
# M.A. PART - II

## Group (I) PAPER-IV

# PUBLIC ECONOMIC

**Dr. Sanjay Deshmukh**
Vice Chancellor,
University of Mumbai

**Dr. Ambuja Salgaonkar**
Incharge Director ,
IDOL, University of Mumbai

**Dr. Dhaneswar Harichandan**
Incharge Study Material Section,
IDOL, University of Mumbai

**Programme Co-ordinator** : **Ms. Rajashri Pandit**
Asst. Prof-cum-Asst. Director, IDOL,
University of Mumbai, Mumbai-400 098

**Course Writer** : **Dr. Sujata Khodilkar**
Head of Departments of Economics,
S.K. Somaiya College,
Vidyavihar, Mumbai

: **Prof. Shashikant Mundhe**
Siddharth College,
Fort, Mumbai

: **Prof. Vitthal Sontakke**
Kirti College,
Dadar, Mumbai

**APRIL 2016 M.A. PART - II, GROUP (I), PAPER IV,**
**PUBLIC ECONOMICS**

**Published by** **Incharge Director**
**Institute of Distance and Open Learning ,**
**University of Mumbai,**
**Vidyanagari, Mumbai - 400 098.**

**DTP Composed** : Ashwini Arts
Gurukripa Chawl, M.C. Chagla Marg, Bamanwada,
Vile Parle (E), Mumbai - 400 099.

**Printed by** :

## CONTENTS


| Unit No. | Title | Page No. |
|---|---|---|
| | **Module - 1- Government in a Market Economy** | |
| 1. | Allocative Efficiency, Market Failures & Govt. Intervention | 01 |
| 2. | Social Welfare Functions & Govt. Failures | 24 |
| 3. | Interest Group Capture and Politico Economic Interactions | 46 |
| | **Module - 2 - Public Expenditure Rationale & Evaluation** | |
| 4. | Public Expenditure : Rationale and Evaluation - I | 59 |
| 5. | Public Expenditure : Rationale and Evaluation - II | 83 |
| 6. | Public Expenditure : Rationale and Evaluation - II | 115 |
| | **Module - 3 - Taxation & Public Sector Pricing** | |
| 7. | Taxation - I | 158 |
| 8. | Taxation - II | 170 |
| 9. | Public Sector Pricing | 181 |
| | **Module - 4 - Deficits, Debt and Solvency** | |
| 10. | Government Budget Constraint | 187 |
| 11. | Financial Programming Model | 196 |
| 12. | Taxation, Inflation and Interest Rates | 203 |
| | **Module - 5 - Fiscal Policy & The Macroeconomy** | |
| 13. | Macroeconomics Effects of Fiscal Deficits | 211 |
| 14. | Twin Deficit Approach to Growth | 227 |
| 15. | Financial Crisis | 253 |
| | **Module - 6 - Public Finance with Many Jurisdictions** | |
| 16. | Fiscal Federalism | 268 |
| 17. | Intergovernmental Grants & Revenue Sharing | 283 |


# 1

## MODULE -1 : Government in a Market Economy

# ALLOCATIVE EFFICIENCY, MARKET FAILURES & GOVT. INTERVENTION

**Unit Structure :**

1.1 Objectives
1.2 Introduction
1.3 Efficiency in Resource allocation
1.4 First Best Economy
1.5 Market Failures
1.6 Rationale for State Intervention
1.7 Regulation and Taxation
1.8 Distributional Objectives of the State
1.9 Questions

## 1.1 OBJECTIVES

1. To understand the concept of Efficiency in Resource allocation.
2. To study and understand the **Market Failures**.
3. To understand **Rationale for State Intervention and Regulation and Taxation**.

## 1.2 INTRODUCTION

**Public economics** or **economics of the public sector** is the study of government policy through the lens of economic efficiency and equity. At its most basic level, public economics provides a framework for thinking about whether or not the government should participate in economics markets and to what extent its role should be. In order to do so, microeconomic theory is utilized to assess whether the private market is likely to provide efficient outcomes in the absence of governmental interference. Inherently, this study involves the analysis of government taxation and expenditures. This subject encompasses a host of topics including market failures, externalities, and the creation and

implementation of government policy. Public economics builds on the theory of welfare economics and is ultimately used as a tool to improve social welfare.

Broad methods and topics include:

- the theory and application of public finance
- analysis and design of public policy
- distributional effects of taxation and government expenditures
- analysis of market failure and government failure.

Emphasis is on analytical and scientific methods and normative-ethical analysis, as distinguished from ideology. Examples of topics covered are tax incidence, optimal taxation, and the theory of public goods.

## 1.3 EFFICIENCY IN RESOURCE ALLOCATION

There is a **scarcity** of almost everything that brings people happiness. Economics is the analysis of how society determines how it will distribute, or **allocate**, its scarce goods to a population and to a set of purposes that has an infinite desire for more. This is the first and most fundamental reason why economics is sometimes referred to as "the dismal science": never can all desires be satisfied.

Society accomplishes its allocation through market system or/and command system:

- Market systems use *prices* as signals to allocate its resources.
- Command systems make use of *political choice* to allocate its resources.

For our purposes, we shall focus on explaining how market allocation functions.

### Productive and Allocative Efficiency

Through these means, society strives to achieve both **productive efficiency** and **Allocative efficiency**.

An example of *productive inefficiency* is when a method of production yields the same as another that uses less of any

resource but does not use more of any other resource. Hence, there would be no reason to use the less productive method.

An example of *Allocative inefficiency* is when a method of production using more of a certain resource and less of another than another method costs more to society overall. Even if it is productively efficient, the resources are not used in the best distribution--the allocation is inefficient. In this situation, the inefficient method should be swapped for the more efficient method by a redistribution of which combination of resources is used for a certain mode of production.

Another interpretation of Allocative inefficiency is the distribution of goods to members of society in a way that yields less than optimal happiness. This interpretation is almost never achieved but is nevertheless the goal toward which economists strive.

Productive efficiency must be satisfied before Allocative efficiency may be.

**Allocative efficiency** is a state of the economy in which production represents consumer preferences; in particular, every good or service is produced up to the point where the last unit provides a marginal benefit to consumers equal to the marginal cost of producing. In the single-price model, at the point of allocative efficiency, price is equal to marginal cost. At this point the social surplus is maximized with no deadweight loss, or the value society puts on that level of output produced minus the value of resources used to achieve that level, yet can be applied to other things such as level of pollution. Allocative efficiency is the main tool of welfare analysis to measure the impact of markets and public policy upon society and subgroups being made better or worse off.

Although there are different standards of evaluation for the concept of allocative efficiency, the basic principle asserts that in any economic system, choices in resource allocation produce both "winners" and "losers" relative to the choice being evaluated. The principles of rational choice, individual maximization, utilitarianism and market theory further suppose that the outcomes for winners and losers can be identified, compared and measured.

Under these basic premises, the goal of maximizing allocative efficiency can be defined according to some neutral

principle where some allocations are objectively better than others. For example, an economist might say that a change in policy increases allocative efficiency as long as those who benefit from the change (winners) gain more than the losers lose.

### Conditions

It is possible to have Pareto efficiency without allocative efficiency. By shifting resources in the economy, a gain in benefit to one individual could be greater than the loss in benefit to another individual (Kaldor-Hicks efficiency). Therefore, before such a shift, the market is not allocatively efficient, but might be Pareto efficient.

When a market fails to allocate resources efficiently, there is said to be market failure. Market failure may occur because of imperfect knowledge, differentiated goods, concentrated market power (e.g., monopoly oroligopoly), or externalities.

### In contract theory

In contract theory, allocative efficiency reflects a contract in which the skill the offering party demands and the skill of the party that agrees to the contract are the same.

### Allocative Efficiency and the Production Possibilities Frontier

Because resources are scarce, a society must decide how to use those resources for its maximum benefit. When resources are used to produce one good or service, those resources become unavailable for any other purpose. Therefore, to understand how an economy allocates resources, it is best to look at a simplified economy that consists of the production of only two goods. In this simple economy, a **production possibilities frontier**, or model, can be constructed that shows every combination of the production of the two goods. Because economic resources are scarce, society has to decide what those resources will be used to produce, since any resources that are used to produce one good reduce the resources available for the production of other good.

To better understand the allocation of resources, there are several assumptions for this simplified economy:

- full employment and productive efficiency,
- fixed resources, and
- fixed technology.

The purpose of the assumptions is to show how the production possibilities of an economy changes with the allocation of its resources, so we must assume that nothing else changes — as economists like to say, *ceteris paribus*, meaning all other factors remaining the same. So if employment was less than full employment or productive efficiency was less than the maximum, then that introduces other production possibilities such as increasing productive efficiency or increasing employment. Thus, making these assumptions simplifies the analysis. Consider the production possibilities between bread and guns.

Any movement along the curve indicates the reallocation of resources from one good to the other, from guns to bread or vice versa. Production outside of the curve is unattainable, and production below the curve is attainable but is not utilizing all of the resources, which violates the assumptions and reduces the aggregate wealth of society.

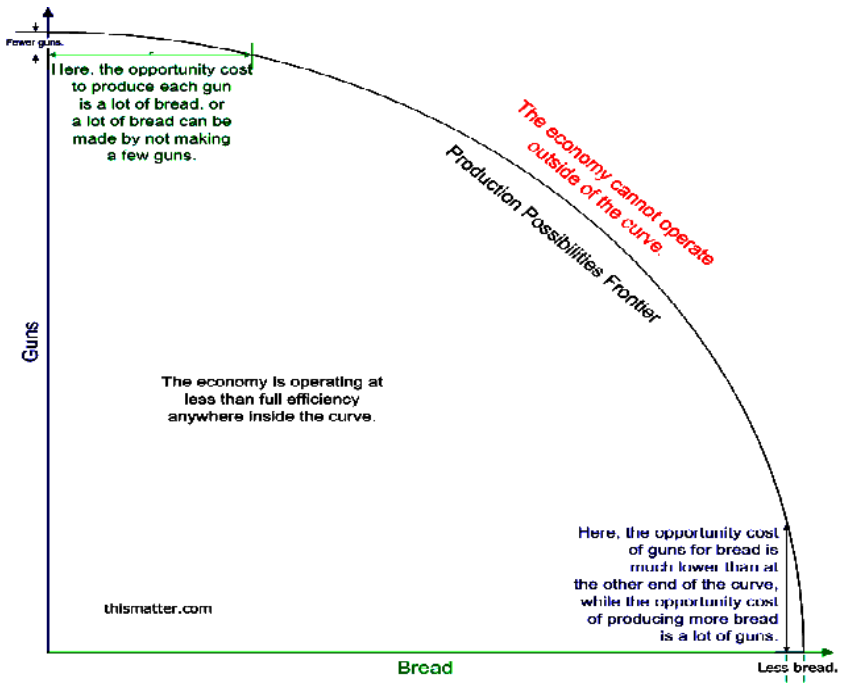



**Figure 1.1**

## Increasing Opportunity Cost

Every specific allocation of resources has an **opportunity cost**, which is what is given up to produce a specific product. Sacrificing the production of guns to make one more unit of bread is the opportunity cost of bread in our simplified economy, which

is the number of guns that must be given up to make one more unit of bread.

Generally, as more resources are allocated to produce one good, the cost of an additional unit of the good usually increases after a certain point. The reason is because when only a few units of the goods are produced, then the most suitable factors of production are used, which lowers the cost of producing the good. But as the number and quantity of goods increases, then other factors of production must be used that will not be as efficient. For example, if a factory is already operating near capacity, then to increase its output, it must hire more labor, train new workers, and utilize space that was devoted to other resources. In our bread example, farmers would first use the best land to grow the grains necessary to produce bread, but if our simplified economy wanted to produce mostly bread, it would have to grow grains on less suitable land and in less suitable climate. Irrigation may even have to be provided. This is sometimes referred to as the **law of diminishing returns**.

**Allocative Efficiency**

Because the opportunity cost of producing extra units of goods increases after a certain point, then there is a point for which the cost of producing the good is less than its benefit to society. Because of economies of scale, opportunity costs at first decline in the production of goods. However, when the production exceeds a certain amount, then the opportunity cost starts to increase. Furthermore, as the supply of one good increases, the demand for the good decreases — in other words, society wants it less. Hence, because the opportunity cost of producing one good increases with the amount produced and society's aggregate desire for the product decreases, then there is a certain amount of each individual good or service that provides the maximum satisfaction to society.

How well economic resources are allocated to satisfy economic wants is known as **allocative efficiency**, which is the production of those goods and services that are most desired by society in the quantity that they desire.

Because the opportunity cost of producing goods or services and the benefit of the output to society depends on how much of each good or service is produced, then how does an

individual producer know how much to produce? Since the producers' costs are variable, then it makes sense to ask if an additional unit should be produced.

Because of economies of scale, the cost of producing an additional unit, known as the **marginal cost**, continually declines until it reaches a minimum. On the other hand, the **marginal benefit**, which is the benefit to society of the last unit, also declines with increasing quantity. Since the amount that consumers are willing to pay for a particular product depends on the marginal benefit of the last unit produced, a producer can maximize profit by producing more units until the marginal cost of producing the last unit is equal to the marginal benefit of that unit to society.

Note that, in economics, economic cost also includes the profit that would satisfy the producer. Therefore, we can derive a simple formula for maximizing allocative efficiency. Allocative efficiency is maximized when the marginal benefit equals the marginal cost of producing one extra unit.

**Marginal Benefit = Marginal Cost**

If the marginal benefit exceeds the marginal cost, then people will be willing to pay a higher price than the economic cost of production, which therefore allows the manufacturer to increase profits by producing more. When allocative efficiency is maximized with respect to the good, then the manufacturer is producing the good in the exact quantities that society desires. On the other hand, if the manufacturer produces more goods than what society desires, then the price that they are willing to pay will be less than the economic cost of production, and will cause profits to drop.

**Expanding the Production Possibilities Frontier**

The assumptions of the production possibility frontier simplify the analysis of the effect of allocating resources differently. Over time, however, the production possibility frontier can be expanded.

Increases in the production possibility frontier will be made possible by increased resources, such as population growth, advances in technology, and also by the application of

technology, such as developing irrigation systems so that more arid land can be used for farming. However, the largest expansion of the production possibility frontier of any economy will probably result from advances in technology, especially computer, network, and robotics technology. Not only do computers reduce the need for resources, but robotics will increasingly reduce the need for human labor. With the aid of computers, networks, and robotics, each person will be able to do more and more, thereby greatly increasing the productivity of each individual. Furthermore, the Internet can quickly distribute information and knowledge to anyone who wants it, which also greatly improves productivity.

Consider the simple elimination of paper. The Internet, computers, e-readers, and other portable devices allow the creation, distribution, and consumption of information and knowledge in an electronic format. Content can be created on computers, distributed by the Internet to anywhere in the world for virtually no cost, then read on portable devices. This reduces the need for the many trees to manufacture paper, the need for delivery vehicles and their requirement for fuel and other maintenance, the need of publishing houses and distribution centers, such as bookstores, and their associated land and labour.

Note that economic growth will always be uneven, that the production of some goods can become more cost-effective faster than other goods. However, when the cost of producing a good decreases, it also frees economic resources for the production of other goods or services.

An economy’s choice of how much to invest in capital and consumer goods will determine how fast future growth will be. A greater investment in capital goods over consumer goods will usually result in faster future growth but at the expense of current consumption.

International trade also allows a nation to benefit from the comparative advantage of other countries. Thus, by using the cheaper labour in China and India, United States manufacturers and service providers can lower the costs of their output, which benefits American consumers.

The production possibility frontier can also be expanded by eliminating anything that causes unemployment, such as

discrimination based on race, sex, or any other quality and by eliminating unnecessary bureaucracies and laws. For instance, by greatly simplifying the tax code, people would be able to do their own taxes, eliminating the need for tax preparers, and many accountants and tax lawyers. It would also minimize the many economic distortions created by a complicated tax code.

So why are not the laws changed to maximize economic efficiency? Because politicians like to cater to special interests and because most of them do not know economics. Hence, we all suffer.

## 1.4 FIRST BEST ECONOMY

It is an Economics concept means if all but one requirement for achieving a most desirable economic situation cannot be satisfied, it is always beneficial to satisfy the remaining ones.

### Theory of the second best

In economics, the **theory of the second best** concerns the situation when one or more optimality conditions cannot be satisfied. The economists Richard Lipsey and Kelvin Lancaster showed in 1956, that if one optimality condition in an economic model cannot be satisfied, it is possible that the next-best solution involves changing other variables away from the values that would otherwise be optimal. Politically, the theory implies that if it is infeasible to remove a particular market distortion, introducing a second (or more) market distortion may partially counteract the first, and lead to a more efficient outcome.

### Implications

In an economy with some uncorrectable market failure in one sector, actions to correct market failures in another related sector with the intent of increasing economic efficiency may actually decrease overall economic efficiency. In theory, at least, it may be better to let two market imperfections cancel each other out rather than making an effort to fix either one. Thus, it may be optimal for the government to intervene in a way that is contrary to usual policy. This suggests that economists need to study the details of the situation before jumping to the theory-based conclusion that an improvement in market perfection in one area implies a global improvement in efficiency.

**Example**

Even though the theory of the second best was developed for the Walrasian general equilibrium system, it also applies to partial equilibrium cases. For example, consider a mining monopoly that's also a polluter: mining leads to tailings being dumped in the river and deadly dust in the workers' lungs. Suppose in addition that there is nothing at all that can be done about the pollution without also reducing production. However, the government is able to break up the monopoly.

The problem here is that increasing competition in this market is likely to increase production (since competitors have such a hard time restricting production compared to a monopoly). Because pollution is highly associated with production, pollution will most likely increase. Thus, it is not clear that eliminating the monopoly increases efficiency. Gains from trade in coal will increase, but externalities from pollution will increase as well, possibly outweighing the gains from trade.

## 1.5 MARKET FAILURE

In economics, **market failure** is a situation in which the allocation of goods and services is not efficient. That is, there exists another conceivable outcome where an individual may be made better-off without making someone else worse-off. Market failures can be viewed as scenarios where individuals' pursuit of pure self-interest leads to results that are not efficient – that can be improved upon from the societal point of view. The first known use of the term by economists was in 1958, but the concept has been traced back to the Victorian philosopher Henry Sedgwick.

Market failure occurs due to inefficiency in the allocation of goods and services. A price mechanism fails to account for all of the costs and benefits involved when providing or consuming a specific good. When this happens, the market will not produce the supply of the good that is socially optimal – it will be over or under produced.

In order to fully understand market failure, it is important to recognize the reasons why a market can fail. Due to the structure of markets, it is impossible for them to be perfect. As a result, most markets are not successful and require forms of intervention.

**Reasons for market failure include**:

1. **Positive and negative externalities:** an externality is an effect on a third party that is caused by the consumption or production of a good or service. A positive externality is a positive spillover that results from the consumption or production of a good or service. For example, although public education may only directly affect students and schools, an educated population may provide positive effects on society as a whole. A negative externality is a negative spillover effect on third parties. For example, second hand smoke may negatively impact the health of people, even if they do not directly engage in smoking.

   Traffic congestion is an example of market failure that incorporates both non-excludability and externality. Public roads are common resources that are available for the entire population's use (non-excludable), and act as a complement to cars (the more roads there are, the more useful cars become). Because there is very low cost but high benefit to individual drivers in using the roads, the roads become congested, decreasing their usefulness to society. Furthermore, driving can impose hidden costs on society through pollution (externality). Solutions for this include public transportation, congestion pricing, tolls, and other ways of making the driver include the social cost in the decision to drive.

   Perhaps the best example of the inefficiency associated with common/public goods and externalities is the environmental harm caused by pollution and overexploitation of natural resources.

2 **Environmental concerns:** effects on the environment as important considerations as well as sustainable development.

3 **Lack of public goods:** public goods are goods where the total cost of production does not increase with the number of consumers. As an example of a public good, a lighthouse has a fixed cost of production that is the same, whether one ship or one hundred ships use its light. Public goods can be under-produced; there is little incentive, from a private standpoint, to provide a lighthouse because one can wait for someone else to provide it, and then use its light without incurring a cost. This problem - someone benefiting from resources or goods and

services without paying for the cost of the benefit - is known as the free rider problem.

4 **Property rights:**
Markets work most effectively when consumers and producers are granted the right to own property, but in many cases property rights cannot easily be allocated to certain resources. Failure to assign property rights may limit the ability of markets to form.

5 **Information failure:**
Markets may not provide enough information because, during a market transaction, it may not be in the interests of one party to provide full information to the other party.

6 **Unstable markets:**
Sometimes markets become highly unstable, and a stable equilibrium may not be established, such as with certain agricultural markets, foreign exchange, and credit markets. Such volatility may require intervention.

7 **Inequality**
Markets may also fail to limit the size of the gap between income earners, the so-called income gap. Market transactions reward consumers and producers with incomes and profits, but these rewards may be concentrated in the hands of a few.

8 **Underproduction of merit goods:** a merit good is a private good that society believes is under consumed, often with positive externalities. For example, education, healthcare, and sports centers are considered merit goods.

9 **Overprovision of demerit goods:** a demerit good is a private good that society believes is over consumed, often with negative externalities. For example, cigarettes, alcohol, and prostitution are considered demerit goods.

10 **Abuse of monopoly power**: imperfect markets restrict output in an attempt to maximize profit.

When a market fails, the government usually intervenes depending on the reason for the failure.

**Remedies**

In order to reduce or eliminate market failures, governments can choose two basic strategies:

**1. Use the price mechanism**

The first strategy is to implement policies that change the behaviour of consumers and producers by using the price mechanism. For example, this could mean increasing the price of 'harmful' products, through taxation, and providing subsidies for the 'beneficial' products. In this way, behaviour is changed through financial incentives, much the same way that markets work to allocate resources.

**2. Use legislation and force**

The second strategy is to use the force of the law to change behaviour. For example, by banning cars from city centers, or having a licensing system for the sale of alcohol, or by penalising polluters, the unwanted behaviour may be controlled.

In the majority of cases of market failure, a combination of remedies is most likely to succeed.

**3. Others measures:**

- Tax on Negative Externalities – e.g. Petrol tax
- Carbon Tax e.g. tax on CO2 emissions
- Subsidy on positive externalities – why government may subsidies public transport
- Laws and Regulations – Simple and effective ways to regulate demerit goods, like ban on smoking advertising.
- Buffer stocks – aim to stabilise prices
- Government failure – why government intervention may not always improve the situation

## 1.6 RATIONALE FOR STATE INTERVENTION

**What are the main reasons for government intervention in markets?**

The main reasons for policy intervention by the government are:

- To correct for **market failures**
- To achieve a more **equitable distribution of income and wealth**
- To improve the **performance of the economy**

| Type of Market Failure | Consequence of Market Failure | Example of Government Intervention |
|---|---|---|
| Factor immobility | Structural unemployment | State investment in education and training |
| Public goods | Failure of market to provide pure public goods, free rider problem | Government funded public goods for collective consumption |
| Demerit goods | Over consumption of products with negative externalities | Information campaigns, minimum age for consumption |
| Merit goods | Under consumption of products with positive externalities | Subsidies, information on private benefits |
| Imperfect information | Damaging consequences for consumers from poor choices | Statutory information / labeling |
| High relative poverty | Low income families suffer social exclusion, negative externalities | Taxation and welfare to redistribute income and wealth |
| Monopoly power in a market | Higher prices for consumers causes loss of allocative efficiency | Competition policy, measures to encourage new firms into a market |

**Rationales for Intervention**

A second One reason that governments impose policies on their agricultural sector is the belief that intervention can accelerate the rate of income growth. Investment policies-the provision of public goods, such as the research and development of new technologies and infrastructural development (roads, schools, health facilities)-are examples of public sector interventions essential for increased economic activity. Sometimes, these investments will not be made by the private sector. Private investors may be unable to capture the full benefit from investment in public goods because it is impossible or too costly to exclude those who do not pay for services created. In other instances, consumption by one consumer does not reduce the availability of the good or service for others. Consumers therefore avoid declaring their willingness to pay for the good or service, and a market does not form. Finally, capital requirements of the investment might

exceed the private sector's capacity to mobilize necessary financial resources. For most of these investments, the public sector has the potential to recover the costs of intervention through user fees or through taxation of the commodities or the regional populations that benefit from the investment.

The correction of market failures represents a second rationale for government intervention in the agricultural sector. If market imperfections are present, the prices of goods or services will not reflect their true scarcity values because the private sector is unable to develop the institutions necessary for efficient market functioning. Rural credit markets, for example, might be hampered by a lack of information on alternative lending and borrowing opportunities in other regions, or by the absence of formal lending institutions that can mobilize savings. Market power is another example of a market failure; private sector suppliers (or consumers) are able to influence prices because their numbers are small and because buyers (or sellers) have no other market outlets. These conditions are asserted to prevail often in factor markets (those for labor, credit, and land) and sometimes in remote rural commodity markets.

Another type of market failure arises because of externalities-costs or benefits from production activities that are not fully reflected in market incentives. Soil erosion, environmental pollution, and overutilization of common property resources are common externalities. Some form of government intervention-a tax, subsidy, or regulatory control-is justified so that user costs (or returns) will reflect fully the effects of the externality. The value of an externality is often difficult to quantify, and in many cases subjective judgments must be made as to whether externality effects are significant. These measurement problems, combined with the administrative costs of tax and subsidy policies, cause quantitative or legislative regulations to be widespread policy responses to externalities.

Although policies to correct market failures or to provide public goods can be important, the most common rationale for intervention in developing country agriculture is the promotion of non-efficiency objectives. The establishment of an efficient economy and the maximization of aggregate income are not the only, or necessarily the most important, goals of economic policy. When policy-makers are dissatisfied with the implications of income maximization, policies will be used to alter the economy. In some cases, these interventions will reflect neutral policymakers acting

on a mandate from society. But more often, policies respond to the desires of special interest groups within or outside agriculture.

Income distribution concerns are often at the top of the list of non-efficiency objectives. Food is the most basic of necessities, and low prices of food are considered an important determinant of the welfare level of poor consumers. Staple food prices influence producer income levels as well, and the manipulation of producer prices may generate a more equitable distribution of income in the economy. Income distribution policies will also reflect the influences of rent-seekers-agricultural commodity producers and input suppliers, consumers of food, and industrialists who view changes in agricultural prices as ways to increase profitability in production or to increase purchasing power in consumption. Government policies can benefit target groups through direct regulation of prices-such as tariffs or subsidies on imports-or through policies that provide market power to the target group, such as the designation of monopoly suppliers of particular agricultural products or the allocation of import and export licenses.

Price stabilization is a second common justification for intervention in agriculture. Dependence on the weather causes agricultural production to exhibit a relatively large degree of random variation. When combined with inelastic demand, supply variations can cause market prices to fluctuate substantially from one production cycle to the next. The consequent potential income fluctuations for poor producers and variations in expenditure for poor consumers are often unacceptable to policy-makers. To avoid substantial fluctuations in domestic market prices, many governments establish a set of policies, choosing among international trade controls, storage schemes, price fixing, and rationing. Elements of market failure are also partially responsible for interventions of this type. In production, for example, crop insurance and futures and options markets are institutions that reduce the uncertainty of future prices and income. However, these institutions are usually absent from developing country markets.

National concern over the appropriate role for agriculture in the economy provides a third set of non-efficiency rationales for government intervention. Food security and self-reliance of staple food supplies are commonly held objectives for agricultural policy. For food-importing countries, the attainment of these objectives requires intervention to increase domestic production. This intervention might involve changes in producer prices of outputs and inputs, investment in infrastructure for production or marketing

activities, or quantitative restrictions on the production of alternative crops. Agriculture also contributes to government revenue and the maintenance of fiscal balance in the public sector. Income taxes are a relatively unimportant revenue source in most developing countries because informal methods of income payment are prominent. As a consequence, the administrative costs of income monitoring and tax collection are often prohibitive, and indirect taxes on commodities are an important source of revenue. Because of its large size, the agricultural sector is usually expected to play a prominent role in the generation of tax revenues.

The relative importance of each justification for intervention in the agricultural sector follows no particular pattern across countries. In part, this variation results from wide disparities in the distribution of political power. The importance and effectiveness of various lobbying groups-domestic producers, consumers, government agencies, and foreign governments and corporations-vary enormously across countries. Consequently, cross-country variations in agricultural sector objectives are large. Differential resource constraints also create cross-country differences in agricultural objectives. Governments have objectives for sectors other than agriculture, which implies that budget constraints are a potential limitation on agricultural sector interventions. Technological limitations also might mean that some objectives cannot be realized at reasonable cost. To some extent, policy-makers can overcome constraints by judiciously selecting policies. Selection of the policy that minimizes budgetary cost allows the furtherance of more objectives than would otherwise be the case. But, ultimately, constraints in most developing countries become binding well before all the objectives of agricultural policy can be achieved.

## 1.7 REGULATION AND TAXATION

We study which policy tool and at what level a majority chooses in order to reduce activities with negative externalities. We consider three instruments: a rule that sets an upper limit to the activity which produces the negative externality, a quota that forces a proportional reduction of the activity, and a proportional tax on it. For all instruments the majority chooses levels which are too restrictive when the activity is performed mainly by a small fraction of the population, and when costs for reducing activities or paying taxes are sufficiently convex. Also a majority may prefer an instrument different than what a social planner would choose; for instance a rule when the social planner would choose a tax.

Three ways of reducing the level of an activity generating negative externalities are routinely used: a rule that sets an upper bound to this activity, a proportional tax on it, a compulsory proportional reduction of the activity for everybody. We should investigate which policy and at what level would be chosen by majority voting. The latter does not deliver the optimal policy choice for two reasons. First, for given policy instrument, majority voting does not yield the optimal level of it. Second, and perhaps more interestingly, when choosing amongst alternative instruments, majority voting in general does not lead to the choice of the optimal one. For instance, the majority may choose a rule instead of a proportional tax because a rule concentrates on the minority the burden of the reduction of the activity which generates negative externalities. A social planner would instead choose a tax and, if he were constrained to choose a rule, he would choose one which was more permissive than the one chosen by the majority. We thus have a "double distortion" caused by voting. This case arises when those who generate a negative externality are a minority. The opposite double distortion occurs when the activity with negative externality is enjoyed by many. In this case a social planner may choose a restrictive rule, while a majority may choose a lenient tax. These insights are consistent with the evidence that in many cases we observe regulation while the optimal policy would be taxation, or vice versa. For example, in agriculture the limits in the use of pesticides are quite frequent whereas taxes on them are less common. In the case of air pollution, there is a sharp contrast between the use of taxes and the use of emission standards. The latter are preferred when polluters are concentrated in specific industries or plants, such as emissions of pollutants by power generation industries or by steel and cement makers. Anti-smoking regulations became very strict as the number of smokers declined. We see low levels of taxation when the polluters are the majority; for instance low taxes on gasoline in the US or on heating fuel.

Policymakers may choose quotas when tax collection is costly or simply impossible, or because they are perceived as a fair method of sharing the sacrifices of curbing externalities (e.g. international agreements, like the Kyoto protocol, or in many cases in the European Union). We label our negative externality "pollution" for brevity. However our discussion of instrument choice applies to many other policy issues, which may include construction rules, speed limits, rules of behavior in communities like condominiums, prohibition (or very strict regulation) of certain

activities, from gambling to selling of organs, to prostitution to free acquisition of guns and many others. Thus we should have a model which is sufficiently general to be applied to a variety of different cases. In some of those, the externality has the straightforward interpretation of monetary costs inflicted on others. In other cases, it may take the form of a negative “utility cost” inflicted on others, who engage in certain activities which they find objectionable, like gambling or prostitution. Baron (2003) claims that “moralistic” goals regarding how others should behave are prominent in how people vote. Roth (2007) in his discussion of organ exchanges argues that repugnance of certain transactions related to trades in organs, implies relevant social costs.

This is why we feel that it is appropriate to use a majority rule voting model. Much of the literature on “pollution” strictly defined adopts lobbying models. While lobbying pressures are clearly important, especially for legislation which affects one particular sector, clearly decisions regarding the list of activities mentioned above, from smoking to gun control etc., involve voting in legislatures or even in private associations, e.g. owners’ associations. Take, for instance, smoking regulations. Clearly the decision regarding smoking age, taxation over cigarettes etc., is influenced by the lobby of the tobacco industry. But the fraction of individuals smoking will also influence the legislative choice regarding regulation and taxation of smoking. The same applies to gun control: the gun lobby is strong but different states in the US have different regulations as a function of the preferences of the voters. We should make clear from the outset that we consider only proportional taxes on the polluting activities. By allowing any type of curvature on the tax schedule, including corners, one could reproduce patterns which approximate a rule, and are quite far from the allocation generated by a proportional tax. In a “positive” politico economic model we need to worry about the existence of a Condorcet winner [Condorcet - French mathematician and philosopher (1743-1794)]. While we can prove its existence with a proportional tax, in general one cannot do that with any curvature of the tax schedule.

Information asymmetries and costs of administration may give rise to imperfections in both externality taxes and direct regulation. Used alone, neither instrument may be capable of achieving the most efficient reduction in external costs, and a more efficient outcome may be achieved by combined use of multiple

instruments. We should look at the economic issues which arise when externality taxes and direct regulation are used in parallel. It explores the properties of two simple models of imperfect tax and imperfect regulation, reflecting different form of imperfection in the tax instrument.

Regulation may affect consumption behaviour in a number of ways. We suggest that for a number of consumer externalities it may be useful to think of regulation as an increase in the cost to consumers of obtaining the good (for example, where the sale of alcoholic drinks is limited to a small number of outlets). Regulation thus has effects which are similar to - but not equivalent to - an increase in price. It will be seen that the representation of regulation here differs sharply from the emission limits or technology mandates typically considered when analyzing the regulation of industrial emission.

## 1.8 DISTRIBUTIONAL OBJECTIVES OF THE STATE

An argument for government intervention can also be made if the economy has widespread inequality of income, opportunity or wealth. Such inequalities can occur even if the economy is efficient in a narrow economic sense. In such circumstances, the level of economic welfare may be raised by the redistribution of resources to alleviate these inequalities. This reasoning underlies the provision of state education, social security programmes and compulsory pension schemes. It should be stressed that the gains from these policies are with respect to normative assessments of welfare, unlike the positive criteria lying behind the concept of economic efficiency. In conducting an economic policy the state will have two conflicting aims. On the one hand, it will aim to raise revenue to finance the policy with the minimum loss to society. The raising of revenue leads to losses due to the resources used in the collection process and from the economic distortions that it causes. Minimising these losses is the efficiency aspect of policy design. Conversely, the state may also feel that it is desirable to intervene in the economy in order to attain a more equitable distribution of the economy's resources. This is often accompanied by a corresponding reduction in the degree of concern for the aggregate level of economic activity. This motivation represents the equity side of policy design. Due to their distinct natures, it is inevitable that the aims of equity and efficiency regularly conflict. The efficient policy is often highly inequitable, whilst the equitable policy may

introduce into the economy significant distortions and disincentives. Given this, the design of optimal policy is concerned with reaching the correct trade-off between equity and efficiency objectives. The optimum trade-off will depend upon the concern for equity that is expressed in the objectives of the policy maker. The resolution of the trade-off between equity and efficiency is the major determinant of the resulting policy program, with aspects of the policy being attributable to one or the other. This distinction is often a helpful way in which to think about optimality problems and their solutions.

**Redistribution of income** and **redistribution of wealth** are respectively the transfer of income and of wealth (including physical property) from some individuals to others by means of a social mechanism such as taxation, charity, welfare, land reform, monetary policies, confiscation, divorce or tort law. The term typically refers to redistribution on an economy-wide basis rather than between selected individuals, and it always refers to redistributions from those who have more to those who have less.

Redistribution tax policy should not be confused with pre-distribution of wealth, where the lower and middle classes pay higher net effective tax percentage rates, as the elite pay regressive tax rates. Itemized deductions, often called tax loopholes, tend to perpetuate pre-distribution preferences in lieu of implementing a neutral tax system, such as a flat tax. Many alternate taxation proposals have been floated without the political will to alter the status quo. The proposed "Buffett Rule" is a hybrid taxation model, a compromise of opposing systems, intended to minimize the favouritism of the special interest tax design.

The effects of a redistribution system are actively debated on ethical and economic grounds. The subject includes analysis of its rationales, objectives, means, and policy effectiveness.

**Modern forms of redistribution**

Today, income redistribution occurs in some form in most democratic countries. In a progressive income tax system, a high income earner will pay a higher tax rate than a low income earner. Another taxation-based method of redistributing income is the negative income tax.

Two other common types of governmental redistribution of income are subsidies and vouchers (such as food stamps). These transfer payment programs are funded through general taxation, but benefit the poor, who pay fewer or no taxes. While the

persons receiving transfers from such programs may prefer to be directly given cash, these programs may be more palatable to society than cash assistance, as they give society some measure of control over how the funds are spent.

The difference between the Gini index for the income distribution before taxation and the Gini index after taxation is an indicator for the effects of such taxation.

Wealth redistribution can be implemented through land reform that transfers ownership of land from one category of people to another, or through inheritance taxes or direct wealth taxes. Before-and-after Gini coefficients for the distribution of wealth can be compared.

**Objectives**

The objectives of income redistribution are to increase economic stability and opportunity for the less wealthy members of society and thus usually include the funding of public services.

One basis for redistribution is the concept of distributive justice, whose premise is that money and resources ought to be distributed in such a way as to lead to a socially just, and possibly more financially egalitarian, society. Another argument is that a larger middle class benefits an economy by enabling more people to be consumers, while providing equal opportunities for individuals to reach a better standard of living. Seen for example in the work of John Rawls, another argument is that a truly fair society would be organized in a manner benefiting the least advantaged, and any inequality would be permissible only to the extent that it benefits the least advantaged.

Some proponents of redistribution argue that capitalism results in an externality that creates unequal wealth distribution.

Some argue that wealth and income inequality are a cause of economic crises, and that reducing these inequalities is one way to prevent or ameliorate economic crises, with redistribution thus benefiting the economy overall. This view was associated with the under consumptionism school in the 19th century, now considered an aspect of some schools of Keynesian economics; it has also been advanced, for different reasons, by Marxian economics. It was particularly advanced in the US in the 1920s by Waddill Catchings and William Trufant Foster. There is currently a great debate concerning the extent to which the world's extremely rich have become richer over recent decades.

**Equity** or **economic equality** is the concept or idea of fairness in economics, particularly in regard to taxation or welfare economics. More specifically, it may refer to equal life chances regardless of identity, to provide all citizens with a basic and equal minimum of income, goods, and services or to increase funds and commitment for redistribution.

Inequality and inequities have significantly increased in recent decades, possibly driven by the worldwide economic processes of globalisation, economic liberalisation and integration. This has led to states 'lagging behind' on headline goals such as the Millennium Development Goals (MDGs) and different levels of inequity between states have been argued to have played a role in the impact of the global economic crisis of 2008–2009.

Equity is based on the idea of moral equality. Equity looks at the distribution of capital, goods and access to services throughout an economy and is often measured using tools such as the Gini index. Equity may be distinguished from economic efficiency in overall evaluation of social welfare. Although 'equity' has broader uses, it may be posed as a counterpart to economic inequality in yielding a "good" distribution of wealth. It has been studied in experimental economics as inequity aversion. Low levels of equity are associated with life chances based on inherited wealth, social exclusion and the resulting poor access to basic services and intergenerational poverty resulting in a negative effect on growth, financial instability, crime and increasing political instability.

The state often plays a central role in the necessary redistribution required for equity between all citizens, but applying this in practise is highly complex and involves contentious choices.

## 1.9 QUESTIONS

1) Explain the concept of efficiency in Resource allocation.
2) Write note on :
   a) First Best Economy
   b) Market Failure
3) Justify the Rationale for State Intervention.
4) Discuss the distributional objectives of the state.

# 2

# SOCIAL WELFARE FUNCTIONS & GOVT. FAILURES

**Unit Structure :**

2.1 Objectives
2.2 Introduction
2.3 Social welfare functions
2.4 Decision making in government
2.5 Arrow Impossibility Theorem
2.6 Government Failures
2.7 Questions

## 2.1 OBJECTIVES

- To understand the concept of Social welfare functions
- To study the decision making process in governments
- To consider the relevance of Arrow's Impossibility theorem
- To study the reasons of Government Failures

## 2.2 INTRODUCTION

**Welfare economics** is a branch of economics that uses microeconomic techniques to evaluate well-being (welfare) at the aggregate (economy-wide) level. A typical methodology begins with the derivation (or assumption) of a social welfare function, which can then be used to rank economically feasible allocations of resources in terms of the social welfare they entail. Such functions typically include measures of economic efficiency and equity, though more recent attempts to quantify social welfare have included a broader range of measures including economic freedom (as in the capability approach).

The field of welfare economics is associated with two fundamental theorems. The first states that given certain assumptions, competitive markets produce (Pareto) efficient outcomes. it captures the logic of Adam Smith's invisible hand. The second states that given further restrictions, any Pareto efficient

outcome can be supported as a competitive market equilibrium. Thus a social planner could use a social welfare function to pick the most equitable efficient outcome, then use lump sum transfers followed by competitive trade to bring it about.Because of welfare economics' close ties to social choice theory, Arrow's impossibility theorem is sometimes listed as a third fundamental theorem.

Attempting to apply the principles of welfare economics gives rise to the field of public economics, the study of how government might intervene to improve social welfare. Welfare economics also provides the theoretical foundations for particular instruments of public economics, including cost–benefit analysis, while the combination of welfare economics and insights from behavioural economics has led to the creation of a new subfield, behavioural welfare economics.

In welfare economics, a **social welfare function** is a function that ranks social states (alternative complete descriptions of the society) as less desirable, more desirable, or indifferent for every possible pair of social states. Inputs of the function include any variables considered to affect the economic welfare of a society. In using welfare measures of persons in the society as inputs, the social welfare function is individualistic in form. One use of a social welfare function is to represent prospective patterns of collective choice as to alternative social states.

The social welfare function is analogous to the consumer theory of indifference-curve/budget constraint equilibrium for an individual, except that the social welfare function is a mapping of individual preferences or judgments of everyone in the society as to collective choices, which apply to all, whatever individual preferences are for (variable) constraints on factors of production. One point of a social welfare function is to determine how close the analogy is to an ordinal utility function for an individual with at least minimal restrictions suggested by welfare economics, including constraints on the amount of factors of production.

There are two major distinct but related types of social welfare functions. A Bergson–Samuelson social welfare function considers welfare for a *given set* of individual preferences or welfare rankings. An Arrow social welfare function considers welfare across *different possible sets* of individual preferences or welfare rankings and seemingly reasonable axioms that constrain the function.

## 2.3 ARROW SOCIAL WELFARE FUNCTION

Kenneth Arrow (1963) generalizes the analysis. Along earlier lines, his version of a social welfare function, also called a 'constitution', maps a set of individual orderings (ordinal utility functions) for everyone in the society to a social ordering, a rule for ranking alternative social states (say passing an enforceable law or not, ceteris paribus). Arrow finds that nothing of behavioural significance is lost by dropping the requirement of social orderings that are *real-valued* (and thus cardinal) in favour of orderings, which are merely *complete* and *transitive*, such as a standard indifference curve map. The earlier analysis mapped any set of individual orderings to *one* social ordering, whatever it was. This social ordering selected the top-ranked *feasible* alternative from the economic environment as to resource constraints. Arrow proposed to examine mapping different sets of individual orderings to possibly different social orderings. Here the social ordering would depend on the set of individual orderings, rather than being *imposed* (invariant to them). Stunningly (relative to a course of theory from Adam Smith and Jeremy Bentham on), Arrow proved the *General Possibility Theorem* that it is impossible to have a social welfare function that satisfies a certain set of "apparently reasonable" conditions.

The concept of 'Social Welfare Function' was propounded by A. Bergson in his article 'A Reformulation of Certain Aspects of Welfare Economics' in 1938. Prior to its various concepts of social welfare had been given by different welfare theorists but they failed to provide a satisfactory solution to the problem of maximisation of social welfare and measurement. Bentham talked of welfare in terms of 'the greatest happiness of the greatest number.'

Neo-Classical welfare theorists discussed the problem of social welfare on the basis of cardinal measurability of utility and interpersonal comparison of utility. Analysis of Pareto optimality maximises social welfare by satisfying various marginal conditions of production, distribution and allocation of resources among products. But unfortunately they are not fulfilled due to the existence of various externalities and imperfections in the market. Moreover, Pareto optimality analysis fails to measure the changes in welfare resulting from any change which benefits one section of society and harms the other.

Compensation principle as given by Kaldor-Hicks-Scitovsky attempts to measure the changes in social welfare resulting from such economic changes which harm some and benefit others through hypothetical compensating payments.

Compensation theorists claimed to give a value-free objective criterion based on ordinal concept of utility but, this is based upon implicit value judgements and does not evaluate changes in social welfare satisfactorily.

By providing the concept of social welfare function Bergson and Samuelson have attempted to provide a new approach to welfare economics and have succeeded in rehabilitating welfare economics. They have put forward the concept of social welfare function that considers only the ordinal preferences of individuals.

They agree to Robbins' view that interpersonal comparison of utility involves value judgements but they assert that without making some value judgements, economists cannot evaluate the impact of changes in economic policy on social welfare.

Thus, according to them, welfare economics cannot be separated from value judgements. According to them, welfare economics is essentially a normative study. But the approach to study it must be scientific despite the fact that the use of value judgements in it is unavoidable.

**Bergson-Samuelson Social Welfare Function:**

Social welfare function is an ordinal index of society's welfare and is a function of the utility levies of all individuals constituting the society.

**Bergson-Samuelson social welfare function can be written in the following manner:**

$W = W(U_1, U_2, U_3 \ldots\ldots\ldots\ldots, U_n)$

Where W If represents the social welfare $U_1$, $U_2$, $U_3$, . .. ., $U_n$ represent the ordinal utility indices of different individuals of the society. The ordinal utility index of an individual depends upon the goods and services he consumes and the magnitude and kind of the work he does. The important thing to note about social welfare function is that in its construction explicit value judgements are introduced.

Value judgements determine a form of the social welfare function; with a different set of value judgements, the form of social welfare function would be different. Value judgements are essentially ethical notions which are introduced from outside economics. The value judgements required to construct a social welfare function may be obtained through democratic process with voting by individuals or it may have to be imposed on the society in a dictation manner.

Whatever the case may be, the form of social welfare function depends upon the value judgements of those who decide about them since it expresses their views regarding the effect which the utility level of each individual has on the social welfare. In the worlds of Prof. Scitovsky. “The social welfare function can be thought of as a function of each individual’s welfare which in turn depends both in his personal well being and on his appraisal of the distribution of welfare among all members of the community”.

Since the value judgements required for the formation of social welfare function are not of the economist himself and instead they are introduced from outside economics they are not obtained through any scientific method.

It has been claimed that social welfare function has solved the basic problem of welfare economics, since it thinks unnecessary for the economists themselves to make value judgements concerning what is a desirable distribution of welfare as between in individuals constituting the society. In other words, economist need not himself decide about what is the most desirable distribution of welfare. He can take value judgements regarding distribution as given from outside economics.

Bergson’s social welfare function is supposed to be dependent on changes in economic events that have a direct effect on individual welfares. The ordinal utility level of an individual is a function of his own consumption of goods and services and not of others.

Moreover, the utility level of an individual depends on his own value judgments regarding the composition of different goods and services consumed which depends upon his tastes. An individual may derive more utility from the consumption of liquor

whereas another individual may derive very nominal utility or no utility at all from it.

**Social Welfare Function and Value Judgements:**

So far we have been mainly concerned with the value judgements of individuals regarding their utility levels. From the view point of social welfare function, the value judgements regarding the welfare of the society as a whole are relevant.

The formulation of a welfare function for the society as a whole is a very difficult task because utility being a mental phenomenon cannot be measured or estimated accurately by any person or institution entrusted to furnish value judgements regarding the changes in social welfare, Moreover, addition and subtraction of utilities of different individuals by an authorised person or institution too is a very difficult task.

The social welfare function and its form depend upon the value judgements of the person or institution that the society has authorised to decide. The authorised person or institution may be anybody but for true value judgements regarding the social welfare he must be unbiased because changes in social welfare will depend upon his value judgements.

“These judgements as to what constitute justice and virtue in distribution may be those of the economist himself or those set up by the legislature, by some other governmental authority or by some other unspecified person or group.” A social welfare function can be attained by common consensus or it may be forced upon the society by a dictator.

Since the forms of social welfare functions are known by value judgements about social welfare, therefore there arises the problem of finding an authority who could give purely unbiased value judgements. Bergson and Samuelson have assumed a “Superman” who provides value judgements about changes in social welfare.

Superman alone can take decisions about the solution of various problems of the economy. What goods and services should be produced and supplied in the society? How much of various goods should be produced? What should be the quality and kind of goods?

What should be capital intensity of producing a particular type of good? What should be the pattern of distribution of national income among different sections of the society? Which wants should be satisfied at present and which at a future date and so on. All these questions can be answered by the superman alone in accordance with his views about the determinants of social welfare.

The society would have to accept the solutions of all these questions provided by him assumption that he will give any value judgements which aim to achieve maximum social welfare rather than maximum self-interest. Thus we are free from the addition, subtraction, measurement and interpersonal comparisons of utilities by assumption the existence of a superman.

In modern age of democratic governments people elect their representatives who constitute the Government. The political party in majority forms the Government and rules the country. The representatives' Government formed by the majority rule formulates various policies on the basis of value judgements and it is expected that all the policy decisions by the Government will aim at maximising social welfare rather than maximising the welfare of an individual or a particular section of the society.

Bergson and Samuelson expressed the view that all value judgements used to construct the social welfare function must be consistent which implies that if in a given situation A is preferred to B and B is preferred to C then A must be preferred to C. This is nothing new to the students of economics as this is the well know assumption of transitivity in social choice among various alternatives.

We can explain the social welfare function with the help of social indifference curves or welfare frontiers. Let us assume a society of two persons. In such a case social welfare function can be represented with the help of social indifference curves.

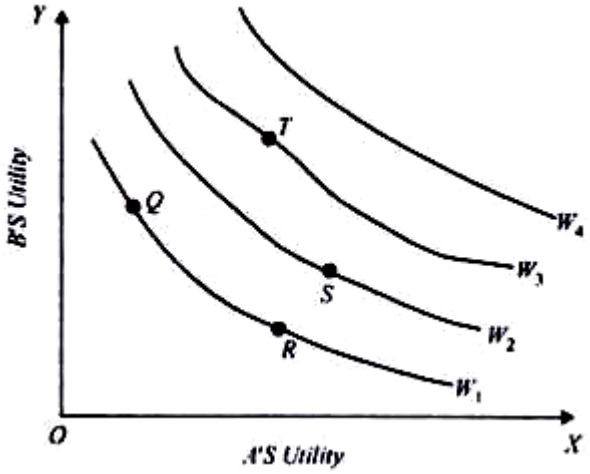



**Figure 2.1 - Social Indifference Curves depicting Social Welfare Function**

In Fig. 42.1 the utilities of individuals A and B have been represented on the horizontal and vertical axes respectively. $W_1$, $W_2$ and $W_3$ are the social indifference curves representing successively higher levels of social welfare. A social indifference curve is a locus of various combinations of utilities of A and B which result in an equal level of social welfare.

The properties of social indifference curves are just like that of individual consumer's indifference curves. Given a family of social indifference curves, the effect of a proposed change in policy on social welfare can be evaluated. In terms of Fig. 42.1 any policy change that moves the economy from Q to T is an improvement.

Similarly, a movement from Q to S or from R to S also represents an improvement in social welfare, and a movement from T to Q or T to S represents a decrease in social welfare. A movement along the same social indifference curve represents no change in the level of social welfare.

Analysis of Pareto optimality failed to provide a 'unique optimum solution' which represents maximum social welfare. There are a large number of solutions which are optimum on the basis of Pareto criterion. In terms of Edgeworth-box diagram every point on the contract curve represents the optimum position. In terms of Grand Utility Possibility Frontier, all points on it are Pareto optimal or economically efficient. But Pareto criterion does not tell us the best of them.

Thus, Paretian analysis leaves us with a lot of indeterminacy in the choice of maximum social welfare point. Now, the significance of social welfare function is that it enables us to obtain a unique optimum position regarding social welfare.

This unique optimum position is best of all the Pareto optima and therefore ensures the maximum social welfare. By including the concept of grand utility possibility frontier along with Bergson-Samuleson social welfare function we are able to obtain a unique optimum position or maximum social welfare position which is explained below.

**Grand Utility Possibility Frontier and Position of Constrained Bliss:**

As shall be explained below, a grand utility possibility frontier is a locus of the various physically attainable utility combinations of two persons when the factor endowments, state of technology and preference orders of the individuals are given.

In other words, every point on the grand utility possibility curve represents the optimum position with regard to the allocation of the products among the consumers, allocation of factors among different products and the direction of production. Thus every point on the grand utility possibility curve represents a Pareto optimum and as we move from one point to another on it the utility of one individual increases while that of the other falls.

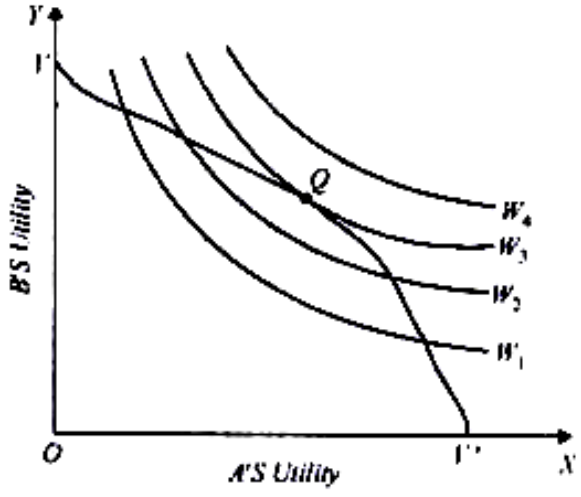



**Figure 2.2 - Social Welfare Function and Position of Constrained Bliss**

Now, let us superimpose grant utility possibility curve on the social indifference curves representing social welfare function to find a unique optimum position of social welfare. In Fig. 42.2 social indifference curves $W_1$, $W_2$, $W_3$ and $W_4$ representing the social welfare function have been drawn along with the grand utility possibility curve VV'.

Social indifference curve $W_3$ is tangent to the grand utility possibility curve VV' at point Q. Thus, point Q represents the maximum possible social welfare given the factor endowments, state of technology and preference scales of the individuals. Point Q is known as the point of constrained bliss since, given the constraints regarding factor endowments and the state of technology.

Q is the highest possible state of social welfare which the society can attain. Social welfare represented by the social indifference curve $W_4$ is higher than social indifference curve $W_3$ passing through Q but it is not possible to attain it, given the technology and factor endowment.

Thus, from among a large number of Pareto optimum points on the grand utility possibility curve, we have a unique optimum point Q at which the social welfare is the maximum. The point of constrained bliss represents the unique pattern of production of goods, unique distribution of goods between the individuals and unique combination of factors employed to produce the goods.

**The following features of the Bergson-Samuelson Social Welfare function are worth noting:**

1. The Bergson-Samuelson social welfare function is based on explicit value judgements and involves interpersonal comparisons of utility in ordinal terms.

2. Bergson-Samuelson social welfare function, the maximum social welfare position is completely determined as a result of the introduction of value judgements regarding distribution of welfare among individuals.

3. The social welfare function is not based on any unique value judgements. Instead, any set of value judgements can be used by a welfare economist to construct a social welfare function. Thus, it is

not any unique function but changes with the change in value judgements.

4. Once the social welfare function has been decided upon by value judgements, the maximisation technique is used to obtain the maximum social welfare position at which allocation of resources is Pareto optimum and also the distribution of goods and services is equitable. Thus, bath efficiency and equity are achieved so that social welfare may be maximised.

5. Used along with the Pareto optimality analysis the concept of social welfare function enables us to find a unique optimum solution which combines economic efficiency with distributive justice.

**A Critical Evaluation of Bergson-Samuelson Social Welfare Function:**

The main aim of welfare economics has been to find an acceptable social welfare function which could measure the changes in social welfare resulting from a change in economic and non- economic variables. Bergson and Samuelson solved this problem by formulating a social welfare function which is based on explicit value judgements.

