassan (2014), suggest that legitimacy and ethical factors (specifically involvement of stakeholders in the process of designing, monitoring and enforcing regulations) is crucial as a process factor that may be more important than mere deterrence measures. The authors advocate for participatory co-management systems that are most likely to be more effective than top-down mechanisms in promoting compliance. They also suggest the need for investments in the education of fishermen, the provision of alternative income and employment opportunities outside of fishing, access to credit to finance the acquisition of legal nets, and the effective regulation of importation of illegal nets.

As required in some of the Review questions at the end of Modules 3.4 and 3.5, it will aid the course of learning and policy formulation for readers/students to investigate what exactly work best and what does not work (and why) in relation to any identifiable critical renewable natural resource in their communities or countries.

ACKNOWLEDGEMENT

The African Economic Research Consortium (AERC) wishes to acknowledge and express its immense gratitude to the following resource persons, for their tireless efforts and valuable contribution in the development and compilation of this teaching module and other associated learning materials.

1. Prof. Thomas Sterner, University of Gothenburg, Sweden. (Email: thomas.sterner@economics.gu.se);
2. Prof. Aderoju Oyefusi, University of Ibadan, Nigeria. (Email: aderoju.oyefusi@uniben.edu; aderojuoyefusi@yahoo.com);
3. Dr. John Mutenyo, Makerere University, Uganda. (Email: jkmutenyo@bams.mak.ac.ug; jkmutenyo@yahoo.com);
4. Prof. Samuel A. Igbatayo, Afe Babalola University, Nigeria. (Email: remisamuel2002@yahoo.com).

Thank you.

Facebook Twitter Website Email Website

Our mailing address is:
African Economic Research Consortium (AERC)
3rd Floor, Middle East Bank Towers, Jakaya Kikwete Road
P. O. Box 62882
00200 Nairobi
Kenya