This function can incorporate the various economic and non-economic determinants of the welfare of individuals. In this function utility or welfare is conceived and measured in ordinal terms. Preferences or utilities of different individuals of the society and decisions about them are taken through a democratic method or by an authorised institution on the basis of its own value judgements. Even according to its bitter critic little, the concept of social welfare function is a brilliant theoretical construct which completes the formal mathematical system of welfare economics.

Pareto optimality analysis does not help us in providing a unique solution to the problem of maximising social welfare. As seen above, with the help of social welfare function we can measure the changes' in social welfare even when one individual becomes better off and another worse off by making some distributional value judgements in the form of social welfare function.

The Bergson- Samuelson's social welfare function incorporating explicit value judgements is an improvement over

earlier attempts such as compensation principle advanced by Kaldor, Hicks and Scitovsky. However, economists have pointed out some important drawbacks in the concept of social welfare function.

**Limited Practical Significance:**

Little, Streeten and Baumol have pointed out that social welfare function is of limited practical significance. According to Little, the social welfare function can neither be used in a democratic state, nor even in a totalitarian one because in them there would be as many vague social welfare functions as there are individuals. Social welfare function, to quote Little is only "a formal device necessary to a perfectly general abstract system of 'welfare', which is devoid of any practical significance."

Likewise, Paul Streeton also thinks that social welfare function is a highly formal concept which has hardly any relation with the important facts of social life and choice. To quote him, "No political programme or individual value standard would fit the model of a social welfare function of the required type" Prof. Baumol is also of the opinion that the concept of social welfare is of limited practical value as it does not tell us how to get the value judgements which it requires for its construction.

Though Bergson criterion of social welfare function, writes Baumol, "provides us with a highly useful frame of reference, unfortunately it does not come equipped with a kit and a set of instructions for collecting the welfare judgements which it requires. Thus it still leaves us with the difficult part of the job unsolved"

Welfare depends on a wider range of variables than those associated with utility. Social welfare function approach is based on the utility which an individual derives from economic variables such as consumption of goods and services. Apart from these economic variables, welfare or well-being of individuals depends on a whole range of political and environmental variables such as enjoyment of human rights, political freedom, pollution-free environment.

Thus, "in comparing different economic systems or in comparing different ways of organising a given economy, the possibility that some of these variables might be affected cannot be ignored. Thus a reorganisation that gives everyone more income and leisure might not improve the welfare of the community if at the

same time it limits individual freedom or requires the abandonment of cherished cultural traditions."

**Impassibility of Constructing a Social Welfare Function from Individual Preferences:**

A highly damaging drawback of social welfare function has been pointed out by K.J. Arrow who has shown that social welfare function cannot be constructed on the basis of value judgements arrived at through democratic process of majority rule in group decision-making.

Arrow has proved that the majority rule leads to contradictory results or intransitivity of social choices when individuals are asked to make a choice from among more than two alternatives available to them. Therefore, Prof. Arrow concludes that a social welfare function which is based on mere ordinal preferences cannot in principle be constructed from the preferences of all the individuals comprising a society. Of course, social welfare function can be set up on the basis of value judgements of an individual who can impose his will on the society but that will reflect the aims and aspirations of an absolute dictator.

**Prof. Amartya Sen's Critique: Judging welfare or well-being in terms of utility is of limited significance:**

Prof. Amartya Sen has criticised modern welfare economics convening both Pareto efficiency and social welfare function on the ground that utility is not a true indicator of well-being. To quote him, "A difficulty with welfarism arises from the particular interpretation of well-being that utility provides. To judge, the well-being of a person exclusively in the metric of happiness or desire-fulfillment has some obvious limitations.

These limitations are particularly damaging in the context of interpersonal comparisons .of well-being." He further adds, "A person who has had a life of misfortune, with very little opportunities, and rather little hope, may be more easily reconciled to deprivations than others reared in more fortunate and affluent circumstances.

The metric of happiness may, therefore, distort the extent of deprivation, in a specific and biased way. The hopeless beggar, the precarious landless labourer, the dominated housewife, the hardened unemployed or the over-exhausted coolie may all take

pleasures in small mercies and manage to suppress intense suffering for the necessity of continuing survival, but it would be ethically deeply mistaken to attach a correspondingly small value to the loss of their well-being because of the survival strategy".

It follows from above that Amartya Sen has criticised the concept of social welfare based on utility which means psychological reactions of individuals to goods and services which they consume. Further, Prof. Sen shifts the focus on promoting positive freedoms of individuals for assessing the change in their welfare following a change in organisation or public policy. He defines freedom as 'capabilities to function' as to what persons can do or cannot do. It is capabilities to function that reflect freedom in the positive sense and determine well-being or welfare of the people.

## 2.4 DECISION MAKING IN GOVERNMENT

Decision-making is central to a government. How those decisions are made is important especially if the whole issue of decision-making might be seen to compromise the accepted standards of politics. At this moment in time, people are questioning the decision-making process within this government, especially who has an input into those final decisions.

Do we have a cabinet decision-making process or a process whereby decisions are made by a small clique both in the executive and out of it?

A Cabinet decision-making process is when all in the Cabinet feel confident enough to make an input into an issue without fear of retribution if their views are at odds with the Prime Minister or other senior members in the Cabinet (Foreign Secretary, Chancellor + Home Secretary). Such a process will be seen as democratic and inclusive. The final decision made will almost certainly be what the majority of the Cabinet feel is required, though the Prime Minister can override a majority Cabinet decision even after a full discussion, as part of his authority. However, this would be very rare and might spark a Cabinet rebellion. This process would still be seen as part of the Cabinet decision-making process as a full and frank exchange of vies have taken place.. The process is open, inclusive and all members should feel part of the process.

Prime Ministerial government is what Tony Blair has been accused of. The claim is that he bypasses accepted standards of decision-making and comes to decisions after consulting just a few people – including unelected people such as his former Director of Communications, Alaistair Campbell. In this way, the claim is that decisions taken have not included an input from all members of the Cabinet. Such a process, it is said, causes mistrust in the Cabinet. The claimed input of non-elected people from outside of the Cabinet also undermines the democratic process as such people are unaccountable to the electorate. Blair, of course, denies that he engages in such a system of decision-making.

Blair is not the only Prime Minister who was accused of this. Harold Wilson had his so-called 'kitchen cabinet' in the 1960's. Margaret Thatcher is said to have dominated her Cabinet to such an extent that she had a Prime Ministerial form of government even though she met with her Cabinet. It is said that her Cabinets were so dominated by her that they simply rubber-stamped her wishes regarding policy.

**Who makes an input into government decisions?**
Using the cabinet decision-making process –

- House of Commons via MP's who would have access to Cabinet members and are in parliamentary committees that examine government policy.
- House of Lords who comment and vote on government bills
- Pressure groups who have access to MP's and Cabinet members
- Civil Servants: senior civil servants ('mandarins') also give an input even if they are non-elected

**Using the Prime Ministerial system:**

- The Prime Minister
- People in the Prime Minister's inner circle, including civil servants and 'others', which might include a few in the Cabinet, non-elected government personnel etc.

In one sense, the ultimate decision-making body is Parliament as all government bills go through the Commons and the Lords. If a bill is voted down, any decision taken by the Cabinet or an inner-circle becomes invalid.

Also, the EU has a major input into government decision-making as all legislation must fit in with EU standards and the Factor tame decision by the Lords that EU legislation is superior to national legislation.

## 2.5 ARROW IMPOSSIBILITY THEOREM

In social choice theory, **Arrow's impossibility theorem**, the **General Possibility Theorem** or **Arrow's paradox** is an Impossibility theorem stating that when voters have three or more distinct alternatives (options), no ranked order voting system can convert the **ranked preferences** of individuals into a community-wide (complete and transitive) ranking while also meeting a pre-specified set of criteria. These pre-specified criteria are called *unrestricted domain*, *non-dictatorship*, *Pareto efficiency*, and *independence of irrelevant alternatives*. The theorem is often cited in discussions of voting theory as it is further interpreted by the Gibbard–Satterthwaite theorem.

The theorem is named after economist Kenneth Arrow, who demonstrated the theorem in his doctoral thesis and popularized it in his 1951 book *Social Choice and Individual Values*. The original paper was titled "A Difficulty in the Concept of Social Welfare".

In short, the theorem state that no rank-order voting system can be designed that always satisfies these three "fairness" criteria:

- If every voter prefers alternative X over alternative Y, then the group prefers X over Y.
- If every voter's preference between X and Y remains unchanged, then the group's preference between X and Y will also remain unchanged (even if voters' preferences between other pairs like X and Z, Y and Z, or Z and W change).
- There is no "dictator": no single voter possesses the power to always determine the group's preference.

Voting systems that use cardinal utility (which conveys more information than rank orders; see the subsection discussing the cardinal utility approach to overcoming the negative conclusion) are not covered by the theorem. The theorem can also be sidestepped by weakening the notion of independence. Arrow rejected cardinal utility as a meaningful tool for expressing social welfare, and so focused his theorem on preference rankings.

The axiomatic approach Arrow adopted can treat all conceivable rules (that are based on preferences) within one unified framework. In that sense, the approach is qualitatively different from the earlier one in voting theory, in which rules were investigated one by one. One can therefore say that the contemporary paradigm of social choice theory started from this theorem.

**Statement of the theorem:**

The need to aggregate preferences occurs in many disciplines: in welfare economics, where one attempts to find an economic outcome which would be acceptable and stable; in decision theory, where a person has to make a rational choice based on several criteria; and most naturally in voting systems, which are mechanisms for extracting a decision from a multitude of voters' preferences.

The framework for Arrow's theorem assumes that we need to extract a preference order on a given set of options (outcomes). Each individual in the society (or equivalently, each decision criterion) gives a particular order of preferences on the set of outcomes. We are searching for a ranked voting system, called a *social welfare function* (*preference aggregation rule*), which transforms the set of preferences (*profile* of preferences) into a single global societal preference order. The theorem considers the following properties, assumed to be reasonable requirements of a fair voting method:

**1**. **Non-dictatorship**

The social welfare function should account for the wishes of multiple voters. It cannot simply mimic the preferences of a single voter.

2. **Unrestricted domain**(or **universality**) For any set of individual voter preferences, the social welfare function should yield a unique and complete ranking of societal choices. Thus:

- It must do so in a manner that results in a complete ranking of preferences for society.
- It must deterministically provide the same ranking each time voters' preferences are presented the same way.

3. **Independence of irrelevant alternatives (IIA)**

   The social preference between x and y should depend only on the individual preferences between x and y (*Pairwise Independence*). More generally, changes in individuals' rankings of *irrelevant* alternatives (ones outside a certain subset) should have no impact on the societal ranking of the subset. For example, the introduction of a third candidate to a two-candidate election should not affect the outcome of the election unless the third candidate wins.

4. **Positive association of social and individual values** (or **monotonicity**)

   If any individual modifies his or her preference order by promoting a certain option, then the societal preference order should respond only by promoting that same option or not changing, never by placing it lower than before. An individual should not be able to hurt an option by ranking it *higher*.

5. **Non-imposition**(or **citizen sovereignty**)

   Every possible societal preference order should be achievable by some set of individual preference orders. This means that the social welfare function is subjective: It has an unrestricted target space.

Arrow's theorem says that if the decision-making body has at least two members and at least three options to decide among, then it is impossible to design a social welfare function that satisfies all these conditions at once.

A later (1963) version of Arrow's theorem can be obtained by replacing the monotonicity and non-imposition criteria with:

6. **Pareto efficiency**(or **unanimity**)

   If every individual prefers a certain option to another, then so must the resulting societal preference order. This, again, is a demand that the social welfare function will be minimally sensitive to the preference profile.

The later version of this theorem is stronger—has weaker conditions—since monotonicity, non-imposition, and independence of irrelevant alternatives together imply Pareto efficiency, whereas Pareto efficiency and independence of irrelevant alternatives together do not imply monotonicity. (Incidentally, Pareto efficiency on its own implies non-imposition.)

**Remarks on IIA**

The IIA condition can be justified for three reasons :

- normative (irrelevant alternatives should not matter),
- practical (use of minimal information), and
- strategic (providing the right incentives for the truthful revelation of individual preferences). Though the strategic property is conceptually different from IIA, it is closely related.

Arrow's death-of-a-candidate example suggests that the agenda (the set of feasible alternatives) shrinks from, say, X = {a, b, c} to S = {a, b} because of the death of candidate c. This example is misleading since it can give the reader an impression that IIA is a condition involving *two* agenda and *one* profile. The fact is that IIA involves just *one* agendum ({x, y} in case of Pairwise Independence) but *two* profiles. If the condition is applied to this confusing example, it requires this: Suppose an aggregation rule satisfying IIA chooses b from the agenda {a, b} when the profile is given by (cab, cba), that is, individual 1 prefers c to a to b, 2 prefers c to b to a. Then, it must still choose b from {a, b} if the profile were, say, (abc, bac) or (acb, bca) or (acb, cba) or (abc, cba).

**Interpretations of the theorem**

Although Arrow's theorem is a mathematical result, it is often expressed in a non-mathematical way with a statement such as *"No voting method is fair," "Every ranked voting method is flawed,"* or *"The only voting method that isn't flawed is a dictatorship"*. These statements are simplifications of Arrow's result which are not universally considered to be true. What Arrow's theorem does state is that a deterministic preferential voting mechanism—that is, one where a preference order is the only information in a vote, and any possible set of votes gives a unique result—cannot comply with all of the conditions given above simultaneously.

Various theorists have suggested weakening the IIA criterion as a way out of the paradox. Proponents of ranked voting methods contend that the IIA is an unreasonably strong criterion. It is the one breached in most useful voting systems. Advocates of this position point out that failure of the standard IIA criterion is trivially implied by the possibility of cyclic preferences. If voters cast ballots as follows:

- 1 vote for A > B > C
- 1 vote for B > C > A
- 1 vote for C > A > B

Then the pair wise majority preference of the group is that A wins over B, B wins over C, and C wins over A: these yield rock-paper-scissors preferences for any pairwise comparison. In this circumstance, *any* aggregation rule that satisfies the very basic majoritarian requirement that a candidate who receives a majority of votes must win the election, will fail the IIA criterion, if social preference is required to be transitive (or acyclic). To see this, suppose that such a rule satisfies IIA. Since majority preferences are respected, the society prefers A to B (two votes for A>B and one for B>A), B to C, and C to A. Thus a cycle is generated, which contradicts the assumption that social preference is transitive.

So, what Arrow's theorem really shows is that any majority-wins voting system is a non-trivial game, and that game theory should be used to predict the outcome of most voting mechanisms.[9] This could be seen as a discouraging result, because a game need not have efficient equilibria, *e.g.*, a ballot could result in an alternative nobody really wanted in the first place, yet everybody voted for.

## 2.6 GOVERNMENT FAILURES

Government intervention to resolve market failures can also fail to achieve a socially efficient allocation of resources. Government failure is a situation where government intervention in the economy to correct a market failure creates inefficiency and leads to a misallocation of scarce resources.

***Examples of government failure include:***

1. Government can award **subsidies** to firms, but this may protect inefficient firms from competition and create barriers to entry for new firms because prices are kept 'artificially' low. Subsidies, and other assistance, can lead to the problem of **moral hazard**.
2. **Taxes** on goods and services can raise prices artificially and distort the efficient operation of the market. In addition, taxes on incomes can create a disincentive effect and discourage individuals from working hard.
3. **Governments** can also fix prices, such as minimum and maximum prices, but this can create distortions which lead to:

- **Shortages,** which may arise when government fixes price below the market rate. Because public healthcare is provide free at the point of consumption there will be long waiting lists for treatment.
- **Surpluses,** which may arise when government fixes prices above the natural market rate, as supply will exceed demand. For example, guaranteeing farmers a high price encourages over-production and wasteful surpluses. Setting a 'minimum wage' is likely to create an excess of supply of labour in markets where the 'market clearing equilibrium' is less than the minimum.

4. **Information failure** is also an issue for governments, given that government does not necessarily 'know' enough to enable it to make effective decisions about the best way to allocate scarce resources. Many economists believe in the *efficient market hypothesis,* which assumes that the market will always contain more information than any individual or government. The implication is that market prices and market movements should be free from interference because markets cannot be improved upon by individuals or governments.
5. **Excessive bureaucracy** is also a potential government failure. This is caused by the public sector when it tries to solve the principal-agent problem. Government must appoint bureaucrats to ensure that its objectives are pursued by the managers of public sector organisations, such as the NHS.
6. Finally, there is the problem of **moral hazard** associated with the payment of welfare benefits. If individuals know that the state will provide unemployment benefit, or free treatment for their poor health, they are less likely to take steps to improve their employability, or to avoid activities which prevent poor health, such smoking, a poor diet, or lack of exercise.

This occurs when government intervention in the economy causes an inefficient allocation of resources and a decline in economic welfare.

**Reasons for government failure**

- **Lack of incentives:** In the public sector, there is limited or no profit motive. Because workers and managers lack incentives to improve services and cut costs it can lead to inefficiency. For example, the public sector may be more prone to over-staffing. The government may be reluctant to make people redundant because of the political costs associated with unemployment.
- **Poor information:** politicians may have poor information about the type of service to provide. Politicians may not be experts in their department, but concentrate on their political ideology.

- **Political interference**:e.g. politicians may take the short term view rather than considering long term effects
- **Administration cost of government:** bureaucracy in running public services
- **Moral hazard:** The government may offer a guarantee to all bank deposits to protect financial system, but this could encourage banks to take risks – because they know they can be bailed out by the government.

**Overcoming government failure**

There are various things the government can try and do to overcome government failure

- Give performance targets / profit incentives
- Competitive tendering – where public sector bodies face competition from the private sector for the right to run a public service.
- Employing outside private sector consultants to make decisions about how to cut costs.

**Evaluation**

It should be remembered many public services are not subject to the same profit goals. It is difficult to give a profit motive in health or education because the goal is not profit but quality of service.

## 2.7 QUESTIONS

1) Critically explain Arrow Social Welfare Function.
2) Explain the role of Government in decision-making.
3) Explain in detail Arrow Impossibility Theorem.
4) Discuss the role of government in resolving market failure.

# 3

# INTEREST GROUP CAPTURE AND POLITICO ECONOMIC INTERACTIONS

**Unit Structure :**

3.1 Introduction
3.2 Interest Group Capture
3.3 Politico Economic interactions
3.4 Privatization
3.5 Institutional Economics and the state
3.6 Politico Economic modelling in India
3.7 Question

## 3.1 INTRODUCTION

**Interest Group capture or Regulatory capture** is a form of political corruption that occurs when a regulatory agency, created to act in the public interest, instead advances the commercial or political concerns of special interest groups that dominate the industry or sector it is charged with regulating. Regulatory capture is a form of government failure; it creates an opening for firms or political groups to behave in ways injurious to the public (e.g., producing negative externalities). The agencies are called "captured agencies".

## 3.2 THEORY OF INTEREST GROUP CAPTURE

For public choice theorists, regulatory capture occurs because groups or individuals with a high-stakes interest in the outcome of policy or regulatory decisions can be expected to focus their resources and energies in attempting to gain the policy outcomes they prefer, while members of the public, each with only a tiny individual stake in the outcome, will ignore it altogether.Regulatory capture refers to the actions by interest groups when this imbalance of focused resources devoted to a particular policy outcome is successful at "capturing" influence with the staff or commission members of the regulatory agency, so that

the preferred policy outcomes of the special interest groups are implemented.

Regulatory capture theory is a core focus of the branch of public choice referred to as the economics of regulation; economists in this specialty are critical of conceptualizations of governmental regulatory intervention as being motivated to protect public good. Often cited articles include Bernstein (1955), Huntington (1952), Laffont & Tirole (1991), and Levine & Forrence (1990). The theory of regulatory capture is associated with Nobel laureate economist George Stigler, one of its main developers.

Likelihood of regulatory capture is a risk to which an agency is exposed by its very nature. This suggests that a regulatory agency should be protected from outside influence as much as possible. Alternatively, it may be better to not create a given agency at all lest the agency become victim, in which case it may serve its regulated subjects rather than those whom the agency was designed to protect. A captured regulatory agency is often worse than no regulation, because it wields the authority of government. However, increased transparency of the agency may mitigate the effects of capture. Recent evidence suggests that, even in mature democracies with high levels of transparency and media freedom, more extensive and complex regulatory environments are associated with higher levels of corruption (including regulatory capture).

**Relationship with federalism**

There is substantial academic literature suggesting that smaller government units are easier for small, concentrated industries to capture than large ones. For example, a group of states or provinces with a large timber industry might have their legislature and/or their delegation to the national legislature captured by lumber companies. These states or provinces then becomes the voice of the industry, even to the point of blocking national policies that would be preferred by the majority across the whole federation. Moore and Giovinazzo (2012) call this "distortion gap".

The opposite scenario is possible with very large industries, however. Very large and powerful industries (e.g. energy, banking) can capture national governments, and then use that power to block policies at the state or provincial level that the voters may want.

**Economic rationale**

The idea of regulatory capture has an obvious economic basis, in that vested interests in an industry have the greatest financial stake in regulatory activity and are more likely to be motivated to influence the regulatory body than dispersed individual consumers, each of whom has little particular incentive to try to influence regulators. When regulators form expert bodies to examine policy, this invariably features current or former industry members, or at the very least, individuals with contacts in the industry.

Some economists, such as Jon Hanson and his co-authors, argue that the phenomenon extends beyond just political agencies and organizations. Businesses have an incentive to control anything that has power over them, including institutions from the media, academia and popular culture, thus they will try to capture them as well. This phenomenon is called "deep capture".

**There are two basic types of regulatory capture**

- **Materialist capture**, also called *financial capture*, in which the captured regulator's motive is based on its material self-interest. This can result from bribery, political donations, or the regulator's desire to maintain its government funding.
- **Non-materialist capture**, also called *cognitive capture* or *cultural capture*, in which the regulator begins to think like the regulated industry. This can result from interest-group lobbying by the industry.

**Thus,** Regulatory capture is a theory associated with George Stigler, a Nobel laureate economist. It is the process by which regulatory agencies eventually come to be dominated by the very industries they were charged with regulating. Regulatory capture happens when a regulatory agency, formed to act in the public's interest, eventually acts in ways that benefit the industry it is supposed to be regulating, rather than the public.

Public interest agencies that come to be controlled by the industry they were charged with regulating are known as captured agencies. Regulatory capture is an example of gamekeeper turns poacher; in other words, the interests the agency set out to protect are ignored in favour of the regulated industry's interests.

## 3.3 POLITICO ECONOMIC INTERACTIONS

**Political economy** is a term used for studying production and trade, and their relations with law, custom, and government, as well as with the distribution of national income and wealth. *Political economy* originated in moral philosophy. It was developed in the 18th century as the study of the economies of states, or *polities*, hence the term *political* economy.

In the late 19th century, the term *economics* came to replace *political economy*, coinciding with the publication of an influential textbook by Alfred Marshall in 1890.Earlier, William Stanley Jevons, a proponent of mathematical methods applied to the subject, advocated *economics* for brevity and with the hope of the term becoming "the recognised name of a science."

Today, *political economy*, where it is not used as a synonym for economics, may refer to very different things, including Marxian analysis, applied public-choiceapproaches emanating from the Chicago school and the Virginia school, or simply the advice given by economists to the government or public on general economic policyor on specific proposals. A rapidly growing mainstream literature from the 1970s has expanded beyond the model of economic policy in which planners maximize utility of a representative individual toward examining how political forces affect the choice of economic policies, especially as to distributional conflicts and political institutions.It is available as an area of study in certain colleges and universities.

In its contemporary meaning, *political economy* refers to different, but related, approaches to studying economic and related behaviours, ranging from the combination of economics with other fields to the use of different, fundamental assumptions that challenge earlier economic assumptions:

- *Political economy* most commonly refers to interdisciplinary studies drawing upon economics, sociology, and political science in explaining how political institutions, the political environment, and the economic system — capitalist, socialist, or mixed — influence each other. The *Journal of Economic Literature* classification codesassociate political economy with three subareas: the role of government and/or power relationships in resource allocation for each type of economic system,international political economy, which studies the economic impacts of international relations, and economic

models of political processes. The last area, derived from public choice theory and dating from the 1960s, models voters, politicians, and bureaucrats as behaving in mainly self-interested ways, in contrast to a view, ascribed to earlier economists, of government officials trying to maximize individual utilities from some kind of social welfare function. An early and continuing focus of that research program is what came to be called constitutional political economy.

Economists and political scientists often associate political economy with approaches using rational-choice assumptions, especially in game theory, and in examining phenomena beyond economics' standard remit, such as government failure and complex decision making in which context the term "positive political economy" is common. Other "traditional" topics include analysis of such public policy issues as economic regulation, monopoly, rent-seeking, protection, institutional corruption, and distributional politics .Empirical analysis includes the influence of elections on the choice of economic policy, determinants and forecastingmodels of electoral outcomes, the political business cycles, central-bank independence, and the politics of excessive deficits.

A recent focus has been on modeling economic policy and political institutions as to interactions between agents and economic and political institutions, including the seeming discrepancy of economic policy and economist's recommendations through the lens of transaction costs. From the mid-1990s, the field has expanded, in part aided by new cross-national data sets that allow tests of hypotheses on comparative economic systems and institutions. Topics have included the breakup of nations,the origins and rate of change of political institutions in relation to economic growth, development, backwardness, reform, and transition economies, the role of culture, ethnicity, and gender in explaining economic outcomes, macroeconomic policy, the environment, fairness, and the relation of constitutions to economic policy, theoretical and empirical.

New political economy may treat economic ideologies as the phenomenon to explain, per the traditions of Marxian political economy. Thus, Charles S. Maier suggests that a political economy approach "interrogates economic doctrines to disclose their

sociological and political premises.... in sum, [it] regards economic ideas and behavior not as frameworks for analysis, but as beliefs and actions that must themselves be explained." This approach informs Andrew Gamble's *The Free Economy and the Strong State* (Palgrave Macmillan, 1988), and Colin Hay's *The Political Economy of New Labour* (Manchester University Press, 1999). It also informs much work published in *New Political Economy*, an international journal founded by Sheffield University scholars in 1996.

- International political economy (IPE) is an interdisciplinary field comprising approaches to the actions of various actors. In the United States, these approaches are associated with the journal *International Organization*, which in the 1970s became the leading journal of IPE under the editorship of Robert Keohane, Peter J. Katzenstein, and Stephen Krasner. They are also associated with the journal *The Review of International Political Economy*. There also is a more critical school of IPE, inspired by thinkers such as Antonio Gramsci and Karl Polanyi; two major figures are Matthew Watson and Robert W. Cox.
- Anthropologists, sociologists, and geographers use *political economy* in referring to the regimes of politics or economic values that emerge primarily at the level of states or regional governance, but also within smaller social groups and social networks. Because these regimes influence and are influenced by the organization of both social and economic capital, the analysis of dimensions lacking a standard economic value (e.g., the political economy of language, of gender, or of religion) often draws on concepts used in Marxian critiques of capital. Such approaches expand on neo-Marxian scholarship related to development and under development postulated by André Gunder Frank and Immanuel Waller stein.
- Historians have employed *political economy* to explore the ways in the past that persons and groups with common economic interests have used politics to effect changes beneficial to their interests.
- Political Economy and Law is a recent attempt within legal scholarship to engage explicitly with political economy literature. In the 1920s and 30s, legal realists (e.g., Robert Hale) and intellectuals (e.g., John Commons) engaged themes related to political economy. In the second half of the 20th

century, lawyers associated with the Chicago School incorporated certain intellectual traditions from economics. Since the crisis in 2007, however, legal scholars especially related to international law, have turned to more explicitly engage with the debates, methodology and various themes within political economy texts

## 3.4 PRIVATIZATION

**Privatization**, also spelled **privatisation** (in British English), may have several meanings. Primarily, it is the process of transferring **ownership** of a business, enterprise, agency, public service, or public property from the public sector (a government) to the private sector, either to a business that operates for a profit or to a non-profit organization. It may also mean the government outsourcing of services or functions to private firms, e.g. revenue collection, law enforcement, and prison management.

**Privatization** has also been used to describe two unrelated transactions. The first is the buying of all outstanding shares of a publicly traded company by a single entity, making the company privately owned. This is often described as private equity. The second is a demutualization of a mutual organization or cooperative to form a joint-stock company.

**Forms of Privatization**

There are four main methodsof privatization:

1. Share issue privatization (sip) - selling shares on the stock market.
2. Asset sale privatization - selling an entire organization (or part of it) to a strategic investor, usually by auction.
3. Voucher privatization - distributing shares of ownership to all citizens, usually for free or at a very low price.
4. Privatization from below - Start-up of new private businesses in formerly socialist countries.

Choice of sale method is influenced by the capital market, political, and firm-specific factors. SIPs are more likely to be used when capital markets are less developed or under developed and there is lower income inequality. Share issues can broaden and deepen domestic capital markets, boosting liquidity and (potentially) economic growth, but if the capital markets are

insufficiently developed it may be difficult to find enough buyers, and transaction costs (e.g. under-pricing required) may be higher. For this reason, many governments elect for listings in the more developed and liquid markets, for example Euronext, and the London, New York and Hong Kong stock exchanges.

As a result of higher political and currency risk deterring foreign investors, asset sales occur more commonly in developing countries.

Voucher privatization has mainly occurred in the transition economies of Central and Eastern Europe, such as Russia, Poland, the Czech Republic, and Slovakia. Additionally, Privatization from below is/has been an important type of economic growth in transition economies.

A substantial benefit of share or asset-sale privatizations is that bidders compete to offer the highest price, creating income for the state in addition to tax revenues. Voucher privatizations, on the other hand, could be a genuine transfer of assets to the general population, creating a real sense of participation and inclusion. If the transfer of vouchers is permitted, a market in vouchers could be created, with companies offering to pay money for them.

In economic theory, privatization has been studied in the field of contract theory. When contracts are complete, institutions such as (private or public) property are difficult to explain, since every desired incentive structure can be achieved with sufficiently complex contractual arrangements, regardless of the institutional structure (all that matters is who are the decision makers and what is their available information). In contrast, when contracts are incomplete, institutions matter. A leading application of the incomplete contract paradigm in the context of privatization is the model by Hart, Shleifer, and Vishny (1997). In their model, a manager can make investments to increase quality (but they may also increase costs) and investments to decrease costs (but they may also reduce quality). It turns out that it depends on the particular situation whether private ownership or public ownership is desirable. The Hart-Shleifer-Vishny model has been further developed in various directions, e.g. to allow for mixed public-private ownership and endogenous assignments of the investment tasks.

## 3.5 INSTITUTIONAL ECONOMICS AND THE STATE

**Institutional economics** focuses on understanding the role of the evolutionary process and the role of institutions in shaping economic behaviour. Its original focus lay in Thorstein Veblen's instinct-oriented dichotomy between technology on the one side and the "ceremonial" sphere of society on the other. Its name and core elements trace back to a 1919 *American Economic Review* article by Walton H. Hamilton.

Institutional economics emphasizes a broader study of institutions and views markets as a result of the complex interaction of these various institutions (e.g. individuals, firms, states, social norms). The earlier tradition continues today as a leading heterodox approach to economics. A significant variant is the new institutional economics from the later 20th century, which integrates later developments of neoclassical economics into the analysis. Law and economics has been a major theme since the publication of the *Legal Foundations of Capitalism* by John R. Commons in 1924. Since then, there is heated debate on the role of law (formal institution) on economic growth, Behavioural economics is another hallmark of institutional economics based on what is known about psychology and cognitive science, rather than simple assumptions of economic behaviour.

Institutional economics focuses on learning, bounded rationality, and evolution (rather than assume stable preferences, rationality and equilibrium). It was a central part of American economics in the first part of the 20th century, including such famous but diverse economists as Thorstein Veblen, Wesley Mitchell, and Commons. Some institutionalists see Karl Marx as belonging to the institutionalist tradition, because he described capitalism as a historically-bounded social system; other institutionalist economistsdisagree with Marx's definition of capitalism, instead seeing defining features such as markets, money and the private ownership of production as indeed evolving over time, but as a result of the purposive actions of individuals.

"Traditional" institutionalism rejects the *reduction* of institutions to simply tastes, technology, and nature. Tastes, along with expectations of the future, habits, and motivations, not only determine the nature of institutions but are limited and shaped by them. If people live and work in institutions on a regular basis, it shapes their world-views. Fundamentally, this traditional

institutionalism (and its modern counterpart institutionalist political economy) emphasizes the legal foundations of an economy and the evolutionary, habituated, and volitional processes by which institutions are erected and then changed.

The vacillations of institutions are necessarily a result of the very incentives created by such institutions, and are thus endogenous. Emphatically, traditional institutionalism is in many ways a response to the current economic orthodoxy; its reintroduction in the form of institutionalist political economy is thus an explicit challenge to neoclassical economics, since it is based on the fundamental premise that neo-classicists oppose: that economics cannot be separated from the political and social system within which it is embedded.

## 3.6 POLITICO ECONOMIC MODELLING IN INDIA

Macroeconomic modelling is generally motivated by two objectives:forecasting and more significantly, policy analysis. In pursuit of both these objectives,every model must ideally satisfy four criteria. *First* and foremost, it must fit into atheoretical framework. *Second*, the actual specification of the model must reflect aclear understanding of the contextual framework within which policies are formulatedand executed along with an envisaged process of adjustment. *Third*, it is essential thatthe model is built on a firm and rich data base and, *finally*, the estimated structuralmodel must adequately utilise the rigors and sophistications of econometric methodology.

Unfortunately this is a tall order which can seldom be met. Typicallyrefinements in one direction can often be achieved only at the cost of those in someother direction. For instance, it may be possible to devise small models that aretheoretically neat and manageable enough to be subjected to econometric refinementsthese would seldom be able to deal with actual policy issues in a meaningful way.Clearly, an operationally useful model has to go well beyond simple illustrativecaricatures of the economic system. How far one may go will depend on the nature ofcompromise between competing requirements. Given his objectives, ingenuity of themodel builder lies in his ability to hammer out the optimal compromise.

Experience shows that models that can deal with policy issues need to beeclectic rather than exclusively pure in their

structure. Since these have to beconsiderably disaggregative imposing a uniform mode of adjustment across marketsand sectors may be unrealistic. Similarly, the ground realities may not be strictlyconsistent with one single paradigm over time and across markets. Moreover, there isalways the difficulty posed by the non-availability of reliable data on top of the factthat certain phenomena may not even be quantifiable. This is not to argue for theabandonment of theoretical considerations. Far from it, models without a clearly speltout analytical frame are useless because results based on such models can never beinterpreted. A model is first and foremost, an assertion of a process of adjustmentrather than an unstructured description of the course of economic movements. Theplea is only for the necessity to depart from pure and prototype textbook models.

## EVOLUTION OF MACRO-MODELLING IN INDIA

Macro-econometric modelling in India has had one of the longest historiesamongst all countries, particularly those in the developing world. While it is not ourintention to go into this history a few observations on the nature of this work would bein order. Nearly all macro-econometric models for India have had a policy focus,sometimes sharp sometimes hazy and, sometimes well formulated and sometimes notso. Most of the models have had only short to medium run character. With varyingemphasis and success models have been concerned with the level of economicactivity, price behaviour, fiscal and monetary policies, inter-sectoral linkages,investment, saving and consumption, resource mobilisation and public sector capitalformation, trade flows and balance of payments. Each of these has posed seriousproblems of analytical significance, many of these remained unresolved even today.

Broadly speaking the sequence of available models can be seen as belongingto four phases. In the first we have a set of exercises during the late fifties and thesixties which were highly aggregative, simple and exploratory; almost all of themhaving been undertaken as doctoral dissertations. In fact it was these that paved theway for studies belonging to the second phase most of which were also undertaken asdoctoral dissertations. But these were somewhat disaggregative and better focussedon policy issues. The third phase has ushered in models, which were under taken independently, and many of these built on earlier experience by the same author.

These were, as expected, much more disaggregative, with a clearly improved policy content and focus. The current fourth phase has ushered in much larger models thatare comparable to those in developed economies and are maintained on an on-goingbasis.

Needless to point out here that modelling an economy has to be a continuouson-going activity not merely because of the need for forecasting but also moreimportantly because it is only a live model that can (a) incorporate new informationby way of data (b) reflect changes in the perception of contemporary economic issues(c) reflect, as far as possible, new developments in theory and in quantitativemethodology. One disturbing aspect however, of macro-modelling in India has beenthat each model turned out to be a one-time exercise. Thus, despite a relatively earlystart, unlike all developed and many developing countries, India did not have amaintained macro-econometric model till very recently. The only macro-model of thistype during the eighties, built by the National Council of Applied Economic Research(NCAER) with support from the Ministry of Finance has largely been of the CGEvariety. Only a few parts of the model are econometrically estimated. Though some models have been built and maintained by the Reserve Bank of India and some othergovernment institutions these have been used only as in-house enterprises. Neithertheir structure nor any results based on these have been publicly discussed. It is onlysince the early nineties that sustained on-going work on a macro-econometric modelbegan jointly at the Institute of Economic Growth (IEG) and the Delhi School of Economics (DSE). The structure of the model has been discussed at various for a and results based on it frequently presented. It is gratifying that some more models havenow come into existence as an on-going activity.

Given the existing state of the art and the need for its continuation, further workon econometric modelling must move in the following directions. *First*, this should be an on-going activity. Models need to be frequently revised and updated in order toremain useful. To identify gaps and limitations of any models they should also beused frequently for policy analysis and forecasting. This has seldom been the case sofar. *Second*, for an in-depth understanding of the functioning of the economy andmeaningful policy modelling effort must go into developing sub models for specificsectors in a way that these can be used on their own as also be able to serve ascomponents of a larger system. Apart from agriculture, industry and some othersectors which have received

attention so far, the harder job of modelling labour andcapital markets needs to be taken up as far as data permit. How to deal with theinformal components in each sector of the economy is a harder problem but one thatmust figure in the future agenda.

*Third*, it is time that an attempt is made to utilise the higher frequency database. A beginning must be made with quarterly models. For specific segments of theeconomy monthly data too can be usefully utilised. This work can proceed on aparallel basis along with that on annual models. *Fourth*, modelling work must nowmake use of the recent developments in econometrics to make the methodology morerigorous. Time series analysis would be the highest priority in this context because itwill considerably enrich these models in terms of both methodology as well as thefinal outcome. A greater effort must be made in refining modelling work in the lightof the available theory. One item that needs to be taken up seriously is the wayexpectations are handled and built into models, wherever relevant and feasible.Clearly, price and exchange rate formation and financial sub models can be taken upright away in this context.

## 3.7 QUESTION

1) Discuss the Theory of Interest group capture.
2) Write a note on Politico - Economic interactions.
3) Explain the concept of privatization.

# 4

# MODULE - 2

# PUBLIC EXPENDITURE : RATIONALE AND EVALUATION PUBLIC GOODS- I

**Unit Structure:**


4.0 Objectives
4.1 Introduction
4.2 Nash - Cournot Equilibrium and Pareto Efficiency
  4.2.1 Nash Equilibrium
  4.2.2 Pareto Optimality
  4.2.3 Public Goods
4.3 Vernon Smith's Auction Mechanism
4.4 Cycling
4.5 The Median Voter Theorem
4.6 Summary
4.7 Questions


## 4.0 OBJECTIVES

After having studied this unit, you should be able -

- To understand the Nash-Cournot Equilibrium.
- To explain Pareto Optimality.
- To know Public Goods.
- To understand Vernon Smith's Auction Mechanism.
- To know the median Voter Theorem.

## 4.1 INTRODUCTION :

This unit will be helpful to you to develop your understanding about public goods, Nash-Cournot Equilibrium and Pareto Efficiency. Vernon Smith's Auction Mechanism examine the procedure of charging an appropriate tax (charges) for public goods. The Median Voter Theorem is an explanation of voter's preference.

## 4.2 PUBLIC GOODS - NASH - COURNOT EQUILIBRIUM AND PARETO EFFICIENCY

### 4.2.1 Nash Equilibrium:

Consider the case where the case with $N=2$ agents, indexed by $i=1,2$. Most of what we consider here is generalizable for larger N but working with 2 agents makes things much easier. Let agent 1's utility depends on his own action $a_1$ ("action" is defined very broadly here) as well as agent 2's action, so we can write $U_1(a_1,a_2)$, and similarly for agent $2\,U_2(a_1,a_2)$.

**Definition :**

A set of actions $(a_1^N, a_2^N)$ constitutes a Nash equilibrium if

$U_1(a_1^N, a_2^N) \geq U_1(a_1, a_2^N)$ for all $a_1$, &

$U_2(a_1^N, a_2^N) \geq U_1(a_1^N, a_2)$ for all $a_2$.

In other words a set of actions is a Nash equilibrium if each agent cannot do better for herself playing her Nash equilibrium action given other people play their Nash equilibrium action.

**Solving for Nash Equilibria:**

Solving the Nash equilibrium requires solving two maximization problems, namely.

$$\max_{a_1} U_1(a_1,a_2) \; and \; \max_{a_2} U_2(a_1,a_2)$$

Where each person takes each other action as given. Oftentimes finding a Nash involves checking all the possible combinations $(a_1,a_2)$ and aksing yourself "is this a Nash equilibrium?" Sometimes it is possible to eliminate dominated actions iteratively (see a book on game theory) to narrow the cases that need to be checked. However, assuming everything is nicely differentiate and $a_1^N$ and $a_2^N$ are both positive, we can take first order conditions. The first order condition far each first agent is just.

$$\frac{\partial U_1(a_1^N, a_2^N)}{\partial a_1} = 0 \text{ and } \frac{\partial U_2(a_1^N, a_2^N)}{\partial a_2} = 0$$

(Nash FOC)

Which is a system of 2 equations in 2 unknowns $a_1^N, a_2^N$, and so usually a little algebra will yield the solution.

**Reaction Curves:**

By the implicit function theorem the FOC for agent 1 defines what she will play given $a_2$ (not just at the Nash), i.e. agent 1's reaction curve $a_1 = r_1(a_2)$ so that $\frac{\partial U_1(r_1(a_2), a_2)}{\partial a_1} = 0$. A similar reaction curve $r_2(a_1)$ can be defined for agent 2. A Nash equilibrium can be seen as where $a_1^N = r_1(a_2^N)$ and $a_2^N = r_2(a_1^N)$.

This is where the reaction curves cross in a graph with $a_1$ on one axis and $a_2$ on the other.

**Strategic Complements and Substitutes:**

It is useful to know how one agent will react if the other agent changes her action. Differentiating totally the expression $\frac{\partial U_1(r_1(a_2), a_2)}{\partial a_1} = 0$ with respect to $a_2$ we get.

$$\frac{d}{da_2}\left[\frac{\partial U_1(r_1(a_2), a_2)}{\partial a_1}\right] = \frac{\partial^2 U_1}{\partial a_1^2}\frac{dr_1}{da_2} + \frac{\partial^2 U_1}{\partial a_2 \partial a_1} = 0$$ and so solving the slope of the reaction curve.

$$\frac{dr_1}{da_2} = -\left(\frac{\partial^2 U_1}{\partial a_1^2}\right)^{-1}\frac{\partial^2 U_1}{\partial a_2 \partial a_1}$$

The sign of this expression depends on the sign of the second derivatives of the utility function. Cases where $\frac{dr_1}{da_2} > 0$, where a greater action by 2 elicits more of a response by 1, identifies a situation which $a_1$ and $a_2$ are called strategic complements. The alternate case where $\frac{dr_1}{da_2} < 0$ is where $a_1$ and $a_2$ are called strategic substitutes.

**4.2.2 Pareto Optimality:**

**Definition:**

The set of feasible actions $(a_1^P, a_2^P)$ is Pareto optimal if there does not exist another of feasible action $(a_1, a_2)$ such that

$$U_1(a_1, a_2) \geq U_1(a_1^P, a_2^P)$$
$$U_2(a_1, a_2) \geq U_2(a_1^P, a_2^P)$$

With at least one above inequality strict. In other words there does not exist an allocation that makes both as well off and making one strictly better off. A logically equivalent condition is that for any feasible set of actions $(a_1, a_2)$

$$U_1(a_1, a_2) \geq U_1\left(a_1^P, a_2^P\right) \Rightarrow U_2(a_1, a_2) < U_2\left(a_1^P, a_2^P\right)$$

A set of actions that makes agent 1 strictly better of must make agent 2 strictly worse off.

Important Note: Except for the trivial case of one person, pareto optima and Nash equilibria do not necessarily coincide; plenty of Nash equilibria that are not Pareto optima and vice-versa (remember the Prisoner's Dilemma!)

**Solving for pareto Optima**

Consider a social planner who attaches a relative weight $\lambda$ to agent 1 relative to agent 2 where $\lambda \geq 1$ depending whether the planner values agent 1 more or less than agent 2. A theorem from mathematics says that "pretty much" any Pareto optimal allocation can be found by maximizing the weighted utilities.

$$\max_{a_1, a_2} \lambda U_1(a_1, a_2) + U_2(a_1, a_2)$$

For some $\lambda$. Different $\lambda$ will give different Pareto optimal allocations. A popular favourite is to choose $\lambda$=1, which corresponds to the utilitarian social welfare function. Assuming everything is smooth and the Pareto optimal actions are positive the following FOC must hold at $\left(a_1^P, a_2^P\right)$

$$\lambda \frac{\partial U_1}{\partial a_1} + \frac{\partial U_2}{\partial a_1} = 0 \;\&\; \lambda \frac{\partial U_1}{\partial a_2} + \frac{\partial U_2}{\partial a_2} = 0$$

(Pareto FOC)

Compare this condition to the Nash FOC and you can see that the Pareto optimal actions take into account $\partial U_2 / \partial a_1$ and $\partial U_1 / \partial a_2$, i.e., that actions of agent 1 have an effect on agent 2 and vice-versa. These externalities are ignored in the Nash equilibrium and so the Nash equilibrium is only optimal if $\partial U_2 / \partial a_1 = \partial U_1 / \partial a_2 = 0$. Solving each FOC equation for $-\lambda$ and rearranging we see

$$-\lambda = \frac{\frac{\partial U_2}{\partial a_1}}{\frac{\partial U_1}{\partial a_1}} = \frac{\frac{\partial U_2}{\partial a_2}}{\frac{\partial U_1}{\partial a_2}} \Rightarrow \frac{\frac{\partial U_1}{\partial a_2}}{\frac{\partial U_1}{\partial a_1}} = \frac{\frac{\partial U_2}{\partial a_2}}{\frac{\partial U_2}{\partial a_1}}$$

So at the marginal rates of substitution between each action for each agent are equal, i.e. $MRS^1_{a_1a_2} = MRS^2_{a_1a_2}$. At the Nash equilibrium the marginal rates of substitution are typically perpendicular as $MRS^1_{a_1a_2} = \infty \; and \; MRS^1_{a_1a_2} = 0$.

**Utility Possibility Set:**

One can imagine the set of all pairs of utility $(U_1, U_2)$ given by all of the different action $a_1$ and $a_2$. The utility possibility set is that collection.

$$\mu = \left\{(U_1, U_2) : U_1 = U_1(a_1, a_2), U_2(a_1, a_2) \; for \; any \; feasible \; a_1, a_2\right\}$$

which can usually be represented by a graph with $U_1$ on the x-axis and $U_2$ on the y-axis.

By its very nature a Pareto optimum should be on the very edge of that set - that is its "frontier." More formally the utility possibility frontier is the set.

$$\mu_F = \left\{(U_1, U_2) \in \mu : there \; is \; no \left(\widetilde{U}_1, \widetilde{U}_2\right) \in \mu \; such \; that \; U_1 \geq U_1 \; and \; \widetilde{U}_2 \geq \widetilde{U}_2\right\}$$

The difference between the utility possibility frontier and the set of Pareto optima, is that the set of Pareto optima refers to an outcome or allocation while the frontier refers only to utilities. Also, Pareto optima require that at least one inequality is strict. All pareto optima will yield utilities on the frontier, however not quite all points on the frontier will relate to a Pareto optimum since it may contain points where one agent (not both) may do better without it costing the other agent.

Say we are at a Pareto optimum. This means that the objective function is given by $\lambda U_1^P + U_2^P$ where $U_i^P = U_i\left(a_1^P, a_2^P\right)$. Just around the optimum $\left(U_1^P, U_2^P\right)$ we can assume that the sum $\lambda U_1^P + U_2^P = U$ is constant. Using the implicit function theorem again we can treat $U_2^P$ as a function of $U_1^P$ and differentiate $\lambda + \frac{dU_2^P}{dU_1^P} = 0$ which gives us the slope of the utility possibility set

$\frac{dU_2^P}{dU_1^P} - \lambda$. Thus we can imagines a social planner with straight, parallel indifference curves, each with slope -$\lambda$, in a graph. A Pareto optimum will be found where an indifference curve is tangent to the utility possibility frontier, with slope $\frac{dU_2^P}{dU_1^P}$, outlining $\mu$.

**Minimum Utility Formulation:**

If you don't like the idea of pulling $\lambda$ out of a hat, consider an alternate formulation where agent 1 is guaranteed a minimum amount of utility $\overline{\mu_1}$, and agent 2 has her utility maximized. In other words.

$$\max_{a_1, a_2} U_2(a_1, a_2) \; s.t. U_1(a_1, a_2) \geq \overline{\mu_1}$$

If we let $\lambda$ be the Lagrange multiplier on the constraint to get.

$$U_2(a_1, a_2) + \lambda\left[U_1(a_1, a_2) - \overline{\mu_1}\right]$$

Then we get the same FOC as the Pareto FOC (it's the same problem!) except that now $\lambda$ has to be solved for rather than imposed. The constraint $U_1(a_1, a_2) = \overline{\mu_1}$ adds a third equation so that we can solve for all three $\left(a_1^P, a_2^P, \lambda\right)$.

**4.2.3 Public Goods:**

Each agent has utility $U_i(G, x_i)$ where $xi$ is private consumption and public good $G = \sum_i^N = 1\, gi$ where $gi$ is agent I's provision of the public good. The public good, by definition is non rival, consumption by one agent does not reduce it's benefit to another agent, and nonexcludable, i.e., it is prohibitively expensive to keep agents from consuming it. Assume that total consumption $X = \sum_i^N = 1\, x_i$ is produced via a production function F from the public good, where the total amount of public good available is G, so $X = F(G - G)$ with $F(O) = O, F^1(\cdot) > OF^{11}(\cdot) < O,$ and so the marginal rate of transformation of public good into private good $MRS_{GX} = \frac{dx}{DG} = F^1(G - G)$

**Pareto Optimal Provision:**

Back to the case where N=2 then we have $x_1 + x_2 = F(G - G)$ or $x_2 = F(F - G) - x_1$. Then we can write for utility for the individuals as $U_1(x_1, G)$ and $U_2(F(G-G) - x_1, G)$. As we saw above we solve for the Pareto optimum by solving.

$$\max_{x_1, G} \lambda U_1(G, x_1) + U_2(G, F(G-G) - x_1)$$

Assuming $x_1^P, G^P > 0$ then the following two first order conditions must be satisfied at the optimum $(x_1^P, x_2^P, G^P)$.

$$x_1 : \lambda \frac{\partial U_1}{\partial x} - \frac{\partial U_2}{\partial x} = 0$$
$$G : \lambda \frac{\partial U_1}{\partial G} - \left[ \frac{\partial U_2}{\partial G} + \frac{\partial U_2}{\partial x} F^1 \right] = 0$$

Solving each equation for $\lambda$ and then solving for $F^1$ tells us that

$$\lambda = \frac{\frac{\partial U_2}{\partial x}}{\frac{\partial U_1}{\partial x}} = \frac{\frac{\partial U_2}{\partial G} + \frac{\partial U_2}{\partial x} F^1}{\frac{\partial U_1}{\partial G}} \Rightarrow F^1 = \frac{\frac{\partial U_1}{\partial G}}{\frac{\partial U_1}{\partial x}} + \frac{\frac{\partial U_2}{\partial x}}{\frac{\partial U_2}{\partial x}}$$

which is the condition that $MRT_{GX} = MRS_{GX}^1 + MRS_{GX}^2$, this is the "Samuclson Rule" that the marginal rate of transformation should equal the sum of the marginal rates of substitution. In the case of constant returns to scale where $F^1 = PG$ where PG can effectively be considered the price of G in terms of x, then $MRS_{GX}^1 + MRS_{GX}^2 = PG$

**Reaction Curve and Nash Equilibrium:**

To ease the notational burden and a few other issue we'll consider the case where $F^1$ is constant at PG=1. Each individual has a budget constraint $x_i + g_i = M_i$, This constraint implies that there is really only one independent solution. Here we let that be $g_i$ and $x_i = \mu_i - g_i$. we an even redefine utility to depend on each person's action $U_1(g_1, g_2) = U_2(g_1 + g_2, M_1 - g_1)$ & $U_2(g_1, g_2) = U_2(g_1 + g_2, M_2 - g_2)$ to fit it into the previous framework.

The reaction curve $r_1(g_2, M_1)$ of the first agent, which depends on $g_2$ as well as personal income $M_1$ is determined by the FOC evaluated at $\left(r_1(g_2, M_1) + g_2, M_1 - r_1(g_2, M_1)\right)$ is

$$\frac{\partial U_1}{\partial G} - \frac{\partial U_1}{\partial x} \leq 0$$

Where of course equality holds it $r_1(g_2, M_1) > 0$. Assuming that both x and G are normal goods, then, a little effort shows $0 \leq \partial r_1 / \partial M_1 \leq 1, \& -1 \leq a r_1 / \partial g_2 \leq 0$ which means that $g_1 \& g_2$ are strategic substitutes: i.e. for each unit of G agent 2 gives, agent 1 will reduce her contribution of G, albeit less than one -for -one. The possibility $r_1(g_2, M_1) = 0$ is more than a triviality for higher values of $g_2$ and lower values of $M_1$. If the solutions from FOC equation is negative, then this means $r_1(g_2, M_1) = 0$.

Assuming, the FOC, holds with equality this implies $MRS_{GX} = \frac{\partial U_1}{\partial G} / \frac{\partial U_1}{\partial x} = 1$. A similar condition holds for agent 2 so that if both contribute $MRS^1_{GX} + MRS^2_{GX} = 2 > 1 = MRT_{GX}$ and hence that the Nash equilibrium is not optimal. The Nash provision is too small. $G^N = g_1^N + g_2^N < G^P$

## 4.3 VERNON SMITH'S AUCTION MECHANISM :

Whereas most of the tatonnement type procedures ask voters to state either a willingness to pay tax price or a desired quantity, Smith's mechanisms requires the voter to announce both, Each individual I announces both a bid, $b_i$, which is the share of the public goods cost that I is willing to cover and a proposed quantity of the public good, ???. The tax price actually charged I is the difference between the public goods costs, C, and the aggregate bids of the other n - 1 voters, $B_i$, that is,

$$t_i G = (c - B_i) G, \qquad (5.1)$$

Where $B_i = \sum_j \neq i b_j$ and $G = \sum K^n = 1 G_k / n$

The procedure selects 0 quantity of public good only when each voter's bid matches his tax price and each voters proposed public good quantity equals the mean:

$b_i = t_i$ and $G_i = G$, for all *j*. (5.2)

After each iteration of the procedure, voters are told what their tax prices and the public good quantity would have been had

(5.2) been achieved at that iteration. If a voter's bid fell short of his tax price he can adjust either his bid or proposed public good quantity to try to bring about equilibrium. Only when all unanimously agree to both their tax price and the public good quantity does the procedure stop.

At an equilibrium (5.2)is satisfied, and is utility can be written as

$$V_i = U_i(G) - t_i G, \qquad (5.3)$$

Where the utility from consuming G is expressed in money units. Maximizing (5.3) with respect to $G_i$ we obtain the condition for is optimal proposed quantity for the public good.

$$dV_i / dG_i = U_i^1 / n - t_i / n = 0$$

$$U_i = t_i$$

Each voter equater his marginal utility from the public good to his tax price summing (5.4) over all voters, we obtain

$$\sum_{i=1}^{n} U_i = \sum_{i=1}^{n} t_i = \sum_{i=1}^{n} \left(C - B_i\right)_i = C \qquad (5.5)$$

Equation (5.4) and (4.5) define the condition for the Lindahl equilibrium.

The auction mechanism induces individuals to reveal their preferences for the public good by charging each. Voter a tax based not on his stated preference for the public good, but on the aggregate of all other stated preference (bids). In this respect, it resembles the demand revealing procedures. Each voter must be willing to make up the difference between the public good's costs at the aggregate bids of the other voters for the good to be provided. The ultimate incentive to state one's preferences honestly is provided by the knowledge that the good will not be provided unless all unanimously agree to a single quantity and set of tax prices.

**Criticisms of the unanimity rule:**

The unanimity rule is the only voting rule certain to lead to pareto-preferred public good quantities and tax shares, a feature that led Wicksell (1896) and later Buchanan and Tullock (1962) to endorse it. Two main criticisms have been made against it. First, a groping search for a point on the contract curve might take considerable time, particularly in a large community of heterogeneous tastes. The loss in time by members of the community in discovering a set of pareto-optimal tax shares might outweigh the gains to those who are saved from paying a tax share

exceeding their benefits from the public good. An individuals who was uncertain over whether he would be so “exploited”. Under a less than unanimity rule might easily prefer such a rule rather than spend the time required to attain full unanimity. The second objection against a unanimity rule is that it encourages strategic behavior. If a knows the maximum share of taxes, B will assume rather than go without the public good, A can force B to point C on the contract curve, by voting against all tax shares greater than $t_c$. All gains from providing the public good then accrue to A. If B behaves the same, the final outcome is dependent on the bargaining strengths of the two individuals. The same is true of the other equilibrium along the contract curve. Bargaining can further delay the attainment of the agreement as each players has to “test” the others willingness to make concessions.

The “bargaining problem” under the unanimity rule is the mirror image of the “incentive problem” in the voluntary provision of a public good. The latter is a direct consequence of the joint supply. Nonexclusion properties of a public good. Given these properties, each individual has an incentive to understate his preferences and free ride, since the quantity of public good provided is largely independent of his single message. The literature on voluntary preference revelation procedures has by and large sidestepped this problem by assuming honest preference revelation in spite of the incentives to be dishonest. The strongest analytic result to justify this assumption has been that sincere message transmittal is a minimax strategy that, sincere revelation of preferences maximizes the minimum payoff that an individual can obtain. But a higher payoff might be obtained through a misrepresentation of preferences, and some individuals can be expected to pursue this more daring option. If to remove this incentive one compels all citizens to vote in favor of a public good quantity tax share proposal before it is provided, the free-rider problem does disappear. Each individuals vote is now essential to the public good’s provision. This reversal in the individual’s position in the collective decision alters his strategic options. Where an individual might, under a voluntary revelation scheme, gamble on the rest of the group providing an acceptable quantity of the public good without his contributing, under the unanimity rule he might gamble on the group’s reducing the size of his contribution rather than risk his continual blocking of the collective outcome. Although the strategy options differ, both solution to the public good problem are potentially vulnerable to strategic behaviour.

Recent experimental results of Hoffman and Spitzer (1986) and Smith indicate that strategic bargaining on the part of individuals in unanimity rule situations may not be much of a problem the Hoffman-Spitzer experiments were designed to see whether the ability of individuals to achieve Pareto-optimal

allocations in coase-type externality situation deteriorates as the number of affected parties increases. Since all affected parties has to agree to a bargain before it could be implemented, the experiments essentially tested whether strategic bargaining by individuals would over-turn pareto-optimal allocation proposals under the unanimity rule. Hoffman and Spitzer found that "if anything efficiency improved with larger groups" (with groups as large as 20 on a side).

Unanimous agreement is required on the final iteration under Smith's auction mechanism, as already noted. Experiments with small numbers of voters were characterized by fairly rapid convergence on the Lindahl equilibrium, with bids also falling near the Lindahl tax prices. Strategic misrepresentation of preferences was not observed.

Even if strategic behavior does not thwart or indefinitely delay the achievement of a unanimous collective decision, one might object to the unanimity rule on the grounds that the outcome obtained depends on the bargaining abilities and risk preferences of the individuals. Such a criticism implicity contains the normative judgment that the proper distribution of the gains from cooperation should not be distributed according to the willingness to bear risks. One can easily counter that they should. An individual who votes against a given tax share to secure a lower one risks, under a unanimity rule, not having the good provide at all or if so in a less than optimum quantity voting in this manner expresses a low preference for the public good, in much the same way as voting against the tax share does, because it is "truly" greater than the expected benefits. Some one not willing to vote strategically might be said to value the public good higher and therefore perhaps ought to be charged a higher price for it.

We are clearly in the realm of normative economics here, as we were in comparing points E and L above, and need criteria as to how the gains from cooperation ought to be shared. Indeed, in a full evaluation of the unanimity rule its normative properties must be considered. Wicksell's advocacy of the unanimity rule was based on its normative properties. The unanimity rule would protect individuals from being coerced by other members of community, he argued. Wicksell used "coerced" not in the sense employed by Breton, who took it to mean having a different evaluation of the public good at the margin from one's tax price, but in the sense of being coerced through a collective decision to pay more for a public good than its benefits are in to. This argument for the unanimity rule stems directly from Wicksell's view of the collective choice process as one of mutually beneficial voluntary exchange among individuals, as is Buchanan and Tullock's. This emphasis on the "voluntary exchange" nature of collective choice underlies the

classic essays by both Wicksell and Lindahl and forms an intellectual bond between them, leading in Wicksells case to the unanimity principle, in Lindahl's to a set of tax prices equal to each individual's marginal evaluation of the public good. It also explains the reference to "just" taxation in the titles of each of their essays. We shall return to these issues in chapter 6

**The Optimal Majority:**

When a less than unanimous majority is sufficient to pass on issue, the possibility exists that some individuals will be made worse off via the committee's decision; Wicksell's coercion of the minority can take place. If the issue is of the public good - prisoner's dilemma variety, and there exist reformulation of the issue that could secure unanimous approval, the use of less than unanimity rule can be said to impose a cost on those made worse off by the issue's passage, a cost that could be avoided through the expenditure of the additional time and effort required to redefine the issue so that its passage benefits all. This cost is the difference in utility levels actually secured, and those that would have been secured under a full unanimity rule. Buchanan and Tullock were the first to discuss these costs and refer to them as the "external costs" of the decision rule.

Were there no cost associated with the unanimity rule itself, it would obviously be the optimal rule, since it minimized these external decision costs. But, the time required to define an issue in such a way as to benefit all may be considerable. In addition to attempting to find a formulation of the proposal benefiting all, time may be required to explain the nature of the benefits of the proposal to some citizens unfamiliar with its merits on top of these costs must be added the time lost through the strategic maneuvering that might take place as individuals jockey for more favorable positions along the contract curve, as described earliest.

Most observes, including those most favorably disposed toward the unanimity rule like Wicksell and Buchanan and Tullock, have considered these latter costs sufficiently large to warrant abandoning this rule. If all need not agree to a committee decision, what percentage should agree? The above consideration suggest a trade off between the external costs of having an issue pass against which the individual is opposed, and the cost of time lost through decision making. At the one pole stands unanimity, under which any individual can block any agreement until he has one with which is satisfied, or which he feels is the best he can obtain. The external decision costs under this rule are zero, but the decision time costs may be infinite. At the other extreme, each individual decides the issue alone. No delays may occur as with a pure private good decision, but the external costs of allowing each

individual to decide unilaterally for the community are again potentially infinitely large.

These various possibilities are depicted in figure 4.1, which is taken from Buchanan and Tullock. The cost of a particular collective decision are presented along the vertical axis; the number of people 0 up to N, the committee size, required to pass the issue are presented along the horizontal axis. Curve C is the external cost function representing the expected loss of utility from the victory of a decision to which an individual is opposed under the committee decision rule. Curve D depicts the decision time costs of achieving the required majority to pass the issue as a function of the size of the required majority. The optimal majority is the percentage of the committee at which these two sets of costs are together minimized. This occur at k, where the vertical addition of the two curve reaches a minimum. The optimal majority to pass the issue, given these cost curves, is K/N. At this percentage the expected gain in utility from redefining a bill to gain one more. Supporter just equals the expected loss in time from doing so.

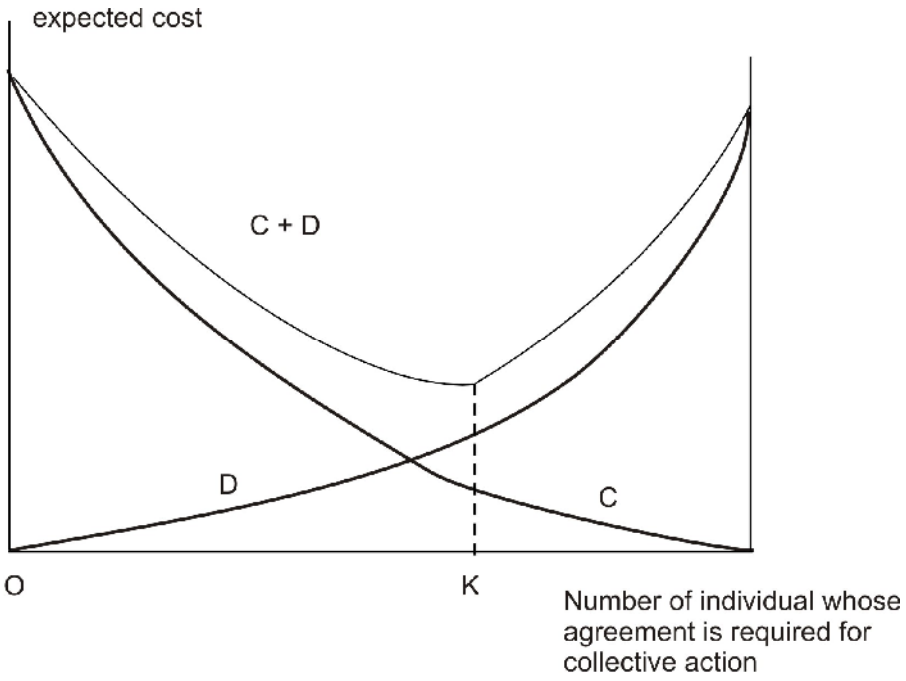



**Figure 4.1 choosing the optimal majority**

Since these costs are likely to differ from issues to issue, one does not expect one voting rule to be optimal for all issues. The external costs will vary depending upon both the nature of the issued to be decided and the characteristics of the community deciding them. Ceteris paribus, when opinions differ widely, or information is scarce, large amounts of time may be required to reach a consensus, and if the likely costs to opposing citizens are not too high relatively small percentage of the community might be

required to make a decision. Again, the extreme example here is the pure private good. In contrast, issues for which large losses can occur are likely to require higher majorities (e.g. issued pertaining to the Bill of Rights). The larger the community, the greater the number of individuals with similar tastes, and thus the easier it is likely to be to achieve a consensus among a given absolute number of individuals. Thus, an increase in N should shift the curve is rightward and downward. But, the fall in costs of achieving a consensus among a given number is unlikely to be fully proportional to the rise in community size. Thus, for issued of a similar type, the optimal percentage of the community required to pass an issue K/N is likely to decreases as the community increases in size.

Individuals whose tastes differ widely from most others in the community can be expected to favor more inclusive majority rules. Individuals with high opportunity costs of time should favor less inclusive majority rules. Buchanan and Tullock assume that the choice of the optimal majority for each category of issued is made in a constitutional setting in which each individual is uncertain over his future position, tastes, and soon. Therefore, each views the problem in the same way, and a unanimous agreement is achieved as to which less than unanimity rule to use for which set of issues. When such a consensus does not exist, the knolty question that must be faced is what majority should be required to decide what majorities are required on all other issues? Having new faced this question, we shall move on.

**A simple majority as the optimal majority:**

The method of majority rule requires that at least the first whole integer above N/2 support an issue before it becomes the committee decision. Nothing we have said so far can indicate why K/N = N/2 should be the optimal majority for the bulk of a committee's decisions; and yet it is. As Buchanan and Tullock note, for any one rule, such as the majority rule, to be the optimal majority for a wide class of decisions, there must exist some sort of a kink in one of the cost functions at the point N/2, causing the sum of two curves to obtain a minimum in a substantial proportion of the cases at this point. A possible explanation for a kink in the decision making cost curve, D at N/2 can be obtained by considering further the internal dynamics of the committee decision process. When less than half of a committees membership is sufficient to pass an issue, the possibility exists for both the issue A and the issue's converse $(\sim A)$ to pass. Thus, a proposal to increase school expenditures by 10 percent might first achieve a winning majority cof, say, 40 percent and a counterproposal to cut expenditures by 5 percent may also receive a winning majority. The committee could, when less than half of the voters suffice to carry an issue, become

deadlocked in an endless series of offsetting proposals absorbing the time and patience of its members. The method of simple majority rule has the smallest possible required majority to pass on issue, which avoids the possibility of self-contradictory issued simultaneously passing.

In figure 4.2 decision costs and external costs curves have been drawn such that their minimum would lie to the left of N/2 were D to continue to decline as it moves leftward from N/2. But the D curve is higher to the left of N/2 owing to the extra decision costs of having conflicting issues pass. This portion of the D curve has been drawn as a straight line, but it could conceivably be U or inverted U shaped to the left of N/2. The discontinuity at N/2 makes this majority the optimal majority for this committee.

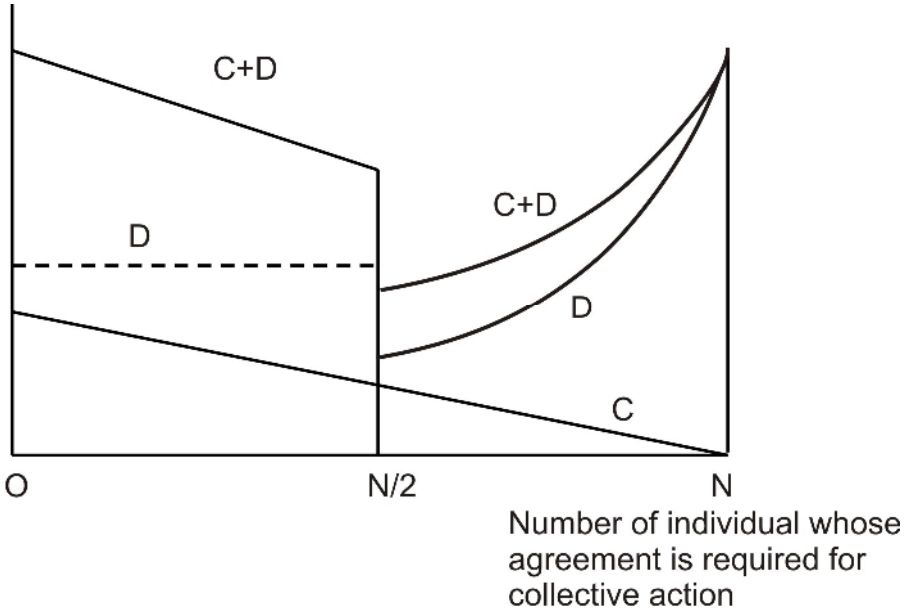



**Figure 4.2 conditions favoring a simple majority as the optimal majority**

Absent a discontinuity, a minimum for C + D occurs to the left of N/2 only when the D curve rises more rapidly as it moves to the right than C does moving to the left; that is, decision costs vary much more over the range of committee sizes than do the external costs of collective decision making. N/2 is the optimal majority for the committee because of the discontinuity in the D curve. Thus, the choice of N/2 as the optimal majority is driven by the shape of the D curve. The method of simple majority rule will be selected as the committee decision rule by a committee whose members place a relatively high value on the opportunity costs of time. Were it not for the loss of time involved in having conflicting proposals like A and $\sim A$ pass, the minimal cost majority for the committee would be less than 0.50. The simple majority is optimal because it is the

smallest majority one can select and still avoid having conflicting proposals both obtain wining majorities.

Speed is not the majority rules only property, however. So important is the simple majority rule as a voting procedure that we shall devote most of the next two chapters to discussing its other properties.

**Majority rule and redistribution:**

As indicated earlier, a committee concerned one with providing public goods and correcting for externalities might nevertheless choose as its voting rule the simple majority rule, if it placed enough weight on solving time. But speed is not the only property that majority rule possesses. Indeed once issues can pass with less than unanimous agreement the distinction between allocative efficiency and redistribution becomes blured. Some individuals are inevitably worse off under the chosen outcome than they would be were some other outcome selected, and there is in effect a redistribution from those who are worse off because the issue has passed to those who are better off.

To see this point more clearly, consider Figure 4.1. The ordinal utilities of two groups of voters, the rich and the poor, as depicted on the vertical and horizontal axes. All of the members of both groups are assumed to have identical preference functions. In the absence of the provision of any public good, representative individuals from each group experience utility levels represented by S and T. The point of initial endowment on the pareto-possibility frontier with only private good production is E. The provision of the public good can by assumption improve the utilities of both individuals. Its provision thus expands the pareto-possibility frontier out to the curve XYZW. The segment YZ corresponds to the contract curve in figure 4.3 CC' under the unanimity rule, both groups of individual must be better off under the provision of he public good for them to vote for it. So the outcome under the unanimity rule must be a quantity of public good and tax share combination leaving both groups somewhere in the YZ segment along the pareto possibility frontier.

But there is no reason to expect the outcome to all in this rang under majority rule. A coalition of the committee members can benefit by redefining the issue to increase their benefits at the expense of the noncoalition members, say, by shifting the tax shares to favor the coalition members. If the each were in the majority they could be expected to couple the public good proposal with a sufficiently regressive tax package so that the outcome wound up in the XY segment. If the poor were in the majority, the taxes would be sufficiently progressive to produce an outcome in ZW. Given the opportunity to redefine the issue proposed through

the alteration of either the quantity of the public good provided, the tax shares, or both, one can expect with certain by that the outcome of the collective choice process will fall outside of the Pareto-preferred segment YZ (Davis, 1970) As long as the issue could be continually redefined in such a way that a majority still benefited, it would pass, and a stable majority coalition could, in principle, push a minority back as far along the Pareto-possibility frontier as their consciences or the constitution allowed.

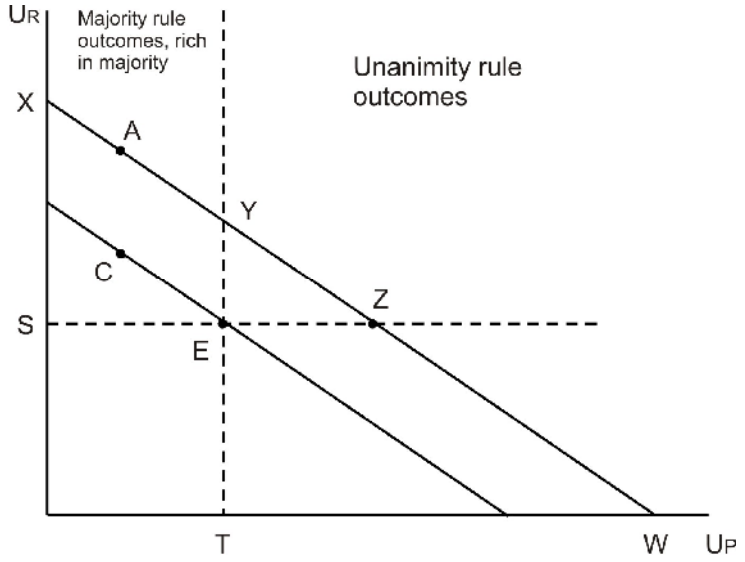



**Figure 4.3**

## 4.4 CYCLING

Given that majority rule must induce some element of redistribution into the collective decision process, we take up next an attribute of majority rule when a pure redistribution decision is to be made. Consider a three-man committee that must decide hew to divide a gift of $100 among them using majority rule. This is a pure distributional issue, a simple zero-sum game. Suppose that $V_2$ and $V_3$ first vote to divide the $100 between themselves, 60/40. $V_1$ now has much to gain from forming a winning coalition. He might propose to $V_3$ that they split the $100, 50/50. this is more attractive to $V_3$ and we can expect this coalition to from a winning coalition. He might new offer $V_1$ a 55/45 split forming a new coalition, and so on. When the issued proposed involve redistribution of income and wealth, members of a losing coalition always have a large incentive to attempt to become members of the winning coalition, even at the cost of less than equal share.

The outcome of a 50/50 split of $100 among a pair of voters is a von Neumann-Morgenstern solution to this particular game. This game has three such solutions, however, and there is no way to predict which of these three if any, would occur. Thus, the

potential for cycles, when issues involve redistribution, seems quite large at is always possible to redefine on issue so as to benefit one or more members and harm some others. New winning coalitions containing some members of the previously losing coalition and excluding members of the previously winning coalition are always feasible. But, as we have seen from the discussion of majority rule, when issues can be amended in the committee, any pure allocative efficiency decision can be converted into a combination of a redistribution and on allocative efficiency change via amendment. That it would seem that when committees are free to amend the issued proposed, cycles must be an eve -present danger.

**Table 4.1**

| Voters | Issues |
|---|---|
| | X Y Z X |
| 1 | > > < |
| 2 | > < > |
| 3 | < > > |
| Community | > > > |

The possibility that majority rule can lead to cycles across issues was recognized over two hundred years ago by the Marquis de condorect, C.L. Dodgson analysed the problem a new one hundred years later, and it has been a major concern of the modern public choice literature beginning with Duncan Black and Kenneth Arrow. Consider the following three voters with preferences over three issues, as in Table 4.1. X can defeat Y, Y can defeat Z, and Z can defeat X. Pairwise voting can lead to an endless cycle. The majority rule can select no winner nonarbitrarily.

If we define Z as a payoff to voters $V_2$ and $V_3$ of 60/140, Y as the payoff (20, 0, 50), and X as (55, 45, 0), the ordinal rankings of issues in figure 4.3 corresponds to the above zero sum pure distribution game. But it is also possible to get ordering as in table 4.1 and figure 4.3 for issues involving allocational efficiency. If X, Y and Z are sequentially higher expenditure on a public good, then the preferences of voters 1 and 3 can be said to be single peaked in the public good - utility space. Voters 2's preferences are doubled - peaked, however, and herein is a cause of the cycle. Change 2's preferences so that they are single peaked, and the cycle disappears.

One of the early important theorems in public choice was proof that majority rule produces an equilibrium outcome when voter preferences are single-peaked. If voter's preferences can be depicted along a single dimension, as with an expenditure issue, this equilibrium lies at the peak preference for the median voter.

Figure 4.4 depicts the single peaked preference for five voters. Voters 3,4 and 5 favor m over any proposal to supply less. Voters 3, 2 and 1 favor it over proposals to supply more. The preference of the median voter decides.

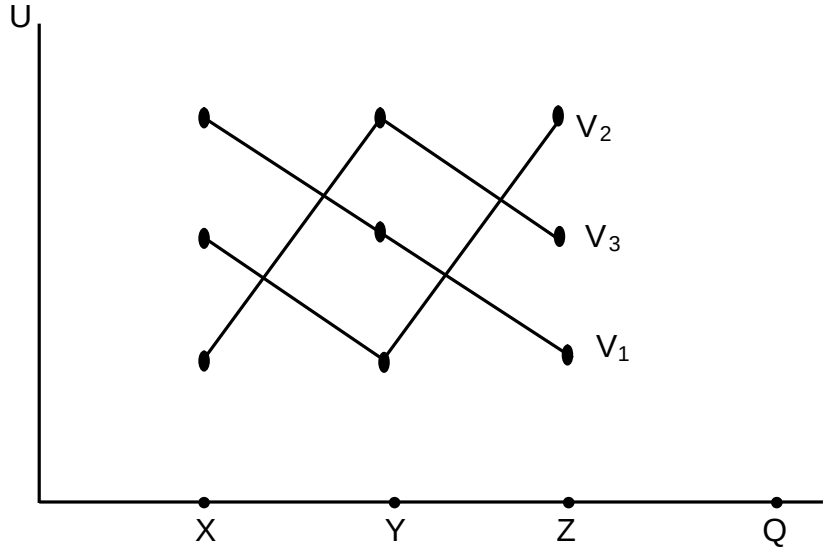



**Figure 4.4 voter preferences that induce a cycle**

## 4.5 THE MEDIAN VOTER THEOREM - ONE - DIMENSIONAL ISSUES:

The proof follows Enclow and Hinich (1984) The two key assumptions for the median voter theorem are (1) that issues are defined along a single dimensional vector x and (2) that each voter's preferences are single peaked in that one dimension. Let voter i's preferences be represented by a utility function $U_i$ ( ) defined over x, $U_i$. Let $x_i$ be voter i's most preferred point along the x vector. Cal $x_i^*$ i's ideal point.

Definition : $X_i^*$ is $i's$ ideal point if and only if (iff) $\left(X_i^*\right) > U_i(x)$ for all $X \neq X_i^*$.

Definition: Let *y* and *z* be two points along the *x* dimension, such that either $y, z \geq X_i^*$ or $y, z \leq X_i^*$. Then voter $i's$ preferences are single peaked iff $\left[U_i(y) > U_i(z)\right] \leftrightarrow \left[\left|y - X_i^*\right| < \left|z - X_i^*\right|\right]$

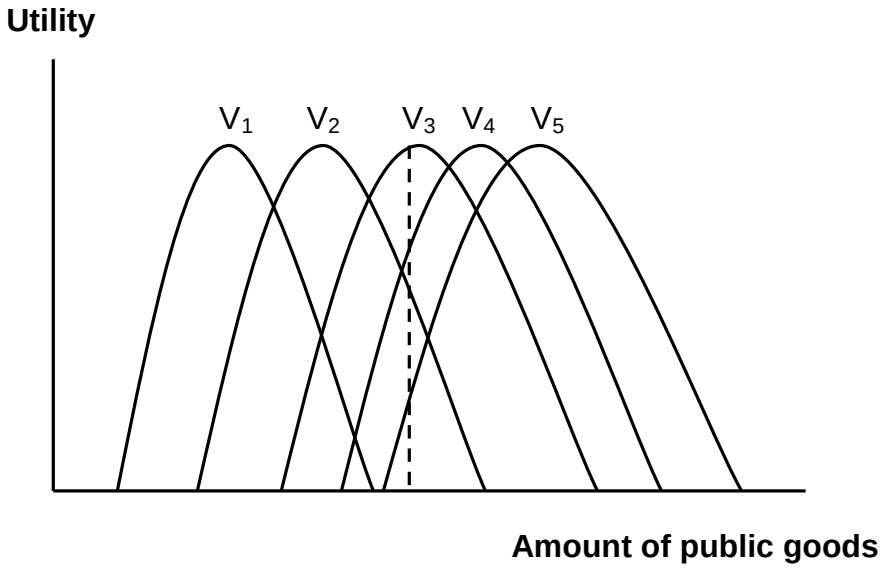



**Figure 4.4 The median voter decides**

In other words, the definition of single peaked preferences says that if y and are two points on the same side of $X_i^*$, then i prefers y to z if and only if y is closer to $X_i^*$ than z. if all preferences are single peaked, then preferences like those of voter 2 in figure 4.3 cannot occur (Note Z is z's ideal point in this figure.)

Definition : Let $\{X_1^*, X_2^*, ..., X_n^*\}$ be the *n* ideal points for a committee of *n* individuals. Let $N_R$ be the number of $X_i^* \geq X_m$, and $N_L$ be the number of $X_i^* \leq X_m$. Then $X_M$ is a median position iff $N_R \geq n/2$ and $N_L \geq n/2$.

Theorem : If *x* is a single -dimensional issue, and all voters have single peaked preferences defined over X, then $X_M$, the median position cannot lose under majority rule.

Proof : Consider any $Z \neq x_m$, say $z < x_m$. Let $R_m$ be the number of ideal points to the right of $X_M$. By definition of single peaked preferences, all $R_m$ voters with ideal points to the right of $X_M$ prefer $X_M$ to z.By definition of median position, $R_m \geq n/2$. Thus, the number of voters preferring $X_M$ to z is at least $R_m \geq n/2$ $X_M$ cannot lose to z under majority rule. Similarly, one can show that $X_M$ cannot lose to any $z > x_m$.

**Alternative Voting Procedure:**

**i) The procedures compared - condorcet efficiency.**

Thus array of procedures is already lengthy and we could easily add to the list the though these cover the most frequently discussed procedures. Each has a certain intuitive appeal. How can one decide which is best?

There are several criteria for defining “best”. First, we might define the economic equivalents to each procedure, as we did with majority rule in earlier chapter and compare the procedures on the basis of their axiomatic properties. These axioms are often rather abstract, however, and thus it may be somewhat difficult to declare procedure. A superior to B just by looking at its momatic properties, we might declare one property most important, and compare the procedures on the basis of their ability to realize this property. The literature has proceeded in both way’s, and we shall discuss the procedure in both ways.

The first of the axioms may (1952) requires of a voting procedure is that it is decisive; that is, it must pick a winner majority rule satisfies this criterion then there are but two candidates, a restriction may imposed on the problem choosing from a pair of alternatives is, however, the simplest choice one can conceptualize, and all of the above procedures select the same winner when m 2 interesting cases involve $m \geq 3$ with m > 2 no candidate may receive a majority of first place votes, and no candidate may defeat all others in pair use contests. Thus, when m>2, bnoth majority rule and the condorect criterion may declare no candidate a winner. Each of the other procedures will make a winner. Thus, for those who, on the basis of the arguments of earlier chapter feel the majority rule ought to be the community’s decision rule, interest in the other procedures arises only when m>2.

**Table 4.2**

| $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|
| X | X | Y | Z | W |
| Y | Y | Z | Y | Y |
| Z | Z | W | W | Z |
| W | W | X | X | X |

**Table 4.3**

| $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|
| X | X | X | Y | Y |
| Y | Y | Y | Z | Z |
| Z | Z | Z | X | W |

Although the other procedures always pick a winner, even when a condorcet winner does not exist, they do not always choose the condorect winner when one does exist. Table 4.2 presents a set of preference ordering for voters in which x is the winner under the plurality rule, although y is condorect winner. Since a single vote for one's most preferred candidate is possible strategy choice for voters under approval voting, x might also win under this procedure with the preferences ordering table 4.1.

In Table 4.3 X is the condorect winner, while Y would be the winner they the Borda count. In Table 4.4 is again the condorect winner while issue wins under the Hare system. Under each of the procedures other than majority rule, a winner may be chosen who is not the condorect winner even when the latter exists.

If one finds the properties of majority rule most attractive, then failure to select the condorect winner when one exists may be regarded as a serious deficiency of a procedure. One way to evaluate the different procedure is to compute the percentages of the time that a condorect winner exists and selected by a given procedure. Merrill has made these percent age calculations and named them condorect efficiencies, that is, the efficiency of a procedure in actually selecting the condorect winner when one exists Table 4.5 reports the result from simulations of an electorate of 25 voters with randomly allowcated utility functions and various number of candidates.

**Table 4.4**

| $V_1$ | $V_2$ | $V_3$ | $V_4$ | $V_5$ |
|---|---|---|---|---|
| Y | W | X | Y | W |
| X | Z | Z | Z | X |
| Z | X | W | X | Z |
| W | Y | Y | W | Y |

**Table 4.5 condorect efficiency for a random society (25 voters)**

| | Number of candidates | | | | |
|---|---|---|---|---|---|
| Voting system | 3 | 4 | 5 | 7 | 10 |
| Runoff | 96.2 | 90.1 | 83.6 | 79.5 | 61.3 |
| Plurality | 79.1 | 69.4 | 62.1 | 52.0 | 42.6 |
| Hare | 96.2 | 92.7 | 89.1 | 84.8 | 77.9 |
| Coombs | 96.3 | 93.4 | 90.2 | 86.1 | 81.1 |
| Approval | 76.0 | 69.8 | 67.1 | 63.7 | 61.3 |
| Borda | 90.8 | 87.3 | 86.2 | 85.3 | 84.3 |
| Social utility maximize | 84.4 | 80.2 | 77.9 | 77.2 | 77.8 |

The first six rows report the condorcet efficiencies for six of the procedures defined in section A. voters are assumed to maximize expected utility under approval voting by voting for all candidates whose utilities exceed the mean of the candidates for the voter. With 2 candidates, all procedures choose the condorect winner with efficiency of 100. The efficiency of all procedures is under 100 percent with 3 candidates. The biggest declines an efficiency in going from 2 to 3 candidates are for the plurality and approval voting procedures. When the number of candidates is as large as 10, the six procedures divide into three groups based on their condorcet efficiency indexes : the Hare, coombs, and Borda procedures all achieve about 80 percent efficiency majority rule with one runoff and approval voting achieve about 60 percent efficiency; and the plurality rule selects the condorect winner only percent of the time.

It is implausible to assume that an electorate would go to the polls nine separate times, as would be required under either the Hare or coombs systems with 10 candidates. Therefore, if either of these procedures were actually used, as a practical matter one would undoubtedly simple ask voters to down their complete rankings of the candidates, and use a computer to determine a winner following the prescribed rule. Thus, the informational requirements of the Hare, coombs, and Borda, procedures are identical; they differ only in how they process this information. Given that they rely on the same information sets, it is perhaps not surprising that they perform about the same.

Of the six procedures listed in Table 4.5, the runoff and plurality procedures are the only ones in common use today. Thus, another way to look the result of Table 4.5 is to calculate the gain in condorcet effiency abandoning the plurality or runoff rule in favor of one of the other for procedures. The biggest gains obviously come in going to the Hare, coombs or Borda procedure, particularly if the number of candidates exceed. But much more information is demanded of the voter at the election approval voting might then be

compared with the runoff and plurality systems a relatively simple procedure with condorcet efficiency properties that excess those of the plurality rule and approach those of the runoff system as the number of candidates expands. An important advantage of approval voting over the majority rule runoff procedure is that approval voting require the voters go to the polls only once.

## 4.6 SUMMARY

- A set of actions is a Nash equilibrium if each agent can not do better for herself playing her Nash equilibrium action given other people play their Nash equilibrium action.
- How to impose an appropriate tax on a public goods is explained by Veron Smith's Auction Mechanism.
- Voter's Preferences are summarized by the median voter Theorem.

## 4.7 QUESTIONS

1) Examine the Nash-Crunot equilibrium and Pareto efficiency.
2) Write a note on Veron Smith's Auction Mechanism.
3) What is Cycling? Explain.
4) Explain in detail the Median Voter Theorem.

# 5

# PUBLIC GOODS - II

**Unit Structure :**

5.0 Objectives
5.1 Introduction
5.2 Public Goods
5.2.1 Private Provision of Public Goods
5.2.2 Public Provision of Public Goods
5.3 Merit Goods
5.4 Public Goods and Prisoner's Dilema
5.5 Public Goods and Chicken's Dilema
5.6 Voluntary Provision of Public Goods with Constant Returns to Scale
5.7 Voluntary Provision of Public Goods with varying Supply Technologies
5.8 The Coase Theorem
5.9 Summary
5.10 Questions

## 5.0 OBJECTIVES

After having studied this unit, you should be able -

- To know the concept of Public Goods.
- To explain the private and Public Provision of public goods.
- To understand the concept of merit goods.
- To know the Coase Theorem.

## 5.1 INTRODUCTION :

A public good is a good or service that can be consumed simultaneously by everyone and from which no one can be excluded. The present unit is devoted to various aspects of public and merit goods. For example, the private and public provision of public goods, Prisoner's dilemma and public goods, chicken's dilemma and public good etc. The Coase Theorem argues against the government's intervention in the form of taxes and subsidies.

## 5.2 PUBLIC GOODS:

A public good is a good or service that can be consumed simultaneously by everyone and from which no one can be excluded. The first feature of a public good is called nonrivalry. A good is nonrival if the consumption by one person does not decrease the consumption by another person. An example is watching a television show. The opposite of nonrival is rival. A good is rival if the consumption by one person decreases the consumption by another person. An example is eating a hotdog.

The second feature of a public good is that it is nonexcludable. A good is nonexcludable if it is imposible, or extremely costly, to prevent someone from benefiting from a good. An example is national defense. It would be difficult to exclude someone from being defended. The opposite of nonexcludable is excludable. A good excludable if it is possible to prevent a person from enjoying the benefits of a good. An example is cable television. Cable companies can ensure that only those people who have paid the fee receive programs.

Figures ----- classifies goods according to these two criteria and gives examples of goods in each category. National defense is a pure public good. One person's consumption of the security provided by our national defense system does not decrease the security of someone lese - defense is nonrival. And the military cannot select those whom it will protect and those whom it will leave exposed to threats - defense is nonexcludable.

Many goods have a public element but are not pure public goods. An example is highway. A highway is norrival until it becomes congested. One more car on a highway with plenty of space does not reduce anyone else's consumption of transportation services.

| | **Pure private goods** | **Excludable and nonrival** |
|---|---|---|
| Excludable | Food | Cable television |
| | Car | Bridge |
| | House | Highway |

| | **Nonexcludable and rival** | **Pure public good** |
|---|---|---|
| Non Excludable | Fish in the ocean | Light house |
| | Air | National defense |

Rival Nonrival

**public Goods and Private Goods**

But once the highway becomes congested, one extra vehicle lowers the quality of the service available to everyone else, it becomes rival like a private good. Also, users can be excluded from a highway by tollgates. Another example is fish in the ocean. Ocean fish are rival because a fish taken by one person not available for anyone else. Ocean fish are also nonexcludable because it is difficult to prevent people from catching them.

**The Free - Rider Problem :**

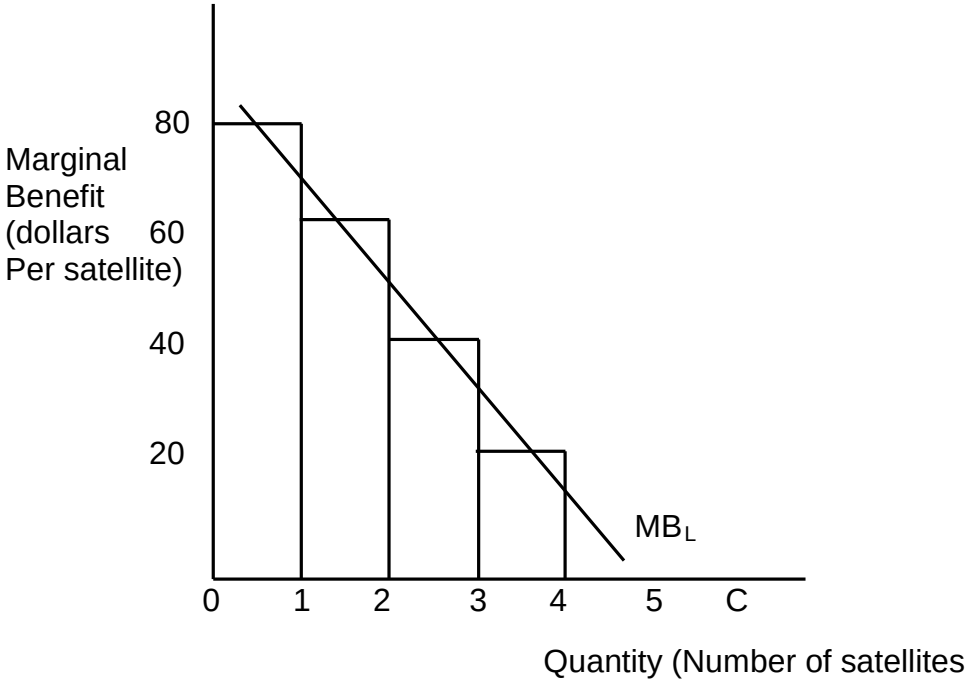
Public goods create a free-rider problem. A free rider is a person who consumes a good without paying for it. Public good create a free-rider problem because the quantity of the good that a person is able to consume is not influenced by the amount the person pays for the good. So no one has an incentive to pay for a public good. Let's look more closely at the free-rider problem by studying an example.

**The Benefit of a Public Good:**

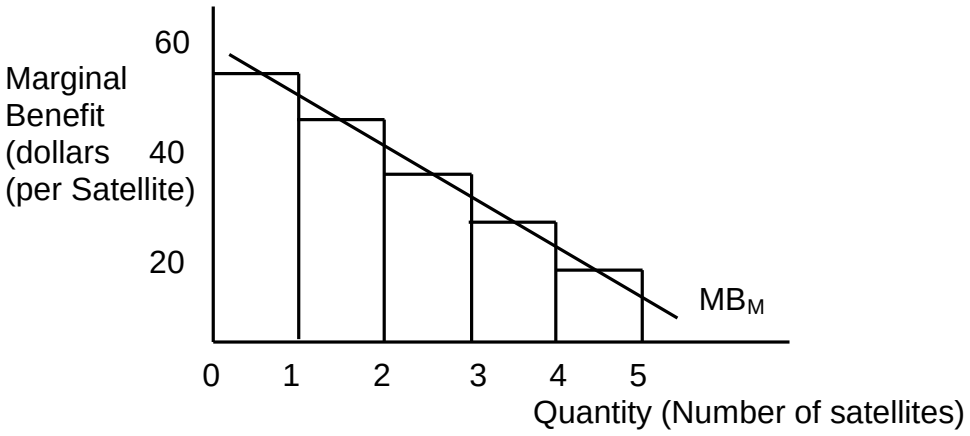
Suppose that for its defense, a country must launch some surveillance satellites. The benefit provided by a satellite is the value of its services. The value of a private good is the maximum amount that a person is willing to pay for one more units, which is shown by the person's demand curve. The value of a public good is the maximum amount that all the people are willing to pay for one more unit of it.

To calculate the value placed on a public good, we use the concepts of total benefit and marginal benefit. Total benefit is the dollar value that a person places on a given level of provision of a public good. The greater the quantity of a public good, the larger is a person's total benefit. Marginal benefit is the increase in total benefit that results from a one-unit increase in the quantity of a public good.

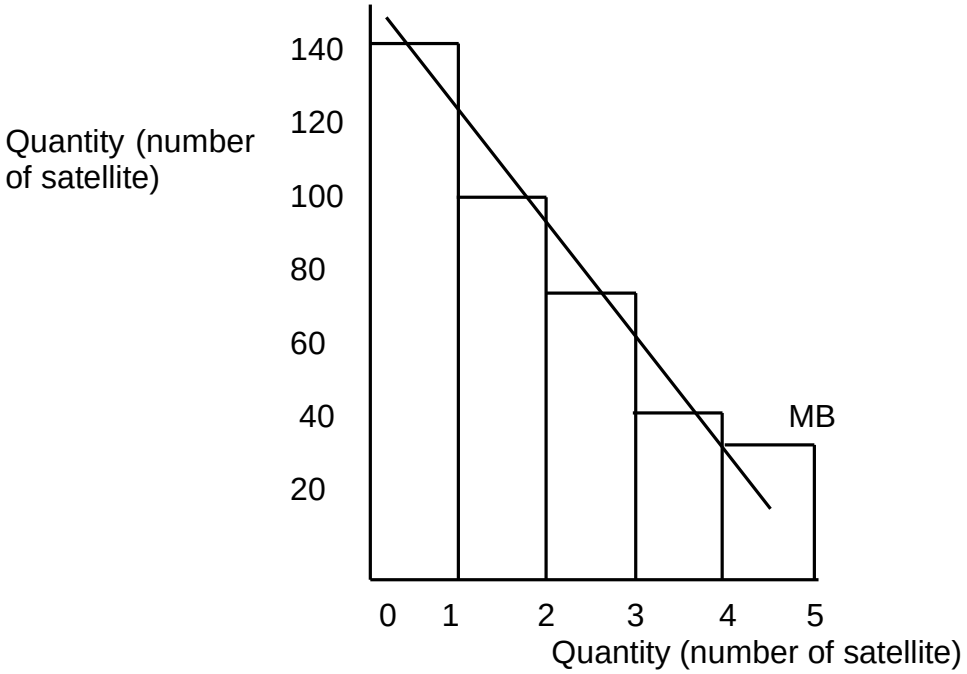
Figure ----- shows the marginal benefit that arises from defense satellites for a society with just two members, Lisa and max. Lisa's and max's marginal benefits are graphed as $MB_L$ and $MB_m$ respectively, in parts (a) and (b) of the figures. The marginal benefit from a public good is similar to the marginal benefit from a private good, its magnitude diminishes as the quantity of the good increases, for Lisa, the marginal benefit from the first satellite is $80, and from the second it is $60. By the time 4 satellites are deployed, Liso's marginal benefit is zero. Fro max, the marginal benefit from the first satellite is $50, and from the second it is $40. By the time 4 satellites are deployed, max perceives only $10 worth of marginal benefit.




**a) Lisa's marginal benefit.**




**(b) Max's marginal benefit**




**(c) Economy's Marginal benefit**

**Figure 5.1**

Part (C) shows the economy's marginal benefit curve, MB. An individual's marginal benefit curve for a public good is similar to the individual's demand curve for a private good. But the economy's marginal benefit curve for a public good is different from the market demand curve for a private good. To obtain the market demand curve for a private good, we sum the quantities demanded by all individuals at each price we sum the individual demand curves horizontally. But to find the economy's marginal benefit curve of a public good, we sum the marginal benefits of each individual at each quantity we sum the individual marginal benefit curves vertically. The resulting marginal benefit for the economy made up of Lisa and max is the economy's marginal benefit curve graphed in part (c) the curve MB. Lisa's marginal benefit from the first satellite gets added to Max's marginal benefit from the first satellite because they both enjoy security from the first satellite.

**The Efficient Quantity of a Public Good:**

An economy with two people would not any satellites because the total benefit falls for short of the cost. But an economy with 250 million people might. To determine the efficient quantity we need to take the cost as well as the benefit into account.

The cost of a satellite is based an technology and the prices of the resources used to produce it (just like the cost of producing sweaters Figure ---- sets out the benefits and costs. The second and third columns of the table show the total and marginal benefits.

The next two columns show the total and marginal cost of producing satellites. The final column shows not benefit. Total benefit, TB, and total cost, TC, are graphed in part (d) of the figure.

Figure ----- The Efficient Quantity of a Public Good

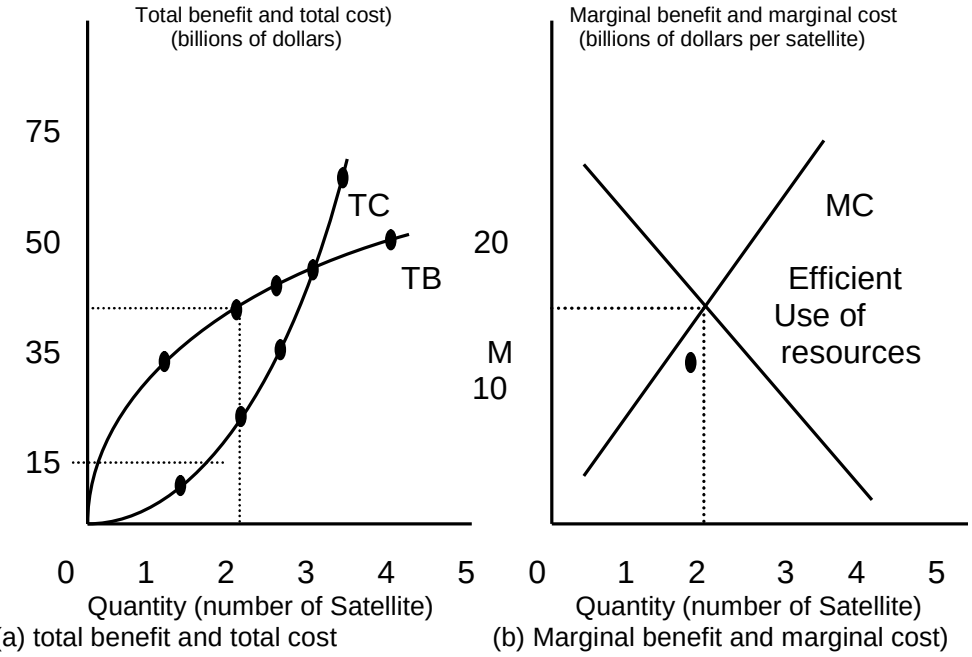



**Figure 5.2**

| Quantity number of satellites | Total benefit (billions of dollar) | Marginal benefit (billions of dollars per satellite) | Total cost (billions of dollar) | Marginal cost (billions of dollars per satellite | Net benefit (billions of dollars) |
|---|---|---|---|---|---|
| 0 | 0 | 20 | 0 | 5 | 0 |
| 1 | 20 | 15 | 5 | 10 | 15 |
| 2 | 35 | 10 | 15 | 15 | 20 |
| 3 | 45 | 5 | 30 | 20 | 15 |
| 4 | 50 | 0 | 50 | 25 | 0 |
| 5 | 50 | | 75 | | -25 |

The efficient quantity is the one that maximizes net benefit - total benefit minus total cost- and occurs when 2 satellites are provided.

The fundamental principles of marginal analysis that you have used to explain how consumers maximize utility and how firms maximize profit can also be used to calculate the efficient scale of provision of a public good. Figure ---- (b) shows this alternative approach. The marginal benefit curve is MB, and the marginal cost curve is MC. When marginal benefit exceeds

marginal cost, net benefit increases if the quantity produced increases. When marginal cost exceeds marginal benefit, net benefit increases if the quantity produced decreases. Marginal benefit equals marginal cost with 2 satellites. So making marginal cost equal to marginal benefit maximizes net benefit and uses resources efficiently.

### 5.2.1 Private Provision:

We have now worked out the quantity of satellites that maximizes net benefit. Would a private firm North Pole. Protection, Inc. - deliver that quantity? It would not. To do so, it would have to collect $15 billion to cover its costs - or $60 from each of the 250 million people in the economy. But no one would have an incentive to buy his or her share of the satellite system. Everyone would reason as follows : The number of satellites provided by North Pole protection, Inc. is not affected by my $60. But my own private consumption is greater if I free ride and do not pay my share of the cost of the satellite system. If I do not pay, I enjoy the same level of security and I can buy more private goods. Therefore I will spend my $60 on other goods and free rise on the public good. This is the free-rider problem.

If everyone reasons the same way, North Pole protection has zero revenue and so provides no satellites. Because two satellites is the efficient level, private provision is inefficient.

### 5.2.2 Public Provision:

Suppose there are two political parties, the Hauks and Doves, that agree with each other on all issues except for the quantity of satellites. The Hawks would like to provide 4 satellites at a cost of $50 billion, with benefits of $50 billion and a net benefit of zero, as shown in Fig. ----. The Doves would like to provide 1. satellite at a cost of $ 5 billion, a benefit of $20 billion, and a net benefit of $15 billion see figure ---.

Before deciding on their policy proposals, the two political parties do a “what - if” analysis. Each party reason as follows. If each party offers the satellite program it wants. Hawks 4 satellites and Doves 1 satellite the voters will see that they will get a net benefit from the Hauks, and the Doves will win the election.

Contemplating this outcome, the Hawks realize that their party is too hawkish to get elected. They figure that they must scale back their proposal to 2 satellites. At this level of provision, total cost is $15 billion, total benefit is $35 billion, and net benefit is $20 billion. If the Doves stick with 1 satellite, the Hauks will win the election.

But contemplating this outcome, the Doves realize that they must match the Hauks. They too propose to provide 2 satellites on exactly the same terms as the Hauks. If the two parties offer the same number of satellites, the voters are indifferent between the parties. They flip coins to decide their votes, and each party receives around 50 percent of the vote.

The result of the politicians' "what - if" analysis is that each party offers 2 satellites, so regardless of who wins the election, this is the quantity of satellites installed and this quantity is efficient. It maximizes the perceived net benefit of the voters. Thus in this example, competition in the political market place result in the efficient provision of a public good. But for this outcome to occur, voters must be well informed and evaluate the alternatives. But as you will see below, they do not always have an incentive to achieve this outcome.

The principle of Minimum differentiation. In the example we have jus studied, both parties propose identical policies. This tendency towards identical policies is an example of the principle of minimum differentiation, which is the tendency for competitor to make themselves identical to appeal to the maximum number of clients or voters. This principle not only describes the behaviour of political parties but also explains why fast food restaurants cluster in the same block and even why new auto models share similar features. If McDonald's opens a restaurant in a new location, it is likely that Burger king will open next door to McDonald's rather than a mile down the road. If Chrysler designs a new van with a sliding door on the drives side, most likely Ford will too.

## 5.3 MERIT GOODS:

The problem of social goods, by its very nature, has immediate application to the government's provision of goods and services. But it is also of interest in relation to transfers. Taxing and rendering transfer payments may be viewed simply as a process of taking by those who benefit. But this is not the entire story. To the extent that A's giving to B is based on A's desire to see B's position improved (rather than to derive pleasure from own giving), A will derive equal satisfaction from similar giving by C or D. Giving thus generates externalities not only for the recipient but also for others who see his position improved. Giving thereby assumes social good characteristics which call for budgetary implementation. In practice, it is, of course, difficult to distinguish between the takings and giving aspects of majority based redistribution, but both elements are presents. The rise of the welfare state over the past fifty years may well be integrated as involving increased radiness to give as well as to take.

In concluding this survey of the problems passed by social goods, we once more turn to their basic nature, this time focusing on the way in which wants for such goods are generated and on the nature of “merit goods”.

**The premise of individualistic Evaluation:-.**

Our distinction between private and social goods was based on certain technical characteristics of social goods, i.e., the non-rival nature of consumption and the inapplicability of exclusion. It did not depend on a difference in psychological attitudes or in social philosophy regarding the two types of goods. Utilities derived from social as well as private goods are experienced by individuals and included in their preference systems. The same individualistic psychology was applied to both types of goods.

The premise that all wants (private or social) are experienced by individuals rather than group entities is quite compatible with the notion that individuals do not live in isolation but an association with others. Human beings are social animals, and A’s preferences will be affected by those of B and C. dominant tastes and cultural values influence individual preferences and in turn are determined by them. Fashions are a pervasive factor in molding tastes, and not only with regard to clothing. To say that wants are experienced individually, therefore, is not to deny the existence of social interaction. Nor can it be argued that social goods differ from private goods because they satisfy the more noble aims of life.

Furthermore, the proposition that wants are experienced individually does not exclude altruism. If A is a socially minded person, he or she will derive satisfaction not only from his or her own consumption but also from consumption by B; or B, who is selfish, may enjoy only his or her own consumption. Utilities are interdependent and this fact broadens the range over which the economics of social goods applies. But granting all this, what matters here is that satisfaction is experienced in the last resort by A and B individually and not by a mysterious third entity called A + B.

Finally, we recognize that the quality of wants may differ. Some are concerned with the noble and others with quite ordinary aspects of life. But this does not bear on the distinction between private and social goods. The wants to be satisfied may be noble or base in either case: social goods may carry high cultural or aesthetic values, such as music education or the protection of natural beauty or they may relate to everyday needs, such as roads and fire protection. Similarly private goods may satisfy cultural needs, such as harpsichord recordings, or everyday needs, such as

bubblegum. Clearly, no distinction between private and social goods can be drawn on this basis.

**Communal Wants:**

The premises of wants, based on the needs and preferences of individuals, appeals to widely held values of Western culture. It also permits one to conduct the analysis of public provision within the same economic framework that applies to the analysis of private goods. The concept of communal needs, on the other hand, is hard to interpret and does not fit such analysis. Moreover, it carries the frightening implications of dictatorial abuse. Yet the concept of community also has its tradition in Western culture, from the Greeks through the Middle Ages and to date, and should be given at least brief consideration.

The central proposition to be examined is that there exists a community interest as such, an interest which is attributable to the community as a whole and which does not invalue a “mere” addition, vertical or horizontal, of individual interests. This community interest then is said to give rise to communal wants, which are generated by and pertain to the welfare of the group as a whole. This raises two basic questions one is to whom and how is the community interest revealed, and the other is over what range of needs should the community concept be applied.

Some observes would view the structure of communal wants as being revealed through a senate of sages, as in plato, or a political leader who, as was once believed in Maoist China, transmits his “insights” to the people. The people after an initial period of compulsion, come to accept these values as their own, thus removing the distinction between private and collective wants. This tenet is clearly inconsistent with our views of democracy; nor can it be defended by arguing that “in the end”, all preferences are socially conditioned social and environmental influences, to be sure, are pervasive, but there remains a considerable degree of freedom (unless suppressed) in individual responses thereto.

**Merit Goods:**

A more attractive interpretation is that by virtue of sustained association and mutual sympathy, people come to develop common concerns. A group of people share and historical experience or cultural tradition with which they identify, thereby establishing a common bond. Individuals will not only defend their home but will join others in defending their territory or in protecting their countryside. Such common interests and values may give rise to common wants i.e. wants which individual feel obliged to support as members of the community. This obligations may be accepted as falling outside the freedom of individual choice which ordinarily applies.

No all situations which at first sight appear to involve such common preferences fall within this category. Thus individual choice may be limited in situation such as these.

1. Interference is needed to guide children or the mentally disabled.
2. Provision for certain services such as education may be imposed to expand information on available options, without continuance of that interference after the information is gained.
3. Corrective action may be needed when consumer choice is based on false advertising.
4. Government subsidies to goods with external benefits do not involve interference with individual choice but permit such choice to be made more efficiently.
5. Budgetary decision by majority rule inevitably involve interference with minority preferences such violations are the inevitable if unfortunate by product of a process basically designed to implement individual preferences.

In situation such as these, society undertakes to correct for failures in the process by which individual choice is implemented effectively. Moving closer to the case of merit goods, let us consider the case of giving in kind. An individual donor may choose to give in kind rather than in cash, because he or she considers certain uses by the recipient as meritorious or taxpayers may prefer social programs which provide in kind aid, such as food stamps or low-cost housing, over cash grants. Supporters of the program feel that such uses are felt to the meritorious. As noted below, this may also enter into what is considered a fair state of distribution.

But acceptance of constraints on individual choice may extend beyond the act of giving and budgetary supports. Individuals as members of their society may feel obliged to share certain costs (e.g., for maintaining the Lincaln Memorials or to accept certain priorities in the use of their own funds because this is called for as a matter of respect for community values. This consideration may apply to the provision of what we have called social as well as private goods. Similar considerations may hold for the case of social bas, or demerit goods, e.g. prostitution. The concept of merit or demerit goods, to be sure, must be viewed with caution because it may serve as a vehicle for totalitarian rule. Yet such common values and concerns do exist in a cohesive society and their existence may place some limitation on the conventional doctrine of individual choice.

## 5.4 PUBLIC GOODS AND PRISONER'S DILEMMA:

Probably the most important accomplishment of economics is the demonstration that individuals with purely selfish motives can mutually benefit from exchange. if a raises cattle and B corn, both may improve their welfare by exchanging cattle for corn with the help of the price system, the process can be extended to accommodate a wide variety of goods and services.

Although often depicted as the perfect example of the beneficial outcome of purely private, individualistic activity in the absence of government (the invisible hand theorem presumes a system of collective choice comparable in sophistication and complexity to the market system it governs. For the choices facing A and B are not merely to trade or not, as implicitly suggested. A can choose to steal B's corn, rather than give up his cattle for it, B may do likewise unlike trading, which is a positive sum game benefiting both participants in an exchange, stealing is at best a zero-sum game. What A gains, B loser. If stealing, and guarding against it, detract from A and B's ability to produce corn and cattle, it becomes a negative sum game. Although with trading each seeks to improve his position and both end up better off, with stealing the selfish pursuits of each leave them both worse off.

The example can be illustrated with strategy matrix 4.1. To simplify the discussion, let us ignore the trading option and assume that each individual grows only corn. Square 1 gives the allocation when A and B both refrain from stealing (A's allocation precedes B's in each box).

**Matrix 4.1 Stealing as Prisoner's dilemma**

| **B** / **A** | Does not steal | Steals |
|---|---|---|
| Does not steals | 1 (10,9) | 4 (7,11) |
| Steals | 2 (12,6) | 3 (8,8) |

Both are better off when they both refrain from stealing, but each is still better off it he alone steals (cells 2 and 4). In matrix 4.1, stealing is a dominant strategy for both players, so defined because it dominates all other strategy options by promising a higher payoff for the chooser than any other strategy, given any choice of strategy by the other player. In an anarchic environment, the independent choices of both individuals can be expected to lead both to adopt the dominant stealing strategy with the outcome cell

3. The distribution of corn in cell 3 represents a “natural distribution” of goods namely, the distribution that would emerge in an Hobbesian state of nature.

From this “natural” state, both individuals become better off by tacitly or formally agreeing not steal, provided that the enforcement of such an agreement cost less than they jointly gain from it. The movement from cell 3 to cell 1 is a pareto move that lifts the individuals out of a Hobbesion state of nature. An agreement to make such a move is a form of “constitutional contract” establishing the property rights and behavioral constraints of each individual. The existence of these rights is undoubtedly a necessary precondition for the creation of the “post constitutional contracts”, which makeup a system of voluntary exchange. problems of collective choice arise with the departure from Hobbesian anarchy, and are conterminous with the existence of recognizable groups and communities.

A system of property rights and the procedures to enforce them are a samuelsonian public good in “that each individually consumption leads to no subtraction from any other individual’s consumption of that good”. Alternatively, a pure public good can be defined as one that must be provided in equal quantities to all members of the community. Familiar examples of pure public goods are national defense and police and fire protection. National defense is the collective provision against external threats; laws and their enforcement safeguard against internal threats; fire department against fires. Nearly all public goods whose provision requires an expenditure of resources, time or moral restraint can be depicted with a strategy box analogous to matrix 4.1. Replace stealing with paying for an army, or a police force, or a fire department, and the same strategy choices emerge. Each individual is better off it all contribute to the provision of the public good than if all do not, and each is still better off if only he does not pay for the good.

A pure public good has two salient characteristics, joint ness of supply, and the impossibility or inefficiency of excluding other from its consumption, once it has been supplied to some members of the community joint ness of supply is a property of the production or cost function of the public goods. The extreme case of joint ness of supply is a good whose production costs are all fixed, and thus whose marginal productions costs are zero (e.g. a public monument). For such a good, the addition of more consumers (viewers) does not detract form the benefits enjoyed by others. Even a good with falling average costs, although positive marginal costs, has elements of joint ness that raise collective provision issues.

The joint supply characteristic creates the potential gain from a cooperative more from cell 3 to 1. Given joint ness of supply a cooperative consumption decision is necessary to provide the good efficiently. If it took twice as many resources to protect A and B from one another as it does not protect only one of them, collective action would be of them, collective action would be unnecessary in the absence of nonexclusion. Each could choose independently whether or not to provide his own protection.

People can be excluded from the benefits from viewing a statue placed within a private gallery if they do not pay to see it. But people cannot be prevented from viewing a statue or monument placed in the central city square. For many public goods, the exclusion of some members of the community from their consumption is impossible or impractical. Failure of the exclusion principle to apply provides an incentive for noncooperative, individualistic behavior, a gain from moving from cell 1 to either cell 2 or cell 4. the impossibility of exclusion raises the likelihood that purely voluntary schemes for providing a public good will break down. Thus, (Together, the properties of public goods provide the raison deter for collective choice. Jointness of supply is the carrot, making cooperative collective decisions beneficial to all, absence of the exclusion principle the apple tempting individuals into independent noncooperative behavior.

Although the purest of pure public goods is characterized by both jointness of supply and the impossibility of exclusion preference revelation problems arise even if only the first of these two properties is present. That is, an alternative definition of a public good is that it may be provided in equal quantities to all members of the community at zero marginal cost. The substitution of “may” for “must” in the definition implies that exclusion may be possible. A classic example of a public good fitting this second definition is a bridge. In the absence of crowding, once built, the services of the bridge can be supplied to all members of the community, but they need not be. Exclusion is possible, As long as the marginal cost of someone’s crossing the bridge remains zero, however, excluding anyone who would experience a marginal benefit from crossing violates the pare to principle. Jointness of supply alone can create the need for collective action to achieve pareto optimality.

Matrix 4.1 depicts the familiar and extensively analyzed prisoners dilemana. The salient feature of this game is that the raw player ranks the four possible outcomes 2 > 1 > 3 > 4, while the column player has the ranking 4 > 1 > 3 > 2. The noncooperative strategy is dominant for both players. It is the best strategy for each player in a single play of the game regardless of the other player’s strategy choice. The outcome, square 3, is a cournot-Nash

equilibrium. It has the unfortunate property of being the only outcomes of the prisoners dilemma game that is not pareto optimal. From each of the other three squares a move must make at least one player worse off, but from 3 a move to 1 makes both better off.

Despite the obvious superiority of the cooperative nonstealing outcome to the joint stealing outcomes, the dominance of the stealing strategies ensures that the nonstealing strategies do not constitute an equilibrium pair, at least for a single play of the game. The cooperative solution may emerge, however, as the outcome of a “super game” of prisoners’ dilemma games repeated over and over by the same players The cooperative solution can arise, even in the absence of direct communication between the players, if each player chooses a super game strategy that effectively links his choice of the cooperative strategy in a single game to the other players choice of this strategy one such super game strategy one such super game strategy is for a player to play the same strategy in the present game as the other player (s) played in the previous game. If both (all) players adopt this strategy, and all being by playing the cooperative strategy, the cooperative strategy, the cooperative outcome will emerge in every play of the game. This “tit-for-tat” strategy beats all others proposed by a panel of go game theory expects in a computer tournament conducted by Robert Avelrod (1984).

An alternative strategy, which achieves the same outcome, is for each player to play the cooperative strategy as long as the other players (s) does, and then to punish the other player (s) for defecting by playing the noncooperative strategy for a series of plays following any defection before returning to the cooperative strategy. Again, if all players begin by playing cooperatively, this outcome continues throughout the game. In both of these cooperative strategies, equilibrium solutions to the prisoner’s dilemma super game, the equilibrium comes about through the punishment (or threat thereof) of the noncooperative behavior of any player, in this case by noncooperation of the other player (s). This idea that noncooperative (antisocial, immoral) behavior must be punished to bring about conformity with group mores is to be found in most, if not all, moral philosophies, and forms a direct linkage between this large literature and the modern theory.

In experimental studies the appearance of cooperative solutions in prisoner’s dilemma games has been found to depend on the number of players number of plays of the game, and size of gain from adopting the cooperative strategy relative to both the loss from the noncooperative outcome and the gain from successful playing of the noncooperative strategy. The latter two need no elaboration. The first tow factors combine to determine the predictability of the response of the other player (s), when the

number of other players is small, it is obviously easier to learn their behavior and predict whether they will respond to cooperative strategy choices in a like manner. It is also easier to detect noncooperative behavior and, if this is possible, single if out for punishment, thereby further encouraging the cooperative strategies. When numbers are large, it is easy for one or a few players to adopt the non-cooperative strategy and either no be detected, since the impact on the rest is small, or not be punished, since they cannot be discovered or it is too costly to the cooperating players to punish them. Thus, voluntary compliance with behavioral sanctions or provision of public goods is more likely in small communities than in large. Reliance on voluntary compliance in large communities or groups leads to free riding and the under or nonprovision of the public good.

In game experiments, cooperative solutions are reached only after a series of plays of the game. In the absence of direct communication and agreement, time is needed to learn the behavior of the other player (s) Generalizing from these findings, one can expect the voluntary provision of public goods and cooperative behavioral constraints to be greater in small, stable communities of homogenous behavior patterns.

In the large, mobile, heterogeneous community, a formal statement of what behavior is mutually beneficial (e.g. how much each must contribute for a public good) may be needed even for individuals to know what behavior is consistent with the public interest. Given the incentives to free rise, compliance may require the implementation of individualized rewards or sanctions. Mancurolson found that individual participation in large, voluntary organization like labor unions, professional lobbies, and other special interest groups was dependent not on the collective benefits these organizations provided for all of their members, but on the individualized incentives they provided in the form of selective benefits for participation and attendance, or penalties in the form of dues, fines, and other individualized sanctions.

Thus, democracy, with its formal voting procedures for making and enforcing collective choices, is an institution that is needed by communities of only a certain size and impersonality. The family makes an array of collective decisions without every voting; a tribe votes only occasionally. A metropolis or nation state may have to make a great number of decisions by collective choice processes, although many of them may not correspond to what we have here defined as a democratic process. Similarly, small, stabled communities may be able to elicit voluntary compliance with group mores and contributions for the provision of local public goods by the use of informal communication channels and peer group pressure. Larger more impersonal communities must

typically establish formal penalties against a social behavior (like stealing), levy taxes to provide for public goods, and employ police force to ensure compliance.

The size of the community, its reliance on formal sanctions and police enforcement, and the breakdown of the prisoner's dilemma may all be dynamically related. Detection takes time an increase in the number of violation can be expected to lead to a further increase in violence but only with a time lag. If, because of an increase in community size, or for some other reason, the frequency of violation in latter periods could be expected increase even further, and with these the need for and reliance on police enforcement of the laws. James Buchanan (1975) has described such a process as the erosion of a community's legal (i.e.) rule-abiding capital. Michael Taylor has presented a similar scenario. Taylor relates the breakdown of the cooperative solution to the prisoners dilemma not to the size of the community, however, but to the level of government intervention itself. Intervention of the state in the provision of community want or in the enforcement of social modes, psychologically :frees" an individual from responsibility for providing for community wants and preserving its mores state intervention leads to increased asocial behavior requiring more state intervention, and so on. The theories of Buchanan and Taylor might constitute one explanation for the rising government expenditures that have occurred in this century. Their theories would link these expenditures directly to the increasing mobility and urbanization that has occurred during the century, and the consequent increase in government intervention this has caused.

The scenarios by Buchanan and Taylor of the unraveling of the social fabric mirror to a remarkable degree the description by John Rawls (1971) of the evaluation of a just society, in which the moral (just, cooperative) behavior of one individual leads to increasingly moral behavior by others, reinforcing the cooperative behavior of the first and encouraging still more. The dynamic process in these scenarios is the same, only the direction of change is reversed.

## 5.5 PUBLIC GOODS AND CHICKENS:

The prisoners dilemma is the most frequently used characterization of the situations to which public goods give rise. But the technology of public goods provision can be such as to generate other kinds of strategic interactions consider the following example.

The properties of two individuals share a common boundary. G owns a goat that occasionally wanders into D's garden and eats

the vegetables and flowers. D has a dog that sometimes crosses into G's property, chasing and frightening the goat so that it does not give milk. A fence separating the two properties could stop both from happening.

Matrix 2.2 depicts the situation with no fence, both D and G experience utility levels of one. The fence costs $1,000 and each would be willing to pay the full cost if necessary to get the benefits of the fence. The utility levels of each (2) are higher with the fence than without it, even when they must pay the full cost alone. This assumption ensures that the utility levels of both individuals are still higher if each must pay only half the cost of the fence (square 1) last of all, each is, of course, best off it the fence is built and he pays nothing (payoffs of 3.5 to G and D, respectively, in squares 2 and 4).

Matrix 2.2 depicts the game of "chicken". It differs from the prisoner's dilemma in that the outcome in which no one contributes (cell 3), which is pareto inferior to the outcome that both contribute (cell 1) is not an equilibrium since each individual is better off even if he must pay for the fence alone, each would be willing to move to square 2 or 4, as the case may be, rather than see the outcome remain at cell 3. cells 2 and 4 are both equilibria in this game, and they are the only two. The ordering of payoffs in a game of chicken for the row player is cell 2 > 1 > 4 > 3, whereas in a prisoners' dilemma it is 2 > 1 > 3 > 4. the interchange of the last two cells for both players causes the shift in the equilibrium.

**Matrix 4.2 Fence building as a game of chicken**

| **G** \ **D** | Contribute to building fence | Does not contribute |
|---|---|---|
| Contributes to building fence | 1 (3,3) | 4 (2,3.5) |
| Does not contribute | 2 (3.5,2) | 3 (1,1) |

In cells 4, 1 and 2 the fence is built. These cells differ only in who pays for the fence and the resulting utility payoffs. In cell 4, G pays the full $1,000 cost of the fence and experiences a utility level of 2. in cell 1, G pays $500 and receives a utility level of 3, while in cell 2G pays nothing for a utility level of 3.5. The lower increment in utility in going from a $500 fall in income to no change in income, compared with going from a $1,000 fall in income to a $500 fall, reflects and assumption of the declining marginal utility of income. If both G and D have declining marginal utilities of income, as assumed in the figures in matrix 4.2, then the solution that they share the cost of the fence is welfare maximizing as well as

equitable. Under alternative assumptions, a stronger higher fence may be built when the cost is shared, and the result may be an efficiency gain from the cost-sharing solution in cell 1. but the outcome in cell 1 is not an equilibrium. Both D and G will be better off if they can convince the other to pay the full cost of the fence. One way to do this is to precommit onself not to build the fence, or at least to convince one's neighbor that one has made such a commitment so that the neighbor, say D, believes that her choice is between cells 2 and 3, and thus naturally chooses cell 2.

The chicken game is often used to depict the interactions of nations. Let D be a superpower, which favor's having other countries install democratic institutions, C a country favoring communist institutions. A civil war rages in small country S between one group seeking to install a communist regime and another group wishing to install a democratic constitution. The situation could easily take on the characteristics of a game of chicken. Each superpower wants to support the group favoring its ideology in S and wants the other superpower to back down. But if the other superpower, say C, is supporting its group in S, then D is better off backing off than supporting its group in S and thereby being led into a direct confrontation with the other superpower. Both powers are clearly better off if they both back off than if the confrontation occurs.

Given this game of chicken configuration of payoffs, each superpower may try to get the other to back off by precommiting itself to defending democracy (communism) wherever it is threatened around the world. Such a precommitment combined with a reputation for "toughness" could force the other superpower to back down each time a clash between communist and noncommunist forces occurs in a small country.

The danger in a chicken situation, however, is that both superpowers may become so committed to their strategy of supporting groups of their ideology, and so committed to preserving their reputations for toughness, that neither side will back down. The confrontation of the superpowers is precipitated by the civil war in S. The fence does not get built.

As in prisoners dilemma situations the joint cooperation solution to the chicken game can emerge from a chicken supergame, if each player recognizes the long-run advantages to cooperation and adopts the tit-for-tat supergame strategy or an analogous one. Alternatively, the two superpowers (neighbors) may recognize the dangers inherent in the non-cooperative, precommitment strategy and thus may directly approach one another and agree to follow the cooperative strategy. Thus although the structure of the chicken game differs from that of the prisoner's

dilemma, the optimal solutions of the game are similar, requiring some sort of formal or tacit agreement to cooperate. As the number of players increases, the likelihood that a formal agreement is required increases. Thus for the chicken game, as for the prisoner's dilemma game, the necessity of having democratic institutions to achieve the efficient, cooperative solution to the game increases in likelihood as the number of players of the game rises.

## 5.6 VOLUNTARY PROVISION OF PUBLIC GOODS WITH CONSTANT RETURNS TO SCALE:

In this section we explore more formally the problems that arise in the voluntary provision of a public good. Consider as the pure public good a levy or dike built of bags of sand. Each member of the community voluntarily supplies as many bags of sand as she chooses. The total number of bags supplied is the summation of the individual contribution of each member. The more bags supplied, the higher and stronger the dike, and the better off are all members of the community.

Letting G; be the contribution to the public good of individual I, then the total quantity of public good supplied is

$$G = G_1 + G_2 + G_3 + ....G_n \quad \ldots\ldots\ldots\ldots \quad 4.1$$

Let each individual's utility function be given as $U_i(X_i, G)$, where $X_i$ is the quantity of private good; consumes.

Now consider the decision of I as to how much of the public good to supply, that is, the optimal $G_i$, given her budget constraint $Y_i = P_x X_i + P_g G_i$ where $Y_i$ is her income and $P_x$ and $P_g$ are the prices of the private and public goods, respectively. In the absence of an institution for coordinating the quantities of public good supplied, each individual must decide independently of the other individuals how much of the public good to supply. In making this decision, it is reasonable to assume that the individual takes the supply of the public good by the rest of the communities as fixed, and chooses the level of $G_i$ that maximizes $U_i$ given the values of $G_i$ chosen by all other individuals j, in the community. Her objective function is thus.

$$Oi = Ui(Xi, G) + \lambda i(Yi - PxXi - PgGi) \qquad (4.2)$$

Maximizing (4.2) with respect to *Gi* and *Xi* yields

$$\frac{\partial Ui}{\partial G} - \lambda i Pg = 0 \qquad (4.3)$$

$$\frac{\partial Ui}{\partial Xi} - \lambda iPx = 0 \qquad (4.4)$$

From which we obtain

$$\frac{\partial Ui / \partial G}{\partial Ui / \partial Xi} = \frac{Pg}{Px} \qquad (4.5)$$

As the condition for utility maximization. Each individual purchases the public good as if it were a private good, taking the purchases of the other members of the community as given. This equilibrium is often referred to as a cournot or Nash equilibrium, as it resembles the behavioural assumption cournot made concerning the supply of a homogeneous private good in an aligopolistic market.

Now let us contrast (4.5) with the conditions for pareto optimality. To obtain these, we maximize the following welfare function.

$$W = r_1U_1 + r_2U_2 + ... + Y_nU_n, \qquad (4.6)$$

Where all $r_i > 0$ given the positive weights on all individual utilities, any allocation that was not pareto optimal that is, from which one person's utility could be increased without lowering anyone else's could not be at a maximum for W. Thus, choosing $X_i$ and $G_i$ to maximize W gives us a pareto optimal allocation.

Maximizing (4.6) subject to the aggregate budget constraint

$$\sum_{i=1}^{n} Y_i = P_x \sum_{i=1}^{n} X_i + P_g G \qquad (4.7)$$

We obtain the first order conditions.

$$\sum_{i=1}^{n} r_i \frac{\partial U_i}{\partial G} - \lambda P_g = 0 \qquad (4.8)$$

and

$$r_i \frac{\partial U_i}{\partial X_i} - \lambda P_x = 0, i = 1, n \qquad (4.9)$$

Where $\lambda$ is the Lagrangian multiplier on the budget constraint. Using the n equations in (4.9) to eliminate the $r_i$ in (4.8), we obtain.

$$\sum_i \frac{\lambda P_x}{\partial U_i / \partial X_i} - \partial U_i / \partial G = \lambda P_g \qquad (4.10)$$

From which we obtain

$$\sum_i \frac{\partial U_i / \partial G}{\partial U_i / \partial X_i} = \frac{P_g}{P_x} \qquad (4.11)$$

Equation (4.11) is the familiar samulsonian (1954) condition for pareto optimality in the presence of public goods. Although independent utility maximization decisions lead each individual to equate her marginal rate of substitution of public for private good to their price ratios, as if the public good were in fact private (4.5), pareto optimality requires that the summation of the marginal rates of substitution over all members of the community be equated to this price ratio (4.11)

That the quantity of public good provided under the cournot - Nash equilibrium (4.5) is likely to be less than the pareto - optimal quantity an be seen by rewriting (4.11) as

$$\frac{\partial U_i / \partial G}{\partial U_i / \partial X_i} = \frac{P_g}{P_x} - \sum_{j \neq i} \frac{\partial U_i / \partial G}{\partial U_i / \partial X_i} \qquad (4.12)$$

If G and x are normal goods in each individuals utility function, then

$$\sum_{j \neq i} \frac{\partial U_j / \partial G}{\partial U_j / \partial X_j} > 0$$

and the marginal rate of substitution of public for private good for individual I defined by (4.12) will be less than that defined by (4.5), which implies that a greater quantity of G and a smaller quantity of $X_i$ are being consumed when (4.12) is satisfied than when (4.5) is.

To gain a feeling for the quantitative significance of the differences, consider the special case where $U_i$ is a Cobb-Douglas utility function, that is, $U_i = X_i^{\alpha} G^{\beta}, O < \alpha < 1$, and $O < \beta < 1$ under this assumption (4.5) becomes

$$\frac{BX_I^{\alpha} G^{B-1}}{\alpha X_i^{\alpha-1} G^B} = \frac{P_g}{P_x} \qquad (4.13)$$

From which it follows that

$$G = \frac{P_x}{P_g}\frac{B}{\alpha}X_i \qquad (4.14)$$

Substituting from (4.1) and the budget constraint yields.

$$\sum_i G_i = \frac{P_x}{P_g}\frac{B}{\alpha}\left(\frac{Y_i}{P_x} - \frac{P_g}{P_x}G_i\right) \qquad (4.15)$$

From which we obtain

$$\left(1+\frac{B}{\alpha}\right)G_i = -\sum_{j\neq 1} G_j + \frac{B}{\alpha}\frac{Y_i}{P_g} \qquad (4.16)$$

OR

$$G_j = \frac{\alpha}{-\alpha + B}\sum_{j\neq 1} G_j + \frac{B}{\alpha + B}\frac{Y_i}{P_g}$$

Equation (4.17) implies that individual voluntarily chooses to supply a smaller amount of the public good, the larger she believes the amount of public good provided by the other citizens to be. With only two individuals in the community, (4.17) defines the familiar reaction curve from duopoly theory. In this situation, it is a negativity sloped straight line.

If al members of the community have identical incomes, Y, then all will choose the same levels of $G_i$ and (4.17) can be used to find the contribution in equilibrium of a single individual:

$$G_i = -\frac{\alpha}{\alpha+\beta}(n-1)G_i + \frac{B}{\alpha+\beta}\frac{Y}{P_g}, \qquad (4.18)$$

From which are obtain

$$G_i = \frac{\beta}{\alpha_n + \beta}\frac{Y}{P_g} \qquad (4.19)$$

The amount of the public good provided by the community through independent contributions the becomes.

$$G = nG_i = \frac{n\beta}{\alpha n + \beta}\frac{Y}{P_g} \qquad (4.20)$$

These quantities can be compared to the pareto optimal quantities. With all individual incomes equal, all individuals contribute the same $G_i$ and have the same $X_i$ left over, so that (4.11) becomes

$$n\frac{\beta X_i^{\alpha} G^{\beta-1}}{\alpha X_i^{\alpha-1} G^{\beta}} = \frac{P_g}{P_x} \qquad (4.21)$$

Using the budget constraint to eliminate the $X_i$ and rearranging yields for the pareto optimal contribution of a single individual.

$$G_i = \frac{\beta}{\alpha+\beta}\frac{Y}{P_g} \qquad (4.22)$$

and

$$G = nG_i = \frac{n\beta}{\alpha+\beta}\frac{Y}{P_g} \qquad (4.23)$$

Let us call the pareto - optimal quantity of public good defined by (4.23) Gpo, and the quantity under the cournot-Nash equilibrium (4.20) $G_{cn}$. Their ratio is then

$$\frac{GCN}{GPO} = \frac{\dfrac{n\beta}{\alpha n+\beta}\dfrac{Y}{P_g}}{\dfrac{n\beta}{\alpha+\beta}\dfrac{Y}{P_g}} = \frac{\alpha+\beta}{\alpha n+\beta}$$

This ratio is less than one, if m > 1, and tends toward zero as n becomes increasingly large. Thus, for all communities greater than a solitary individual, voluntary, independent supply of the public good leads to less than the pareto - optimal quantity being supplied, and the relative gap between the two quantities grows as community size increases.

The extent of under provision of the public good at a cournot - Nash equilibrium depends on the nature of the individual utility functions. For the cobb-Douglas utility function, the extent of under provision is smaller, the greater is the ratio of $\beta$ to $\alpha$. With $\alpha$=0 that is when the marginal utility of the private good is zero $G_{CN} = G_{PO}$. This equality also holds with right angled indifference curves, where again the marginal utility of the private good, holding the quantity of the public good fixed, is zero. But with the familiar, smooth, convexto the origin indifference curves, one can expect an under provision of a voluntarily provided public good, and an under provision whose relative size grows with the size of the community.

To achieve the pareto - optimal allocation same institution for coordinating the contributions of each individual is needed.

## 5.7 VOLUNTARY PROVISION OF PUBLIC GOODS WITH VARYING SUPPLY TECHNOLOGIES:

Many public goods might be depicted using the summation technology of the previous section. Public goods of a prisoner's dilemma type for example, community order, environmental quality are provided by each individual contributing to the "production" of the public good by not stealing or not polluting. For the typical public good of this kind, the quantity supplied is to some degree additive with respect to each individual's contribution. The more people there are who refrain from stealing, the more secure is the community and the greater the benefits consumed by all members.

There are other public goods, however, for which the participation of all members is necessary to secure any benefits. The crew of a small sailboat two-man rowboats, and bobsleds are examples. For the rowboat to go in a straight line each rower must pull the oar with equal force. Under - or over contribution are penalized by the boat's moving in a circle only the equal contribution of both rowers is rewarded by the boat's moving forward with such goods, cells 2, 4 and 3 collapse into one and cooperative behavior is voluntarily forthcoming.

Goods such as these are produced by what Jack Hirshleifer has named the "weakest - link" technology. The amount of public good provided is equal to the smallest quantity provided by any member of the community. At the other pole from weakest link technology one can conceive of a best shot technology for which the amount of public good provided is equal to the largest quantity provided by any one member of the community. As an example of the best shot technology, one can think of a community first having each member design a boat (bridge) for crossing a given body of water, and then the best design being selected and constructed.

The weakest link technology is like a fixed coefficients production function for public goods. Individual is marginal contribution to public good supply $\partial G / \partial G_i$ is zero, if his contribution exceed that of any other member of the community ($G_i > G_j$ for some *j*). But $\partial G / \partial G_i$ equals the community supply function when $G_i < G_j$ for all *j*. the summation technology assumes additive and separable production function, whenever as the best shot technology assume a sort of discontinuously increasing return. The latter seems the least plausible of the three, so, we consider only

the cases falling in the range, between the weakest link and summation production technologies.

Consider a community of two Australian farmers whose fields are adjacent to one another and border on a segment of the bush. Each night the kangaroos come out of the bush and destroy the farmer's crops. The farmers can protect their crops, however, by erecting fences along the border between their property and the bush. Each farmer is responsible for buying fence for his own segment of the border. The following technologies can be enuisaged.

Weakest link Kangaroos adopt quickly to changes in their environment and discover the lowest point in the fence. The number of Kangaroos entering both farmers fields is determined by the height of the fence at its lowest point.

Unweighted summation Kangaroos are very dumb and probe the fence at random. The number of Kangaroos entering the two fields varies inversely with the average height of the two fences.

Diminishing returns : if one farmer's fence is lower than the others some Kangaroos learn to probe only the lower fence, but not all learn, and the higher fence stops some Kangaroos from going over.

Now consider the following general formulation of public good supply; let G be the number of units of public good provided, defined in this case the number of Kangaroos prevented from entering the fields. Let the units fence purchased at price $P_f$ be defined so that

$$G = F_4 W F_2, O \le F_1 \le F_2, O \le W \le 1 \qquad (4.25)$$

Where $F_I$ is farmer is purchase of fence. If *W=0,* we have the weakest link and $G = F_1$, the smaller of the two contributions. The larger W is, the more 2.5 contribution beyond 1's contributes to the supply of G, until with W=1, we reach the unweighted summation supply function examined above to simplify the problem, assume that both farmers have identical utility functions and both G and the private good X are noninferior. Then the farmer with the lower income will always choose to purchase the smaller quantity of fence, so that farmer 1 is the farmer with the smaller income of the two. He maximizes his utility $U_1(X, G)$ by choosing a level of private good consumption $X_1$ and contribution to the public good $F_1$ satisfying his budget constraint.

$Y_1 = P_x X_1 + P_f F_1$. The solution is again equation (4.5) with the price of the public good now $P_f$.

The solution to the utility maximization problem for farmer 2 is, however,

$$\frac{\partial U_2 / \partial G}{\partial U_2 / \partial X} = \frac{P_f}{wPx} \qquad (4.26)$$

As long as $F_2 > F_1$. In effect, farmer 2 faces a higher relative price for the public good F, since his contribution do not contribute as much on the margin as 1's oluing to the technology defined by (4.25) the smaller w is, the less fence 2 buys (the smaller his optimal contribution to the public good) with small enough w, the solution to (4.26) would require $F_2 < F_1$. But then 2 would be the smaller contributor and his optimal contribution would be defined by (4.5) since 2 favors a greater contribution than 1, he simply matches 1's contribution if satisfying (4.26) violates $F_2 > F_1$.

To determine the condition for the pareto-optimal level of G, we choose of $X_1$, $D_2$ and G to maximize 1's utility, holding 2's utility constant and satisfying (4.25) and the individual budget constraints, that is we maximize

$$L = U_1(X_1, G) + r\left[U_2 - U_2(X_2, G)\right] + \lambda\left[G - F_1 - \omega F_2\right] \qquad (4.27)$$

From which it follows that

$$\frac{\partial U_1 / \partial G}{\partial U_1 / \partial X} + W\frac{\partial U_2 / \partial G}{\partial U_2 / \partial X} = \frac{P_f}{P_x} \qquad (4.28)$$

Only in the extreme weakest link case, where W = 0, is the condition for pareto-optimality for the community (4.28) satisfied by the two individuals acting independently, for then (4.28) collapse to (4.5) and both farmers purchases the amounts of fence satisfying (4.5) with W = 1, on the other hand, we have the unweighted summation supply of public good and (4.28) becomes (4.11), the samuelsonian (1954) condition for pareto optimality, and too little public good is being supplied.

Moreover, the difference between the quantity of public good supplied voluntarily when each farmer acts independently and the pareto optimal quantity increases with w. To illustrate this, again let both individuals have identical income Y, and identical utility functions $U = X^{\alpha} G^{\beta}$. Both then purchase the same quantity of fence F and private good X. From (4.5) and (4.25) we obtain the cournot -= Nash equilibrium quantity of public good supplied

through the independent utility - maximizing decisions of the two farmers.

$$G_{CN} = \frac{\beta Y(1+w)}{P_i\left[\alpha(1+w)+\beta\right]} \qquad (4.29)$$

In the same way (4.28) can be used to obtain pareto-optimal G :

$$G_{PO} = \frac{B}{\alpha+\beta}\frac{Y}{P_f}(1+w) \qquad (4.30)$$

Dividing (4.29) by (4.30) we obtain the ratio of inhdependently supplied to pareto- optimal quantities of public good;

$$\frac{G_{CN}}{G_{PO}} = \frac{\alpha+\beta}{\alpha(1+w)+\beta} \qquad (4.31)$$

With W = 0, the ratio is one, but it falls as W increases.

With n individuals, (4.28) generalizes to

$$\frac{\partial U_1/\partial G}{\partial U_1/\partial x} + W_2\frac{\partial U_2/\partial G}{\partial U_2/\partial x} + W_3\frac{\partial U_3/\partial G}{\partial U_3/\partial x} + ..... + W_N\frac{\partial U_n/\partial G}{\partial U_n/\partial x} = \frac{P_f}{P_x} \qquad (4.32)$$

And (4.31) generalizes to.

$$\frac{G_{CN}}{G_{PO}} = \frac{\alpha+\beta}{\alpha(1+w_2+w_3+....+w_n)+\beta} \qquad (4.33)$$

The gap between the independently provided and pareto-optimal quantities of public good increases as the number of members of the community increases, and the weights on the additional contribution increase.

Experiments by Harrison and Hirshleifex (1986) with two players indicate that individuals will voluntarily provide nearly the pareto-optimal quantity of public good in weakest link (w=0) situations but underprovide in summation and best shot situations. Experimental result by van de kragt, Orbell, and Dawex (1983) with small groups also indicate that efficient public good provision is forthcoming in situation resembling the weakest link technology. Thus, voluntary provision of public goods without coordination for coercion at - pareto-optimal levels is possible when the technology of public good provision conforms to the weakest link condition.

Unfortunately, with large communities it is difficult to think of many public goods for which voluntary provision is feasible, and all $w_i$ for contributions greater than the minimum are zero or close to it. In large communities, therefore, some institutional mechanism for coordinating and coercing individual contributions to the supply of public goods seems likely to be needed.

## 5.8 THE COASE THEOREM:

Ronald Coase in a classic article Rubbished in 1960 challenged the conventional wisdom in economics regarding externalities, taxes and subsidies. Coase argued that the existence of an external effect associated with a given activity did not inevitably require government intervention in the form of taxes and subsidies. Pareto-optimal resolution to externality situations could have been and were often worked out between the affected parties with the help of the government. Moreover, the nature of the outcome was independent of the assignment of propery rights, that is, in the case of a negative externality associated with E whether the law granted the purchaser of F the right to purchase E in unlimited quantities, or the law granted B the right to be protected from any adverse effects from A's consumption of E.

Although coase develops his argument by example, and neither states nor proves any theorems, the main results of the paper are commonly referred to as the coase theorem. The theorem can be expressed as follows.

**The Coase Theorem :**

In the absence of transactions and bargaining costs, affected parties to an externality will agree on an allocation of resources that is both pareto-optimal and independent of any prior assignment of property rights.

Pigou was wrong, government intervention is not needed to resolve externality issues.

Consider first a discrete case of the theorem. Let A be a factory producing widgets with a by product of smoke. Let C be a laundry whose costs are raised by A's emissions of smoke. Given that A is in business, C's profits are $24,000, but if A were to cease production altogether, C's profits would rise to $31,000. A's profits are $3,000. Assuming A's factors can be cost lessly redeployed, society is better of if A ceases production, C then earns a net surplus over costs of $31,000, while the combined surplus when A and C both operate is only $27,000.

But suppose that there are no laws prohibiting smoke emissions. A is then free to produce, and the socially inferior outcome would appear to ensure. It would, however, pay C to tribe the owners of A to cease production by promising to pay them $ 3,000 per annum. Alternatively, C could acquire A and close it down. If I is the cost of capital, and the market expects A to earn $3,000 profits per year in perpetuity, then the market value of A is $3,000 li. The present discounted value to C of shutting A down is $7,000 li, however. The owners of C will realize an increase in wealth of $4,000 li by acquiring and closing A.

To see that the socially efficient outcome arises regardless of the assignment of property rights, assume that A's annual profit is $10,000 and the figures for C are as before. Now the efficient solution requires that A continue to operate. Suppose, however, that the propery right's lie with C. strict air pollution laws exist and C can file a complaint against A and force it to cease production. The profits of A are now such, however, that it can offer C a bribe of $7,000 + $\alpha, 0 \leq \alpha \leq$ $3,000, not to file a complaint. The owners of both firms are as well or better off under this alternative than they are if A closes, and the socially efficient outcome can again be expected to occur.

Note that under the conditions of the first example, where A's profits were only $3,000, it would not pay A to bribe C to allow it to continue to produce, and the socially efficient outcome would again occur.

When the externality-producing activity has a variable effect on the second party as the level of the activity changes, the coase theorem still holds. If A's marginal rate of substitution of E for $X$ ($MRS^{A}_{Ex}$) falls as E increases, then $MRS^{A}_{Ex} - P_e / P_x$ is negative, sloped, as in Figure 4.1. The point where $MRS^{A}_{Ex} - P_e / P_x$ crosses the horizontal axis, $E_1$, is the level of E that A chooses when she acts independently of B. It is the level of E satisfying equation (4.35).

If E creates a negative externality on E, then - $MRS^{B}_{Ex}$ is positive. In figure 4.1 - $MRS^{B}_{Ex}$ is drawn under the reasonable assumption that B is willing to give up an increasing amount of X to prevent A from consuming another unit of E, the higher E is, $E_{po}$ is the pareto - optimal level of E, the level satisfying equation (4.41).

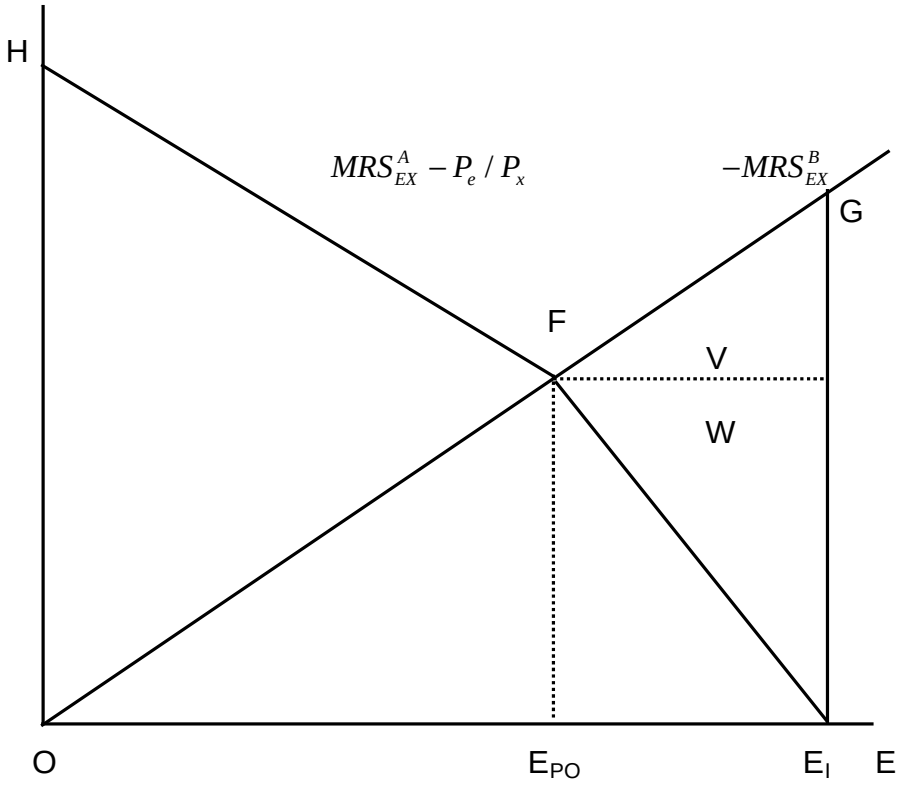



**Figure - 5.3 Pareto-optimal quantity of a good with external effects.**

The area $E_{po}$ $FGF_I$ measures the utility loss to B from A's consumption of $E_I$ instead of $E_{po}$. $E_{po}FE_I$ measures A's utility gain from these extra units of E. Both B and A will be made better off if A accepts a bribe of Z from B to consume $E_{po}$ rather than $E_I$, where $E_{po}FE_I < Z < E_{po}FGE_I$. In particular, if B were to offer A a bribe of $E_{po}F$ for each unit of E she refrained from consuming, A would choose to consume exactly $E_{po}$ units of E and A would be better off by the area W, and B by the area V as against the independent action outcome at $E_I$.

With the property rights reversed, B could forbid A from consuming E and force the outcome at 0. But then A would be forgoing $OHFE_{PO}$ benefits, while B gains only $OFE_{PO}$, as opposed to the pareto-optimal allocation $E_{PO}$. Self interest would lead A to propose and B to accept a bribe Z', to allow A to consume $E_{PO}$, where $OFE_{PO} < Z' < OHFE_{PO}$.

Coase demonstrated his theorem with four examples drawn from actual cases. Several recent experiments have been run with student subjects in which the students are given payoff tables that resemble those one would observe in an externality situation. Pareto - optimal outcomes are observed in well over 90 percent of the experiments. The coase theorem offers a logical and empirically relevant alternative to government action in externality situation.

But does it hold up as the number of parties involved in the externality increase? We now turn to this question.

## 5.9 SUMMARY

- A public good is a good or service that can be consumed simultaneously by everyone and from which no one can be excluded.
- Merit goods are those goods and services that the government feels that people left to themselves will under consume and which therefore ought to be subsidized or provided free at the point of use.

## 5.10 QUESTIONS

1) Define and explain the concept of Public Goods.
2) What is a Merit Good? Explain the concept in detail.
3) Explain in detail the concept of the Coase Theorem.

# 6

# PUBLIC EXPENDITURE

**Unit Structure:**

6.0 Objectives
6.1 Introduction
6.2 Wagner's Law
6.3 The Leviathan Hypothesis
6.4 Classes and Interest Groups
6.5 Fundamentals of Project Evaluation
6.6 Principles of Shadow Pricing
6.7 Summary
6.8 Questions

## 6.0 OBJECTIVES

After having studied this unit, you should be able -

- To understand the Wagner's Law.
- To explain the Leviathan Hypothesis.
- To know classes and Interest groups.
- To tell fundamentals of Project evaluation
- To understand the Principle of Shadow Pricing.

## 6.1 INTRODUCTION

The present unit discusses the rationale behind the public expenditure and presents various methods of evaluating it. We will study and understand the "Law of Rising Public Expenditure" putforth by German economist Adalph Wagner. The Leviathan Hypothesis about reason of systematic bias in the fiscal system. We will also try to understand the fundamentals of Project evaluation along with the concept of Shadow Pricing.

## 6.2 PUBLIC EXPENDITURE GROWTH: WAGNER'S LAW

Writing in the 1880s, the German economist Adolph Wagner advanced his "law of rising public expenditure." He felt, perhaps in anticipation of trends to be realized fifty to a hundred years later,

that the development of modern industrial society would a give rise to increasing political “pressure for social progress” and call for increased assurance for “social consideration” in the conduct of industry. In consequence, continual expansion of the public sector and its share in the economy should be expected. Has this law been borne out over the years, and just how should it be defined?

**Absolute Growth:**

Public expenditures, not surprisingly, have risen vastly in dollar terms. As shown in Table, line 1, such expenditures (including all levels of government) have increased by a multiple of nearly 2000 over the past ninety years. But this is not a meaningful way of looking at expenditure growth prices over the same period (line 13) rose by a multiple of 13, so that the multiple in terms of constant dollars (line 2) drops to 135. also population (line 12) more than tripled, so that the constant dollar multiple, measured on a per capita basis (line 3), falls to 33.

**Growth in Relation to GNP:**

Assurance for population and price changes are obvious correction, but they are not enough. One must also note that there has been a vast increase in productivity over the period, leading to a nearly six fold rise in per capita income in constant dollars. There is every reason to expect that part of this gain should have been spent on the goods and services provided by the public sector. In other words, focus should be on the share of government in total expenditures, with the law of rising public expenditures defined in terms of a rising public sector share.

**Expenditure-to-GNP Ratio:**

Beginning with the most global measure, we find that the ratio of public expenditures (all levels of government) to GNP rose from 6 to 35 percent over our nearly ninety-year period, with a nearly sixfold increase by the relative size of the public sector. This leaves us with a substantial increase, but by no means so drastic a rise as is suggested by the record of growth in total dollar terms.

The path of overall expenditure growth, as measured by the ratio of total public expenditures to GNP, is shown in line 4 of table - and is further plotted in figure - with years selected so as to avoid wartime peaks, we note a 2.4 percentage point growth in the U.S. ratio from 1890 to 1913, and a rise of 7.5 points from 1913 to 1929. This was followed by a 10 percent increase in the depression years of the 1930s and the post World War II adjustments. The rise continued in the 1960s and 1970s but at a declining rate, reaching a constant ratio in the 1980s. Even though the rate of increase has varied by subperiods, it is evident that Wagner’s law of a rising expenditure share was borne out if we take the longer sweep though at a slowing rate, and has come to a halt in recent years.

**Expenditure Elasticity:**

Another view of the same development is taken in Table - where the data of table - are recast in terms of expenditure elasticities. The table shows the GNP elasticity of total and civilian expenditures over selected years. We note that both elasticities were substantially above unity on through the 1960s, reflecting the rising expenditure to GNP ratios. However, we also note that the elasticity fell in recent decades. The table also shows the economy's marginal propensity to spend in the public sector, defined as the increase in expenditures as a percent of the increase in GNP, while the marginal propensity to spend on civilian outlays moved up on through the 1970s, this tendency was reversed during the 1980s.

**Growth by type of expenditure:**

To explain the growth in the overall expenditure share in GNP it is helpful to consider a breakdown by expenditure categories.

**Table 6.1 Growth of Government Expenditure in the United States**

| | 1890 | 1902 | 1913 | 1922 | 1929 | 1940 | 1950 | 1960 | 1970 | 1980 | 1987 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Total expenditures | | | | | | | | | | | |
| 1) Current Dollar (billions) | 0.8 | 1.6 | 3.2 | 9.3 | 10.7 | 20.4 | 70.3 | 151.2 | 332.9 | 953.0 | 1550.0 |
| 2) 1985 dollars (billions) | 10.9 | 18.7 | 33.5 | 69.8 | 75.2 | 158.3 | 298.4 | 192.4 | 825.8 | 1211.0 | 1476.0 |
| 3) per capita, 1985 dollars | 17.3 | 23.6 | 34.5 | 63.6 | 616.0 | 1196.0 | 1960.0 | 2718.0 | 4024.0 | 5133.5 | 5721.0 |
| 4) As percentage of GNP | 6.1 | 8.0 | 7.8 | 12.5 | 1.04 | 20.4 | 24.4 | 29.4 | 32.5 | 35.1 | 35.0 |
| By type, percentage of GNP | | | | | | | | | | | |
| 5) Purchases | n.a.+ | n.a. | n.a. | n.a. | 8.6 | 14.2 | 13.5 | 19.5 | 21.8 | 19.4 | 21.1 |
| 6) Transfers | n.a. | n.a. | n.a. | n.a. | 1.8 | 6.2 | 10.9 | 5.9 | 11.3 | 15.7 | 13.3 |
| By function, percentage of GNP | | | | | | | | | | | |
| 7) Civilian | 5.0 | 5.8 | 6.8 | 10.7 | 9.2 | 18.1 | 19.4 | 20.6 | 25.2 | 29.9 | 28.3 |
| 8) Defense | 1.5 | 1.5 | 1.0 | 2.9 | 1.2 | 2.3 | 5.0 | 8.8 | 7.6 | 5.2 | 6.7 |
| By level of government, percentage of Total | | | | | | | | | | | |
| 9) Federal | n.a. | 34.5 | 30.2 | 40.3 | 31.5 | 49.3 | 63.7 | 64.3 | 62.5 | 66.1 | 64.5 |
| 10) State | n.a. | 10.8 | 11.6 | 13.4 | 16.8 | 22.3 | 18.2 | 16.5 | 19.4 | 19.7 | 35.5 |
| 11) Local | n.a. | 54.8 | 58.3 | 46.4 | 51.7 | 28.4 | 18.1 | 19.1 | 18.1 | 14.2 | |
| Related Statistics | | | | | | | | | | | |
| 12) Population (millions) | 63 | 79 | 97 | 110 | 122 | 132 | 152 | 181 | 205 | 227 | 258 |
| 13) Price index (1985-100) | 8.0 | 8.5 | 9.8 | 7.8 | 13.0 | 11.6 | 21.4 | 27.7 | 37.5 | 76.8 | 105.5 |
| 14) GNP, current dollars (billions) | 13 | 20 | 41 | 74 | 103 | 100 | 288 | 515 | 1015.5 | 27.32 | 4500 |

**Figure 6.1 :** U.S. civilian and defense expenditures as percentage of GNP call levels of government.

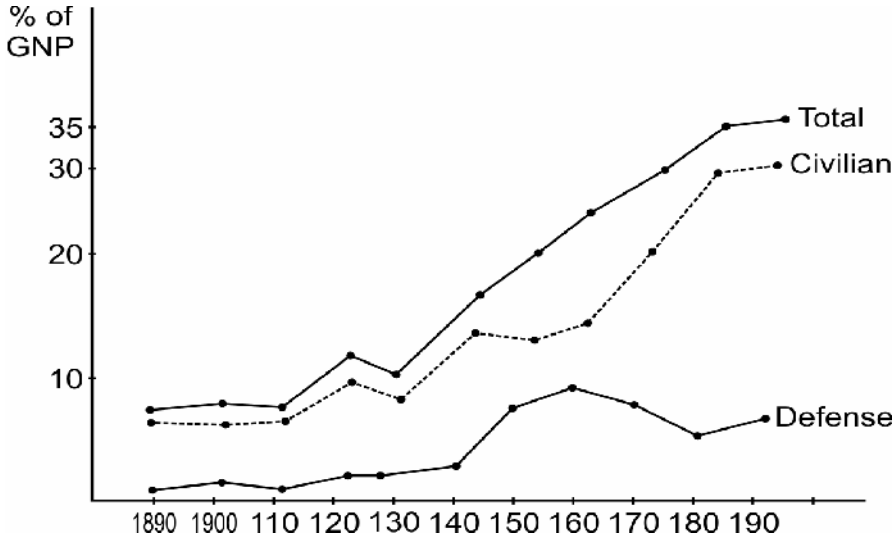



**Table : 6.2**

Expenditure Elasticities and propensities (All levels of Government)

| | GNP Elasticity | | Marginal Propensity | |
|---|---|---|---|---|
| | **Total** | **Civilian** | **Total** | **Civilian** |
| 1890-1929 | 1.7 | 1.8 | 10.9 | 9.0 |
| 1929-1950 | 2.9 | 2.2 | 30.4 | 18.4 |
| 1950-1970 | 1.6 | 1.8 | 35.9 | 26.5 |
| 1970-1980 | 1.1 | 1.2 | 34.6 | 30.7 |
| 1980-1987 | 1.1 | 1.0 | 36.6 | 27.5 |

**By level of Government**

As shown in lines 9 through 11 of table 6.1, the trend over the century was toward increased expenditures centralization. Although the federal share in 1929 was about the same as at the beginning of the century, the depression decade of the thirties brought a substantial step-up. The same happened during the forties, with the federal share emerging substantially above its pre-World War II level. Since 1950, however, the federal share has been fairly stable. Also note that the rising federal share was accompanied by a decline in the local and a gain in the state share. As we will see later, these ratios tend to overstate the shift toward centralization, be cause intergovernmental grants are included at the grantor level. Given the increasing importance of such grants in the 1960s and 1970s. Centralization as measured by shares in expenditures to the public has been less pronounced.

**Defense versus civilian Expenditures:**

Has expenditure growth been driven by rising expenditures for defense or by civilian expenditures as well? The ratios for the two shares are shown in lines 7 and 8 of table 5.1 and are plotted figure 6.1. We note that for our ninety year period the civilian expenditure ratio has increased somewhat faster than the defense ratio, but both have risen substantially. However, the pattern by subperiods differs sharply. The rise in the civilian ratio explained almost the entire increase for 1890 to 1940, whereas from 1940 to 1950 the defense ratio rose sharply while the civilian ratio showed little change. From 1960 to 1980 the defense ratio actually fell, while the civilian ratio rose sharply, a trend which came to a halt and was slightly reversed in the 1980s.

Although such comparisons have their shortcomings, it is evident that expenditure growth has not been primarily a matter of rising defense expenditures viewed over the longer run, the civilian expenditure ratio has been the driving force. As against a ratio of a percent in the pre-depression year 1929, it stood at about 23 percent of GNP in 1970 and had risen to nearly 28 percent by 1978.

**Purchase versus Transfer payments:**

Table 6.1, lines 5 and 6, shows a further breakdown of U.S. expenditure growth, this time between purchases and transfers (including interest). We find that both purchases and transfers have contributed to the rising expenditure share but the transfer share has been of increasing importance since the 1930s. Reflecting the rise of social security and the growing importance of welfare payments, transfer payments have accounted for three - quarters of the growth in the civilian expenditure ratio since that time. Over the decade of the 1960s, the purchase ratio showed little change, and during the 1970s it declined. Taking the overall picture, we see.

## 6.3 THE LEVIATHAN HYPOTHESIS

The theory of representative democracy, as described in the preceding section, has been subjected to severe criticism. The theory, like that of perfect markets, establishes a normative model which does not necessarily reflect its real application. Thus markets can function efficiently only if consumers are well - informed, if competition prevails, it prices are flexible, if no externalities are to be dealt with, and so forth. Not all these conditions prevail, and situation arise where markets do not work as the normative model suggests. Much the same holds for the model of fiscal democracy presented in the preceding pages. For the system to function

efficiently, voters must be informed, the vying for votes on the part of politicians must be competitive, the formation of party plat forms must be based on broad coalitions, voting systems must be sensitive to preferences, distortion through strategic behaviour must be minimized, and so forth. In reality, these conditions are rarely met. Defects in the fiscal process must thus be considered and have been viewed from a variety of perspectives Marxist critics as noted below have seen the fiscal process as an instrument of class struggle, shaped by the diverse interests of capital and labor. In recent years, conservative critics have viewed the growth of the public sector as expressing a systematic bias in the fiscal system toward over expansion. A modern leviathan is said to arise and threaten free institutions. Leaving the record of public sector growth and its economic determinants to the following chapter, we here consider the reasons why such bias is said to prevail. These reasons are found to lodge in both the voting process and the way in which the agents of government (bureaucrats and politicians) impose their own wishes on the public.

**Voting Bias:**

As we have seen in our earlier analysis, social goods and goods the benefits of which are largely external will be in undersupply without public provision. This leaves open the questions of whether the scope of public provision will be deficient or excessive, given our institutional selting. We take this to mean whether is above or below what would be provided in line with consumer evaluation.

**Cost to Minority:**

One basis plank of the overexpansion hypothesis is that majority voting by its very nature will result in oversupply. While only 51 percent of the voters may join in legislating a particular program which meets their interest, the tax cost is borne by all the members of the group. Assuming finance by a head tax, the cost to the majority will be only 51 percent of the total and the majority will disregard the 49 percent borne by the others who have no interest in the project. Oversupply thus results because the majority will consider only that part of the cost which it must bear.

Such may indeed be the case, but we should also note that opponents of projects do not consider the loss of benefits to proponents as projects are denied. To establish a general bias toward overexpansion, it must be shown that proponents are in a better position to organize than are opponents; or that proponents, feeling strongly about their project, find it more worth their while to spend money and effort to secure a majority vote. Perhaps so, but

a distortion may arise in either direction and the a priori conclusion of excess bias is at best a shaky one.

**Underestimation of Tax Burden:**

A further cause of oversupply is the fact that voters tend to underestimate the cost of taxation which they actually bear. Voters are seen to support expenditure legislation without being fully aware that an opportunity cost is involved, or they may assume that the cost will be borne by someone else. This will particularly be the tendency if taxes are invisible. Thus an increase in property or income tax is felt more directly and therefore meets more opposition than on increase in indirect taxes, especially if such taxes are added to cost at earlier stages of production rather than appear as an addition on the retail bill. The less visible the taxes the more likely it is that expenditures will be considered costless and that overexpansion will result. Under conditions of deficit finance in particular, an increase in expenditures seems costless.

Similar considerations apply where tax revenue rises due not to a legislate a rate hike but to an automatic increase such built in revenue gain may come about due to economic growth and inflation and may permit additional outlays which might not have been agreed to if a tax increase had to be condoned by specific legislation indexing of the income tax has eliminated this problem.

Fiscal illusion exists, but once more the argument has two sides while tax payers may underestimate their burden, they may also underestimate expenditure benefits. Benefits which one derives from private purchases are visible and ratified by the purchase price. If I want my car repaired I must pay the garage, which tells me the value of benefits derived, but the roads are there for me to enjoy, like sunshine, and I may take their benefit for granted. Moreover, it has also been argued that the political process leaves a deficiency in the provision for social goods because the consumer - voter is subject to intensive advertising pressures from the producers of private goods, so that his or her perceived needs are distorted in the latter’s favor. This may well be the case, but it should also be noted that private producers who produce public goods (whether the defense or construction industry or teacher unions) spend much effort and funds to persuade legislators and voters that their services are needed. As usual, there is a cross - current of forces and the net effect is by no means evident.

**Deficit Finance:**

A vote to raise expenditures when matched by a vote to raise taxes carries a visible opportunity cost to the taxpayer. But

this is not the case if the additional outlays are to be deficit - financed. Voters will tend to overlook the future cost of debt service and view the increase in programs as being more or less costless. Thus deficit finance, even through it may of times be needed for purposes of stabilization, tends to expand the budget. Surplus finance by the same taken generates a curbing effect.

**Public Employee Voting:**

Overexpansion, finally, is said to result because public employees as voters support large budgets simply because they create jobs for them and quite independent of the benefits derived from public services. Perhaps so, but it may also be held that employees of firms which produce goods sold to private consumers vote with the opposite interest in mind. Moreover, it may be noted that recent decades of public sector expansion have been associated with a declining share of the public sector in total employment.

**Monopoly Government:**

Voting bias is not the only cause, so the critics argue, which leads to overexpansion of the budget. “Bureaucrats” and politicians also contribute thereto. They do not serve to implement the wishes of the voter as the theory of representative government assumes, but strive to impose their own will. They find it in their interest to expand the budget and they are in a position to do so.

**Bureaucrats:**

Consider first the case of bureaucrats, the term now commonly applied to government officials and employees, a group which in an earlier social climate was referred to more kindly as civil servants. The bureaucrats central objective, so the arguments postulates, is to maximize the size of their bureaus so as to raise their salaries or extend their power.

The empire - building bureaucrat will submit a budget request which (1) asks for more funds than needed to perform a given function, (2) overstates the benefits to be derived from a given level of services, and (3) inflates the total in anticipation of expected out - backs. The granting agency may be duped by these tactics, but an excessive level of activity may result even if such cheating is ruled out. Thus it has been postulated that the sponsor of an activity, who decides on the budget request, will accept any proposal provided only that “the project is worth the money” in the sense that total benefits do not fall short of total costs. The bureaucrat will then propose the largest budget compatible with this

condition. As shown in figure - this budget will be in excess of the efficient level.

DD' in figure - represents the sponsor's marginal evaluation of successive units of service, and Os gives the marginal cost of providing them.

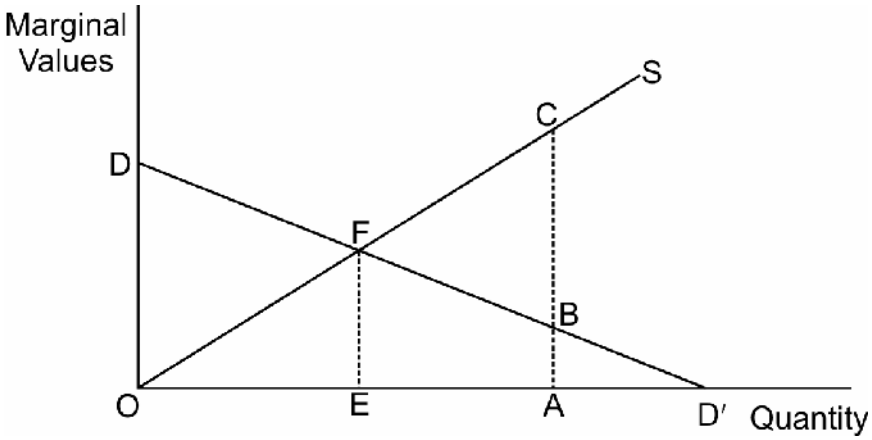



**Fig. 6.2 : Maximizing behaviour of bureaucrats**

At output OA, the total benefit, or area OPBA, matches total costs, or area OCA. For lower levels of outlay, benefits exceed costs, and for higher levels costs exceed benefits service level OA and the corresponding budget of OCA are thus the largest which the sponsor will grant, and this is the budget which the bureau head will offer. This budget, however, exceeds the efficient budget output OE and expenditure level OFE, the level at which marginal costs and benefits are equal. While budget additions involving quantities from OE to OA still appear worthwhile to the grantor, since total benefits continue to exceed total costs, extension beyond OE is inefficient. Beyond that point each successive unit costs more than the benefits it yields, are worth.

How realistic is the model of Figure will the grantor be that naïve and will the typical bureau head in fact maximize bureau size? Or may other motivations enter such as to serve the public interest and to expedite efficiency? While self - interest may well be a factor, it is hardly a fair reading of human nature to postulate that it is the only mode. Moreover, even if the bureau head intends to maximize his or her bureau, he or she may not be fell to do so. Budgetary requests are examined within the department before they are presented to the office of management and budget (OMB). They must then pass OMB scrutiny before they go to the Congress, where they must further pass congressional judgment. Although this procedure is not perfect, it does impose a constraint which the model of fig overlooks. The bureau head is hardly in a monopoly like situation, free to impose on all - or - nothing offer. It may even

be to his or her advantage to establish a reputation for prudence. In all, the monopoly bureau provides an interesting analogy to the private sector, but it does not tell the entire story viewed from a different perspective, public employees function as civil servants who fulfill an important task in society. They are needed (1) to provide technical expertise in the design of programs, so as to enable decision makers (the elected representative) to make intelligent choices; and (2) to implement and operate programs once they are enacted. In this role, they provide an element of continuity to the governmental process and introduce a sense of rationality into its operation. Their services are crucial to the functioning of the modern state and to the design as well as implementation of public policy.

At the same time, civil servants not only function as aids to elected representatives but they themselves affect the outcome. In the conduct of government, as anywhere else, knowledge is power. Public programms are complex and elected officials may have neither the time nor the expertise to analyze them. That branch of the government which is backed by technical experts is thus at a great advantage. Moreover, in rendering advice, the technician can hardly avoid (and may not wish to avoid) introduction of his or her own policy judgments, thereby influencing policy outcome.

**Politicians:**

Similar considerations are applied to the role of the politician. According to the theory of representative democracy the politician functions as an entrepreneur who endeavors to maximize votes so as to stay in power. He or she does so by promoting the provision of a bundle of public services which reflects the wishes of the voters. Thereby the politician serves the interest of the voter, just as the profit maximizing entrepreneur serves the interest of consumers. Critics hold that the politician, like the bureaucrat, wishes to maximize the size of the budget. The politician does so because a larger budget serves his or her interest, whether to gain in power, influence, or (indirectly or by way of Kickbacks) in income. Given this objective, the politician will not present a program which maximizes votes. Instead, he or she will advocate the largest possible program which can still secure a majority, and this budget may well exceed that desired by the median voter. This is shown in figure, where the budget size is measured on the horizontal axis and net benefits (excess of benefits over tax price) are measured on the vertical axis.

The OA, OB, … OE curves then record the net gains or losses which various voters A,B, C, D and E will derive from various budget sizes, it being assumed that a given tax burden distribution applies. Under unrestrained voting, where voters are presented

with all available options, budget size OM will win, reflecting the preference of the median voter, or C. but the largest budget for which a majority can still be achieved is given at OL. This budget size will have the support of D and E, who would be willing to vote for even larger budgets since at L their net gain is still positive. Budget size OL will also have the support of C, who stands to gain from budget expansion up to that level. The politician thus permits voters to choose only among budgets of OL or larger, and OL wins.

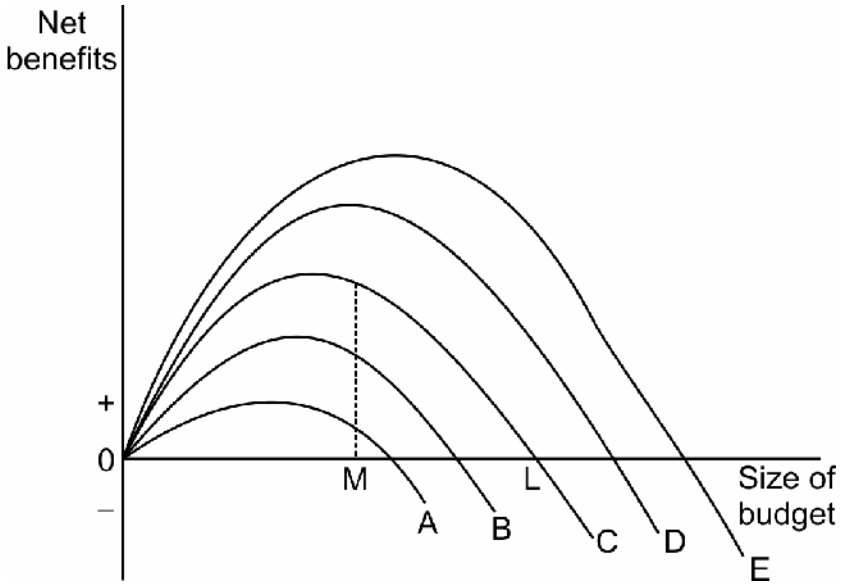



**Fig. 6.3 : Net benefit curves and agenda setting**

Once more we are left with the question of whether politicians do in fact have the power to constrain the 'voters' choice in this fashion, or whether their power to do so is limited by the loss of votes to rival politicians with more attractive platforms. Given the guarantee of free and periodic elections, it would seem that gross departures from the preferences of the voters cannot be sustained for long and that corrections will occur. The "tax revolt" of the late seventies is a case in point. A distinction must be drawn also between politicians imposing their will upon the public and political leadership which sets directions of public policy. The latter may enrich the democratic process rather than impede its performance.

**Campaign Financing:**

It remains to note the role of campaign financing as a distorting factor in the fiscal decision process. Although given relatively little attention in the body of literature reviewed in this section, it is surely a major source of bias in the system. With the rise of media campaigning, campaign costs and with them the importance of political contributions have vastly increased. The recent rise of single issue oriented PACs (political action

committees) and their lobbying activities has further added to the dependence of political candidates on the support of well funded interest groups. With vast sums spent to commit legislators in key committees such as ways and means and to rally voter support, independent leadership and non partisan judgment have become increasingly difficult. These pressures bear on tax and expenditure policy alike, and little is known about whether more funds are spent to promote expansion or restriction of the public sector. It is evident, however, that an efficient setting of tax and expenditure structures is impeded. Reform of campaign finance is thus a priority item.

**A political Business Cycle:**

The view of government as manipulating the public rather than implementing its wishes bears not only on the size of government but also on the conduct of macropolicy. The hypothesis here is that policy makers realize the importance of economic conditions for election success and therefore conduct macro policy so as to create favourable conditions at election time.

Considerable empirical work has been done to explore the relationship between election outcomes and economic variables, specifically, attempts have been made to establish the extent to which the outcomes of presidential and congressional elections depends on economic variables such as unemployment, inflation, and changes in real income. The question here is whether the party in power will be blamed for poor economic performance and rewarded for success. One would expect this to be the case, but the results of such analysis depend on how the problem is formulated. Thus, it matters a great deal how much of a response lag is allowed for (does only the record of the election year matter, or does the voter have a longer memory?) and just how the index of economic performance is defined.

How ever this may be, politicians will expect favorable economic conditions to have beneficial election effects, and they may therefore be expected to time policy actions accordingly. That is to say elections will be preceded by expansionary policies to stimulate employment or by structural measures (say, farm policies) to please particular sections of the electorate. In this way, government may generate a politically based business cycle.

**Budget Limitations:**

The view that our political process overextends the public sector has generated proposals for rule changes which will render expansion more difficult. Where as state legislation to limit the fiscal powers of local government is nothing new, such practice greatly gained in momentum during the seventies, especially the late

seventies after california's "tax revolt" led to the passage of proposition 13. this series of amendments, moreover, was extended to also limit the fiscal powers of state legislatures, and more recently introduction of an amendment to the U.S. constitution has been under consideration by the congress.

**State - Local level:**

The tax revolt of the late seventies, as signaled by California's proposition 13, was above all a protest against the property tax. Even though property taxation had risen less rapidly during the sixties and early seventies than most other state and local taxes, inflation shifted an increasing share of the tax burden from business to residential real estate. Rapidly rising real estate values, moreover, had increased property tax liabilities relative to income, thus leaving the taxpayer with the perception of an increased tax burden. California's proposition 13 rolled back assessed values and limited future increase to 2 percent a year while imposing a rate ceiling of 1 percent. in addition, a two-thirds majority was required for other taxes to be increased. Under California's proposition 4, passed in 1979, an expenditure limit was imposed on the state budget, restricting the inflation adjusted growth of state expenditures to that of population. Since then, adoption of legal checks to fiscal expansion has grain widely. Nearly twenty states imposed constitutional limitations on the growth of state finances, and most states also limit the permissible increase in property taxation. The implications of these changes for state and local finances are still emerging and their impact will be considered later, when state and local finances are examined.

**Federal Level:**

The U.S. constitution, as noted earlier, does not impose on overall limitation on the taxing and spending powers of the federal government. Limitations on taxing power apply to permissible types of taxes and the preservation of due process only, whereas expenditures are limited only by the requirement that they must serve the public welfare. Congressional legislation on fiscal matters proceeds under the ordinary rule of absolute majority, with only a two-third majority required to override presidential vetoes. There is no constitutional provision requiring a balanced budget or limiting the public debt.

In recent years, various constitutional amendments have been proposed to limit the fiscal powers of the government. To become law, they must be passed by a two-thirds majority in both houses and must be ratified by three-fourths of the states. The leading amendment, as passed by the House in 1982, contained two major provisions. First, the congress would be required to plan

for a balanced budget. That is to say, revenue estimates (as based on the average income of the preceding three years) must match planned expenditures. A three-fifths majority would be required for a planned deficit, and a simple majority for a planned surplus. Second, tax receipts would not be permitted to grow more rapidly than national income. A more rapid increase would require endorsement by a simple majority of all members in both houses of congress. While the amendment drive has bogged down as unrealistic in face of the large deficits during the 1980s, it remains on the agenda for potential action.

In the meantime, congress has attempted to deal with the problem by legislation rather than by constitutional amendment. Thus the balanced budget Act of 1986 (also referred to as the Gramm Rudman - Hollings Act) established targets for deficit reduction leading to balance by 1911, requiring across - the - board expenditure cuts lest other action be taken. It remains to be seen whether this goal will be reached.

## 6.4 CLASSES AND INTEREST GROUPS:

The critique of the democratic model, outlined in the preceding sections, derives largely from on analysis of the behaviour of single individuals, whether they are voters, official, or politicians. An alternative approach emphasizes that individual action is constrained by membership in classes and groups, so that the fiscal process is seen as a matter of group interaction.

**The Marxist Model:**

Such as alternative approach is in line with the Marxist view, where by the state (prior to the revolution) is to be seen as an instrument by which the ruling (capitalist) class exploits the subjugated (working) class Action of the state must be interpreted as part of the class struggle, which transcends the political as well as the individual sphere of social relations.

Fiscal history may be seen in this perspective. In the middle Ages, the feudal lord extracted payments in cash or kind from his serfs to sustain his rule and the military establishment needed to maintain or improve his position. Thus it was in the interest of the ruling class to have as strong and rich a state as possible. With the rise of democratic government, the ruling class lost its tight control over the state, and power went increasingly to popular majorities who shifted the costs of maintaining the state to the hither to ruling class. As a result, the ruling class changed it, view of the state. Its interests were now served better by a weak state, and it thus came to favor small budgets, law taxes, and general noninterference with

the private sector. Marx in turn advocated a highly progressive income tax, listed in the communist manifesto as one of the means to hasten the breakdown of the capitalist system.

More recently, Marxist writers have emphasized the interdependence between “monopoly capital” and the fiscal state. The need to absorb surplus output is said to call for expanding public outlays, especially on defense; and a rising level of transfer payments is seen as necessary to maintain social peace. At the same time monopoly capital is said to oppose the necessary financing, thus creating a fiscal crisis of the state.

This view of fiscal politics reflects the Marxist framework in which the social process 16 seen in terms of class struggle. It is not surprising, then, that tax and expenditure decisions will be a major instrument of that struggle, dissatisfaction with taxation has indeed been a major factor in the history of revolutions, and redistributive fiscal measures have to a degree expropriated the “capitalist class.” But by the same token, the role of budgetary activity may change from a means of struggle to a tool of social accommodation once a less divisive view of society is taken. Budget policy then becomes on instrument of gradual reform and cooperation. Looking back at the history of the last century, we see that there can be little doubt that fiscal action played a key role in this growth of social cohesion. Indeed, the rise of the modern welfare state, with its emphasis on transfers and progressive taxation, has placed the public budget at the hub of the social system. The recent shift of political attitudes and critique of the welfare state in turn have focused on a critique of its fiscal components.

**Multiple Groupings:**

Although the fiscal process as on instrument of class struggle is too partial a view, fiscal interest groups are a powerful factor. The structure of groupings, however, is multidimensional, cutting across the Marxist categories of class capital and labor in the construction industry will combine to promote highway programs, while capital and labor in the defense industry will combine in favor of defense. Consumers receiving both wage and capital income will combine to support programs the benefits of which the value highly. Thus the actual interest structure is much more complex than a simple division into capital and labor would suggest.

A similar picture may be drawn with respect to taxation various taxpayer groups organize to represent their interests, and the congressional tax committees are under great pressure from such groups, whether it be the oil industry orgering for depletion allowances, the real estate lobby wanting faster depreciation,

governors advocating exemption of interest, or university representatives calling for deductibility of tuition payments. Consumers of product will combine in opposing its taxation, whether their income is derived from capital or from labor, and they will be joined by both workers and capitalists deriving their income from the production of X. The distinction between capital and labor income becomes relevant, however, when it comes to the treatment of the two income sources under the income tax. But even here alignment by income level, independent of source, is as or more important.

By offering a well-organized reflection of voter concerns, interest groups can make a constructive contribution. But they also distort. Some groups are organized more easily than others, and some have more financial resources to press their views, and automatic development of a neatly balanced structure of countervailing power cannot be relied upon. It is thus important for a public policy to develop an institutional setting in which a more balanced representation of group interests prevails.

However this may be, a realistic view of the fiscal system must take account of the strategic role of multiple interest groups, economic, demographic, and regional. A positive theory of fiscal behaviour centered on the interaction of interest groups and their impact on fiscal institutions and decisions may well be more realistic than those based on preferences of individual voters, or on their disregard by self - interested bureaucrats.

**Decision Rules :**

Project evaluation, like all issues in allocation economics, involves determination of the ways in which the most efficient use can be made of scarce resources. In its simplest form, the issue is how to determine the composition of the budget of a given size or how to allocate a total of given funds among alternative projects. There is also the more complex question of determining the appropriate size of the budget. Further complications arise when projects are not divisible but in lumpy form. In taking a first look at these various situations, we assume that benefits and costs are known. The identification and measurement of costs and benefits are considered in later sections of this chapter.

**Divisible Projects:**

We begin with a setting in which all projects are finely divisible, i.e., may be increased or decreased by small amounts. As will be noted later, this is not a very realistic assumption, but it permits us to bring out the basic rationale of project selection.

**Budget Size Fixed:**

Suppose that the budget director is to advise the legislature either congress or a city council how best to allocate a given sum, say $ 1 billion, between two expenditure projects, X and Y. The problem may be likened to that of the head of a consumer household who must allocate the family budget first, the director must determine the cost C involved in providing each service and the benefit B to be derived there from. Then outlays must be allocated between X and Y so as to derive the greatest total benefit from the budget, i.e. to maximize the sum of net benefits $\left(\sum,\text{NB}\right)$ or the excess of total benefits over costs $\sum,\left(\text{B}-\text{C}\right)$. With $\sum\text{C}$ given by the size of the budget, the task is simply to maximize $\sum,\text{B}$.

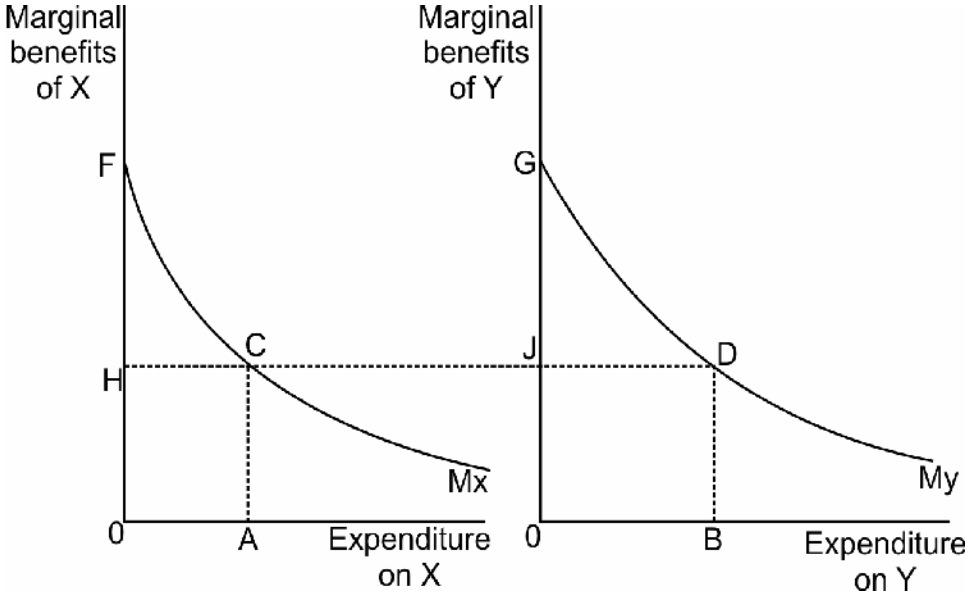



**Figure 6.4 : Expenditure allocation with fixed budget.**

This is shown in figure, where the Mx and My schedules show the value of the marginal benefits (additions to total benefits) derived from spending successive dollars and X and Y. The opportunity cost of spending a dollar on x is the cost loss of benefits due to not spending it on Y. total expenditures should therefore be distributed between X and Y so that the benefit derived from spending the last dollar on X will equate that derived from spending the last dollar on Y. Thus OA is spend on X and OB on Y such that AC = BD and OA + OB equals total permissible outlays. By equating the benefits derived from the marginal dollars on X and Y, we maximize the sum of total benefits derived from X (as measured by the area OFCA) and from Y (as measured by the area OGDB)

**Budget Size Variable:**

A more global view of budgeting indicates that the problem is not simply one of dividing up a budget of given size but also one of determining the size of the budget itself. The government must thereby decide how resources are to be divided between private and public use. We must therefore drop the assumption of a fixed budget and reconsider project choices along with determination of total budget outlays. Within the fixed budget, the opportunity cost of pursuing one public project consists of the benefit lost by not pursuing another public project. But in the open budget the opportunity cost of public projects must be redefined as the lost benefits from private projects which are forgone because resources are transferred to public use.

The task now is to maximize $\sum (B - C)$, including benefits and costs of both public and private projects. This condition is met by equating marginal benefits for the last dollar spent on alternative public and private projects. Public projects are expanded and private projects are restricted until the benefit from the last dollar spent in either sector is the same. Interpreting X as "the" public project and Y as "the" private project, we find that the solution of figure again applies. Given perfect markets, the marginal benefit from spending $1 in the private sector or BD equals $1, and the same must hold on the public side. Thus public expenditures are extended until the last dollar spent yields a dollar's worth of benefits.

**Table 6.3 :**

**Project choice with lumpy projects and fixed budget**

| Project | Costs* C | Benefits B | Net Benefits B - C | B/C | B/c Ranking |
|---|---|---|---|---|---|
| I | 200 | 400 | 200 | 2.0 | 2 |
| II | 145 | 175 | 30 | 1.2 | 5 |
| III | 80 | 104 | 24 | 1.3 | 4 |
| IV | 50 | 125 | 75 | 2.5 | 1 |
| V | 300 | 420 | 120 | 1.4 | 3 |
| VI | 305 | 330 | 25 | 1.1 | 6 |
| VII | 125 | 100 | -25 | 0.8 | 7 |

**Lumpy Projects:**

We have assumed so fax that expenditures may be divided finely between project X and Y, so that benefits may be equated for the marginal dollar spent on each. Where we deal with the allocation of funds between broad expenditure categories, this marginal approach is more or less applicable. But when it comes to

specific allocation within departments, choices must be made among particular projects which are indivisible, involve lump-sum amounts, or are not smoothly expandable. If a choice has to be made between a road connecting cities X and Y and another connecting X and Z, where the X to Y distance is twice the X to Z distance, no marginal adjustment is possible.

**Budget Size Fixed:**

We begin once more with the fixed budget case. Suppose that we have $7,00,000 to spend, say, on alternative highway projects, and that we may choose among projects I to VII, as shown in table. The cost of each project is measured by the dollar amount required. The benefit valuation gives the total benefit for each project. Returning to figure - we find that the total benefit for a project, involving cost OA corresponds to the area OACF.

In dealing with this case, let us consider various decision rules. Let rule require us to rank projects in line with their benefit cost ratio and move down the line until inclusion of a further project would exceed the budget constraint. We then choose projects. IV, I, V and III. Total cost is $ 6,30,000, benefits are $1,049,000, net benefits equal $419,000, and $70,000 of the available budget is left. As on alternative, let rule 2 call for the mix of projects which yields the largest net benefit. By trying various combinations, we find that net benefits are maximized by choosing IV, I, V and II. Here total cost is $ 695,000, benefits are $1,120,000, and net benefits equal $425,000. An amount of $ 50,000 remains unspent. Rule 4, finally, might be to minimize the amount left over, subject only to the constraint that projects must have a benefit cost ratio in excess of 1. In this case, the choice is for I, II, IV, and VI, with a cost of $ 700,000, benefits of $1,030,000, and net benefits of $313,000, nothing is left over.

Comparing the merits of the three rules, we find it is evident that both 1 and 2 are superior to 3, since both buy more benefits at a smaller cost. The choice between 1 and 2 is more difficult. Rule 1 is reasonable, because it calls for select on of projects which yield the highest return per dollar of the constrained resource, the available budget. Rule 2 offends this principle by choosing project II over III. Yet by moving from rule 1 to rule 2, additional benefits of $71,000 are bought at an additional cost of $65,000. net benefits rise by $6,000, and even though the marginal benefit cost ratio is only 1.09, this may still be considered a paying proposition. Rule 2 will clearly be preferred if we interpret the fixed budget case rigidly so as to consider turned back funds as worthless. Taking a broader view and allowing for a possible transfer to another budget, we note that rule 2 will be superior only if other budgets cannot after projects with a benefit cost ratio above 1.09.

**Budget Size Variable:**

If there is no fixed limit to the budget size, the problem is once more one of the weighing public against private uses of resources. Since we are now dealing with lumpy projects, this can no longer be done by balancing the benefits derived from marginal outlays on both uses. We now proceed by the rule that a public project is worth undertaking so long as the benefits derived there from exceed its costs. The justification for the rule is that the cost of spending n dollars in the public sector is the loss of n dollars of benefits, a loss which results from not spending n dollars in the private sector. The rule may be stated by saying that a project should be undertaken so long as (B - C) > 0.

## 6.5 FUNDAMENTALS OF PROJECTS EVALUATION:

The problem of project evaluation is linked closely to that of consumer surplus and the change therein.

**Consumer Surplus:**

This linkage is shown in figure. Suppose that the demand curve for a given product, say automobiles, is given by AB. The demand curve shows the maximum amounts which consumers are willing to pay for successive automobiles. Thus they would be willing to pay a price of $P_1$ for the first car, of $P_2$ for the second, and so forth. The utility of the first car would be measured by the block $OP_1$ $NQ_1$, that of the second car by the block $Q_1$ $RVQ_2$, and so forth. If the blocks are drawn sufficiently small and are added up, they sum to the area under the demand curve utility or benefit derived from various levels of consumption as indicated by the consumer's willingness to pay.

Consumers will extend purchases to the point where the marginal value of the last unit equals marginal cost or the price which they must pay. If the product were available at a zero price, they would consume OB. The benefit would equal OAB, and with price equal to zero, this entire area would measure their "consumer surplus". If the price were to equal OC, OD units would be bought and total benefits would equal OAED. With cost equal to OCED, the consumer surplus would be OAED minus OCED or CAE. This surplus, to repeat, is the excess of what consumers would be willing to pay for D units over what they must pay to obtain them.

**Net Benefit of Projects:**

We can now apply the concept of consumer surplus to measuring the benefit derived from a public project. The demand for the services of the project is again given by AB, and the project

is introduced with a unit cost of OC. Returning to the tabulation of benefits and costs in table ----. Suppose that we have an indivisible project of size of. Total benefits as recorded in the table correspond to area OAHF, with AB reflecting the vertically added “demand curves” of the consumers. Total costs correspond to area OCKF, and net benefits, equal to consumer surplus, correspond to CAHK. Project choice in the fixed budget maximizes the sum of these consumer surplus areas. For divisible projects provision should be carried to OD, the point where marginal evaluation equals marginal cost, i.e., the marginal gain in consumer surplus becomes zero and total surplus, equal to CAE, is maximized.

**Types of benefits and costs:**

In identifying various types of benefits and costs, these major categories may be distinguished.

Benefits and costs may be real or pecuniary real benefits and costs may be :

Direct or indirect.
Tangible or intangible
Final or intermediate
Inside or outside.

Illustrations of various types of benefits and costs are given in table.

**Table 6.4 :**
Illustrations of project benefits and costs.

| | | Benefits | Costs |
|---|---|---|---|
| Irrigation project | | | |
| Real Direct | Tangible<br><br>Intangible | Increased farm output<br>Beautification of area | Cost of pipes<br><br>Loss of wilderness |
| Indirect | Tangible<br><br>Intangible | Reduced soil erosion<br>Preservation of rural society | Diversion of water<br>Destruction of wildlife |
| Pecuniary | | Higher real income of farm equipment industry | |

| MOON SHOT PROJECT | | | |
|---|---|---|---|
| Real direct | Tangible | As yet unknown | Cost of inputs |
| | Intangible | Joy of exploration | Pollution of universe |
| Indirect | Tangible | Technical progress generated | |
| | Intangible | Gain in world prestige | |
| Pecuniary | | Relative increase in land values at cape kennedy | |
| EDUCATION PROJECT | | | |
| Real direct | Tangible | Increased future earnings | Cost of teacher's salaries, cost of buildings and books |
| | Intangible | Enriched life | Forgone leisure time |
| Indirect | Tangible | Reduced cost of crime prevention | |
| | Intangible | More intelligent electrorate relative increase in teachers incomes | |

The most important distinction is that between real and pecuniary aspects. Real benefits are the benefits derived by the final consumers of the public project. They reflect an addition to the community's welfare, to be balanced against the real cost of resource withdrawal from other uses. Pecuniary benefits and cost come about because of changes in relative prices which occur as the economy adjusts itself to the provision of the public service and pattern of resource demand changes. As a result, gains or losses accrue to some individuals but are offet by losses or gains which are experienced by others. They do not reflect net gains or costs to society as a whole.

As labor is hired and a road is constructed, the wage rates for construction workers may rise because the relative scarcity of their skills is increased. At the same time, increased taxes needed

to pay for the road may result in reduced amount for other services, and a loss of income elsewhere in the system. Such pecuniary changes do not reflect net gains or losses to society because they are matched by offsetting losses or gains. They must be distinguished from real costs and benefits which do. The latter must be allowed for and pecuniary changes should not enter into the evaluation. Such at least is the case unless distributional weights are to be attached to the particular gains or losses which accrue to various individuals, or unless such changes occur outside the jurisdiction within which the project is evaluated.

**Types of Real Benefits:**

As noted before, all real benefits should be allowed for in cost benefit analysis, but various types of benefits may be distinguished.

**Direct Versus Indirect:**

Real benefits and costs may be direct or indirect or, which is the same, primary or secondary. Direct benefits and costs are those related closely to the main project objective, whereas indirect benefits are in the nature of by - products. This distinction has a common - sense meaning but cannot be defined rigorously. The most useful interpretation is in terms of legislative intent. Thus, a river development program may have flood control as its immediate objective but may also have important bearing on the supply of power, on irrigation, or on soil erosion in adjacent areas. Development of defense technology, while aimed primarly at increase defense capacity, may have important side effects on improving technology in the private sector. The space program may be undertaken primarily to explore the stars, but it may also lead to gains in defense technology or technological improvement, in the automobile industry. An education program may be directed primarily at raising the earning power of the student but it may also reduce the need to combat delinquency. In all these cases, indirect or secondary results may be distinguished from the direct or primary objective obviously, the former should be included along with the latter in assessing project benefits. Tracing of the more indirect benefits may be difficult, but they should be included.

**Tangible versus Intangible:**

The term “tangible” is applied to benefits and costs which can be valued in the market, whereas others which cannot are referred to as “intangible” social goods and social costs, as shown in earlier chapter, typically fall into the category of intangible. Thus, the beautification of an area which may result from an irrigation project is on intangible benefit, whereas the increased farm output

is tangible. Moreover, intangible features may arise with regard to certain benefits or costs, such as health or loss of life which are private in nature but which cannot be readily assessed in money terms. Even though intangible costs and benefits are more difficult to measure, they should nevertheless be included in the analysis.

**Intermediate versus final:**

Another significant distinction is between projects which furnish benefits to consumers directly (since they in value the provision of final goods) and projects which enter into the production of other goods and are thus of an intermediate type. A particular project may in fact provide for both types of goods. Thus weather forecaste may be considered as a consumer good for those who plan an outing, as well as an intermediate good in servicing aviation.

**Inside Versus Outside:**

A final distinction is between benefits and costs which accrue inside the jurisdiction in which the project is undertaken and others which accrue outside. Thus, flood-control measures undertaken on the Connecticut river by Vermont may not only be helpful in Vermont but also prevent floods farther down in the state of Connecticut. The former benefits are internal and the latter are external. They constitute a “spillover” from one jurisdiction to another. Both benefits should be included in assessing the project, but interstate cooperation is needed to do so. This is a matter which we will pursue further when dealing with the economics of fiscal federalism.

**Measurement of benefits and costs:**

In section A the principle of project selection was introduced, based on the simplifying assumption that the dollar value of benefits and costs is known. We must now take a more careful look at the problem of measurement. We consider for the time being the valuation of costs and benefits “when they occur, leaving question of their valuation over time by discounting for later consideration. The question of measurement would be simple if all values could be observed in terms of market prices. But such is not the case costs and benefits are frequently in intangible forms, and even where market prices are observable these may be in need of adjustment because markets are not perfect and distortions must be allowed for.

**Valuation of intangible items:**

We begin with the valuation of intangible (nonmarket) items, a problem which must be solved for many public projects before cost benefit analysis can be applied to them.

**Social Benefits and Costs:**

Project benefits may be essentially intangible, as with the case of national defense, or both tangible and intangible benefits may result. Thus, education yields intangible benefits via cultural enrichment and improved functioning of the democratic process. At the same time, there is a tangible benefit of increased earning power. Similarly, costs may be partly tangible, (e.g. the cost of the resource input into the construction of a superhighway) and partly intangible (e.g., the resulting damage to the beauty of a wilderness area).

Wherever intangible benefits and costs are involved, measurement takes us back to the central problem of social good evaluation. The value of such benefits and costs cannot be derived readily from market prices, and a political process is needed to determine them. Voters must decide how much they value clean air or water or the protection afforded by an addition to national defense cost - benefit analysis is no substitute for this process; it is only a way of choosing among projects after the value of a benefit has been determined. Thus it is most easily applied in those areas where benefits are tangible and there is least need for public provision to begin with.

**Intangible Private Benefits or Costs:**

Related problems arise in connection with benefits and costs which are private in nature (the problem not being one of externalities) but which do not tend themselves to market evaluation. If the government undertakes a cancer research project with resulting reduction in suffering, how can the benefits be valued? How should one evaluate the cost of death and injury which result from highway accidents? What about the benefits of crime prevention? The benefits and costs of some of the most important public project may encounter these more or less insoluble difficulties of evaluation. Yet they must be faced before the mechanisms of benefit - cost analysis can be applied.

In certain cases, indirect valuation methods of a more or less satisfactory nature may be applied to these intangible items and economists have shown considerable ingenuity in developing such procedures. This is illustrated by the following cases.

i) Highways enter as an intermediate good in the services of the trucking industry. As the highway is improved, the cost of trucking falls and so should the prices charged by the trucking firms. The reduced charge to truck users may then serve as a basis for estimating the capital value of the road.

ii) Highway improvement, similarly, will reduce travel costs for individuals, and time saved thereby offers a basis on which to measures the benefit obtained. The personal value of time, in turn, may be derived by observing the differential prices paid for under system of transportation involving differences in travel time.

iii) A school program aimed at reducing absenteeism will be reflected in reduced delinquency and thus save costs of law enforcement. Other gains may be measured in terms of increased earnings due to improved training.

iv) A medical program may result in reducing the death rate from a particular disease. The resulting benefit (or at least part thereof) may be measured by the loss of earnings which is avoided thereby.

v) The value of a park may be measured by the travel-related and other costs which visitors are willing to undergo.

vi) The value of a noise-abatement program for aircraft may be measured by the observed increase in property values adjoining airports.

In these and other forms, “hedonistic prices” may be observed and used to appropriate the market value of apparently intangible project benefits.

**Shadow pricing of market items:**

Returning to projects whose tangible costs & benefits are recorded directly in the market via sale or purchase, no such difficulties arise, provided we deal with competitive markets. In this case, the tangible benefit is measured by the price which the public service fetches in the market, or the price at which a similar service is purchased by consumers from private suppliers. The cost is similarly measured by the price which the government must pay for the product (if the government purchases it from private firms) or by the cost which it must incur (the factor prices which it must pay) if it undertakes the production itself. The cost thus determined will measure the opportunity cost incurred in forgoing the alternative private use of resources.

**Monopoly:**

Matters are more difficult, however, in the case of imperfect markets. Here market prices of outputs do not reflect true resource costs and adjustments are needed. Such adjusted values are referred to as “shadow price”. Thus rental incomes or monopoly profits should not be counted suppose that the market cost of a given product is $1 million but that in a competitive market it would have cost only $900,000, equal to the marginal resource cost of its production. The social opportunity cost in this case is $900,000, not $ 1 million, even though the government pays the higher price. The profit of $100,000 is a pecuniary gain to the monopolist, but not a real resource cost to society. A problem of shadow pricing may also arise in competitive markets where the transfer of a factor to public use raises its price in private use, and the question arises about the price (before or after reduction in private activity) at which the opportunity cost should be measured. A midway value offers a reasonable approximation to the proper result.

**Taxes:**

A further need for adjustment arises in connection with taxes. If the government purchases input needed in the construction of a project, the market price may include sales or excise taxes. This tax component of the price does not reflect a social cost (being merely a transfer from purchases to the government) and should therefore be disallowed in computing the cost of the project. Another major tax-related problem arises in determining the social opportunity cost of capital and, as we will se in the discussion of discounting, the appropriate treatment of taxes on capital income. Once more shadow pricing is needed to correct for the tax.

**Unemployed Resources:**

Another aspect of shadow pricing relates to the costing of otherwise unemployed resources. The cost to be accounted for in public resource use is the lost opportunity for putting these resources to alternative uses, whether they are other public projects (in the fixed budget context) or private projects (in the open budget setting). This reasoning breaks down if the resources are otherwise unemployed and the opportunity cost is zero. Thus, it might be argued that public works are costless in a period of unemployment or may even be beneficial beyond their own value in that they create additional employment via multiplier effects.

This argument is correct as far as it goes. Using unemployed resources poorly may indeed be better than not using them at all. But it is not as good as using them for a superior purpose. Unless

there are political constraints which permit only one use, cost benefit analysis should apply the concept of opportunity cost even where resources are unemployed. Otherwise their employment in a superior alternative is impeded.

But though unemployment is no excuse for failing to evaluate the merits of alternative uses, employment effects of particular projects becomes relevant to benefit evaluation if alternative policies to deal with unemployment are not available. The resulting gain in employment is then an additional benefit, or the opportunity cost of labor is zero. Project A may be preferred to project B even though its intrinsic merit is less, provided that the superior effect on employment outweighs the latter short fall. Thus, building a road in location X may be superior to doing so in location Y if x has a high unemployment rate while Y does not, even though benefit calculus in the absence of employment effects would point to Y. such is the case provided that alternative ways of dealing with unemployment in X are not available. This may be so because unemployment is of a regional nature and not amenable to reduction by stabilization policy on a national scale. If alternative approaches, such as relocation, are available, cost-benefit analysis should compare policy packages, e.g. road construction in Y plus relocation of manpower from x, with road construction in X. to put it differently, efficient policy planning has to be on a isolated consideration of specific policy tools or projects.

**Developing Economics:**

The problem of shadow pricing assumes particular importance in developing economies where government investment and project evaluation frequently play major roles. Consider the pricing of labor in a labor surplus economy. Whereas labor is typically unemployed or underemployed in the traditional sector of the economy, labor costs in the developed sector may be subject to institutional forces which push them well above their competitive level. In such a situation, it becomes desirable in project evaluation to use a shadow price for labor substantially below its market price.

Another aspect of shadow pricing which is often important in developing countries relates to the exchange rate. If the local currency is overvalued, as is frequently the case, both imports and exports will be undervalued relative to that of domestic goods. One of the implications is that imported capital goods are cheap relative to domestic inputs, especially where labor is overvalued. In consequence, an excessively capital - intensive method of production is encouraged. Once more, proper project evaluation will apply a corrected or shadow price for the market rate of exchange, reflecting its value in the absence of measures to support it.

**Cost Effectiveness Analysis:**

In some instances use of cost-benefit analysis may not be feasible simply because a precise measure of benefits cannot be obtained. This still leaves open the more limited task of cost effectiveness analysis. That is to say, a comparison may be made between the costs of achieving the same outcomes by different procedures. As noted below, this technique proved of special importance in the evaluation of weapons systems.

## 6.6 PRINCIPLES OF SHADOW PRICING:

Much of applied welfare economics as it has developed over the past several decades in values the study and use of shadow prices. Since (as we will argue) shadow pricing means different things to different people, we start with a rather broad definition and spend some time talking about various alternative meanings of the term. Shadow pricing will refer to the study and use of first order welfare impacts associated with changes in the levels of particular goods or groups of goods.

This definition accords with the common view of what a shadow price stands for. Suppose we are talking about some input to a government project. Then the shadow price ought to be opportunity cost per unit of increasing this input, and the correct way to measure this opportunity cost is by the marginal welfare foregone. Ambiguity in the definition arises over the ceteris paribus assumptions. When an input level is changed, what other things are allowed to change with it? Obviously something must change in equilibrium or else an infeasibility develops. And it would be silly to ignore this infeasibity here since the shadow price would be zero if we do.

On the other hand, issues of feasibility frequently are ignored when economists discuss shadow pricing of outputs. We ask, for example, what would an extra unit of public good be worth if ‘we had it (and never mind how would we get it?) Here shadow price becomes a hypothetical construction. Having obtained such a price, we proceed by comparing it with the least cost way of actually getting on extra unit.

Clearly, it is useful to think differently about the ceteris paribus assumption in different contexts. But to avoid confusion, it is important to be precise about the assumption in each context. We can achieve precision by specifying which variables are to be treated as independent in each instance. We may choose these so that there are feasibility constraints linking the independent variables in which case the associated shadow prices will be

hypothetical. Having chosen the independent variables, the model must be specified in such a way that equilibrium determines all remaining (dependent) variables as functions of the independent ones.

We will look at several different ways of making these specifications momentarily. However, before doing so, we contrast the general approach of shadow pricing from that of decomposition and point out some potential pitfalls in the interpretation of shadow prices. Let us refer back to the fundamental decomposition equation (for simplicity, we assume in this chapter that collective goods do not affect firms) :

$$\frac{\nabla_\alpha Z}{\beta} = \Omega \nabla_\alpha g = Q \nabla_\alpha b + \sum_{n}{}' \frac{\beta^h - \bar{\beta}}{\beta} \left[ \frac{\nabla_\alpha v^h}{\lambda^h} \right] + Q \nabla_\alpha Y + t \nabla_\alpha C \qquad (1)$$

Since the change in government inputs is weighted by producer prices in this expression, one might be tempted to say that the shadow price of such government inputs should be producer prices. But that statement would not have any operational meaning here; there is no consistent specification of independent and dependent variables under which b is independent with Q as its shadow price vector. For those statements to be true, the variables C and Y would have to be independent, was derived under the assumption that C, Y, and b were linked through market equilibrium conditions.

We can see that Q has no operational significance as a shadow price another way. Using market identities, we can write.

$$t \nabla_\alpha C = t \nabla_\alpha Y - t \nabla_\alpha b \qquad (2)$$

Substituting (2) into (1) and recombining terms, we have the alternative decomposition.

$$\frac{\nabla_\alpha Z}{\beta} = \Omega \nabla_\alpha g = P \nabla_\alpha b + \sum_{h}{}' \frac{\beta^h - \bar{\beta}}{\bar{\beta}} \left[ \frac{\nabla_\alpha v^h}{\lambda^h} \right] + Q \nabla_\alpha Y + t \nabla_\alpha Y \qquad (3)$$

Now the government inputs are weighted by consumer prices. Should we conclude that they now serve as shadow prices? Obviously not unless we enuisage a context in which b changes whereas none of the other terms in the decomposition will. Again, it is clear that there is no such consistent context.

None of this is meant to denigrate the decomposition approach. Indeed, it may be particularly helpful when we are concerned with practical measures (we have reasonable information on changes in tax revenue and profits). Also, it serves

the purpose of breaking complex expressions into simple component parts each of which has a natural interpretation. And as we shall see momentarily, legitimate shadow prices frequently can be expressed in terms of such component parts.

**Categories of shadow prices :**

Let us distinguish several potentially useful contexts for shadow pricing. First, think of the "most hypothetical" situation already mentioned. We ask what the extra unit would be worth if we had it, and nothing else in the system changed call this the detached context. As examples, the detached shadow price of 9 will be $\Omega$, that for b will be 0, and that for agenth's leisure (in a quantity - constrained situation) will be $\partial V^h / \partial L$. At the other extreme, suppose we think of a planning problem in which all relevant constraints have been built in. The parameters of such a problem are the independent variables, and arguments chosen by the planner are the dependent variables, call this the programming context. Each parameter has a programming shadow price equal to the rate at which the objective will go up per extra unit of that parameter. When the programming problem appears in the form.

$$\max_x Fo(x) \qquad (4)$$

Subject to

$$Fi(x) \leq ri, \quad i = 1,.....,k,$$

Where the r's denote exogenous parameters, the interpretation of a programming shadow price should be very familiar. For a particular ri it will be the Lagrange multiplier associated with the ith constraint more generally, programming shadow prices can be computed from the envelope theorem, as indicated in earlier chapter.

It is worth noting that the welfare measures of the previous chapter can be interpreted as programming shadow prices. The context is one in which project scale $\alpha$ is thought of as the only parameter. Assuming that have could set up a programming problem that determined outcomes as a function of scale, $\nabla_\alpha Z$ would be the programming shadow price of $\alpha$.

Among the many other possible consistent frameworks for defining shadow prices, there is one that occupies an important place in public economics. It treats project variables (g, b) as independent with all other variables dependent on them through equilibrium conditions. Note carefully the hypothetical element present here. With g and b thought of as independent, the

equilibrium must be definable as functions of all combinations, not just those that are technologically possible. The shadow price of g; denotes the social worth of an extra unit of g; assuming that it could be had without changing b or any of the other g's.

Since welfare now is to be thought of a function of both b and g, the new context (call it the project context) is that of

$$Z(g,b)=W\left[v^1(g,P(g,b),T^1(g,b),\pi(g,b),\ldots V^N(g,P(g,b),T^N(g,b),TT_N(g,b))\right] \quad (5)$$

Consequently, the project shadow price vector for g is $\nabla_g Z$, whereas that for b is $-\nabla_b Z^2$. The following discussion will concentrate primarily on finding formulas for evaluating project shadow prices.

**Formulas based on second best decomposition:**

How can we calculate partial derivatives of Z in (10 5)? Fortunately, most of the constructions of the previous chapter are still available to us. To see this recall our procedure $\alpha$ and totally differentiated with respect to $\alpha$. Now we simply want to differentiate with respect to g and b separately. And since all budget and material balance conditions are assumed to hold as functions of g and b separately, all substitutions that we made using these conditions before can still be made here. In this process, all demands and supplies are interpreted as reduced - form functions of g and b. We indicate these functions with an asterisk. Thus, $C^{h*}(g,b)=C^h\left[g,P(g,b),T^h(g,b),\pi_h(g,b)\right]$, and similarly for all other behavioral functions. Formulas must look the same as before except that terms like $\nabla_g b$ disappear since g and b are chosen independently. Thus, we have immediately the following two analogs which decompose shadow prices in much the same way as we decomposed the project before.

$$\frac{\nabla_g Z}{\bar{\beta}}=\Omega+\sum_h \frac{\beta^h-\bar{\beta}}{\bar{\beta}}\left[\frac{\nabla_q V^{*h}}{\lambda^h}\right]+Q\nabla_g Y^*+\nabla_g C^* \quad (6)$$

$$\frac{-\nabla_b Z}{\bar{\beta}}=Q-\sum_h \frac{\beta^h-\bar{\beta}}{\bar{\beta}}\left[\frac{\nabla_b V^{*h}}{\lambda^h}\right]-Q\nabla_b Y^*-t\nabla_b C^* \quad (7)$$

These two expressions give us the relevant shadow prices normalized so that they are measured in numeraire units ($\bar{\beta}$ is the average marginal welfare of a dollar). Such a normalization is obviously appropriate if we want to compare shadow prices with market prices (as we do). Given this desired choice of units, the

notation can be simplified somewhat by transforming the welfare function so that the average marginal welfare of a dollar is unity ($\overline{\beta}$ = 1) at the status quo. Then, the welfare shadow prices will be automatically normalized.

Here, we will discuss primarily the shadow price equation for b, leaving the other case to the reader as an exercise. Obviously, there will not be much to say beyond the general sorts of remarks found in previous chapter. Unless we find some simplifying principles, look at special cases, or both. Let us start with some interesting special cases.

Suppose that all changes in exogenous elements of income must be uniform across people. This condition naturally holds if there are no such elements (no lump-sum taxes and no profit income) but could hold more generally (as, e.g., if we have a negative income tax with uniform guarantee). Actually, what we require is slightly weaker, namely, that changes in exogenous income do not correlate with the welfare weights. Such lack of correlation allows us to simplify distributive terms considerably.

To see this, let us write individual utility change in the following additive form:

$$\frac{\nabla_b V^{*h}}{\lambda^h} = -\left\{\nabla_b Pc^h\right\} - \nabla_b T^h + \nabla_b \pi_h \qquad (8)$$

Given our assumptions, the last two terms in this expansion do not depend on h. Therefore, since a covariance is additive in either term and takes the value zero whenever one of its arguments is constant, the distributive term in (7) reduces to

$$\sum_h \left(\beta^h - 1\right)\left[\frac{\nabla_b V^{*h}}{\lambda^h}\right] = -\sum_h \left[\left(\beta^h - 1\right)\left\{\nabla_b Pc^h\right\}\right] \qquad (9)$$

Thus, the distributive term is measured by the correlation between welfare weights and individual consumption of a particular “market basket” of goods; the weights used in making up the market basket are to be determined by the relative price changes induced by increased demand for b.

Let us suppose further that firms behave competitively so that we can ignore the “monopoly profit” term. (The reader is invited to generalize the following discussion so as to include this term). Then the shadow price vector for b takes the form.

$$-\nabla_b Z = Q + \sum_h \left[\left(\beta^h - 1\right)\left\{\nabla_b Pc^h\right\}\right] - t\nabla_b C^* \qquad (10)$$

We are left with two relatively simple “gaps” between shadow price and market price (Q). Both of these have been

discussed in the literature though usually only under even more restrictive assumptions. Verbally the two gaps are characterized as follows. An increase in the government demand for good; will result (through equilibrium adjustments) in a change in the cost of consuming C equal to $\nabla_{b_j} PC$. The shadow price for j should be below the market price to the extent that (cereris paribus) changes in consumer cost correlate negatively with the welfare weights and vice versa. Thus we attach a relatively law shadow cost to inputs whose use drives up the cost of necessities less than it drives up the cost of luxuries.

In addition, shadow price should be below market price to the extent that (ceteris paribus) the induced changes in private consumption increase tax revenue at fixed (exante) tax rates and vice versa. The reasons here are much as in the previous chapter. Expanding a commodity tax base is welfare improving (to a first order), and we want to encourage use of an input that generates such an "externality" of course, to separate out these shadow price influences commodity by commodity, we would need detained information on how the incidence of demand for one good affects equilibrium prices (and demands).

It is tempting to simplify further by making some assumptions concerning the nature of these equilibrium relationships. For example, suppose that increases in the demand for good j led to changes in the price (S) only of good j. Then, in evaluating the gaps, we need look only at private net demands for good j. Unfortunately, such "separability" assumptions are always strong in the present context and are frequently inconsistent. Indeed, if good j is untaxed, increased demand for it by the planner requires that some tax rate (and associated consumer and / or producer price) change elsewhere in the system.

One fairly common interpretation of the second gap involves these in a consistent assumptions. When only one good is taxed (say j), this interpretation would have it that the shadow price of good, should fall short of its market price whenever good j substitutes for the taxed good in consumption. But will substitutability necessary imply $t_i \partial C_i / \partial b_j > 0$? Not necessarily. Recall that $C^*$ (.) is a reduced form demand function that depends on b through all induced price effects. Now, extra public demand for j probably will increase $P_j$, and this would cause an increase in $C_i$ if nothing else happened; but the requirements for new tax revenue will also force an increase in t so P; is likely to rise as well. Since the second consideration will tend to lower demand for I, the net effect is indeterminate.

The likelihood that prices of taxed goods will go up regardless of what else happens (when the government expands marginally) suggest that we should observe a general tendency for project shadow prices to exceed market prices. Indeed, we gave an intuitive arguments for this position earlier in the context of optimal taxation. We will attempt to quantity this bias in sections 4

**Case of tradeable goods:**

Given our discussion at the outset of the chapter, the reader should realize that when new choice variables are “added” to a problem, resultant shadow prices are likely to change. An important example that illustrates this point is provided by the case of tradeable goods. Suppose the planners are always free to export and / or import tradeable goods at fixed world prices (lable these O). Then, regardless of the remaining specification, we can argue that the relative shadow prices of two tradeables must equal their relative world price. And the argument will illustrate a useful way to reason about shadow prices.

Assumes that goods 1 and 2 are tradeable and that the planner is thinking about substituting good 2 for good 1 as public input. He can take one unit of good 1 (that he had been planning to use), export it, and use the proceeds to import some of good 2. At the given world prices, he will be able to import $O_1 / O_2$ units of good 2. note carefully that there transaction can be done without changing any of the dependent variables (such as domestic prices) of the system. Therefore, the only effect of these transaction on Z will occur through the direct changes in b. consequently, he will want to make the substitution unless

$$\frac{\partial Z}{\partial b_1} \geq \frac{O_1}{O_2}\frac{\partial Z}{\partial b_2} \qquad (11)$$

However, he will want to substitute in the opposite direction unless this inequality is reversed, so we must have.

$$\frac{\partial Z / \partial b_1}{\partial Z / \partial b_2} = \frac{O_1}{O_2} \qquad (12)$$

That is, relative shadow prices do equal relative world price. Note that this conclusion requires virtually no assumptions about the nature of distortions in the domestic economy. All that is required is the capacity to make independent choices in the trade dimensions. Of course, as to absolute shadow prices, all we could assert is

$$-\nabla_b Z = kO \text{, some constant k,} \qquad (13)$$

For the vector of tradeable goods. And the determination of k required information about the rest of the system, as we will see.

**Formulas based on optimal taxation:**

A justifiably famous result from the work of Daimond and Mirrless (1971) tells us that when commodity taxation is optimal, firms behave competitively, and all profits are taxed, the relative valuation of private goods by planners should be the same as that by private firms. This result sometimes is interpreted as saying that planners should use Q for shadow prices although this interpretation is not correct given our definition of shadow price.

We derive the diamond - mirrless result by using the structure of the optimal tax framework to simplify the shadow pricing formula (7) to that end, we will utilize one degree of freedom in this (and other) equations, which was noted in chapter but has been suppressed since our initial decomposition was formed by adding and subtracting an arbitrary constant times the sum of the normalized marginal utility changes. Subsequently, we set this constant equal to the average welfare weight in order that the "distribution" term be zero when the welfare weights happen to be equal. However, new that we no longer care about the structural decomposition, there is no reason to stick to this choice : Therefore we rever to the more general form of

$$-\nabla_b Z = k\,Q - \sum_h \left(B^h - k\right)\left[\frac{\nabla_b V^{*h}}{\lambda^h}\right] - k\,Q\nabla_b\, Y^* - kt\,\nabla_b C^* \qquad (14)$$

Where k is the arbitrary constant.

With no profit income accruing to consumers, b affects private consumption evels only through its effect on consumer prices.

Consequently,

$$\nabla_b C^* = \nabla_p C\left\{\nabla_b P\right\} \qquad (15)$$

Further, our current assumption imply that the next to the last term in (14) and the last two terms in (8) are zero. Substituting (8) and (15) into (14), our shadow - pricing formula becomes.

$$-\nabla_b Z = k\,Q + \left\{\nabla_b P\left[\sum_h \beta^h C^h - k\,C - k\left\{t\,\nabla_p C\right\}\right]\right\} \qquad (16)$$

The last term in our new expression represents a pecuniary externality involving welfare effects through price changes. Since it looks pretty complicated, we have not made much progress yet. But now, we use the fact that k is a free choice and that tax rates are set optimally. The term in brackets is quite familiar from earlier chapter. Recall that optimal taxation can be characterized by the condition that the ratio.

$$\mu^{k} = \frac{\sum_{h}\left(\beta^{h} C_{k}^{h}\right)}{C_{k} + \sum_{j} \ell_{j} (\partial C_{j} / \partial P_{k)}} \qquad (17)$$

Is independent of K for all taxed goods. Thus, if we set k equal to the common ratio ($\mu$ in the notation of chapter) the vector in brackets vanishes identically except in the untaxed numeraire component. Since prices never change in that component / the inner product vanishes, yielding.

$$-\nabla_{b} Z = \mu Q \qquad (18)$$

We have derived the fundamental result of Diamond and Mirrless. The relative shadow prices of any pair of government inputs are equal to their relative producer prices. There is a useful characterization of this result in terms of production efficiency. The government should act to ensure that the marginal rate of technical substitution between all pairs of private goods is the same “inside”. The government as it is elsewhere, consequently achieving aggregate production efficiency.

Among important propositions in economics this is one of the most elegant but least intuitive. Indeed, there is a plausible intuitive. Indeed, there is a plausible intuitive argument against it deriving from the general theory of second best. When second best distortions exist, we except to observe trade - offs among efficiency criteria; therefore, should we not give up a little production efficiency in order to lessen the distortion from indirect taxation? The general rule is yes, but the answer here (as we just proved) is no. The key to this puzzle is an element of independence. With the power to set separate tax rates on each good, the government exerts control over consumer prices independent of the levels of producer prices. Therefore, it is free to pursue first best efficiency in production without paying any price elsewhere. There is a close analogy with the pollution tickets scheme of chapter 5 where control over the aggregate pollution level justified a first best distribution of pollution rights whether or not the aggregate level was chosen correctly.

When shadow values for the collective goods are computed similarly, we again find that all indirect effects through price changes are eliminated so that.

$$\nabla_g Z = \sum_h \beta^h \Omega^h + \mu t \nabla_g C . \qquad (19)$$

Here, there is an extra correction to the pseudoprice reflecting the direct influence of collective goods on the consumption tax base.

Sometimes the “production efficiency” result is taken to mean that the government shadow prices for private goods should be equal to producer prices. However, this is going too far. The absolute shadow price vector is not Q but a constant multiple of Q. We should ask. What does this constant represent, and how should it influence decision making? We can get an illuminating interpretation of it by seeing where it comes from and what it means in the analysis of previous chapter.

Refer back to the original tax problem ( ) since g and b were treated as parameters, this problem defines an indirect welfare function : Z (g, b). and the derivatives of Z with respect to b are precisely the shadow prices we have been discussing (for comparison purposes, we assume a normalization such that $\overline{\beta} = 1$). Thus, we can use the envelope theorem to provide an independent derivation of the shadow price formula. Differentiating the Lagrangian for problem ( ) we find.

$$-\nabla_b Z = \mu Q + \left\{ \nabla_b P \left[ \sum_h \beta^h C^h - \mu C - \mu \left\{ t \nabla_p C \right\} \right] \right\} \qquad (20)$$

Observe that this formula has exactly the same from as (16). The term in square brackets vanishes due to optimal tax rules and $\mu$ (since it represents the lagrange multiplier for the government budget constraint) can be interpreted as the value of a numeraire unit in the government budget relative to the value of a lump-sum unit (which has been normalized to 1), Naturally, if the government raised its money through lumpsum methods, this number would be unity. However, we suggested earlier that money in the government budget may be more costly because the alternative way of raising it involves second - best taxation.

Indeed, the early literature took this result for granted and argued that the government always should hold itself to a higher standard than private firms because government spending imposes an added deadweight loss (which is frequently referred to as excess burden in the literature). Although this conclusion may be generally true empirically, it cannot be proved theoretically. The

difficulty is a common one in problems involving second best (indeed, one that we have encountered before) starting from a position that is already second best, a new project may have a positive or negative effect on marginal deadweight loss; only if the effect is positive will the previous argument be correct.

How can we evaluate the magnitude of $\mu$? Well, we can actually calculate it using formulas (10.17), where the commodity k could be chosen arbitrarily. But for both computational and expositional reasons it is best to multiply each equation in (17) by the associated tax rate of that good, sum them, and then solve for $\mu$.

$$\mu = \frac{\sum_h \left(\beta^h t C^h\right)}{t C + <t, \nabla_p C, t>} \tag{21}$$

Where we have used our notation for a bilinear form in the denominator (see preface for details). This expression for the excess burden markup agrees with measures of marginal deadweight loss derived in the optimaltax literature, except that most authors confined themselves to the case of identical consumers so that the welfare weights were absent. Some went on to generate further decomposition using suctsky conditions, a procedure that is not particularly instructive here since aggregate demand generally fails to satisfy slustsky conditions.

Suppose, for the moment, that welfare weights do not correlate with consumption so that the numerator in (21) becomes tax revenue. Since tax revenue better be positive $\mu$ must then exceed unity if the quadratic form (in the denominator) were negative definite. Several plausible arguments can be given for the required negativity. For example, it is sufficient that all own -price effects are negative and “dominate” cross-price effects. Also, it is enough that substitution effects tend to dominate income effects (the slustsky substitution matrix is always negative definite).

However, it is theoretically possible to find exceptions. Suppose that the optimal tax structure is dominated by a tax on labor the aggregate supply of which is back. Ward bending (due to a strong income effect). Then the own-price contribution of labor to the quadratic form is positive (consumption of leisure goes up as the wage rises). Further, if all extra potential wage income is consumed as leisure, there will be no cross price effects, and $\mu$ can be positive.

Increased taxation actually improves the cost of second best in this example since more labor is supplied (getting us closer to first - best supply) as the tax rate is raised (and the after tax wage

falls). This observation explain why the government would rather raise marginal revenue through the indirect tax system than through lump-sum methods in the example. Actually, our example still is not completely convincing. Since the optimal tax structure cannot be totally dominated by a wage tax-un less labor is in perfectly inelastic supply (in which case it is not backward pending). Rather than try to construct a convincing general equilibrium example it is more instructive to generalize our approach to cover nonoptimal taxation where examples such as the preceding are valid.

However, before leaving the optimal tax framework, we comment on distributional considerations. When welfare weights correct with net tax revenues, marginal excess burden must be adjusted accordingly. The correlation will be negative whenever (1) absolute tax correlate positively with income and (2) income levels correlate negatively with welfare weights. The first condition holds even for quite tax structure (and certainly for all progressive ones); only if we collect more from the poor than the rich, does it fail. The second condition. Obviously involves a value judgment (if our discussion in chapter but one that is commonly held when the conditions do hold, the correct marginal excess burden to apply will be lower than what we would get if we ignore distributional considerations.

The appropriate intuition for this factor is as follows : when costs are measures in numeraire aggregates, they tend to overstate welfare costs. Since a bigger than average share of these costs is borne by households with lower than average welfare weights of course, a similar statements can be made about marginal benefits measured in numeraire aggregates we will return to these issue again in the next section.

**General expressions for marginal cost of government spending:**

We showed in the previous section that when commodity taxation was optimal, the departures of government shadow prices from producer price (on private goods) could be captured by a single number : $\mu$. This observation suggests a different decomposition for evaluating shadow prices in general; if will involve a “markup” number (analogous to $\mu$) that is common to all projects, and a correlation due to the nonoptimal nature of taxation that will differ depending on particular characteristics of the project. To see how this decomposition can be obtained, return to (10-16) (Note that we retain the normalization $\bar{\beta}=1$) :

$$-\nabla_b Z = k\,Q + \left\{\nabla_b P\left[\sum_h \beta^h C^h - k\,C - kt\,\nabla_p C\right]\right\} \qquad (22)$$

This expression is valid whether or not taxation is optimal, although it does require that all pure profits are taxed can assumption we retain here. Recall that k is a free parameter. When taxation is nonoptimal no choice of k will make the vector in square brackets. $[\wedge t(k)]$ identically zero; however, for any particular set of price change direction (SP), we could choose k to make SP $[\wedge t(k)]$ equal to zero. We want SP to be average in some appropriate sense unfortunately, there is no natural concept of "average" for this purpose. However it is not unreasonable to choose SP proportional to the tax vector t; this is the price change we would get, for example, if all taxes were increased proportionally and tax incidence fell entirely on consumers. The practical advantage of this choice is that it leads to the same algebraic representation of the markup as before (21).

We return the focus in the following discussion to the shadow price of the project as a whole. The reader is invited to obtain parallel expressions for shadow pricing government inputs or outputs separately. Further, we drop terms involving the private firms (implicitly assuming no collective goods effects in production and competitive behaviour); they can be added back in with no difficulty. The price change resulting from our $\alpha$ project $(\nabla\alpha P)$ probably will not be exactly proportional to t, in which case the measures we have been developing will need to be corrected with a term that will depend on the difference; $\Delta P = \nabla\alpha P - t$. Indeed, simple substitutions yield the following expression for the marginal value of the project.

$$\nabla_\alpha Z = \sum_h \beta^h \Omega^h \nabla_\alpha g - \mu Q \nabla_\alpha b + \mu t \nabla_g C \nabla_{\alpha g} + \Delta P[\wedge_t(\mu)]. \quad (23)$$

For an average project, the last term in our new decomposition vanishes, and marginal cost $(Q\nabla\alpha b)$ enters negatively with weight $\mu$. All projects are marked up at this common rate, though particular ones may require further "cost corrections". This correction will be positive for projects that increase prices on relatively undertaxed goods [a k for which $\wedge tk(\mu)$ > 0) more than proportionally to the associated tax rate and vice versa.

Marginal benefits haw are measured by welfare weighted Samuelson prices with a correction for complementarities or substitutabilities between public goods and taxed goods since $\nabla_g C$ is computed with prices (and hence all private opportunity sets) fixed, there is always some private goods (s) that is complementary with a given g and at least one that is a substitute; only if the

complementary ones tend to be taxed more heavily than the substitutes do we get a positive correction. Clearly, a project whose public goods outputs were separable in demand from private goods would have no correction. It is note worthy that any complementarity correction must be marked up in the same way as marginal cost.

Let us return briefly to the question of whether there is a bias for or against projects. Since we are thinking here of a general bias, the last term in (23) is irrelevant. Further, since a general project has no obvious net complementarity, the third term is also irrelevant so, using superscript n to index a project that is "neutral" in this sense, we can decompose its net benefits into efficiency and equity terms as follows.

$$\nabla_\alpha Z^n = \sum_h \Omega^h \nabla_\alpha g^n - \gamma^n Q \nabla_\alpha b^n + \sum_h (\beta^h - 1)\left[\Omega^h \nabla_\alpha g^n - \gamma^n \frac{tc^h}{tc} Q \nabla_\alpha b^n\right], \quad (24)$$

Where $r^n = tC/(tC + <t, \nabla pC, t>)$, and it measures the "distribution - free" markup factor. Note that in the new distributional term, marginal project cost is allocated among consumers using weights $\left(tc^h / tc\right)$. These weights represent the relative (statutory) tax payments.

Now, in both distributional and the non distributional terms, $\gamma^n$ acts as a markup to marginal cost. Therefore it seems reasonable to think of $\gamma^n$ as the "marginal cost of government spending". Each unit of actual spending must be treated as if it were $\gamma^n$ units in benefit cost calculations.

We can get some feeling for the relative magnitude of $\gamma^n$ in the case where a single good (say the first) is taxed. In this case, $\gamma^n = 1(1 + x)$, where

$$x = \frac{(t_1)^2}{t_1 c_1} \frac{\partial C_1}{\partial P_1} = \frac{t_1}{P_1} \in_1 \quad (25)$$

Where $\in_1$ is (as usual) the own price elasticity of good 1, (Recall our sign convention, which makes elasticities negative in the normal case.

The "perverse case" discussed earlier now occurs whenever the net demand function is positively slope (or the net supply is backward bending); x will be positive and $\gamma^n$ will be less than unity. However, in the normal case, $\gamma^n$ may be quite than unity. To see this, expand 1 / (1 + x) in a second order Taylor series around x = 0

$$\gamma^n \approx 1 - \frac{t_1}{p_1}\epsilon_1 + 2\left[\frac{t_1}{p_1}\epsilon_1\right]^2 \qquad (26)$$

Suppose the elasticity is - 1 and the tax rate is 0.3. Then the marginal cost of government spending is approximately 1.43; a neutral project would need to show a benefit - to - cost markup at least this large in order to be desirable. Of course, higher (lower) elasticities would lead to a higher (lower) markers, but there seem, to be a strong presumption that government projects should indeed be held to a higher standard than those in the private sector.

Finally, we note that if taxation is sufficiently nonoptimal, $\gamma^n$ might turn out to be much larger. Indeed, if the tax rate somehow approached 100 percent on a commodity having unitary demand elasticity, the preceding calculations would yield $\gamma^n = 4$)

## 6.7 SUMMARY

- Wagner's Law is about the rising public expenditure, presented by German economist Adalph Wagner.
- The Leviathan Hypothesis discusses the reasons of systematic bias in the fiscal system.
- Individual action is constrained by membership in classes and groups, so that the fiscal process is seen as a matter of group interaction.
- Social cost benefit analysis of public projects is very essential for better governance.
- Shadow Pricing refers to the study and use of first-order welfare impacts associated with changes in the levels of particular goods or groups of goods.

## 6.8 QUESTIONS

1) Explain in detail the Wagner's Law of rising public expenditure.
2) Write a detailed note on the Leviathan Hypothesis.
3) What are the fundamentals of project evaluation?
4) Examine the principles of Shadow Pricing.

# 7

# MODULE - 3

# TAXATION AND PUBLIC SECTOR PRICING
# TAXATION - I

**Unit Structure:**

7.0 Objectives
7.1 Meaning and principles of taxation
7.2 Horizontal Equity
7.3 Haig-Simons Income
7.4 Young's prescription for vertical equity
7.5 Excess Burden of Taxation: Partial Equilibrium Analysis
7.6 Excess Burden of Taxation: General Equilibrium Analysis
7.7 Questions
7.8 References

## 7.0 OBJECTIVES

- To understand the basic principles of taxation.
- To discuss the ideal tax base (Haig-Simons income).
- To elaborate the principles of taxation in the context of horizontal and vertical equity.
- To analyze the concept of excess burden applying partial and general equilibrium approaches.

## 7.1 MEANING AND PRINCIPLES OF TAXATION

### 7.2.1 Meaning and characteristics of tax

Tax is a compulsory contribution imposed by a public authority in order to fund the production and distribution of goods and services undertaken by the government in common interest. According to Dalton, "a tax is a compulsory contribution imposed by

a public authority; irrespective of the exact amount of service rendered to the tax payer in return, and not imposed as a penalty for any legal offense."

We can thus enlist the characteristics of tax as:

i) It is compulsory and it is a legal imposition.

ii) Absence of quid-pro-quo, that is, the amount of tax paid may not be equal to the benefit/service received; for example, public roads/parks are used by all, irrespective of the amount of tax they pay.

iii) It involves sacrifice.

iv) Tax revenue collected serves a collective or common interest. It is spent for satisfying broad economic objectives and for social welfare.

v) Tax is collected at regular time intervals.

vi) There is no direct relationship between tax payer and the public authority that imposes the tax, that is, although the tax is levied by the government, it is actually administered by the civil servants or various tax departments.

vii) Tax is a major source of public revenue.

### 7.2.2 Principles of taxation

Principles of taxation refer to the parameters or criteria for arriving at an optimal tax structure. There are 3 principles of taxation and they can be briefly discussed as follows:

i. Absolute equality- It implies equality in tax payment. In this approach, each individual or tax paying unit pays equal absolute amount of tax. The individual's tax liability is simply computed as the total amount of government spending divided by the total number of tax paying units. But equality does not imply equity, that is, fairness in the distribution of tax burdens.

ii. Ability to pay- It determines equity on an equal sacrifice basis. It suggests that all tax payers should bear an equal sacrifice in the payment of tax. Tax should be levied according to an individual's capacity to pay, given by his/her income, wealth etc. This principle leads to 2 tenets or types of equity, namely, horizontal equity and vertical equity.

   Horizontal equity implies that individuals having the same ability to pay, say given by income, must bear equal burden of the tax. Vertical equity implies that people having different incomes, that is, abilities to pay must share different burden of the tax.

iii. Benefit principle- It is a major alternative to the ability-to-pay principle. According to this principle, tax should be related to the extent of benefit/utility derived from the goods or services provided by the government. This principle has the advantage of directly relating the revenue and expenditure sides of the budget. It involves combining efficiency (allocational) and equity (distributional) considerations. This principle is based on quid-pro-quo relationship. A person voluntarily exchanges purchasing power in the form of taxes for the purchase of public goods that provide a certain utility.

## 7.2 HORIZONTAL EQUITY

There are 3 principles of tax design to satisfy horizontal equity and achieve an ideal tax base. These are as follows:

a. People bear the tax burden- The first principle of tax design is that people ultimately bear the burden of any tax no matter what is actually taxed. The interesting question in terms of equity is, which people finally bear the burden of the tax. If the initial point of tax (impact) is different from the final point of burden of the tax (incidence), it implies tax shifting. If tax is passed from producers to consumers of the final product through prices, it is called 'forward shifting' and if it is passed by the firm to labour or stockholders of the firm, it is called 'backward shifting'.

b. Individuals sacrifice utility- The second principle is that individuals ultimately sacrifice utility when they pay taxes, so that the ideal tax base would be individual utility levels. Feldstein thus defined horizontal equity as: Two people with the same utility before tax must have the same utility after tax. In the same way, vertical utility would mean that if a person 'i' has greater utility than another person 'j' before tax, then person 'i' must have greater utility than 'j' after tax.

c. Ideal tax base as a surrogate measure of utility- Taxing utility is impossible, but it serves as a goal to be achieved. So in place of utility, another tax base that will be a practical best substitute of utility must be chosen. So it implies that if two people have the same "ideal" tax base, then they must pay the same amount of tax.

**Vertical equity**

Once the ideal tax base is determined, a tax structure to attain vertical equity can be designed. In this context, alternative tax rate systems can be discussed. Let:

$Y_i$ = value of the ideal tax base for individual i

$T_i$ = burden of the ideal tax on individual i.

The tax rate system is progressive if $T_i/Y_i$ increases as $Y_i$ increases.
It is proportional if $T_i/Y_i$ remains constant as $Y_i$ increases.
It is regressive if $T_i/Y_i$ decreases as $Y_i$ increases.

Degressive system is a combination of progressive and proportional, where $T_i/Y_i$ initially increases and then remains constant as $Y_i$ increases. It is accepted that the progressive tax rate system is suitable to achieve vertical equity, given that income or wealth are the tax base.

The two traditional principles of vertical equity in taxation are as follows:

i) Minimum aggregate sacrifice- This was put forward by the utilitarian school led by Jeremy Bentham, which believed that the economic goal of society should be to maximize aggregate happiness or utility. The corresponding utilitarian tax policy was thus, to minimize the aggregate sacrifice from collecting the taxes. Under the assumptions of identical tastes and diminishing marginal utility of income, aggregate sacrifice is minimized by levying taxes top-down, that is, in highly progressive manner till the required total tax revenue is collected.

Suppose there are three groups of consumers whose pretax (before tax) incomes are $Y_1$, $Y_2$ and $Y_3$, with $Y_1 < Y_2 < Y_3$. Assume that their pretax marginal utilities of income are- $\partial \delta U^1 / \partial Y^1 = 10$, $\partial U^2 / \partial Y^2 = 9$ and $\partial U^3 / \partial Y^3 = 8$, reflecting diminishing marginal utility.

If the government wants to collect a given amount of tax revenue, the minimum aggregate sacrifice principle requires that the government tax people in the third group until either their marginal utility rises to 9 or the required tax revenue has been collected. Alternatively, the government taxes both second and third groups until either their marginal utility rises to 10 or the required tax revenue has been collected. This pattern of tax collections is highly progressive in terms of tax burdens.

ii) Equal sacrifice- The other suggestion is equal sacrifice in terms of utility. The question is whether the government should require equal absolute or equal proportional sacrifice. Letting $Y_h$ be pretax income and $T_h$ be the tax for person h, then:

Equal absolute sacrifice will be given by: $U(Y_h) - U(Y_h - T_h) = c$, all $h = 1, \ldots\ldots, H$

And equal proportional sacrifice can be given as: $[U(Y_h) - U(Y_h - T_h) / U(Y_h - T_h) = k$,
all $h = 1,\ldots, H$.
Conclusion: Mainstream economists have not reached a consensus about what constitutes the best surrogate of utility that can be used as an ideal tax base. Income and consumption are the two possible considerations.

## 7.3 HAIG-SIMONS INCOME

Smith or Mill could not come out with an objective ideal tax base from their ability-to-pay principles. Robert Haig and Herbert Simons independently concluded that a certain broad-based measure of income was the ideal tax base. They gave a broad-based measure of income to act as an ideal tax base. It is a comprehensive tax base. Their proposal called "Haig-Simons income" came to be regarded as the best surrogate measure of utility until 1960s, when consumption began to be considered as an alternative.

Haig- Simons argued that purchasing power is the best surrogate measure of utility. This led them to propose that income defined as increase in purchasing power during the year as the ideal tax base for a tax levied annually.

Haig- Simons income Ξ consumption + the increase in net worth.

Consumption is the additional purchasing power actually taken and the increase in net worth is additional potential purchasing power that has been deferred (postponed) for future consumption. Net worth can be increased either by new saving or by increase in the value of the individual's assets existing at the beginning of the year, the individual's capital gains. Thus,

Haig- Simons income Ξ consumption + savings + capital gains

**Or**

Haig- Simons income Ξ personal income + capital gains.

Having determined that Haig- Simons income is the best surrogate measure of utility, horizontal equity is defined as: Two people with identical Haig- Simons income are equals and should pay the same tax. Vertical equity can be defined as: Two people with different Haig- Simons income are unequals and should pay different taxes.

We now elaborate further on the constituents of Haig-Simons income

**Sources of income are-**

i. Income is derived from personal income or capital gains.

ii. Personal income may be earned (wages, rents etc) or unearned (transfer payments, grants).

iii. Income may be received in cash or kind.

iv. Earned income may be derived from labour, capital or land.

**Uses of income could be-**

i. Income could be used for consumption or savings. Both these increase utility. In terms of tax policy, the relevant consideration is increase in purchasing power, whether realized currently as consumption or postponed through saving.

ii. Capital gain may be derived by selling an asset or simply accrues in value without a sale. It is a form of saving. Capital losses must be deducted from income.

iii. Consumption choices are irrelevant, as all types of consumption are increasing utility and viewed as voluntary. These will therefore also include contributions to private charities and tax payments to other governments.

The only legitimate deduction from Haig- Simons income is expenditures necessary for earning the income in the first place, example, business expenditures.

Limitations- Haig- Simons income is a perfect surrogate measure of utility if people have the same tastes, abilities and opportunities; otherwise it is a poor surrogate/substitute. It considers only one of the two variables, namely, labour and leisure that confer utility.

## 7.4YOUNG'S PRESCRIPTION FOR VERTICAL EQUITY

H. Peyton Young introduced a modern and new view about the problem of tax design. He revived the equal sacrifice ability-to-pay principle of vertical equity. Young viewed that if society views broad-based taxes as a necessary evil, a sacrifice made for the common good, then the levying of these taxes ought to be viewed as a cooperative game played by all members of the society. The design problem was meant to put forward a set of sacrifice principles that the society could agree to in levying of a broad-based tax and identify the principles that will apply to such a tax structure.

Young has given six principles that he thought a democratic society could agree to in the levying of a broad-based tax. Young's six principles as the bases for an equitable tax structure can be discussed as follows:

i. The consistency principle- It implies that if a method of taxation is considered to be fair for the entire group of taxpayers, then it must be also considered for any subgroup of the taxpayers. The force of this principle is to ensure that people cannot alter their tax liabilities simply by joining different subgroups.

ii. Monotonicity- The strong version of this principle says that if the government is forced to increase total tax revenues, then everyone's tax liability must increase. The weak version is that if total tax revenues increase, then no individual's tax liability can decrease. This principle captures the spirit of ability-to-pay as a sacrifice.

iii. The composition principle- The method used to raise a given amount of tax revenue must also be used to raise any increment in tax revenue. In other words, society should follow the method that it believes is fair.

iv. Horizontal equity- Two people with equal utility before tax should have equal utility after tax.

v. Vertical equity- There should be no utility reversals. That is, for any two people, the person with higher utility before tax must have higher utility after tax. Principles of horizontal and vertical equity can be stated in terms of Haig- Simons income since it is assumed to be an appropriate surrogate measure of utility.

vi. Scale invariance or the homogeneity principle- Suppose everyone's incomes and the revenue requirement increase by a scalar $\theta$, then everyone's tax liability must also increase by $\theta$. This principle is standard in income distribution theory, where it is applied to measure income inequality.

Conclusion: These principles form a fundamental basis for devising a tax structure that aims at vertical equity.

## 7.5 EXCESS BURDEN OF TAXATION: PARTIAL EQUILIBRIUM ANALYSIS

**Introduction-** Burden of taxation has been a highly discussed topic by economists and policymakers. The burden of tax can be

measured at two points: 1) at the initial point on person/entity it is legally imposed and 2) at the final point on a person/entity the burden ultimately falls. Therefore, the burden of the tax depends on whether shifting is possible. The extent of burden can be represented by the concept of excess burden of a tax.

**Definition-** The 'excess burden' of taxation or deadweight loss from a tax has been defined as an amount that is lost in excess of what the government collects. While the taxpayer experiences a welfare loss when making a tax payment, this loss usually exceeds the value of the tax payment and this is an excess burden.

We can represent the excess burden with the help of the following diagram 7.1.

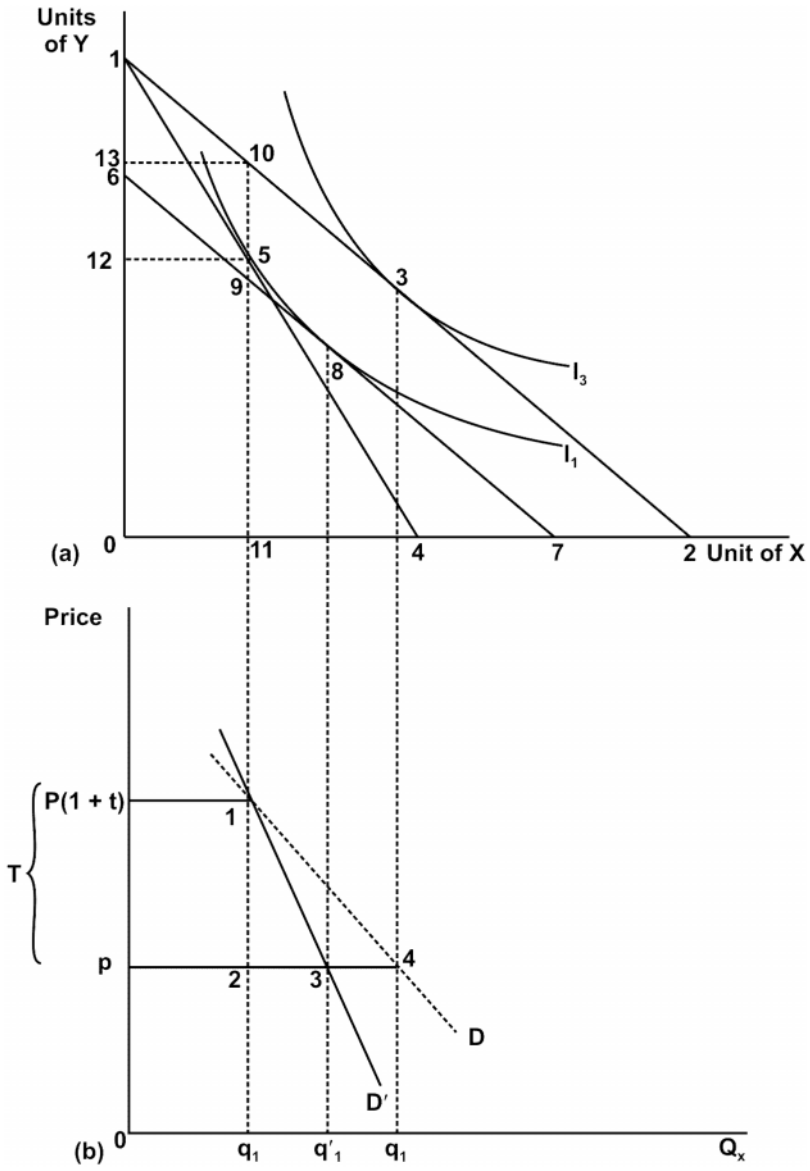



**Diagram 7.1: Excess burden of tax under partial equilibrium analysis**

A selective excise tax (of rate t) is levied on consumption of one particular good X only. Thus we are analyzing the concept of excess burden in a partial equilibrium framework. (a) Panel of the diagram indicates that an individual has a fixed budget which is allocated between goods X and Y. the initial budget line is 12 and the slope of the budget line reflects the relative price of the two goods $-P_x/P_y$. Before taxation the individual chooses a combination of the two goods shown by tangency point 3. This leads to maximization of welfare and is reached at $I_3$. At this tangency point the slope of the indifference curve given by the marginal rate of substitution ($-MU_x/MU_y$) is equal to the slope of the budget line ($-P_x/Py$).

When the excise tax is introduced on good X, at rate t, the relative price of X rises and the budget line shifts along the X axis from 12 to 14. The individual's welfare is reduced to $I_1$ and the new equilibrium is at tangency point 5. At equilibrium point 5, the individual equates the slope of indifference curve $I_1$ ($-MU_x/MU_y$) with the slope of budget line 14 ($-P_x$ (1+ t)$/P_y$.

The equivalent variation of the tax change can be estimated by shifting the budget line 12 backwards in a parallel fashion to 67 until a new tangency point, i.e. 8, located on indifference curve $I_1$. The sum 16 (=10-9) is the amount that can be taken from the individual in the absence of a change. It means that imposition of a selective excise tax leaves the individual with exactly the same welfare as if the change in question had taken place.

In the diagram 7.1 given above, panel (a) if distance 16 units of Y were taken from the individual it would leave the individual with exactly the same welfare as if the selective excise tax has been imposed. This distance 16 units of Y is the same as distance 10-9. However, the tax raised is only 10-5 units of Y and thus by comparison with the equivalent variation of the price change, there is an excess burden equal to 59 (i.e. 10-9 minus 10-5). This excess burden measured in units of Y is the loss of welfare (10-9) from the price change in excess of the tax payment (10-5). This loss, 59, is a 'dead weight loss'.

Now, in order to estimate the welfare loss it is necessary to consider the compensated demand curve. In the diagram above, panel (b), the dashed line D represents the uncompensated demand curve and D' represents the income compensated demand curve. This shows how much individuals would demand of the good at the same real income.

Panel (b) also shows that the excise tax raises the price of good X from P to P (1+t). in the absence of this tax, the amount of money possible to take away from the individual to make the

individual as well off as if tax were levied (the equivalent variation) is the area P(1+t)13P. This is a measure of the welfare loss (consumer surplus loss) that the individual experiences from the imposition of the tax. The triangle 123 is not a transfer; it is a 'dead weight loss'. There is no offsetting gain.

## 7.6 EXCESS BURDEN OF TAXATION: GENERAL EQUILIBRIUM ANALYSIS

Introduction- It considers the resource allocation costs of taxation. This framework also compares the dead weight loss created by selective excise tax and lump sum income tax. 2 goods (X and Y) or markets are considered instead of one in the case of general equilibrium analysis.

**Explanation-**

Consider the diagram 7.2 given below.

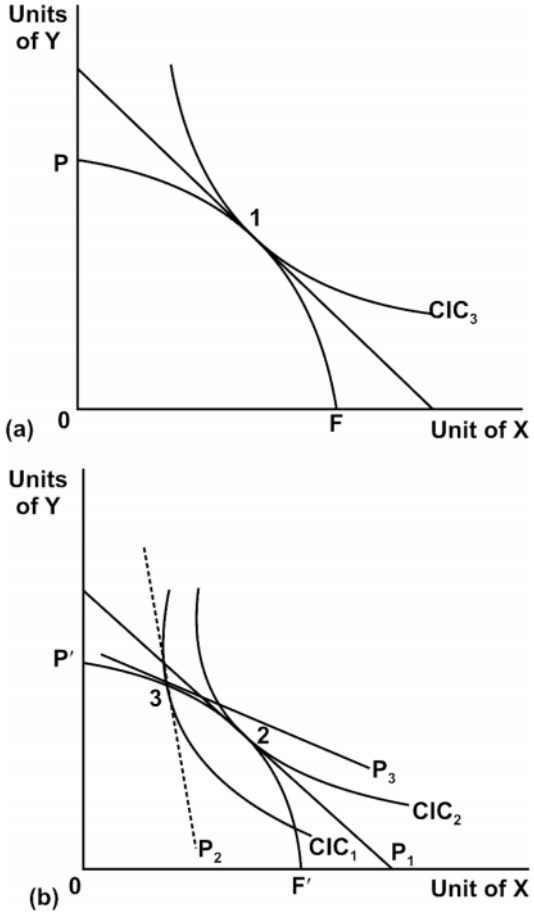



**Diagram 7.2: Excess burden of tax under general equilibrium analysis**

Consider an economy that produces two goods X and Y. in part (a) of diagram 7.2, the production possibility frontier or transformation curve for the economy is given by PF. FOR Pareto optimality, the best output level for the two goods is shown to be 1. This is a combination of the amounts of the two goods produced at which the marginal rate of transformation ($MRT_{xy}$) is equal to the marginal rate of substitution ($MRS_{xy}$), i.e. the slope of the production possibility curve PF is equal to the slope of the community indifference curve $CIC_3$. If there was perfect competition, this would be the position (given by point 1) before the introduction of taxation.

The arrows in part (a) of diagram 3.2 denote that the production possibility curve for the two goods is moved inwards as resources are taken from the private sector. In panel (b) of the diagram, P'F' shows the available resources after taxation, for the provision of private goods. We now need to answer the question that whether a form of tax is used to raise the revenue will affect the extent of excess burden.

If the revenue were raised by a selective excise tax, then the relative prices of the two goods would be affected. The set of relative prices shown by the slope $P_1$ would be altered. The prices that consumers face would be equal to $P_x$ (1+t)/$P_{y.}$ The slope of price line $P_2$ is obviously steeper and as individuals equate the marginal rate of substitution to this set of prices, the new equilibrium would be at 3. At this point the community is at a lower level of welfare, as can be seen from the fact that it is on community indifference curve $CIC_1$ rather than $CIC_2$. The producers set prices equal to the ratio of marginal costs ($MC_x/MC_y$) and this is shown by the transformation curve. Producer prices are thus shown by $P_3$. The added distortion of the tax increases the burden felt by the community, i.e. of being on $CIC_1$ rather than $CIC_2$.

Conclusion- This analysis shows that the selective excise tax created an additional burden compared to lump sum income tax. This conclusion is however subject to certain conditions like, the two goods are substitutes of one another, before tax situation was Pareto optimal supported by perfect competition and there are no externalities operating. Income tax has lower excess burden when the elasticity of compensated supply of labour and savings is low.

Thus the general income tax is superior to the selective excise tax in terms of the lower excess burden it puts on the society.

## 7.7 QUESTIONS

1. Discuss in detail the taxation principles of horizontal and vertical equity.
2. What is Haig- Simons Income? What are the criticisms levied against it as a candidate for being an ideal tax base?
3. Discuss the concept of excess burden in a partial equilibrium framework.
4. Discuss the concept of excess burden in a general equilibrium framework.

## 7.8 REFERENCES:

1. Richard W. Tresch (2002), 'Public finance: a normative theory', Academic Pr.
2. John Cullis, Phillip Jones (2009), 'Public finance and public choice: Analytical perspectives', OUP Oxford.
3. Atkinson A.B. and J.E. Stigliz (1987), 'Lectures on public economics', McGraw Hill.

# 8

# TAXATION - II

**Unit Structure :**

8.0 Objectives
8.1 Optimal Taxation and Income Distribution
8.2 Harberger's tax incidence model :
8.3 Effects of taxation on savings and labour supply
8.4 Questions
8.5 References

## 8.0 OBJECTIVES

- To explore the concept of optimal tax in the context of linear and non-linear income tax.
- To investigate the incidence of corporate income tax and the distortions it leads to in the Paretian framework (Harberger's tax incidence model).
- To evaluate the effects of a tax on various aspects like savings, incentives to work, investment, labour supply etc.

## 8.1 OPTIMAL TAXATION AND INCOME DISTRIBUTION

**Meaning**

Optimal taxation is the term used to describe the design of tax systems to minimize excess burdens and also achieving a socially desirable redistribution of income. The problem is to strike a correct balance between equity and efficiency.

One possibility would be to have lump sum taxes that varied with the ability. Those with great ability who had the capacity to earn large incomes would be faced with a high lump sum tax. Those with lower ability would be required to pay a smaller lump sum tax and those with the lowest ability would be receiving a lump sum transfer from the government.

Assuming that the government does know the distribution of ability in society and is certain about other information like elasticity of labour supply, then it can recommend a schedule of tax rates that provides the optimal balance between equity and efficiency. This approach has its origin in a paper of Ramsey and was revived

in seminal papers by Mirrlees and Diamond. In this approach there are only two goods: a composite consumption good, which can be called net income and the other, is leisure. The approach does not consider the possibility of savings and it applies to both income as well as expenditure taken as tax bases.

### 8.1.1 Optimal linear income tax

This discussion makes the following assumptions or is based of the following framework:

i. There is only form of taxation is an income tax with a constant (linear) marginal rate t.

ii. The only type of government expenditure is on lump sum transfer, LST to households. Thus there are no government expenditures on goods and services.

iii. The government must balance its budget.

iv. There are only two individuals, Mr. Low and Mr. High.

v. There are two goods, a generalized current consumption good called net income, NY and leisure, L.

vi. Both individuals have the same preferences between NY and L.

vii. Mr. High has more ability than Mr. Low and as a consequence Mr. High's wage rate exceeds Mr. Low's wage rate.

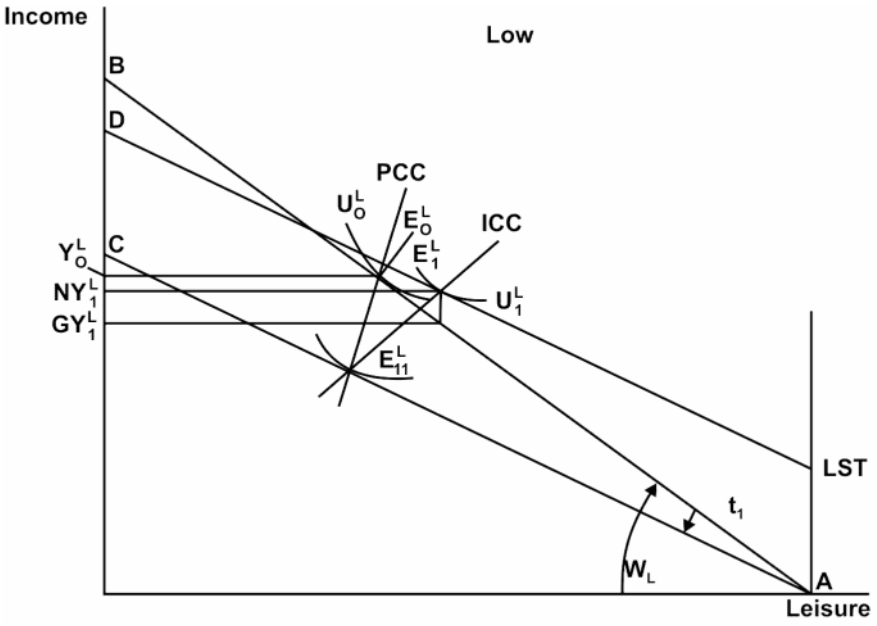
viii. The individuals maximize their welfare subject to their budget constraint.

ix. The government knows the preference function of both individuals.

Each individual's gross income, GY is his gross wage W times the number of hours H that he works, i.e. GY = WH. Net income is gross income less the income tax paid plus the lump sum transfer,

That is, NY = (1 – T) GY + LST. So GDP will be the sum of Low's and high's gross income (i.e. $GY_L$ and $GY_H$) and the total government revenue T is GDP times the marginal tax rate: T = tGDP. As the budget is balanced, tax revenue equals the sum of lump sum transfer, T = LST = $LST_L$ + $LST_H$.

The two person society is represented in the diagram 3.3 that follows. The upper panel in the diagram refers to Mr. Low and the lower panel to Mr. High. If there were no government, each budget constraint, shown by lines AB, would be determined entirely by the wage rates, and the equilibrium would be at $E_0^L$ and $E_0^H$.



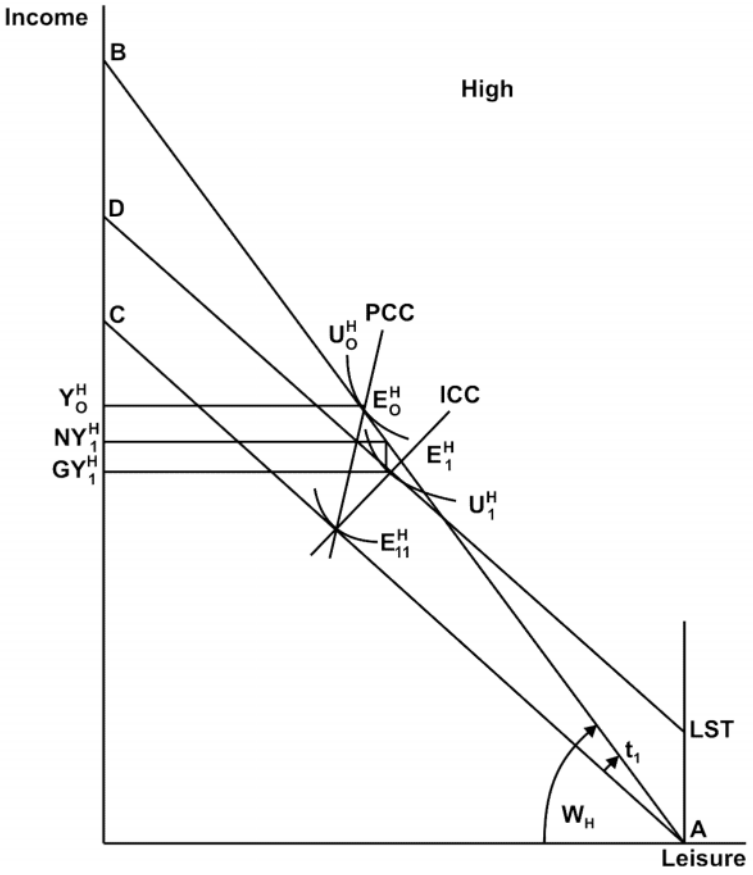




**Diagram 8. 1 Optimal linear income tax**

In the diagram 8.1, AB, AC and DLST are budget constraints without tax, with tax and with tax along with lump sum transfer respectively.

If a tax were imposed at a constant rate without a lump sum transfer, the two individuals would move down their labour supply curves or price-consumption curves PCC, so that if a tax rate t were imposed equilibrium would be at $E_{t1}^{L}$ and $E_{t1}^{H}$.

An increase in the lump sum transfer with a given wage rate will have a pure income effect, causing a movement up an income-consumption curve ICC. If, as the empirical evidence suggest, income and leisure are both normal goods, an increase in the LST will result in an increase in leisure, that is an increase in work and higher LSTs financed by these taxes cause a movement up an ICC decreasing work.

The diagram 8.1 shows the effect of introducing a tax transfer system with a tax rate of $t_1$, which is sufficient to finance a transfer of LST. The after tax budget constraints are thus DLST. Low's grass income falls from $Y_0^{L}$ to $GY_1^{L}$ as he works less, but his net income is now higher than his gross income. Low's welfare has increased and his net income might either fall or rise. The net effect on High's work and his gross income is less certain. In the diagram, the net effect is the reduction in work and in gross income. With only two people in our society the balanced budget condition amounts to requiring that High's net tax paid ($GY_1^{H} - NY_1^{H}$) must just equal Low's net handout ($NY_1^{L} - GY_1^{L}$).

As we increase tax rates we increase work (as drawn in the diagram) and hence we increase tax revenue. But, as we use this revenue to finance high lump sum transfers we decrease work which decreases tax revenue. This suggests that as we raise tax rates, revenue rises at first but that there comes a point when tax receipts start to fall. When this occurs, we have reached the Rawls point- the maximum welfare for Mr. Low. Simulation studies show that total welfare is maximized with lower tax rates than are required by Rawls criterion. This means that the optimal rate depends on society's view of desirable amount of redistribution.

Conclusion- The optimal linear tax thus depends on two factors: the shapes of the individuals' preference maps for net income and leisure, and the society's view about redistribution. The shapes of the preference maps are important because they determine how responsive people are to changes in their budget constraints, and the society's view about redistribution is important because it determines how far redistribution should be carried out.

## 8.1.2 Optimal non-linear income tax

We now examine the arguments for optimal income tax in the more general case where there can be a number of different marginal rates of income tax. This approach is represented in the diagram 8.2 that follows.

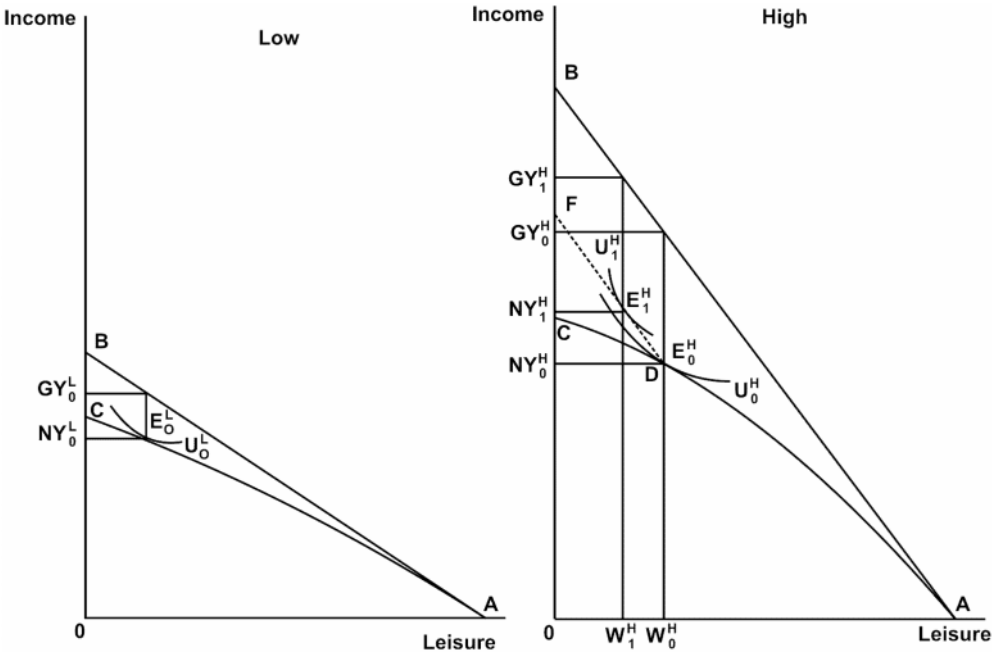



**Diagram 8.2 Optimal non-linear income tax**

The two people are assumed to have the same preferences for income (consumption) and leisure and these preferences are fully revealed to the authorities. However, as a result of the differences in the abilities, Mr. High has a relatively high wage rate, while as Mr. Low has relatively low wage rate. In each case, their gross wage is given by the slope of the line AB. It is also assumed that there is a progressive income tax rate structure which ahs low marginal tax rates on low incomes and steadily rising marginal tax rates as income rise. Mr. Low and Mr. High are in equilibrium at $E_0^L$ and $E_0^H$, given their budget constraints and their preferences.

The total income of the society is GDP = $GY_0^L + GY_0^H$
Disposable income of the society is DY = GDP – T = $NY_0^L + NY_0^H$.
The total tax revenue of the government is: T= $T^L + T^H = GY_0^L - NY_0^L + GY_0H - NY_0^H$= GDP – DY.

If we give equal weightage to High's and Low's welfare, the total welfare of the society is $U = U_0^H + U_0^L$ and the distribution of welfare is given by the relative position of $U_0^L$ and $U_0^H$.

One of the robust findings of the work on optimal taxation is that marginal tax rates on the highest income should be zero.

High's post-tax budget constraint is ADF. With the new segment DF shown as a broken line. It should be noted that with a zero marginal tax rate the new segment DF is parallel to AB, i.e. the gross and net marginal wage rates are equal. It should be also noted that High's total tax liability will be the same on any point on DF as it is at $E_0^H$. However, with the new budget constraint ADF, High's equilibrium would be at $E_1^H$ rather than at $E_0^H$. High's gross income has increased at $E_1^H$ and GDP has risen to $GY_0^L + GY_1^H$. Disposable income has risen to $NY_0^L + NY_1^H$. Tax receipts are unchanged. Total welfare has risen to $U_0^L + U_1^H$. So, using Pareto's criterion we have a clear improvement in welfare.

**Conclusion-** If society places no weight or negative weight on High's welfare and a great weight is placed on Low's welfare then it would not be attracted to a Pareto improvement that increased inequality. Lowering High's marginal tax rate to zero would be more attractive to most people if it were accompanied by more rather than less equality.

## 8.2 HARBERGER'S TAX INCIDENCE MODEL :

The now classic article by Harberger (1962) on the incidence of the corporate income tax introduced the two sector general equilibrium model to the field of public finance. since then his analysis has been extended to various other branches of public finance and to labor economics.[1]

In the Harberger model an imposition of a corporate income tax creates two distortions in Paretian conditions; the one between the marginal rates of technical substitution of the two sectors, and the other between the marginal rate of transformation and the marginal rate of substitution. The first may be called the factor price distortion and the second the commodity price distortion. If we imagine that the tax creates these two distortions in succession, we can decompose the tax impact upon the rate of return into the "factor price distortion (or w)" effect and the "commodity price distortion (or $\pi$)" effect corresponding to the respective distortions.

Among the two effects, the $\pi$ effect is a more indirect and a more subtle concept. Perhaps for this reason all of Harberger's predecessors had neglected the $\pi$ effect in their estimation of the corporate tax incidence. Introduction of the demand side into the two sector general equilibrium model for the first time enabled Harberger to take into account the $\pi$ effect in his estimation of the tax incidence. The $\pi$ effect is a theoretically interesting concept, because its sign is explained by the celebrated theory of Stolper and Samuelson. But its empirical importance in the total tax incidence is unknown.

Let $K_x$, $L_x$ and x denote the amounts of capital, labor, and the output level, respectively, of the x industry. Let $r_x$ be the price of capital this industry faces. Define $K_y$, $L_y$, Y and $r_y$, similarly for the industry y. Harberger assumes that the two industries face the same undistorted price of labor, and chooses the labor as the numeraire so that its price is equal to one. Thus the compensated labor demand function of x industry is written as

$$L_x = L_x (r_x, X) \qquad (1)$$

In similar fashion, we can define the functions $K_x$ ($r_x$, X), $L_y$($r_y$, Y), $K_y$($r_y$, y). The tax we are considering is a tax on the capital used in the corporate sector, and hence we have

$$r_x = r + t \qquad (2)$$

$$r_y = r, \qquad (3)$$

where r is the market price of capital that capital owners face and t is the corporate tax rate.

Assuming that all the factors are fully employed, we have $K_x + k_y = \bar{K}$ and $L_x + L_y = \bar{L}$, where $\bar{K}$ and $\bar{L}$ denote the fixed total amount

Of capital and labor existing in the economy. Substituting (1) and similar equations for $K_x$, $L_y$, and $K_y$, and noting (2) and (3), we have

$$K_x(r+t, X) + K_y(r, Y) = \bar{K} \qquad (4)$$

$$L_x(r+t, X) + L_y(r, Y) = \bar{L} \qquad (5)$$

On the other hand, in the long run equilibrium the profit of each industry must be zero. Thus we have

$$(r+t).K_x(r+t, X) + L_x(r+t, X) = P_x.X \qquad (6)$$

$$r.K_y(r, Y) + L_y(r, Y) = P_y.Y \qquad (7)$$

where $P_x$ and $P_y$ are the prices of x and y, respectively.

In the Harberger model, the government will spend the tax proceeds in the identical manner that the private sector would have spent the taxes, and the pattern of demand remains unchanged by the redistribution of income consumers. Thus he specifies the demand for x as a function of relative commodity prices.

$$X = D(p_x / p_y) \qquad (8)$$

This specification implicitly assumes that the income effect for the commodity x is zero.

The set of five equations (4) - (8) describes a complete model for the five variables $P_X$, $P_Y$, X, Y and r. The solutions for

each of these can be expressed as a function of t. In particular, we write the solution functions for Y and r as

$$Y = Y^*(t) \text{ and} \quad (9)$$

$$r = r^*(t) \quad (10)$$

Our main concern is in characterizing the derivative of the last equation.

Equations (6) and (7) above represent that Cost = Revenue and thus profit = 0. The diagrammatic definition of the decomposition of the tax incidence under Harberger's model into commodity price distortion and factor price distortion is shown in the diagram 3.5 given below.

Diagram 8.3 Decomposition of incidence of tax under Harberger's model

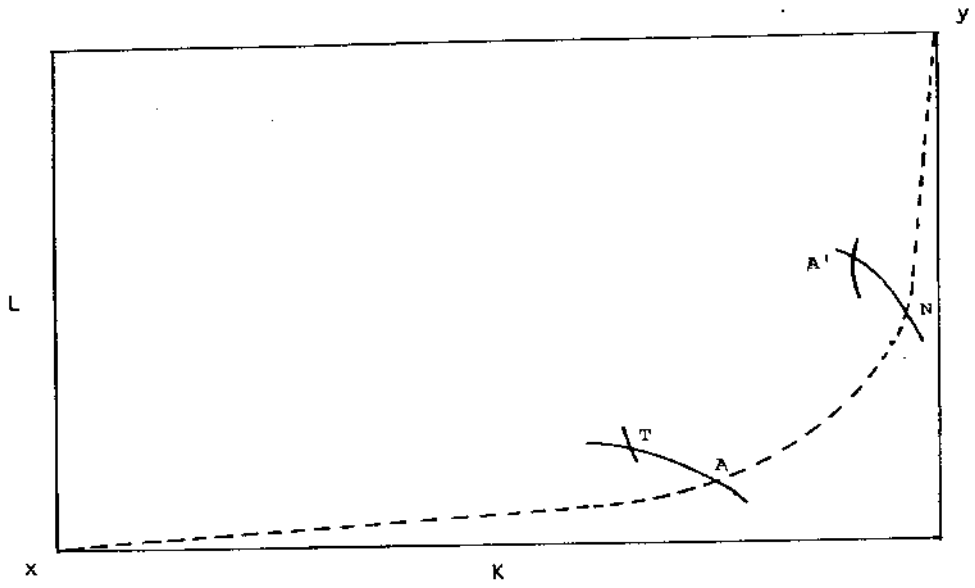



**Diagram 8.3**

The above box-diagram represents the two-sector economy. Assume that point N on the contract curve represents the initial equilibrium with no tax and that point T represents the new equilibrium after corporate tax is imposed on x. When the corporate tax is imposed, the relative price of the capital that the x industry faces is higher than that the y industry faces. Thus, at T the isoquant curves are steeper for x than for y. on the other hand, the output level of y at T is drawn at the level higher than N reflecting the consumer's response to the reduced relative price of the commodity y.

Our concern is in the change in r caused by the shift of equilibrium from N to T. The change in r can be measured by comparing the slopes of the y-isoquant at N and T. To compare

them, it is convenient to decompose the movement NT into two parts. Let A be the intersection of the efficiency locus and the y-isoquant at the output level of the new equilibrium. Then the movement NT may be conceptually decomposed into the two components NA and AT.

We call the change in r associated with the movement from N to A the commodity price distortion (or simply $\pi$) effect, and that accompanied by the movement from A to T the factor price distortion (or simply $\omega$) effect. The $\pi$ effect is equal to the impact upon r of an excise tax on x that would increase the consumption of y by the same amount as the corporate tax t would. Thus, the $\pi$ effect may be considered as caused by the commodity price distortion created by the corporate tax. On the other hand, the $\omega$ effect measures the effect upon r of the factor price distortion created by the tax when the consumption of y is hypothetically kept constant.

The sign of the $\omega$ effect is unequivocally negative as the y isoquant is convex to its origin. The sign of the $\pi$ effect, on the other hand, can be determined by the following proposition :

### 8.2.1 The Stolper-Samuelson Lemma

In the standard two-sector model, an increase in the output of an industry along the contract curve increases (reduces) the relative price of the factor which the industry uses more (less) intensively.

Thus, the $\pi$ effect will be negative or positive according as the x sector is more or less capital intensive. Putting these observations together, we can conclude that the total effect is unequivocally negative when the x industry is the more capital intensive sector, as in the case of Figure 1, while it cannot be signed a priori when the x industry is the less capital intensive sector.

Equations (4) and (5) can be combined to eliminate x yielding:

$$G(Y, r, t) = 0 \qquad (11)$$

From (11), r can be solved for as a function of Y and t :[5]

$$R = R(Y, t).$$

Equation (11) represents the relationship that variables, Y, r, and t have to satisfy if the economy is in full employment, while the y industry is minimizing cost under the prevailing factor prices, and the x industry is doing so under the tax distorted factor prices. Equation (12), therefore, gives the slope of the y-isoquant in the box diagram when t and y are known. When t = 0, this equation identifies the relationship between Y and r on the contract curve.

Let us examine the level of r when the corporate tax is t and the output of the y industry is the one corresponding to A in Figure 1, which is $Y^*$ (t) as defined in (9). Move along the y-isoquant curve passing through A until the angle between the isoquants of the two industries corresponds to t. Then the slope of the isoquant of the y industry at that point is given by R($Y^*$ (t), t). This must be identical to the r given by (10); hence we have the identify.

$r^*$ (t) = R($Y^*$ (t), t) for all t.

Differentiating this with respect to t we have :

$$\frac{dr^8}{dt} = \frac{\partial R}{\partial Y} \cdot \frac{dY^8}{dt} + \frac{\partial R}{\partial t} \qquad (13)$$

[the total effect] [the $\pi$ effect] [the $\omega$ effect]

Where the π effect and the ω effect are is the commodity price and the factor price distortion effects respectively.

The derivative dY*/ dt in equation (13) gives the output change due to the shift from N to T and hence the accompanying change in r is represented by product of δR/δy and dY*/ dt. On the other hand, the movement along the y-isoquant from A to T corresponds to δR/δT. The first and the second terms of the right hand side of equation (13), therefore, give the π effect (commodity price distortion effect) and ω effect (factor price distortion effect). The sign of the π effect depends upon the relative factor intensity of the two industries x and y.

## 8.3 EFFECTS OF TAXATION ON SAVINGS AND LABOUR SUPPLY

Taxation affects the decision-making by households or individuals. It has the following 3 major effects that are discussed as under:

**i.** **Income effect-** Individuals typically make different decisions when their incomes change. Because they are poorer, they postpone their retirement, they cannot enjoy as much leisure, etc. this is known as the income effect. Taxation diverts economic activity from taxed to untaxed areas or from areas with higher taxes to areas with lower taxes. These include leisure, production within the household sector and consumption within the firm sector.

**ii.** **Substitution effect-** The attempts to avoid taxes by substituting non-taxed for taxed activities are called the substitution effects of the tax. These arise where the same real activity can correspond to several different forms of payment, which are taxed at different rates.

iii. **Financial effect-** By providing executives with stock options, firms are able to lower the tax imposed on their managers. The evasion of tax through the cash or the "hidden" economy is a similar rearrangement. The tax system may therefore lead to changes in the form of financial organization and the structure of transactions. These are examples of the financial effect.

iv. **Income taxation and labour supply-** This refers to decisions about hours of work and participation. The basic model of labour supply postulates that an individual's labour supply is a function of after-tax wage and after-tax income from other sources. The labour supply curve is often assumed to be such that, for low wages, an increase in wage rate increases labour supply, but for high wages the labour supply curve bends backwards.

A wage tax reducing the after-tax wage has the effect of decreasing labour supply at low levels and increasing it higher up the scale. On the other hand, a tax on other income (leaving the after-tax wage unchanged) is normally postulated to increase labour supply. Poorer individuals consume fewer goods and less leisure, but consuming less leisure implies supplying more labour. Hence, the supply curve of labour, irrespective of its shape, shifts to the right, i.e. more labour is supplied. An income tax which reduces after-tax wage and other income combines both effects.

## 8.4 QUESTIONS

1. Discuss the relationship between optimal taxation and income distribution for:
   a. Linear income tax and b. non-linear income tax.
2. Discuss the Harberger model of incidence of corporation tax.
3. Examine the effects of taxation on savings and work effort.

## 8.5 REFERENCES

1. John Cullis, Phillip Jones (2009), 'Public finance and public choice: Analytical perspectives', OUP Oxford.
2. Atkinson A.B. and J.E. Stigliz (1987), 'Lectures on public economics', Mcgraw Hill.
3. Tatsuo Hatta and John Haltiwanger (1980), 'A decomposition of the Harberger Expression for Tax Incidence', working paper #221, University of California, Los Angeles.

# 9

# PUBLIC SECTOR PRICING

**Unit Structure**

9.0 Objectives
9.1 Value-added tax (VAT)
9.2 Ramsey Pricing
9.3 Questions
9.4 References

## 9.0 OBJECTIVES

- To overview the meaning and effects of the system of value-added tax (VAT).
- To comprehend the rationale and method of Ramsey pricing (public sector pricing).

## 9.1 VALUE-ADDED TAX (VAT)

**Meaning and need-**

The term “tax on value added” is quite descriptive and sometimes highly misleading. It describes the way in which tax is collected. Each firm pays a tax on the increase in the value of goods that occurs because of the firm’s productive activities; hence the term tax on value-added. If the tax is levied on all firms in the productive process, including retailers, it will be collected upon the total value of the final product. Thus, in its pure form the tax is essentially equivalent to a general single-stage sales tax levied at the retail level. In this sense VAT is not a tax on some new and different tax base. It is simply a different way of collecting a general tax on consumption.

The need for a VAT lay in its features which are as follows:

i. VAT is supposed to be a tax on consumption by firms.

ii. It considers the change in the value of inputs at various stages of manufacturing activity.

iii. It is a tax on sale of goods at various points from raw materials, intermediate and to final goods.

iv. VAT aims at improving the administrative efficiency and economy of collecting taxes on goods.

v. It is a measure to avoid double taxation.

vi. If effective it could also reduce the incentive to avoid paying taxes, i.e. it will reduce the extent of tax evasion.

vii. VAT is said to ne revenue neutral, i.e. it will not decrease the amount of tax mobilized by the government compared to the pre- VAT scenario.

**Effects of VAT-**

Taxes do not just provide revenue to the government; they also affect the economic behavior of the people and the level of economic activities in a significant way. VAT was introduced in India as a part of tax reforms before 1991 in the form of MODVAT (Modified value-added tax). It was simplified under the suggestions of Chelliah committee in 1991. In the post 1991 efforts at tax reforms, VAT was further refined and converted to CENVAT (central value-added tax).

Effects of VAT are felt all over the economy. This is because the tax influences several variables such as savings, investment, employment, distribution, prices and efficiency of resources. Some of these variables are directly affected by VAT, whereas others are indirectly affected.

**The effects of VAT can be discussed as follows:**

**i. Price effects-** The effect of VAT on prices is tremendous and direct. The effect, however, depends upon whether VAT is a new levy (intended to mobilize additional resources) or simply a replacement for the existing taxes to recover the lost revenue from other taxes reduced or replaced by VAT. If VAT is adopted as a replacement for prevailing commodity taxes then the analysis must consider the effects of reduction in prices due to the abolition of the existing taxes. At the same time, it should take into account the price increase due to the introduction of VAT.

VAT would be inflationary if it is shifted forward as the consumer maintains real consumption and accommodative credit policy follows. In fact; it would be necessary to have sufficient increase in wages in order to offset the increase in prices due to VAT so that the consumers can maintain their consumption level. Such an increase in wages will cause inflationary spiral, for the business cost will trigger rise in prices.

Various studies relating to the impact of VAT on prices confirm that the exact effect would depend upon whether it is a new tax

or a substitute for another. The overall experience suggests that this tax is not inflationary. Some studies also suggest that VAT does not have automatic impact on prices, but the degree of impact depends upon the general economic situation and on the other measures taken by the government.

**ii.** Distributional effects- These effects refer to incidence and equity aspects of the tax. The issue centering the distributional effect is that who bears the burden of the tax. In case of VAT, this depends upon various possibilities of shifting, avoidance and tax evasion, as well as on the composition of the users of the commodity.

VAT in comparison to other forms of sales tax stands the test of proper administration because it involves a self-policing mechanism. The buyer of a commodity would take care to obtain an invoice so that he could claim credit for the tax due to the seller. Hence, the equity aspects related to avoidance and evasions are better taken care of by VAT.

The distributional effect, however, depends upon the possibility of shifting of VAT. Normally, profit maximizing firms will shift all commodity taxes forward as these taxes affect marginal costs. But firms which have monopsony in the market for a factor or input will shift VAT backward.

**iii.** Neutrality and Efficiency- The concept of neutrality is concerned with the extent to which the tax avoids distorting the working of the market mechanism. It is thus related with the decision of consumer as to what to buy and how much to buy. Similar issues are related for producers as to how much to produce and what to produce. It is important that these decisions should not be affected by VAT if it is to be neutral.

VAT is designed to be neutral between capital and labour or investment and consumption. Even in respect of foreign trade, the zero-rating keeps VAT neutral. However, Hicks and Joseph, using the technique of ordinal welfare economics, demonstrated that VAT is likely to distort the comparison of benefits and thereby choices via increase in costs and thereby influence the pattern of resource allocation. Such a distortion of choices is known as the excess burden of taxation.

**iv.** Effects on Growth- One of the important objectives of tax policy in a developing country are to increase the rate of savings and investment and to achieve a higher rate of growth. VAT can be viewed as a tax on articles of mass consumption to meet the costs of common benefits. Also, it

can efficiently curb the consumption of luxuries as well as socially undesirable goods. Since capital goods and depreciation on capital are exempt under VAT, this tax does not have any adverse effects on investment. A higher tax rate on luxuries and socially undesirable commodities works as a suitable device for restraining consumption. Thus, VAT is an ideal tax to achieve higher incremental saving ratio and thereby attain higher rate of growth in the economy.

**Conclusion-** VAT being a system of commodity taxation designed to avoid double taxation on inputs being used in the production process as well as on output; it is suitable for adoption of optimal tax structure.

## 9.2 RAMSEY PRICING

**Introduction-** Ramsey pricing is a rule for efficient public utility pricing. It is an analysis of efficient prices in case of regulated firm. First best prices are fully optimal marginal cost prices. Ramsey prices are called the second best optimal prices. They can be defined as uniform prices which maximize total surplus (consumer's and producer's surplus) subject to a breakeven constraint. Uniform prices are those which do not vary with the level of consumption.

**Explanation-**
Symbolically, the most efficient uniform second best prices are those which:
Maximize [CS + PS]
$\{P_1, P_2, \ldots\ldots, P_m\}$
Subject to PS = F where,
CS = Consumer surplus, PS = Producer's surplus, F = fixed input costs.

The question now is how the analyst can compute prices which maximize the total surplus. One way of course is to ask whether or not a given set of price changes is likely to increase total surplus. The price change may be specified by the regulator, the management of the regulated firm or by the analyst.

The methods of finding the Ramsey price that maximizes the total surplus are as follows:

i. **Residential customers-** The monopolist sells M services to a market of residential consumers. The services could be electricity in different times of the day, telephone calls in different mileage bands or different classes of mail service. The utility which we shall assume is a monopolist who changes prices for these services and effects price changes such as $\Delta P_1$, $\Delta P_2$, ….., $\Delta P_\mu$ that are to be considered. The utility is

constrained to break even. The induced change in total surplus of the monopolist is zero and the change in total surplus is simply the change in consumer surplus. For the small price changes, the net effect of a set of price changes is equal to the change in the consumer's bill.

$$\Delta TS = \Delta CS = - X_1 \Delta P_1 – X_2 \Delta P_2 - \ldots\ldots - X_\mu \Delta P_\mu$$

**II**. **Business customers-** It is more difficult to construct such a test when the monopolist is selling to business customers. In such a case, changes in $P_1$, $P_2$, $\ldots P_\mu$ will show up as changes in the business customers' costs of doing business and will affect output prices in the industries in which they sell. These changes in output prices will induce demand changes in the output markets.

To understand the principle of second-best pricing, we begin with a situation in which all prices are set to marginal cost. Where a fixed cost exists the firm will fail to break even. It would seem reasonable to impose high markups in those markets where they will matter least, i.e. where quantity demanded is not too sensitive to price changes. This implies that the regulated firm must increase prices in markets where price elasticities of demand are relatively low. Markups should be relatively lower in those markets with relatively high price elasticities. By following this strategy, markets are altered as less as possible from the price-equal-marginal cost equilibrium, which provided the highest possible value of total surplus.

This suggests that a reasonable formula for the second-best markup of price over marginal cost in each market can be given by:

$$\text{Markup} = P_i – C_i / Pi = \lambda / \varepsilon_i$$

Where, $P_i$ = price in market i,

$C_i$ = marginal cost and $\varepsilon_i$ = price elasticity of demand.

In this formulation, the low elasticity markets get high markups and the highly price-elastic markets get low markups. The proportionately constant λ adjusts markups in all markets uniformly to the point where the firm breaks even.

This pricing rule is also known as the Inverse Elasticity Rule (IER). An alternative way of expressing the IER is that:

$\lambda = [P_i – C_i / Pi]. \mu = [ P_j – C_j / P_j]. \varepsilon_j, j \neq i.$

In other words, for any pair of markets served by a regulated firm, the percentage deviations from marginal cost, weighted by the price elasticities of demand, should be equal for both the markets to the markup λ.

**Conclusion-** The IER was anticipated in 1926 by Frank Ramsey. His work was concerned with optimal excise taxation, and showed that when the effects of change in taxes on the government budget can be ignored, then the tax on each commodity should be inversely proportional to that commodity's price elasticity of demand. Because Ramsey's work is so closely related to public utility pricing theory, prices which maximize total surplus subject to a break even constraint are often called Ramsey prices and the constant λ is called the Ramsey number.

## 9.3 QUESTIONS

1. Write a detailed note on value-added tax (VAT).
2. Elaborate on Ramsey rule in the context of public utility pricing.
3. Write a note on Ramsey number and pricing.

## 9.4 REFERENCES

1. Musgrave, Richard A. & Peggy B. Musgrave (1976), 'Public finance: In theory and practice', McGraw Hill, New York, 2nd edition.
2. Mahesh C. Purohit (2001), 'Value added tax: Experiences of India and other countries', Gayatri publications.
3. Stephen J. Brown & David S. Sibley (1986), 'The theory of public utility pricing', Cambridge University Press.

# 10

# MODULE - 4

# DEFICITS, DEBT AND SOLVENCY GOVERNMENT BUDGET CONSTRAINT

**Unit Structure :**

10.0 Objectives
10.1 Government Budget Constraint (GBC)
10.2 Incorporating GBC in traditional IS-LM model (Christ Multiplier)
10.3 Wealth effect & instability under constrained IS-LM model (Silber model)
10.4 Policy implications of GBC (Solow-Blinder model)
10.7 Questions
10.8 References

## 10.0 OBJECTIVES

- To understand the concept of Government Budget Constraint (GBC).
- To analyze the impact of GBC on traditional IS-LM model.
- To analyze the impact of GBC on traditional IS-LM model including the wealth effect.
- To view the policy implications of money-financed and bond-financed deficit.

## 10.1 GOVERNMENT BUDGET CONSTRAINT (GBC)

**Introduction**

Budget refers to a summarized statement of the government's financial operations or its fiscal activities planned in the future. It is a statement of estimated receipts and expenditures of the public authority to be undertaken in the next, ensuing year. The budget includes two accounts: revenue (recurring items of income & payments) and capital (non-recurring items of income & payments. The budget has two sides: revenue/receipts and expenditure/payments. Budget deficit arises when estimated total expenditure is greater than the total receipts. Symbolically, it can be given as: TE > TR. It is in this context that we are looking at the concept of GBC.

**Meaning of GBC**

The GBC is just the financing rule that a government/budget deficit must be 'covered' by issues of government debt, either demand debt ("money") or deferred obligations ("bonds"). It is one of the gentler paradoxes of literature that recognition that a deficit must be accompanied by some mode of financing, led initially, to the result that the equilibrium fiscal policy multiplier was independent of the mode of financing. Such a fiscal policy multiplier was believed to have a greater value than the one under the traditional IS-LM model.

Certain assumptions made to support this exposition are: i) Prices in the economies considered are fixed, ii) Output is demand determined, iii) Effects of variations in the interest rates on the valuation of wealth is ignored and iv) Effect of incorporating debt service costs in the budget constraint is excluded.

**Explanation**

In a closed economy without a banking system, government expenditure, G, must be financed either by way of taxes or by issuing bonds or creating money. This is represented by the equation:

$G - T = B + M$

Where B = number of bonds issued at a particular face value.

B + M = change in bonds and money supply per period.

With such a constraint, it is clear that long-run equilibrium in the model with wealth effects would require the budget deficit G – T to be zero so that b = m =0. Since we may assume that total tax receipts are a function of national income:

$T = t(Y)$ and $0 < dT/dY < 1$.

Where T = Total tax receipts and Y = national income.

The implication is that, following an initial disturbance or policy change such as an increase in G, the economy reaches long-run equilibrium only when Y has reached a new level where T = G. this is the essence of the GBC.

We now present an algebraic representation of the analysis of fiscal policy under the traditional IS-LM framework.

In the traditional IS-LM model, the IS schedule is derived from the equations representing the real sector and in case of LM schedule from those representing the monetary sector. Full or general equilibrium is derived from the interaction between the IS and LM schedules. The algebraic reduced form for Y at the point of full equilibrium leads to the identification of multipliers.

The real sector giving the IS chedule is given by the following set of behavioural equations:

$C = c_0 + c_1(Y - T)$ ……consumption function (1)

$T = T_0 + T_2Y$ …….. tax function (2)

$G = \overline{G}$ ( $\overline{G}$ with a bar sign on top indicates it is a constant) ……. Government spending function (3)

$I = a - br$ …… Investment function (4)

Where Y = National income, r = rate of interest.

$Y = c + I + G$ …….. Goods market (5)

$S + T = I + G$ …… Equilibrium (5a)

Substituting from (1)-(4) into (5) we get:

$Y = c_0 + c_1 (Y - t_0 - t_1Y) + a - br + \overline{G}$ ……..(6)

Which gives an equation for the IS schedule with Y as the left-hand side variable as

$$Y = \frac{w - br}{Z}, \quad [Y = \text{IS schedule}] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Where z = reciprocal of the unconstrained multiplier, i.e.

$z = 1 - c_1 (1 - t_1)$

and w is an expression in the autonomous factors of the expenditure functions,

$w = c_0 - c_1t_0 + a + G^{-}$

Alternatively, the IS schedule with r as the left-hand side variable is given by:

$$r = \frac{w - Yz}{b} \qquad [\text{IS schedule (r)}] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

The monetary sector is represented by the following behavioural equations and the money market equilibrium:

$M^d = h + dY - er$ money demand function …… (9)

$M_s = \overline{M}$ (M with a bar sign indicates that money supply is a constant) ..money supply function (10)

$M_s = M_d$ money market equilibrium …… (11)

$M_s$ the money stock, is treated as the government's demand debt here.

Substituting from (9) and (10) into (11) gives us

$$\frac{y = \overline{M} - h + er}{d} \qquad \text{LM schedule (Y)} \ldots\ldots\ldots\ldots\ldots\ldots (12)$$

$$r = \frac{h + dy - \overline{M}}{e}$$ LM schedule (r) …….. (13)

Full equilibrium : Full equilibrium of the system requires simultaneous satisfaction of (7) and (12) and equivalently of (8) and (13).

Solving for Y, we obtain by equating (8) and (13), the equation:

$$\frac{y = w}{db = e + z} + \frac{\overline{M} - h}{d + ze = b} = wk_f + \left(\overline{M} - h\right)k_m \text{ …….. ……} \quad (14)$$

Where, $k_f$ is fiscal multiplier and $k_m$ is money multiplier.

Implications of the traditional IS-LM analysis can be spelt out as:

i. The greater the ratio b/e, the greater the value of the money multiplier (km) and less the value of the fiscal multiplier (kf).

ii. In exteme cases, involving 'extreme' values for b and e, one or the other multiplier converges to zero. When $k_f$ converges to zero, $k_m$ goes to 1/d and when , $k_m$ converges to zero, $k_f$ goes to 1/z, the simple unconstrained multiplier.

iii. 'High' value for the ratio b/e implies in a 'high' degree of crowding out of fiscal policy; complete crowding out follows in the traditional IS-LM model. However, this happens only with extreme assumptions about the parameters b or e, i.e. $b \rightarrow \infty$ and $e \rightarrow 0$.

**Conclusion-**

**GBC thus:** i. explains the links between deficits, money and debt, ii. It is essential for understanding the relatrionship between monetary and fiscal policies and iii. Indicates the macro economic effects of fiscal deficit.

## 10.2 INCORPORATING GBC IN TRADITIONAL IS-LM MODEL (CHRIST MULTIPLIER)

**Introduction**

Christ Multiplier describes the effect s of GBC on traditional IS-LM model without considering the wealth effect. Christ's analysis was confined to a money financed deficit, i.e. deficit covered by making additions to money supply.

**Explanation**

Christ (1968) explored the effects on traditional IS-LM analysis of recognising the GBC. Formally, this involves adding the condition that deficit adds to the net total of money plus bonds as given in equation (1) below. Considering the behavioural equations

for the real sector , monetary sector and the simultaneous equilibrium with the financial constraint, we get,

$G - T = \partial (B + M)$ OR $G - T = \Delta B + \Delta M$ …… (1)

Alternately, starting from the initial balance we have,

$\partial G - \partial T = \partial (B + M)$…… (1a)

Bonds donot appear elsewhere in the model, so there is no reason why (1) should be binding for the case of bond financing. Money financing is a different matter. A deficit covered by additions to the money supply will involve shifts in the LM schedule, an effect demonstrated in this Christ's model.

Money supply will involve shifts in the LM schedule, an effect comprehanded by the Christ's model.

Ruling out new open market operations so that if,

$\partial G - T > 0$, then $\partial B \geq 0$ & $\partial M \geq 0$.

And writing p for proportion of the deficit covered by addition to the money stock, we have,

$\partial M = P (\partial G - \partial T)$ …….. (2)

The impact multiplier-

The impact multiplier for a change in government spending financed as to 100p percent, addition to the money stock can be written as:

$\partial Y = k_f . \partial G + k_m . \partial M_s$ …… (3)

Where, $k_f$ and $k_m$ are the fiscal and money multiplier respectively, $\delta M_s$ is the deficit-associated increase in money supply. Substituting for $\delta M_s$ from (2), we obtain:

$\partial Y = k_f . \partial G + k_m . p (\partial G - t_1 \partial Y)$ …….. (4)

Rearranging and dividing the right-hand side top and bottom by $k_f$ gives,

$$\frac{\partial y}{\partial G} = \frac{1 + \frac{b}{e} p}{z + \frac{b}{e}\left(d + p^{t1}\right)} \quad \text{……... ……..} \quad (5)$$

Which with the restriction p = 1 is identical to the result obtained by Christ for the impact effect of a money-financed increase in government spending.

However, the money- financed impact multiplier may or may not exceed the unconstrained multiplier 1/z. the impact effect is more likely to be bigger than the unconstrained multiplier:

i. the smaller the income elasticity of demand for money (d)
ii. The smaller the marginal propensity to consume
iii. The smaller the marginal tax rate.

**Conclusion-** Christ's principal point was that long run equilibrium in the money financing case requires a balanced budget. This holds, no matter what proportion of the deficits is covered by the issue of new money, so long as this is positive. Since money supply will be increasing while the deficit remains and the LM curve will be shifting outwards. Hence, income will nor reach a new long-run equilibrium until

$\partial M = P (\partial G - t_1 \partial Y) = 0$ …….. (6)

**OR**

$$\frac{\partial G}{\partial Y} = \frac{1}{t1}$$ …… (7)

This long run multiplier is unambigiously larger than the impact multiplier or the unconstrained multiplier.

## 10.3 WEALTH EFFECT & INSTABILITY UNDER CONSTRAINED IS-LM MODEL (SILBER MODEL)

**Introduction**

The recognition of the GBC highlights an asymmetry in the IS-LM model, viz., the LM curve accounts for stock of money but does not account for stock of bonds. Ott and Ott (1965) and Silber (1970) attempted to remove this asymmetry by incorporating wealth (that will be given by the term gW) in the money demand function. Thus, the changes in the stock of bonds not offset by an opposite change in the stock of money appear as a change in wealth, and so this affects the demand for money and thereby the LM schedule.

**Explanation**

A wealth term in the demand for money function also requires incorporating it in the consumption function.

We thus, replace the original consumption and money demand functions by the equations:

$C = c_0 + c_1(Y - T) + jW$ …… (1)'

$M_d = h + dY - er + gW$ …… (9)'

(refer original equations 1 & 9 in topic 1 of this unit).

Wealth is defined as

$W = B + M$ ……(a)

The system with behavioural equations augmented by wealth effects, and retaining the budget financing constraint (given by equation 1 in topic 2 of this unit) may now be solved for the impact multiplier.

The differential equation for income may be written as:

$\partial Y = k_f \partial G + k_m \partial M_g - k_m g \partial W_g + k_f j \partial W_g$ …… (b)

where, $\partial M_g$ and $\partial W_g$ are the changes in money stock and in wealth associated with the policy of increasing the government spending, from a position of initial balance, by $\partial$ G. since the change in wealth is identical with the deficit ($\partial G - t_1, \partial Y$) and the deficit is financed as to 100p percent by money creation, (b) can be rewritten as :

$$\frac{\partial Y}{\partial G} = \frac{K_f + K_m P - K_m g + K + j}{1 + K_m p t_1 - K_m g t_1 + K_f j t_1} \text{ ……..…. (c )}$$

Dividing by $k_f$ and noting that $k_m / k_f = b/e$, this reduces to:

$$\frac{\partial Y}{\partial G} = \frac{1 + \frac{b}{e}(p - g) + j}{\frac{b}{e} d + z + t_1 \left[\frac{b}{e}(p - g) + t\right]} \text{ …….. (d)}$$

With $t_1 = 0$, this expression is directly comparable with one of the results quoted by Silber (1970).

If the system is stable, the long run multiplier is again $1/t_1$; adjustment towards this is now assisted by an outward movement of the IS schedule as long as a deficit persists, but hampered by an inward movement of the LM scheduledue to wealth effects.

There are now four elements concerned in the adjustment and in the impact multiplier, viz:

i. The direct (expansionary) effect of the increase in government spending on the IS schedule;
ii. The indirect (expansionary) effect (if p > 0) on the money supply through the financing of the deficit.;
iii. The indirect (concretionary) effect on the demand for money due to the impact of the deficit in increasing wealth and the demand for money;
iv. The indirect (expansionary) effect on the IS schedule of the wealth effect in the consumption function.
v. In view of the opposing effects on the LM schedule ( as noted under ii. & iii. Above), it is possible to imagine, as a special case, a policy of expansionary money financing just sufficient

to offset the contractionary wealth-induced effects on the demand for money. As can be shown from equation (d) money financing provides a larger impact multiplier, and given the long run multiplier, a faster rate of adjustment also.

**Instability-** The impact multiplier (d) need not be positive, and even if it is positive the system need not be stable. The system will not be stable if the balance of the deficit-related effects is negative, i.e. if the sum of the terms in ii, iii and iv mentioned above is negative. If they are negative, any deficitremaining after the initial 'impact' period must tend to reduce income, and any further reduction in income will increase the deficit so adding further net negative (wealth-cum-financing) effect.

**Conclusion-** In the absence of a wealth effect on consumption (j = 0), instability ensues if the increase in demand for money associated with a deficit exceeds the increase in the supply of money associated with it.

## 10.4 POLICY IMPLICATIONS OF GBC (SOLOW-BLINDER MODEL)

**Introduction-**

This model examines the policy implications of wealth effects on GBC. It arrives at a strong paradox that bond-financed deficits lead to greater expansion of income than money-financed deficits. This the model claims happens due to the deficit expanding effects of the service costs of bond finance and due to the incorporation of these costs in the GBC. Solow-Blinder (1973) additionally argue that, the bond-financed fiscal multiplier will exceed the value of $1/t_1$.

**Explanation-**

If the system is stable, long run equilibrium can be found from the government budget financing constraint, rewritten to account for the net (after tax) coupon payments on the bonds issued as a result of the rise in government spending. For the sake of convenience bonds are measured in units such that the coupon is equal to unity (one). Accordingly the budget financing constraint can be stated as:

$$\partial G + (1 - t_1)\frac{\partial B}{\partial G} - t_1 \partial Y = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

so that,

$$\frac{\partial Y}{\partial G} = \frac{1 + (1 - t_1)\partial B_g}{t_1} > 0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Since under money financing $\delta B/\delta G$ (and $\delta B_g$) =0, bond-financed government spending appears to have more expansionary effects than money-financed government spending.

**Conclusion-**

Blinder and Solow (1973), as do others, additionally consider that the inclusion of bond service payments in the financing constraint should be balanced by a corresponding change in the tax function so that disposable income, and hence consumption, will be increased. This would in turn affect the IS schedule and the impact multiplier.

## 10.5 QUESTIONS

1. Discuss the nature of fiscal policy multipliers when the traditional IS-LM model is extended to include: i. The government budget constraint, ii. Wealth effect.
2. "The incorporation of government budget constraint in a closed economy IS-LM model yields fiscal policy multipliers that are larger than their traditional counterparts". Discuss.
3. Discuss the nature of fiscal policy multipliers when the traditional IS-LM model is extended to include wealth effect and comment on the stability of the system.
4. Discuss the Silber model.
5. Write a note on Christ Multiplier.
6. In the context of the Solow-Blinder theorem show how stability of the system implies that bond-financed deficits are more expensionary, not only in the long- run, but also in the short-run.

## 10.6 REFERENCES

1. Michael Artis, Christopher J. Green, Derek Leslie, Graham W. Smith (edited), (1982), 'Demand management, supply constraints and inflation', Manchester University Press.
2. Dornbusch R., S. Fischer and R. Startz (1998), 'Macroeconomics', 7th edition, Mcgraw Hill, New York.
3. Rao M.J.M and R. Nallari (2001), 'Macroeconomic stabilization and adjustment', Oxford University Press.

# 11

# FINANCIAL PROGRAMMING MODEL

**Unit Structure :**

11.0 Objectives
11.1 Financial Programming model
11.2 Seignorage and Inflationary Finance
11.3 The Optimal Inflation Tax
11.4 Olivera-Tanzi Effect
11.5 Questions
11.6 References

## 11.0 OBJECTIVES

- To analyze the interaction between monetary and fiscal policies.
- To examine the effect of inflation on the tax system
- To look at the techniques to reduce monetization of deficit.

## 11.1 FINANCIAL PROGRAMMING MODEL

**Introduction-**

In the context of an independent monetary authority (central bank) that does not give in by monetizing the fiscal deficits, when there is a ceiling on domestic credit and no foreign sources are available, then the government had no option but to adjust (i.e. reduce) the deficit. It can do this either by increasing revenues or decreasing expenditure or both. Financial programming techniques are used to achieve this.

An extension that can be made to the basic financial programming approach is to link the monetary and fiscal accounts through expanding the underlying balance sheet relationships. This is done by discriminating between the expansion of credit to the private sector and that to the public sector, and taking into account the connections between the government budgetary position and official foreign borrowings on one hand and the growth of domestic credit on the other.

The financial programming exercise is often regarded as the one involving the most crucial decisions. The inclusion of fiscal deficits which is presented by this model is an extension of the

basic financial programming approach. The underlying importance of these deficits is that they have to be tackled through fiscal adjustments only.

**Explanation-**

Fiscal policy can be grafted into the financial programming framework in a fairly strightforward manner. To do so, requires three additional ex post identities: first, that the total change in net foreign indebtedness of the country (Δ F) is the sum of changes in the private sector's ($\Delta F_p$) and public sector's ($\Delta F_g$) net foreign debt position, that is

$\Delta F = \Delta F_p + \Delta F_g$ …….. (1)

Second identity is a similar decomposition between the private and public sectors, can be made with respect to changes in domestic credit:

$\Delta DC = \Delta DC_p + \Delta DC_g$ ……. (2)

Where, $\Delta DC_p$ = change in credit channeled to the private sector and $\Delta DC_g$ = change in credit going to the government.

Finally, the government budget constraint (GBC) is introduced, which states that the overall fiscal deficit (Ig – Sg) can be financed either by borrowing from the banking system ($\Delta DC_g$), or by selling debt to the private (non-bank) sector (ΔB), or by borrowing from abroad ($\Delta F_g$), that is

$Ig - Sg = \Delta DC_g + \Delta B + \Delta F_g$ …… (3)

Where, Ig = public sector investments, Sg = public sector savings which is defined as the excess of total government revenue (T) over total government expenditures (G), which includes interest payments, that is Sg = T – G.

These three equations not only directly establish the relationship between monetary expansion and the fiscal position of the government, but indirectly provide the rationale for determining the impact of fiscal policy on growth.

Now, in order to draw the links between fiscal policy and growth, the overall savings-investment balance, can be written in the form of the following equation:

$I = Ip + Ig = Sp + (T - G) + (Z - X)$ …… (4)

This equation (4), indicates that total investment, that is the sum of private investment (Ip) and public investment (Ig), is financed by domestic savings, that is the sum of private savings (Sp) and public savings (Sg = T – G), and external savings (Z – X).

**Empirical evidence-**

While equation (4) suggests a relationship between growth (via investment) and changes in government spending (G) or taxes (T), direct evidence of such a link in developing countries is quite scarce. In standard Keynesian models, a reduction in government expenditure or an increase in taxation is expected to have a multiplier effect on the level of real income, at least in the short run.

While this proposition is well-known, studies have generally found the effect to be statistically insignificant. The lack of positive results is a reflection of the fact that the relationship between the fiscal variables and the level of output in the developing countries is more complicated than basic. Empirical tests tend to suggest that fiscal variables have only a relatively modest role to play.

Besides the demand side, fiscal policy can influence the output through the effects of public sector investment on private sector investment. Despite the many difficulties, both conceptual and data- related, involved in modelling private sector investment behaviour in developing countries, recent studies have identified a positive effect of public investment on private capital formation.

**Conclusion-**

However, the issue of whether a contractionary fiscal policy, taking the form of a cut in real public sector investment, will tend to 'crowd in' private capital formation is not clear. Although the direction of the effect is uncertain, it is apparent that, by varying the level and composition of public investment, the government can alter private investment, and thereby influence the long run growth rate of the economy.

## 11.2 SEIGNORAGE AND INFLATIONARY FINANCE

**Introduction-**

Additional or new money creation causes inflation and affects the real value of the nominal assets. In this context, Seignorage can be viewed as a tax on nominal money holdings of the private sector.

**Meaning and explanation-**

Seignorage is an important implicit tax levied by the government. It consists of the amount of real resources appropriated by the government by means of money creation. With the increase in money supply being denoted by Δ M and the price level by P, real seignorage revenue (as a fraction of real output), is denoted by S*. it can be defined as:

$$S^* = \frac{\Delta M/P}{y} = \frac{\Delta M}{Py} = \mu m = \Delta m + gm + \pi m \quad \ldots\ldots\ldots\ldots\ldots (1)$$

Where: μ = Δ M / Py, that is the rate of growth of money supply or tax rate

m = M / Py, that is real money balances as a fraction of real output or tax base

g = growth rate and

π = inflation rate.

The first two terms in the equation (1), namely, (Δ M / P)/ y and Δ M / Py, are identical and define real seignorage revenue as a fraction of real output. The third term (μm) equates this fraction with the product of the rate of nominal money growth and real balances (as a fraction of real output) held by the public. By analogy with public finance literature, μ is often referred to as the tax rate and m, which is equal to the demand for money balances under the assumption of money market equilibrium, as the tax base. The last expression specifies the value of resources (as a fraction of income) extracted by the government as the sum of three components:

i) The increase in the money-income ratio (Δm)
ii) The increase in nominal money supply needed to maintain a constant money-income ratio in the face of real growth (gm); and
iii) The change in nominal money supply needed to offset inflation effects and maintain a constant money-income ratio (πm).

The last term in this expression (iii) represents the inflation tax as a fraction of real output, ρ. Therefore:

P = πm …… (2)

So that,

S* = Δm + gm + P …… (3)

This implies that in a stationary state, with Δm = g = 0, seignorage will be equal to the inflation tax revenue.

**Conclusion-**

Data in the recent years shows that seignorage has been a negligible source of revenue in almost all industrial countries except Italy. In developing countries, by contrast, seignorage accounts for a substantially higher share of government revenue, especially in India, Pakistan and almost all Latin American countries. Seignorage and the inflation tax also amount to a large fraction of output.

## 11.3 THE OPTIMAL INFLATION TAX

**Introduction-**

While the inflation tax has been long recognized as an important source of government revenue, Phelps (1973) was the

first to emphasize that the inflation rate can be determined optimally in such a context.

**Explanation-**

To prove that inlation rate can be determined optimally, assume that expectations are fulfilled, that is $\pi = \pi^e$, and that the demand for real money balances (M/P) follows a generalized Cagan specification given by:

$M/P = Ay^{\alpha} e^{-\beta\pi}$, $\alpha, \beta > 0$ …….. (1)

In equation (1) above, A = constant,

α= the income elasticity of real money demand

β= the (semi) elasticity of the demand for money with respect to the expected- in this case, actual inflation rate.

By assuming α=1, equation (1) can be rewritten as:

$M/Py = m = A e^{-\beta\pi}$ …… (2)

Combining the equation- P = πm [refer equation (2) in the previous topic of this unit] and equation (2) just above yields:

$P = A \pi e^{-\beta\pi}$ …….. (3)

Equation (3) can be depicted as a curve in π – P space and it is seen that when π = 0, the revenue from inflation tax is also zero. Thereafter, with an increase in the inflation rate, inflation tax revenue rises initially (at a decreasing rate) and then, beyond a certain point, starts falling (at an increasing rate). Maximum revenue is therefore reached when $\partial P / \partial \pi = 0$ and that unique point, the (inflation tax) revenue-maximizing rate of inflation is:

Π(1) = 1 / β …… (4)

Which is the inverse of the inflation rate semi-elasticity of the demand for money.

The above analysis can be easily extended to determine the seignorage-maximizing rate of inflation. Suppose that the economy is in steady-state equilibrium, that is Δm = 0, with a constant growth rate (g), then from the above equations we have:

$S^* = A(g + \pi) e^{-\beta\pi}$ …… (5)

In such a framework, the revenue-maximizing rate of inflation works out to be: π(2) = (1/β) – g. however, as Δm = 0, this implies that the rate of money growth must be equal to the sum of the inflation rate and the real growth rate, that is μ = π + g. Substituting the value of π(2) into this relationship yields: μ* = 1/βwhich is the unique (seignorage) revenue maximizing rate of money growth.

**Conclusion-**

Thus, given the specific assumptions regarding the formation of inflationary expectations, the parameter β can be estimated for individual countries which could provide guidelines for determining an 'optimal' revenue-maximizing monetary policy.

## 11.4 OLIVERA-TANZI EFFECT

**Introduction-**

The Olivera-Tanzi effect investigates the effects of inflation on tax system and points out to a collection lag in tax revenue caused by inflation, in the context of developing countries. An important element that ought to be considered in the debate over the optimal use of inflationary finance relates to the effects of inflation on the tax system, particularly on its links to collection lag in the conventional tax revenue.

This factor, which was first emphasized by Olivera (1967) and more so by Tanzi (1978), has come to be known as the Olivera- Tanzi effect. It plays an important role in the analysis of fiscal, monetary and inflationary dynamics in developing countries.

Taxes are collected with a lag in almost all countries. In developing countries, the average collection lag, defined as the time between the moment taxes are due and the moment that they are actually paid to the fiscal authority, appears to be about 6.5 months for total revenue.

Under such conditions, an increase in the inflation rate will bring a fall in real conventional tax revenue, the extent of which will depend on the average collection lag and the prevalent tax burden, that is, the initial ratio of taxes to aggregate output.

**Explanation of the Olivera- Tanzi effect-**

Formally, let v denote the average lag in the collection of conventional taxes (measured in years), and let π denote the annual inflation rate. The real value of conventional tax revenue (as a fraction of income) is therefore given by:

$t(\pi, v) = t(0)e^{-1\pi}$ …… (1)

Where t(0) denotes the conventional tax rate at zero inflation. In effect therefore, t(0) corresponds to exante tax rate set by the government, while as t(π, v) corresponds to the expost tax rate actually observed as a result of fiscal erosion arising out of inflation and collection lags.

Total government revenue (inclusive of seignorage revenue) measured as a fraction of nominal income, denoted by t, is obtained by combining the equation: $S^* = A(g + \pi)\, e^{-\beta\pi}$ [refer

equation (5) from the previous topic of this unit] and equation (1) given above together yielding:

$t = S^* + t(\pi, v) = A(\pi + g)\, e^{-\beta\pi} + t(0)e^{-v\pi}$ …….. (2)

Setting the derivative of equation (2) with respect to π equal to zero yields the following expression:

$\frac{\partial t}{\partial \pi} = A(1 - \beta\pi)\, e^{-\beta\pi} - Ag\beta\, e^{-\beta\pi} - vt(0)\, e^{-v\pi} = 0$ …….. (3),

which is being a transcendental function of π, does not yield any analytical solution to estimate the unique (total) revenue-maximizing rate of inflation, π(3). However, because $\partial t/\partial \pi < 0$ at, both, π(1) = 1/β as well as at π(2) = (1/β) – g, it is obvious, that this total revenue maximizing rate of inflation, π(3), has to be even lower than the (seignorage) revenue-maximizing rate of inflation. $\pi(2)$, in turn, is lower than the (inflation tax) revenue-maximizing rate of inflation, π(1).

**Conclusion-**

This implies that the fall in the conventional tax rate as a result of inflating the economy beyond, π(3) < π(2)< π(1), will be large enough to outweigh the increase in the seignorage or inflation tax revenue rate, yielding an overall decline in the total tax rate (Tanzi, 1988).

## 11.5 QUESTIONS

1. Explain the techniques to reduce fiscal deficits under the financial programming model.
2. Explain the following concepts:

a. Seignorage and inflationary finance.
b. The optimal inflation tax.
c. The Olivera-Tanzi effect.

## 11.6 REFERENCES

1. Rao M.J.M and R. Nallari (2001), 'Macroeconomic stabilization and adjustment', Oxford University Press.
2. Dornbusch R., S. Fischer and R. Startz (1998), 'Macroeconomics', 7th edition, Mcgraw Hill, New York.

❖❖❖❖

# 12

# TAXATION, INFLATION AND INTEREST RATE

**Unit Structure**

12.0 Objectives
12.1 Tax, Inflation and Interest rates
12.2 Sustainability of public debt or fiscal policy (Ricardian Equivalence)
12.3 Questions
12.4 References

## 12.0 OBJECTIVES

- To overview the relations between tax, inflation and interest rates.
- To discuss the question of sustainability of public debt or fiscal policy.

## 12.1 TAX, INFLATION AND INTEREST RATES

**Introduction-**

In an inflationary environment, the typical tax treatment of interest income and expense- taxing nominal interest received and allowing a deduction for nominal interest paid, can have a significant impact on the level of interest rates and can also affect their volatility. As inflation distorts the base for the tax on interest income and increases the effective tax rates on income in general, this impact is likely to be magnified over time as long as the rate of inflation continues to be high. The tax treatment of interest income and expense may also alter the redistributive impact of monetary restraint and may affect capital flows across countries.

In recent years, the potential effect of taxation on the level of interest rates in an inflationary situation has attracted the attention of some economists. Two major conclusions have come out of the resultant literature. In a situation, where there is inflation and the income tax is imposed on nominal interest income, while nominal interest payments are deductible expenses, taxation should have a positive effect on the nominal rate of interest. To the extent that part

of the interest received is not taxed or part of the interest paid is not tax deductible, this tax effect would naturally be reduced.

**Discussion-**

The various links between taxation, inflation and interest rates can be explained as follows:

**i. Effects of inflation on interest rates in the absence of income taxes**

This discussion is centred aroud the Fisherian equation, which simply states that the nominal or market rate of interest, i equals the sum of the expected real rate, r, and the expected rate of inflation, π. That is,

$i = r + \pi$ …….. (1)

In the absence of expected inflation, the real rate of interest and the nominal rate of interest will be the same. As expected inflation acquires a positive value, the Fisherian hypothesis asserts that, if the expected real rate is constant and therefore independent of the expected inflation, then each percentage point rise in the expected rate of inflation results in a percentage point rise in the nominal rate of interest. This hypothesis is usually expressed as:

$i = r + \beta\pi$, where $\beta = 1$ …… (2)

It must be however, emphasized that equation (2) represents a quite rigid or extreme view of how inflation is likely to affect interest rates.

**ii. Real balance effect**

The real balance effect is associated with Robert Mundell and James Tobin. It postulates a negative relationship between the real rate of interest (r ) and the expected rate of inflation (π). In Tobin's formulation, a rise in expected inflation causes a shift out of money balances and into real capital, thereby depressing the marginal product of capital and the equilibrium real rate of interest. In Mundell's formulation, a rise in the expected rate of inflation reduces the real cash balances of individuals, making them feel poorer. They react by raising the steady state level of saving, thus pushing down the real rate of interest.

**iii. Liquidity effects**

As additional money is injected into the economy, individuals may for a time experience excess liquidity. Thus, before prices and inflationary expectations fully adjust upward, the impact of excess money may lead to a lowering of interest rate. In an economy where inflation has been rampant for some time, this liquidity effect is not likely to be significant.

**iv. Economic activity effect**

The demand for loanable funds is likely to be lower during recessions or during the periods of low economic activity. During this phase many new investments are postponed. On the other hand, it is higher during booms when optimism is prevalent and investment is high. Thus, other things remaining the same, a slowdown in economic activity is likely to pull the rate of interest below the level that would exist if economic activity remained at a "normal" level.

**v. Effects through fiscal deficits**

A more direct effect is through the demand for loanable funds. As the government sells bonds to finance the deficit, the supply of bonds ceteris paribus, increases. The prices of bonds falls and the rate of interest rises.

**vi. Effects of inflation on interest rates in the presence of income taxes**

It is now assumed that nominal rather than real interest income is fully taxed at a marginal rate equal to r and that the nominal rather than real interest expense is fully deductible from the taxpayer's income before the tax is assessed on his taxable income.

It also becomes necessary to make a distinction between a before- tax real rate of interest,r, and an after-tax real rate of interest r*. for simplicity, it is assumed that the tax rate, r, is the same for all tax-payers, that is, the income tax is a proportional tax. If, given the tax rate, r, the net-of-tax expected real rate of interest, r*, is to remain unchanged in the face of a rise in the expected rate of inflation, π, the nominal rate of interest must rise by more than π.

The Fisherian equation must be then modified and rewritten as follows:

$$i = r + \frac{\pi}{1-r} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

In this equation, the effect of π (expected rate of inflation) on i is magnified by the existence of taxes. The higher r (before-tax real interest rate) is, the greater is the impact of π on i (nominal interest rate); r can range between 0 and 1.

In considering the combined impact of expected inflation and income taxation, it is assumed that:

1. Lenders and borrowers agree on a 4 percent real interest rate, r, in the absence of inflation.
2. The effective income tax rate is 25 % and the real rate of interest is independent of the expected inflation and
3. Expected inflation rises from 0 % to 6%.

Equation (3) implies that the nominal rate of interest would have to rise to 12 % in order to maintain the purchasing power of a 4 % interest rate without expected inflation. To be more specific, the lender is paid 12 % of which he pays one fourth, or 3 %, in income taxes and "loses" another 6 % to inflation. He is thus left with an expected after-tax real interest rate of 3 %, which is the same as he would have received in the absence of inflation, but in the presence of income tax.

**Conclusion-**

The various factors that make the nominal rate of interest less responsive to changes in expected inflation are taxes on other assets, taxes on borrowers, tax evasion and capital inflows.

## 12.2 SUSTAINABILITY OF PUBLIC DEBT OR FISCAL POLICY (RICARDIAN EQUIVALENCE)

**Introduction-**

The proposition of Ricardian Equivalence was developed by the classical economist Robert Barro in 1974. The proposition states that private spending can be reduced if an increase in public spending gives rise to an equal tax liability for the private sector in the present time period through tax-financing and in the future due to the need to retire additional public debts caused by the increased public spending.

**Meaning-**

Loosely speaking, the Ricardian equivalence theorem amounts to the following: for a given path of government spending the particular method used to finance the expenditures does not matter, in the sense that real consumption, investment and output are unaffected. Specifically, whether the expenditures are financed by means of taxation or debt, the real consumption and investment plans of the private sector are not influenced. In that sense government debt and taxes are equivalent.

Under Ricardian Equivalence, government debt in the hands of the public should not be counted as net wealth since it is exactly matched by the offsetting liability in the form of future taxation.

**Operation of the theorem of Ricadian Equivalence**

We now examine Ricardian Equivalence under a simple two-period optimizing model of consumption behaviour. Suppose that historical time from the present to indefinite future is split up into two segments. The first segment (called period 1) is the present, and the second segment (called period 2) is the future. There is perfect foresight on the part of households as well as the

government. The households achieve utility by consuming goods in both the periods. Labour supply is exogeneous.

**(a) Household behaviour**

Lifetime utility v is given by:

$$V = U(C_1) + \left(\frac{1}{1+P}\right)U(C_2), P > O \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where $C_t$ is consumption in period t (= 1,2), U(.) is the instantaneous utility function and p is pure rate of time preference, representing the effects of "impatience". The higher the p, the heavier the future utility is discounted, and more impatient the household. At the end of period 0 (i.e. the "past"), the household has financial assets amounting in real terms to $A_0$ over which it also receives interest payments at the beginning of period 1 equal to $rA_0$, where r is the real rate of interest, which is assumed fixed for convenience. The exogeneous non-interest income payments are denoted by Y1 and Y2 , respectively, so that the budget restrictions in the two periods are:

$A_1 = (1 + r) A_0 + (1 - t_1) Y_1 – C_1$, ……. (2)

$A_2 = (1 + r) A_1 + (1 – t_2) Y_2 – C_2 = 0$, ……. (3)

Where $t_1$ and $t_2$ are the proportional tax rates on income in the two periods, and $A_2 = 0$ because it makes no sense for the household to die with a positive amount of financial assets ($A_2 \leq 0$), and it is also assumed that it is impossible for the household to die in debt ($A_2 \geq 0$). Equations (2) and (3) incorporate the assumption that interest income is untaxed.

Now, if the household can freely borrow or lend at the going interest rate r, $A_1$ can have either sign and equations (2) and (3) can be consolidated into a single lifetime budget restriction. Technically, this is done by substituting out $A_1$ from (2)-(3):

$$A_1 = \frac{c_2 - (1-t_2)y_2}{1+r} = (1+r)A_0 + (1-t_1)y_1 - c_1 \Rightarrow c_1 + \frac{c_2}{1+r} = (1+r)A_0 + h$$

…….. (4)

Where the right hand side of (4) represents total wealth, which is the sum of initial financial wealth inclusive of interest received, $(1 + r)A_0$, and human wealth, H:

$$H \equiv (1-t_1)Y_1 + \frac{(1-t_z)y_2}{1+r} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Equation (5) says that the present value of consumption expenditure during life must equal total wealth.

**(b) Government considerations**

In order to demonstrate the Ricardian equivalence theorem, we need to introduce the government and its budget restriction. We assume that the government buys goods for its own consumption ($G_1$ and $G_2$), and finances its expenditure by taxes and / or debt. There is no money in the model and so money financing is impossible. The government, like the household, exists for two periods, and can borrow and lend at the interest rate r.

In parallel with equations (1) to (3) from section (a), i.e. pertaining to households, the government's budget identities are:

$(D_1 \Xi)\ rB_0 + G_1 - t_1Y_1 = B_1 - B_0$, …… (6)

$(D_2 \Xi)\ rB_1 + G_2 - t_2Y_2 = B_2 - B_1 = -B_1$, …… (7)

where $D_i$ and $B_i$ respectively, denote the deficit and government debt in period, (i = 1,2), respectively, and $B_2 = 0$ because the government, like the households cannot default on its debt and is assumed to remain solvent.

Then we can now consolidate equations (6) and (7) into a single government budget restriction:

$$(1+r)B_0 + G_1 - t_1y_1 = \frac{t_2y_2 - G_2}{1+r} \Rightarrow (1+r)B_0 + G_1 + \frac{G_2}{1+r} = t_1y_1 + \frac{t_2y_2}{1+r}, \text{--(8)}$$

Where the left-hand side of (8) represents the present value of the net liabilities of the government, and the right-hand side is the present value of net income of the government (i.e. the tax revenue).

Since government bonds are the only financial assets in the economy, household borrowing (lending) can only take the form of negative (positive) holdings of government bonds. Hence, equilibrium in the financial capital market implies that,

$A_i = B_i$ …….. (9)

For i = 0,1,2.

The first demonstration of the Ricadian equivalence theorem is obtained by solving the government budget restriction for (1 + r)$B_0$, and substituting the result into the household budget restriction (4) taking (9) into account as:

$$C_1 + \frac{C_2}{1+r} = (1+r)B_0 + \left[(1-t_1)y_1 + \frac{(1-t_2)y_2}{1+r}\right]$$

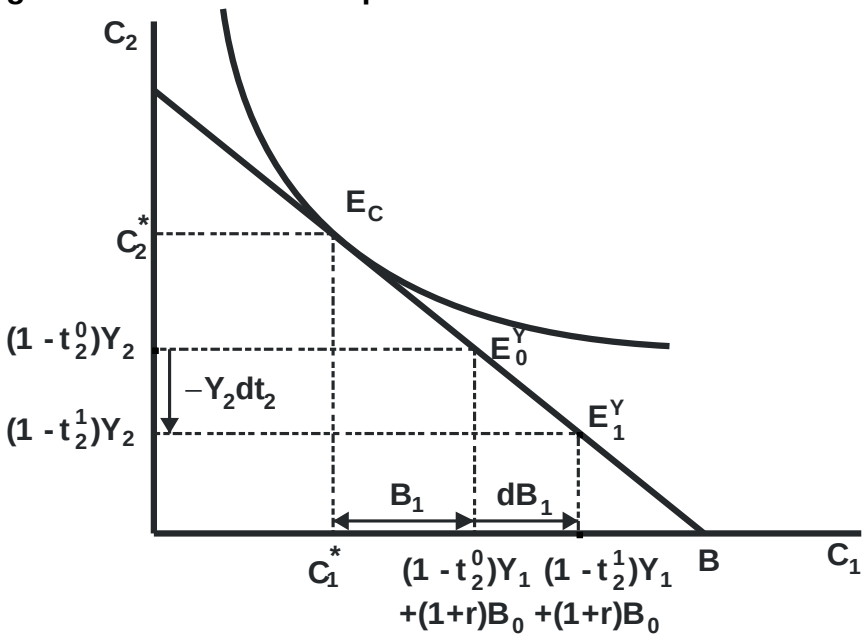
$$= t_1y_1 + \frac{t_2y_2}{1+r} - G_1 - \frac{G_2}{1+r} + (1-t_1)y_1 + \frac{(1-t_2)y_2}{1+r}$$

$$= Y_1 - G_1 + \frac{Y_2 - G_2}{1+r} = \Omega \qquad \ldots\ldots (10)$$

This final expression shows that the tax parameters drop out of the household's budget restriction altogether. Only the present value of (exogeneously given) government spending affects the level of net wealth of the household. Consequently, the choice of $C_1$ and $C_2$ do not depend upon the tax parameters $t_1$ and $t_2$ either. The way in which the government finances its expenditures has no real effects on consumption.

Thus, consumption plans are unaffected by the timing of taxation, but savings are. The reaction of the household is to increase its savings in the first period ($dS_1 > 0$) in order to be able to use the extra amount saved plus interest in the second period to pay additional taxes. Diagram 12.1, illustrates this.

**Diagram 12.1: Ricardian Equivalence**




In the diagram 12.1, the initial endowment point is $E_0^Y$. it represents the point at which the household makes no use of debt in the first period (i.e. $B_1 = 0$) and consumes according to equations (2) and (3). Since the household can freely lend/borrow at the going rate of interest r, however, it can choose any ($C_1$, $C_2$) combination along the budget line AB. Suppose that the optimal consumption point is at $E^C$, where there is tangency between an indifference curve (dv = 0) and the budget line. The optimal consumption levels are given by $C_1$* and $C_2$*, respectively. As a result of the Ricardian experiment, income rises in the first period and falls in the second period, but the net wealth of the household (Ω) is unchanged. Hence, the income endowment point shifts along the given budget

line in a south-easterly direction to $E_1^Y$. The optimal consumption point does not change however, since nothing of importance has changed for the household. Hence, the only thing that happens is that the household increases its savings in the first period and it does so by purchasing more bonds from the government.

**Conclusion-**

Some of the most important theoretical reasons for Ricardian equivalence to fail are:

i. Distorting effects of taxes on labour income
ii. Existence of non-interest incomes
iii. Disparities between lending and borrowing interest rates for government and private sector
iv. Net population growth, due to which burden of tax per capita becomes lower for the future than the present generations.

## 12.3 QUESTIONS

1. Evaluate the inter-relationships between tax, inflation and interest rates.
2. Write in detail on Ricardian equivalence.
3. Examine the sustainability condition of fiscal policy with reference to Ricardian equivalence theorem.

## 12.4 REFERENCES

1. 'Tax treatment of interest' IMF fiscal affairs department, working paper, March 1997. www.imf.org/external/pubs/ft/wp
2. Ben J. Heijdra and Federick van der Ploeg (2002), 'The foundations of modern macroeconomics', Oxford University Press.
3. Dornbusch R., S. Fischer and R. Startz (1998), 'Macroeconomics', 7th edition, Mcgraw Hill, New York.

# 13

# MODULE - 5

# FISCAL POLICY AND THE MACROECONOMY - MACROECONOMIC EFFECTS OF FISCAL DEFICITS

**Unit Structure:**

13.0 Objectives
13.1 Introduction
13.2 Inflation, Money and Fical Deficits
13.3 Adaptive Expectations
13.4 The High Inflation Trap
13.5 Sargent - Wallace Hypothesis
13.6 Summary
13.7 Questions

## 13.0 OBJECTIVES

After having studied this unit, you should be able -

- To understand the relation between inflation, money and fiscal deficit.
- To know the mechanism of adaptive expectations
- To study the concept of High Inflation Trap
- To understand the Sargent - Wallace Hypothesis

## 13.1 INTRODUCTION

The primary fiscal deficit affects the equilibrium rate of growth of the money stock and hence the equilibrium inflation rate. This relation between the inflation, money and fiscal deficit is explained in this unit.

Money financing of fiscal deficit may lead, depending on the mechanism through which expectations are formed and the speed of adjustment of the money market, to multiple steady state equilibrium.

Many economies in the process of liberalization find themselves in serious internal or external macroeconomic imbalances. The high inflation trap is one of them.

Sergeant-Wallace hypothesis led down the rules of implementing monetary policy.

## 13.2 INFLATION, MONEY AND FISCAL DEFICITS:

Consider a closed economy with exogenous output. Suppose that the demand for money function takes the cagan semi logarithmic form used in analyzing inflationary finance in earlier chapter.

$$m = \exp\left(-\alpha \pi^{a}\right), \alpha > 0 \qquad (1)$$

Where $m \equiv M/P$, with M representing the base money stock and P the price level. The expected inflation rate is $\pi^{a}$. The government cannot issue bonds to the public & finances its primary budget deficit d entirely through seigniorage.

$$d = M/P = \mu m, \qquad (2)$$

Where $\mu \equiv M/M$. Combining (1) and (2) implies.

$$d = \mu \exp\left(-\alpha \pi^{a}\right). \qquad (3)$$

Equation (3) specifies how the primary fiscal deficit affects the equilibrium rate of growth of the money stock, and hence the equilibrium inflation rate. However, to the extent that the demand for real money balances is inversely related to the expected rate of inflation, the possibility of multiple solutions to (3) arises. As shown below, and in line with our discussion in earlier chapter, the existence of a “seigniorage laffer curve” implies that there are two steady - state rates of inflation that generate any given amount of seigniorage.

Equation (3) is plotted in figure 1, which is adapted from Bruno and Fischer (1990). Curve D depicts the combination of $\mu$ and $\pi$ for which the primary deficit is constant. Because equation (3) indicates that d = $\mu$ when the expected inflation rate is zero, the deficit is measured by the distance between the origin and the intercept of the D curve on the $\mu$ axis. The government budget constraint is binding at any given moment in time, so that the economy is always located on the D curve. Differentiating equation (1) with respect to time yields, since

$m \equiv m/p - \pi m,$

$$\mu - \pi = -\alpha \pi^{a} \qquad (4)$$

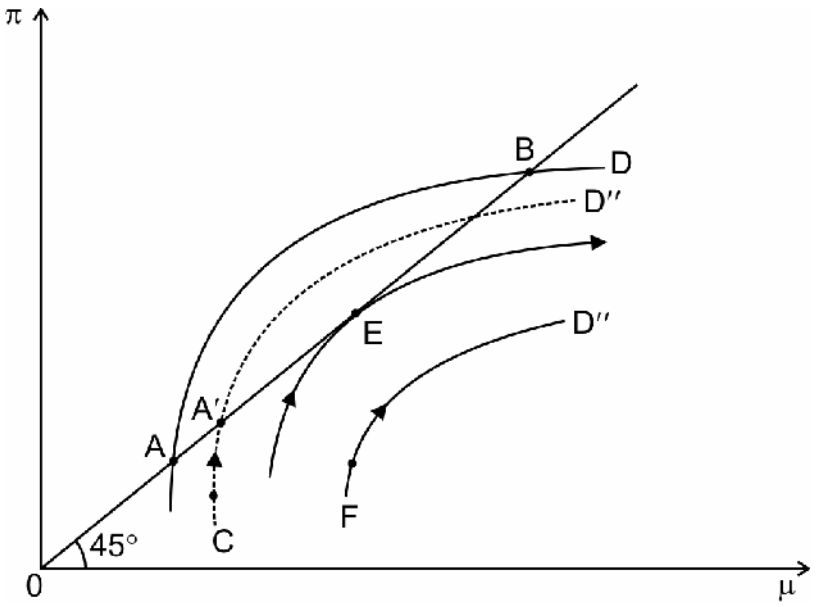



**Fig 13.1 : Seigniorage and Dual Inflation Equilibria so that in the steady state.**

$$\pi = \pi^{a} = \mu \qquad (5)$$

Equation (5) is represented by the 45° line in figure 1. As depicted in the figure, the D curve and the 45° line intersect twice. There are therefore two potential steady - state positions, that is, two inflation rates at which the primary fiscal deficit is financed through revenue from the inflation tax. A low inflation equilibrium (point A) and a high inflation equilibrium (point B). At point A the elasticity of the demand for real money balances is less than unity, whereas at point B it is greater with than unity.

Suppose for a moment that the size of the primary deficit is constrained by the amount of revenue that can be generated through money creation. As shown in earlier chapter, the inflation rate that maximizes steady-state seigniorage revenue is equal to $\pi^{s} = 1/\alpha$ and the corresponding level of revenue is given by

$$d^{s} = \exp(-1)/\alpha$$

Assume now that the primary deficit that the government wishes to finance is fixed at an arbitrary level d. depending on the size of the deficit target, there may be zero, one, or two equilibrium. Because the government cannot obtain more than $d^{s}$ in the long-run equilibrium, there is no steady state if $d > d^{s}$. For $\tilde{d} = d^{s}$ or $\tilde{d} < 0$, there is a unique steady state. For $0 < \tilde{d} < d^{s}$, there are two steady states, and the economy may be "stuck" at the high-inflation equilibrium (point B). To see under what conditions these longrun

outcomes obtain, we consider two alternative assumption about the formation of inflation expectations.

## 13.3 ADAPTIVE EXPECTATIONS:

Consider first the case where inflation expectations are adaptive :

$$\pi^{a} = \beta\left(\pi - \pi^{a}\right), \beta > 0 \qquad (6)$$

Combining equations (4), (3) and (6) determines - together with an appropriate initial condition - the time path of actual and expected inflation, for a given primary fiscal deficit. From (4) and (6), changes in expected inflation are determined by

$$\pi^{a} = \beta\left(\mu - \pi^{a}\right) / \left(1 - \alpha\beta\right)$$

Whereas the actual rate is

$$\pi = \left(\mu - \alpha\beta\pi^{a}\right) / \left(1 - \alpha\beta\right)$$

Which implies that in the steady state

$$\pi = \pi^{a} = \mu.$$

With an adaptive expectational scheme, point A is a stable equilibrium whereas $\beta$ is unstable, if the speed of adjustment $\beta$ is low enough $\left(\text{B.1}/\alpha\right)$. Points located to the right of point B lead to a hyperinflation path. The government prints money at an ever - increasing rate, preventing the expected inflation rate from ever conciding with the actual rate of increase in prices. Although real money balances (the inflation tax base) are reduced at an increasing rate, the pace at which the government is printing money is so rapid that it is still able to finance its deficit.

Suppose that the economy is initially at the stable low-inflation equilibrium (point A), and consider the effect of an increase in the fiscal deficit. Suppose first that the increase is “small” so that curve D shifts to the right to D but continues to interest the 45° line twice. The increase in the fiscal deficit thus leads to an instantaneous jump in the rate of money growth as well as the actual inflation rate from point A to C, and from then on to a gradual increase in both the actual and the expected inflation rate from point C to A. Once expectations begin to adjust, the demand for real money balances which depends, as shown in (1), only on $\pi^{a}$ - starts falling. To compensate for the reduction in the inflation tax base, the government must print money at an accelerated pace, until the new equilibrium is reached. A similar result obtains if the

shift in the D curve is such that there exists only one point of intersection with the 45° line (point E). By contrast, if the increase in the fiscal deficit is large, curve D may not interest the 45° line at all (curve D). There is thus no steady state, and inflation will keep increasing continually. The economy jumps from point A to point F and follows a hyper-inflationary path, moving to the northeast along the curve D"

If bonds can be used as an additional source of financing of the fiscal deficit, dual equilibria will still obtain if the government fixes the interest rate, but a unique steady state inflation rate is attained when the government sets a nominal anchor for the economy for instance, by fixing the rate of growth of the nominal money stock. The existence of dual equilibria is thus a consequence of the government's choice of monetary and fiscal policy rules, given the process through which inflationary expectations are formed. This result has implications for the choice of a nominal anchor in disinflation programs.

Consider now the case where inflation expectations are rational, an assumption that can be implemented here by setting $\beta \to \infty$ in (6) and allowing expected and actual prices to jump on impact. In this case, it can be shown that point B is a stable equilibrium and A is unstable. More important, however, because the initial expected rate of inflation can now jump on impact, all points located on curve D are potential short run equilibria. An increase in the fiscal deficit leads in this setting to an instantaneous jump to a new equilibrium, but there is no guarantee that the economy will be at any particular position on the curve $D^1D^1$ Inflation, without displaying any sign of instability may thus be unnecessarily high under perfect foresight.

The above discussion seems to suggest that large budget deficits may lead to hyperinflation only when private agents have adaptive expectation, that is, when they make systematic errors in predicting future inflation. Because the assumption of adaptive expectations is difficult to defend in situations where inflation is high or tends to follow an unstable path, this would seem to make hyperinflation unlikely in the orthodox model. Bruno and Fischer (1990) and Kiguel (1989), however, have shown that large budget deficits may lead to hyperinflation even under perfect foresight, if there is sluggish adjustment toward equilibrium in the money market.

Following Kiguel, assume that the money market adjusts gradually according to

$$m/m = k\left(1nm^d - 1nm\right), k > 0 \qquad (7)$$

Where $m^d$ denotes desired real balances, given by Equation (1) and k the speed of adjustment. Equation (7) can equivalently be written as

$$\pi = \mu - k\left(\ln m^d - \ln m\right) \qquad (8)$$

Which indicates that the inflation rate adjusts one-for-one with the rate of growth of the nominal money stock, but adjusts only partially in response to differences between the desired and actual levels of real money balances. The inflation rate is therefore sticky (but not predetermined), whereas real balances are predetermined at any point in time.

Solving for the logarithm of money demand from Equation (1) and using the identity $m \equiv M/P - \pi m$ in Equation (8) yields.

$$m = \frac{k}{\alpha k - 1}\left(\alpha d + m - \ln m\right) \qquad (9)$$

Equation (9) is plotted in figure 2 for a value of the deficit equal to $d_0$ and $k < 1/\alpha$. There are two equilibria, one unstable (point A) and one stable (point B). When the speed of adjustment is very high $\left(k \to \infty\right)$, equation (9) becomes

$$m \simeq d + \alpha^{-1} m - \ln m \ ,$$

Which, for $m \simeq 0$, gives a curve similar to D in figure 1.

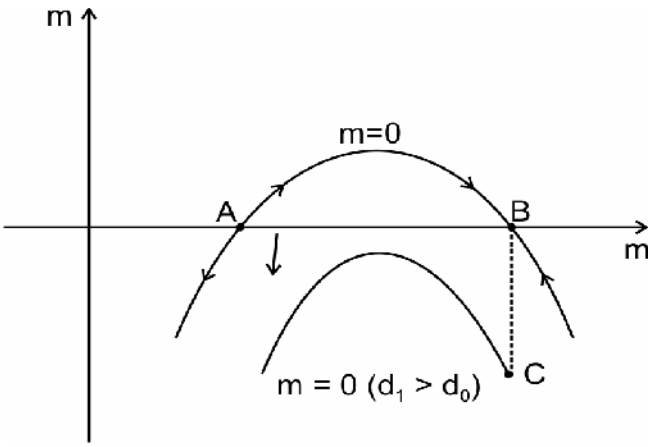



**Figure 13. 2 Fiscal deficits and Inflation with Gradual Adjustment of the money market**

Consider now what happens when the policymaker increases the primary deficit to $d_1 > d_0$. The schedule (m = 0) moves down, so much so that it may no longer interest the horizontal axis put differently, there may be no stationary value of the inflation rate that ensures adequate revenue from the inflation

tax to finance a deficit equal to $d_1$. In such conditions the behaviour of the system will be unstable, characterized by decreasing real money balances and rising rates of inflation. Too large a deficit can therefore lead to a hyperinflationary path, as argued above in the case of adaptive expectations. Under perfect foresight, the potential instability in the inflation process depends crucially on the assumption of sluggish adjustment in the money market. The increase in money growth required to finance a higher deficit creates a temporary excess supply in the money market, which leads to an increase in inflation. The higher inflation rate exerts two conflicting effects on the equilibrium of the money market. On the one hand, it reduces the supply of real money balances, which tends to reequilibrate the market on the other hand, it leads to a fall in the demand for real money balances, which tends to amplify the initial - disequilibrium. When the system does not possess a stable long-run equilibrium, the latter effect dominates the former, and the resulting outcome is accelerating inflation, with a continuous increase in the rate of expansion of the nominal money stock. As shown by Kiguel (1989), the possibility that the economy may follow an unstable inflationary path becomes even more likely if, as a result of the Olivera - Tanzi effect, discussed in previous chapter, the erosion in tax revenue results in a positive relation between the primary fiscal deficit and the inflation rate. The importance of the Tanzi-Olivera effect in hyperinflation episodes has been emphasized also by Dornbusch (1993).

To summarize, money financing of fiscal deficits may lead, depending on the mechanism through which expectations are formed and the speed of adjustment of the money market, to multiple steady state equilibria. Governments can there fore find themselves operating at an unnecessarily high inflation rate. The key message of the analysis, however, is that hyperinflation is an unstable process that emerges as a result of large, unsustainable fiscal deficits financed by money creation. Consequently, an essential feature of stabilization programs n countries undergoing hyperinflation must be a significant fiscal adjustment.

In small, open, developing countries, an additional factor that may affect inflation directly in the short run is the exchange rate. A nominal depreciation affects directly the domestic - currency price of import competing goods and exportables. An indirect effect may also results, as indicated in earlier chapters, if the cost of imported inputs (such as oil and semifinished goods) affects pricing decisions directly. Even if the official exchange rate is fixed, fluctuations in the unofficial exchange rate may affect inflation if domestic price setters take into account the behaviour of the marginal cost of foreign exchange when setting prices. In addition, a depreciation of the exchange rate may also affect inflation by

raising nominal wages, through implicit or explicit indexation mechanisms.

In such conditions, a real exchange rate depreciation is likely to lead to inflationary pressures. The evidence provided by Darrat and Arize (1990), Dornbusch etal (1990), Jorgensen and paldam (1986), and Montiel (1989) supports the view that the exchange rate plays on important rate in the short-run behaviour of inflation in some chronic - inflation countries of Latin America. However, it is worth emphasizing that such evidence is not inconsistent with the presumption that fiscal deficits play a key rate in the long run, as argued by the orthodox. “fiscal view”. For instance, in the results presented by Dornbusch et al (1990), while almost 46 percent of the variability of inflation in Mexico over the period 1982-1987 is accounted for by exchange rate fluctuations in the short run (against about 12 percent for the fiscal deficit proxy) the proportions change to 40 and 55 percent at a longer horizon, similarly, for Bolivia during the period 1982-1986, the proportion of the variance in inflation accounted for by innovations in the real exchange rate and the fiscal deficit are 36 percent and 31 percent respectively, in the short run while in the long run these proportions become 4 percent and 61 percent.

The model developed by Rodriguez (1978) provides a theoretical framework for explaining this type of result. If the fiscal deficit is financed through credit creation by the central bank, as is often the case in developing countries, the monetary expansion will lead to an increase in prices and a progressive erosion of foreign reserves, which will eventually trigger a devaluation if the central bank has limited access to borrowing in international capital markets. A devaluation inflation spiral may develop, in the absence of corrective measures aimed at reducing the deficit. Thus, while the “proximate” cause of inflation may appear to be exchange rate adjustment, the “ultimate” factor responsible for both inflation & exchange rate depreciation may stem from fiscal rigidities.

## 13.4 THE HIGH - INFLATION TRAP

Many economies in the process of liberalization find themselves in serious internal or external macroeconomic imbalance. A double digit, or even triple - digit, inflation is one indicator of such an imbalance. A balance of payments (BOP) crisis and a run on foreign exchange reserves are other indicators policies adopted to tackle one set of problems frequently exacerbate another set. For example, a devaluation may worsen inflation, whereas a temporary price freeze may accentuate a foreign exchange crisis. The liberalization process often begins at a time of simultaneous internal and external disequilibrium, and even

if it does not, these problems are bound to arise at some stage or the other during the process unless appropriate stabilization measures are in place. ‘One is thus invariably led to ponder the necessary links between the choice of stabilization policies and the maintenance of the liberalization process’

When an economy, in the course of opening up, encounters prolonged period of inflation, the problem in most cases lies in the government budget deficit. A lack of a sufficiently broad tax base leads governments to rely on the inflation tax mechanism, and even if a broad tax base exists, high inflation, by causing fiscal erosion (via the Olivera - Tanzi effect) increases the budget deficit. In addition, if the capital account is opened up prematurely, as was the case of Israel in 1977 and Mexico in 1991, inflation is bound to accelerate because of the loss of control over the monetary base.

More importantly, the destabilizing effects of a budget deficit in on open economy are not confined to inflation alone. The budget deficit, which constitutes negative public sector savings, increases the current account deficit of the BOP (when it is viewed as the difference between domestic investment and savings). Thus inflation and BOP crises often go hand in hand with budget deficits. It is thus a necessary condition for both stabilization as well as an orderly conduct of the liberalization process to close the government budget deficit as rapidly as possible.

However, the elimination of budget deficits is not a sufficient condition for rapid stabilization from an initially high inflation trap because, although the source of prolonged inflationary pressures is in most cases a large budget deficit, elements of inertia in the dynamics of inflation often give inflation a line of its own after a certain period of high inflation has elapsed. Thus, inflation may accelerate in response to certain other factors, for example, external price shocks, even when the government budget deficit has been reduced or has not risen.

The dynamics of such a high inflation process usually manifests itself in stepwise or discrete jumps in the inflation rate. The new enhanced rate may then persist in a more-or-less stable fashion for a considerable length of time before yet another (price) shock results in another stepwise jump in the inflation rate.

The inflationary processes in Argentina, Brazil, and Israel in the 1970s and early 1980s are a good example of such high-inflation traps. While such inflation rate jumps are influenced by the size of the budget deficit, they may not be directly correlated with it in effect, on economy may be stuck at a high inflation equilibrium because of a given high budget deficit although, with the same budget deficit, it could have been at a lower rate.

Such a phenomenon was modeled by Bruno and Fischer (1990) who, in the process, highlighted the role of inflationary expectations and the potentially destabilizing effects of fiscal rigidities. To explain their model, we consider a closed economy with exogenous real output. Assume that the money demand function takes the cagan semi-logarithmic form, with unitary income elasticity, similar to the one used in analyzing inflationary financing in previous chapter, that is

$$m = e^{-\beta\pi(e)}, \beta > 0 \qquad (1)$$

Where m (= m / py) denotes real money balances (M/P) as a fraction of real output (y) and $\pi(e)$ is the expected rate of inflation. Assuming that the government cannot issue bonds to the public and runs a budget deficit (BD) that is a constant proportion (d) of output which is financed only through seignorage, then this financing rule implies (with a dot for the time derivative) that.

$$d = BD/PY = M/PY = (M/M)(M/PY) = \mu m = \mu e^{-\beta\pi(e)} \qquad (2)$$

Where $\mu(=m/m)$ is the growth rate of nominal money supply.

Equation (2) reveals how the budget deficit affects the equilibrium growth rate of money supply and, hence, the equilibrium inflation rate. However, to the extent that real money demand is non-linearly related to the expected rate of inflation, the possibility of multiple solutions to equation (2) arises which is in the line with the existence of the so-called 'seignorage Laffer curve'.

Equation (2) is plotted in figure 1 which is based o9n Bruno and Fischer (1990). The curve $B_0D_0$ represents all combinations of $\mu$ and $\pi(e)$ for which the budget deficit is constant at, say, d (0); hence $B_0D_0$ represents an iso-deficit line. From equation (2), it is seen that $d(o) = \mu$ when $\pi(e) = 0$, and therefore the deficit is measured by the distance $OB_0$ from the origin to the intercept of the $B_0D_0$ curve on the $\mu$ axis. The economy is always located on the $B_0D_0$ curve, since the government is arithmetically bound by its budget constraint.

Logarithmic differentiation of equation (1) with respect to time yields.

$$m/m - p/p - y/y = \mu - \pi - g = -\beta\pi(e) \qquad (3)$$

Where $\pi(=p/p)$ is the inflation rate and g (= y/y) is the real growth rate.

In steady - state, therefore we have :

$$\pi = \pi(e) = \mu - g \qquad (4)$$

Equation 4 is represented by the 45° line EE in figure 1, with intercept equal to - g on the $\pi(e)$ - axis. As depicted in the figure, the $B_0D_0$ curve and the EE line intersect twice, implying two potential steady state equilibria the low-infaltion equilibria at 1, and the high-inflation equilibria at H. the dual equilibria are a reflection of the seignorage Laffer curve; the same amount of seignorage can be obtained at either a low or high inflation rate.

The maximum steady, state seignorage revenue is obtained by setting the inflation rate at : $\pi^* = \pi^*(e) = 1/\beta - g$; and the corresponding maximum seignorage revenue ($d^*$) is equal to : $d^* = (1/\beta)e^{(\beta g - 1)}$.

Thus, depending upon the actual size of the deficit which the government wishes to finance, there may be zero, one, or two equilibria. Because the government cannot obtain more than $d^*$ in steady state, there is no steady state solution if $d > d^*$ (as is the case with the curve $B_1D_1$). For $d = d^*$, there would be a unique steady state at M (as is the case with the curve $B_2D_2$). The existence of two steady state equilibria in the case of $d < d^*$ (as is the case with the original curve $B_0D_0$) thus suggests that an economy may find itself at a higher than necessary inflation rate over extended periods of time, that is, at the high inflation trap H, rather than at the law inflation equilibrium. Whether this is likely to happen would depend upon the stability of the respective equilibrium points.

The following important question therefore immediately emerges in this context : Is the reduction of the budget deficit a sufficient condition for stabilization to a lower level of inflation.

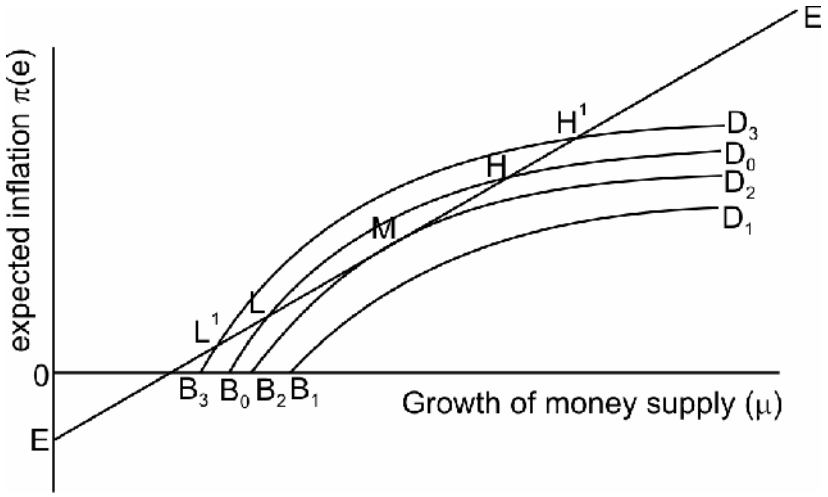



**Figure 13.3 : The high-inflation trap.**

The answer is a qualified negative. For example, in figure 1, reducing the deficit from d(o) to, say, d(3) < d(o), that is shifting the original curve $B_0D_0$ toits new position $B_3D_3$ which is closer to the origin with $OB_3 = d(3) < OB_0 = D(O)$, could switch the economy on to a still higher inflation rate given by $H' > H$, if this high-inflation trap is a stable one compared to the lower inflation rate $L^1 < L$. Thus, as stated earlier, while the source of an inflation could be a large budget deficit, the dynamics of inflation may be such that it could refuse to respond to budget cuts unless accompanied by special stabilization measures.

## 13.5 SARGENT - WALLACE HYPOTHESIS

There is no longer any serious debate about whether monetary policy should be condected according to rules or discretion. Quite appropriately, it is widely agreed that monetary policy should obey a rule, that is, a schedule expressing the setting of the monetary authority's instrument (e.g., the money supply) as a function of all the information it has received up through the current moment. Such a rule has the happy characteristic that in any given set of circumstances, the optimal setting for policy is unique. If by remote chance, the same circumstances should prevail at two different dates, the appropriate setting for monetary policy would be identical. The central practical issue separating monetarists from Keynesians is the appropriate form of the monetary policy rule Milton Friedman has long advocated that the monetary authority adopt a simple rule having no feedback from current and past variables to the money supply. He recommends that the authority cause the money supply to grow at some rate of X percent per year without exception. In particular, the Fed Ought not to try "lean against the wind" in an effort to attenuate the business cycle. Within the context of macro econometric models as they are usually manipulated, Friedman's advocacy of a rule without Feed back seems indefensible.

Friedman's Simple x-percent growth rule is suboptimal. Its logic carries over to larger stochastic difference equation models, ones with many more equations and with many more lags. It is also applies where criterion functions have more variables. The basic idea is that where the effects of shocks to a goal variable (like GNP) display a stable pattern of persistence (serial) correlation), and hence are predictable, the authority can improve the behavior of the goal variable by inducing offsetting movements in its instruments. The notion that the economy can be described by presumably a large system of stochastic difference equations with fixed parameters underlies the standard Keynesian objections to the monism of monetarists who argue that the monetary authority should ignore other variables such as interests rates and

concentrate on keeping the money supply on a steady, growth path. The view that, on the contrary, the monetary authority should look at (and respond to) everything," including interest rates, rests on the following propositions:

a) The economic structure is characterized by extensive simultaneity, sot that shocks that impinge on one variable, e.g., an interest rate, impinge also on most others,

b) due to lags in the system, the effects of shocks on the endogenous variables are distributed over time, and so are serially correlated and therefore somewhat predictable; and

c) The "structure" of these lags is constant over time and does not depend on how the monetary authority is behaving. These propositions imply that variable that the authority observes very frequently, e.g., daily, such as interest rates, carry information useful for revising its forecasts of future value of variables that it can't observe as often, such as GNP and unemployment. This follows because the same shocks are affecting both the observed and the unobserved variables, and because those shocks have effects that persist. It follow then from (c) that the authority should in general revise its planned setting for its policy instruments each time it receives some new and surprising reducing on a variable that is.

Determined simultaneously with a variable, like GNP or unemployment, that it is interested in controlling. Such an argument eschewing a simple x-percent growth rate rule in favor of "looking everything" has been made by Samuelson (1970).

There can be little doubt about the inferiority of an x-percent growth rule for the money supply in a system satisfying propositions (a), (b), and (c) above. A reasonable disagreement with the "look at everything, respond to everything" view would seemingly have to Stem from a disbelief of one of those three premises. In particular, proposition (c) asserting the invariance of lag structures with respect to changes in the way policy is conducted would probably not be believed by an advocate of a rule without feedback.

Thus, combining the natural rate hypothesis with the assumption that expectations are rational transforms the former from a curiosity with perhaps remote policy implications into an hypothesis with immediate and drastic implications about the feasibility of pursuing countercyclical policy. As indicated above, by a countercyclical policy we mean a rule with feedback from current and past economic variables to the authority's instrument, as in a regime in which the authority "leans against the wind." While the present model suggest reasons for questioning even the possibility,

of a successful countercyclical policy aimed at improving the behaviour of the unemployment rate or some closely related index of aggregate activity, the model is compatible with the view that there is an optimal rule for the monetary authority, albeit one that need incorporate no feedback. Such an optimal rule could be determined by an analysis that determines the optimal rate of expected inflation, along the lines of Bailey (1956) or Tobin (1968). If there is an optimal expected rate OR inflation, it seems to imply restrictions on the constant and trend teams (and may be the coefficients on some slowly moving exogenous variable like the labor force) of a rule for the money supply, but is not a cause for arguing for a feedback rule from endogenous variable to the money supply. The optimal rate of inflation, if there is one, thus has virtually no implications for the question of countercyclical policy. Furthermore, there in hardly any theoretical agreement about what the optimal rate of expected inflation is, so that it seems to be a weak read for a control rule to lean on. The simple models utilized above illustrate the implications of imposing the natural rate and rational expectations hypotheses in interpreting the statistical correlations summarized by the reduced form of macro-econometric models, reduce forms that capture the correlations between monetary and fiscal variables on the one hand, and various real variables on the other hand. What is there do recommend these two hypotheses? Ordinarily, we impose two requirements on an economic model. First, that it be consistent with the theoretical core of economics - optimizing behaviour within a coherent general equilibrium frame work; and second that it not be refuted by observations. Empirical studies have not turned up much evidence that would cause rejection at high confidence levels of models incorporating our two hypotheses. Furthermore, models along these lines seem to be the only existing ones consistent with individuals maximizing behavior that are capable of rationalizing certain important correlations.

Such as the Philips curve, that exist in the data and are summarized by the reduced forms of macroeconometric models. The key feature of models that imply hypotheses has been described by Lucas (1973); “All formulations of the natural rate theory postulate rational agents, whose decisions depend on relative prices only, placed in an economic setting in which they cannot distinguish relative from general price movements.” Their inability separately to identify relative and overall nominal price changes is what gives rise to reduced forms like (1). But their rationality implies that only the surprise components of the aggregate demand variables enter. And this has the far - reaching policy implications described above. Several reasons can be given for using the hypotheses of rational expectations. An important one is that it is consistent with the findings that large parts of macro econometric models typically fail tests for structural change

(essentially versions of chow tests). As equation (10) illustrates, if expectations are rational and properly take into account the way policy instruments and other exogenous variables evolve, the coefficients in certain representations of the model (e.g. reduced forms) will change whenever the processes governing those policy instruments and exogenous variables change. A major impetus to work on rational expectations is thus that it offers reason, but probably not the only reason, that macro econometric models fail tests for structural change. Indeed, the hypothesis of rational expectations even offers some hope for explaining how certain representations of the model change out of the sample. A second reasons for employing the hypothesis of rational expectations is that is estimating econometric models it is a source of identifying restrictions. The usual method of modeling expectations in macro econometric models - via a distributed lag on the own variables - leaves it impossible to sort out the scalar multiplying the public's expectations from the magnitude of the weights in the distributed lag on own lags by which expectations are assumed to be formed. Therefore, the coefficients on expectations are generally under identified econometrically. The way out this has usually been to impose a unit sum on the distributed lag whereby expectations are formed. The problem is that this is an ad hoc identifying restrictions with no economic reasons to recommend, it. It is generally in compatible with the hypothesis of rational expectations, which can be use to supply an alternative identifying restriction. A third reason for using the rational expectations hypothesis is that it accords with the economist's usual practice of assuming that people behave in their own best interest. This is not to deny that some people are irrational and neurotic. But we have no reason to believe that those irrationalities cause systematic and predictable deviations from rational behavior.

That a macroeconomist can model and tell the monetary authority how to compensate for. In this regard, it should be noted that the rational expectations hypothesis does not require that people's expectations equal conditional mathematical expectations, only that they equal conditional mathematical expectations plus what may be a very large random team (random with respect to be the conditioning information).

The conundrum facing the economist an be put as follows. In order for a model to have normative implications, it must contain some parameters whose values can be chosen by the policymaker. But if these can be chosen, rational agents will not view them as fixed and will make use of schemes into for predicting their values. If the economist models the economy taking these schemes into account, then those parameters become endogenous variables and no longer appear in the reduced form equations for the other

endogenous variables. If he models the economy without taking the schemes into account, he is not imposing rationality.

## 13.6 SUMMARY

- There is close connection between money, fiscal deficit and inflation.
- The primary deficit affects the equilibrium rate of growth of money stock and hence the equilibrium inflation rate.
- Money financing of fiscal deficit may lead, through adaptive expectations, to multiple steady state equilibrium.
- Process of liberalization may lead to many internal and external macroeconomic imbalances like high inflation trap.
- Sergeant - Wallace hypothesis is about rule of implementing monetary policy.

## 13.7 QUESTIONS

1) Discuss the relation between inflation, money and fiscal deficit.
2) Write a detailed note on adaptive expectations.
3) Explain the concept of high inflation trap.
4) Examine the Sargent - Wallace Hypothesis.

# 14

# TWIN DEFICIT APPROACH TO GROWTH

**Unit Structure :**

14.0 Objectives
14.1 Introduction
14.2 The Twin Deficit Approach to Growth
14.3 The IMF Financial Programming Model
14.4 Financial Crisis
14.4.1 Types of Crisis
14.4.2 Origins of Crises
14.4.3 Identifying Crises
14.4.4 Signals of Crises
14.5 Analytical After thoughts on the Asian Crisis
14.5.1 Contagion
14.5.2 Balance sheet and Transfer Problems
14.6 Summary
14.7 Questions

## 14.0 OBJECTIVES

After having studied this unit, you should be able -

- To know the twin deficit approach to growth.
- To explain the IMF Financial Programming model.
- To understand the Financial Crises.
- To examine the analytical afterthoughts on the Asian crises.

## 14.1 INTRODUCTION:

The twin deficit approach to growth is about the links suggested by these two approaches between the twin (fiscal and trade) deficits and growth.

Developing countries borrow from IMF. This borrowing is conditional on borrowing country's compliance with a set of quantitative policy performance criteria. The design of such a programme and specification of such criteria rely on a conceptual framework referred to as "Financial Programming".

This unit also discusses the financial crises along with the analytical afterthoughts on the Asian Crises.

## 14.2 THE TWIN DEFICITS APPROACH TO GROWTH

While the structuralist approach to macro-economics is after presented as a challenge to 'orthodox' macroeconomics, there are many areas in which the orthodox thinking has provided much insight and, ironically even strengthened structuralist arguments. This is especially true as regards the links suggested by these two approaches between the twin (fiscal and trade) deficits and growth, which demonstrate the fundamental complementarities that often exist between these two rival schools of thought. The 'orthodox' framework developed below. (see Fischer and Easterly 1990) shows how fiscal sustainability, external viability and growth are interrelated and indicates some of the trade-offs involved in alternative strategies of financing fiscal and trade deficits with special reference to their implication for growth.

Following the logic of the three-gap approach, the savings – investment constraint, given by (eqn. 1) can be rewritten as:

$$I = Sp – BD + CAD, \qquad (1)$$

Which states that total investment (I) is equal to national savings, given by the sum of private sector savings (Sp) and public sector disserving (BD = G-T), plus external savings given by the current account deficit (CAD = Z – X).

Following the logic set out in previous chapter, the mode of financing these fiscal and trade deficits is useful in understanding how the deficit debt dynamics relates with total investment and, thus, real economic growth.

From the mode of financing the budget deficit (BD), we have the following relationship:-

$$BD = D + iB = \Delta M + \Delta B, \qquad (2)$$

Where D is the primary deficit, I is the domestic interest rate, B is the stock of internal debt, iB refers to interest payments on internal debt, and $\Delta M$ and $\Delta B$ are the amounts of money-financing and debt-financing, respectively, needed to finance the budget deficit (exclusive of public investment).

From the mode of financing the current account deficit, i is the domestic interest rate, B is the stock of internal debt, (CAD), we have the following relationship (where, as before, $\Delta R = \theta$);

$$CAD = NICA + if\ EF^* = E\Delta F^* \quad (3)$$

Where NICA is the non-interest current account deficit, if is the foreign interest rate, E is the nominal exchange rate, F* is the stock of external debt (in foreign-currency units), if EF* refers to interest payments on the external debt (in domestic currency units), and $\Delta F^*$ is the amount of foreign borrowing (in foreign currency units) needed to finance the current account deficit.

Substituting (eq. 2) and (3) into (1) yields

$$I = S - D + NICA, \quad (4)$$

Where : $S = Sp - iB + if\ EF^*$ (5)

However, from (eqn2), we also have:

$$D = \Delta M + \Delta B = iB \quad (6)$$

Substituting this expression into (eg. 4) and dividing throughout by nominal income (Y=Py) yields :

$$I / Y = S - (\Delta M / Py) - (\Delta B/Py) + ib + n, \quad (7)$$

Where S(=S/Y) is the propensity to save of the private sector; b(=B/Py) is the nominal debt – income ratio; and n (= NICA /Y) is the ratio of the non-interest current account to nominal income.

$$\Delta M / Py = \Delta m + (\pi + g) M, \quad (8a)$$
$$\Delta B / Py = \Delta b + (\pi + g) b, \quad (8b)$$

Where m(=M/Py) denotes real money balances as a fraction of real output, and $\pi$ and g are the inflation rate and real growth rate, respectively.

Similarly, from previous equation we have

$$n = Df^* + (g - \gamma f - e) f^*, \quad (9)$$

Where f* (= EF*/Y) is the external debt – income ratio, $\gamma_f (= i_f - \pi)$ is the real rate of interest on foreign debt, and e(= $\Delta$ E/F) is the rate of depreciation of the nominal exchange rate, substituting (eq 8) and (eq 9) into (eq 7) yields :

$$I/Y = s - \Delta m - (\pi + g)m - \Delta b + (r - g)b + \Delta f^* + (g - rf - e)f^* \quad (10)$$

where $r (= i - \pi)$ is the real rate of interest on domestic debt.

Thus, it is seen that the investment-income ratio is determined not only by the private savings rate, but also by the evolution of total (public and external) debt which, in turn, is

determined by the twin (fiscal and trade) deficits as well as their modes of financing (eq 10) states that while net external capital inflows ($\Delta f^* - r_f F^*$) complement the domestic private savings rate (s), additional domestic public debt ($\Delta b$) can be viewed as a transfer of resources from private savings to the public sector. In a similar vein, increased money – financing via seignorage revenue [$\Delta m + (\pi + g)m$] transfers resources to the public sector through the 'forced savings' mechanism.

With the eruption of the debt crisis in the 1980s, external borrowing was severely curtailed. This, coupled to the fact that in many developing countries, especially in sub-saharan Africa and, to some extent, Latin, America, domestic bond markets were not fully developed, led governments to resort to inflationary financing of their fiscal deficits. The result was either complete seignorage financing [$\Delta m + (\pi + g)m$] of the budget deficit, or quasi-seignorage financing via the inflation tax ($\pi M$), both of which resulted in considerably higher rates of inflation.

Given that the real growth rate (g) is related to the incremental capital –out put ratio (ICOR) and the investment rate in the following manner :

$$gv = I / Y \qquad (11)$$

where V is the ICOR, and I/Y is the ration of investment to GDP in real (assumed equal to nominal) terms. Substituting (eq 11) into (eq. 10) above and solving uniquely in term of the growth rate yields:

$$g = k (s - \Delta m - \pi m - \Delta b + rb = \Delta f^* - r_f f^* - ef^*), \qquad (12)$$

where:

$$k = 1 / (v + m + b - f^*) \qquad (13)$$

(Eqn 12) states that long-term economic growth, a part from being influenced by the savings rate (s) and the ICOR (v) in the traditional Harrod – Domar fashion, is positively related to fresh capital inflows ($\Delta f^*$) and inversely related to the stock of outstanding domestic obligations (m+b). Other things remaining the same, the real growth rate will increase if injections into private savings in the form of government borrowings ($\Delta b$), real interest payments on the external debt ($r_f f^*$), and, of course, inflationary financing in the form of seignorage revenue (($\Delta m + \pi m$).

The adverse impact of inflationary financing on the real growth rate is obvious. However, policy makers continue to resort to it for as Keynes (1923) pointed out, 'Inflationary tax is a form of taxation which the public finds hardest to evade and even the

weakest government can enforce, when it can enforce nothing, else.' If domestic debt financing is used as an alternative, then private investment, through a reduction in private savings, is crowded out. If distortionary taxes, such as trade taxes, are used to finance the deficit, then the efficiency of investment would be lower and, thus, the rate of growth would also be lower on the other hand, an increase in corporate taxation or taxes on capital goods would lower private investment directly. Thus, the last alternative, that is, external borrowing, seems to be the only solution, at least in the short term, although this could have sever repercussions in the long run in terms of an increasing interest burden.

The following is a summary of the results based on the growth equation:

- The higher the average savings rate, the higher would be the growth rate, for a given level of the ICOR.
- Additional public debt and increases in seignorage revenue which are used to transfer resources away from private savings in order to finance public sector expenditures reduce the real growth rate.
- External resource flows complement private sector savings and thereby increase real economic growth in the short run; although the rising burden of interest payments on these external borrowings, coupled to possible increases in real foreign interest rates, would reduce net external capital inflows in the future, and thus adversely affect the long-run real growth rate.
- The larger the burden of interest payments on external debt, the more will be the leakage of resources from the domestic economy, resulting in lower, and perhaps even negative, growth rates. This situation is particularly applicable to the heavily indebted countries (HICs) where the external debt to GDP ratios are relatively much higher and, consequently, so are the debt-servicing costs As such, debt relief measures proposed under the Brady plan could go a long way in enhancing economic growth
- Sustainable economic growth requires external financing accompanied, sooner or later, by some debt relief such that there is a net inflow of external resources.
- An increase in domestic debt (including monetary obligations) would increase interest payments on the domestic debt and thereby transfer resources to the private sector, However, the rise in the domestic debt ratio would more than offset the increase in interest payments on the domestic debt, and therefore the net result would be to lower the real growth rate.

- Increasing the efficiency of investment will enhance the growth rate.

The Brady initiative correctly linked additional external borrowings with growth inducing policies. Furthermore, the bush administration's emphasis on more multilateral lending to the private sector (rather than directly to the government sector) of developing countries is also on the right track because such loans will complement the average private savings rate and thereby positively impact itself upon the real gross domestic products (GDP) growth rate. However, this emphasis on private sector lending for stimulating economic growth in the developing countries is not without pitfalls. In most cases, the governments of the developing world are reluctant to cut public expenditures because they are genuinely interested in satisfying social demands through increased government spending, particularly in the areas of education, health, and infrastructure. In such a situation, the government, when faced with reduced external assistance on account of most foreign capital being directly earmarked for the private sector, can be justified in raising taxes on the private sector (including income and profit taxes) and / or resorting to seignorage revenue in order to finance their already high fiscal deficits. Alternatively, the governments may coerce the private sector to buy government – issued bonds as a quid pro quo for various public services, in which case the domestic debt to GDP ration cannot be stabilized. All these factors may lead to an uncertain macroeconomic environment with no clear rules of the game, thus thwarting private investment of its legitimate role in economic development. Therefore, lending to the developing world by international financial institutions should contain a proper mix between loans to the private and public sectors. In summary, sustainable long-run economic growth in developing countries would, within the framework developed above, require a combination of moderate external financing (accompanied by some debt relief), a higher savings rate, and above all, an increase in the efficiency of investment.

The above 'orthodox' results which clearly highlight the expansionary (contractionary) role of debt relief (devaluation), apart from strengthening the 'structuralist' arguments in these respects, demonstrate the fundamental complementarities that exist between these alternative viewpoints, thereby suggesting the need for a balanced approach to development macroeconomics.

## INERTIAL INFLATION, HETERODOX PROGRAMMES, AND GROWTH.

Two explanations are possible for accelerating inflation. The first is the 'orthodox' one, suggested above, which is obtained by linking inflation to the mode of financing fiscal deficits. In such a

context, money financing and the inflation tax provide a direct link between fiscal deficits and high inflation. On the other hand, there is also the explanation based upon 'inertial' inflation. Several countries in Latin America during their high and chronic America during their high and chronic inflationary phase in the 1980s had nominal wages and other payments indexed to inflation and the resulting staggered wage contracts contributed to such an inertia. Moreover, inflation in most open developing countries is party due to 'imported inflation' as it depends upon the cost of imported consumer, intermediate, and capital goods which play a major role in the functioning of the domestic economy. Therefore, the 'heterodox' approach attempts to link inflation with the rate of growth in unit costs of production, the rate of devaluation of the local currency, imported inflation, and the excess demand (or slack) in the goods market.

Under the circumstances, the rate of inflation ($\Pi$) is given by:

$$\pi = \alpha\ (\Delta W/W) + \beta\ (\Delta E/E + \pi f\ ) + \Upsilon\ (y/y^*), \qquad (14)$$

where $\Delta W/W$ is the role of wage-inflation (that is the rate of growth of nominal wages), $\Delta E/E$ is the rate of depreciation of the nominal exchange rate, $\pi f$ is the foreign inflation rate which is a proxy for imported inflation, y is actual real output, and y* is capacity (potential) output, such that the ratio (y/y*) provides an estimate of the degree of excess demand (y > y*) or slack (y<y*)in the goods market It there is neither (y = y*), then it is assumed that output is close to its full employment level, implying that y/y*= 1.

In most high inflation economies, such as the now reforming transition economies, such as the now reforming transition economies, the rate of wage inflation ($\Delta w/w$) is indexed to inflation in the past period, implying that:

$$\Delta w/w = \Theta\ (1)\ \pi\ (-1)\ ,\ 0 < \Theta\ (1) <1$$

Similarly, the nominal exchange rate is also indexed to inflation under a managed floating exchange rate regime so as to prevent an appreciation of the real exchange rate.

Therefore, we have:

$$\Delta E/E = \Theta\ (2)\ \pi\ (-1),\ 0 < \Theta\ (2) < 1 \qquad (16)$$

Substituting (eq 15) and (eq 16) into (eq 14) above yields.

$$\pi = \Theta\ \pi\ (-1) + \beta \pi f + \Upsilon\ (y/y^*), \qquad (17)$$

Where we have replaced

$[\alpha\Theta(1) + \beta\Theta\ (2)]$ by $\Theta$. Thus, it is seen that the inflation rate is not dependent upon either money supply growth or the fiscal deficit and has an intrinsic inertia of its own especially when $\Theta > 1$, which is usually the case. For example, Jamaica, during 1994-6,

despite consistently generating adequate fiscal surpluses, just could not lower inflation because of such on inertial component between past and present inflation, exacerbated by nominal wage increases and exchange rate changes, combined with low real GDP growth.

The heterodox programmes, such as the Austral Plan in Argentina in 1985-6 and the Guzado Plan in Brazil in 1986, consisted basically of nominal exchange rate targeting and wage controls either by government fiat or a pact with trade unions. However, as shown above, such a stabilization programme puts continuous pressure on the nominal exchange rate to depreciate, thereby fueling further inflation. Therefore, under the circumstances, active government intervention in the form of an incomes policy and / or price controls is needed to slow down the rate of wage inflation for any heterodox programme to be successful. It has often been suggested that because of inertial inflation, which is a structural feature of some Latin American countries, orthodox stabilization programmes, based upon tight fiscal and monetary policies, are not successful in bringing down the inflation rate rapidly. This is however, not exactly true because in order to keep the nominal exchange rate constant, monetary policy must be willing (and able) to offset rising expectations of further devaluation by raising domestic interest rates through a tight money policy. As this would imply a rising interest burden, the logical corollary is that fiscal policy also has to be tightened in order to reduce domestic borrowings.

**BANK FUND MODELS:**

Among the most parsimonius models aimed at quantifying the effects of stabilization programs and medium-term growth policies are those of the international monetary Fund (IMF) and the world Bank After reviewing the basic features of the model developed by each multilateral institution, we examine the issue of whether the main characteristics of the two approaches can be combined.

## 14.3 THE IMF FINANCIAL PROGRAMMING MODEL

Providing advice to developing countries on macroeconomic policy is an important responsibility of the IMF. In addition, the Fund extends financial support to stabilization programs that meet certain criteria, they must be consistent with the principles set out in the institutions articles of agreement and must offer a convincing prospect of repayment. This assistance is conditioned on the borrowing country’s compliance with a set of quantitative policy performance criteria drawn up in consultation with the Fund and embodied in a financial program. The design of such a program

and specification of such criteria rely on a conceptual framework referred to as “financial programming”.

The simplest financial programming model is designed to determine the magnitude of domestic credit expansion required to achieve a desired balance of payments target under a predetermined exchange rate - the regime that, as argued in chapter, has characteristically been adopted in developing countries. The model is in effect a variant of the monetary approach to the balance of payments (MABP) which, indeed, was pioneered at the Fund The First equation of the model is the balance sheet identify for the banking system, which equates assets - in the form of credit to the nonbank sector D and claims on foreigners R - to monetary liabilities M :

$$M = D + ER, \qquad (1)$$

Where E is the nominal exchange rate. In this relationship, R and M are endogenous and D is an exogenous policy variable under the control of the monetary authorities. The second equation is the definition of velocity V as nominal GDP Y divided by the money stock.

$$V = Y / M \qquad (2)$$

In the Fund’s version of the MABP, the money market is required to be in flow (but not necessarily stock) equilibrium.

$$\Delta M = V^{-1}Y - V_{-1}^{-1}Y_{-1} \qquad (3)$$

On the assumption that the nominal exchange rate and velocity are both constant and that nominal output is exogenous, the model can be solved for the change in the stock of international reserves R as a function of V and Y, as well as of the monetary policy instrument D.

$$E\Delta R = V^{-1}\Delta Y - \Delta D \qquad (4)$$

Alternatively, given a target value for the change in reserves (the balance of payments) and projection for V and Y, the required expansion in the stock of credit can be derived from

$$\Delta D = V^{-1}\Delta Y - E\Delta R \qquad (5)$$

As indicated above, in this version of the model nominal output is exogenous. An expanded version, reffered to as the “Polak model”, makes nominal output endogenous as well we write the balance of - payments identity as

$$\Delta R = X - \alpha\left(Y_{-1} + \Delta Y\right) + \Delta F, O < \alpha < 1 \qquad (6)$$

Net exports are taken to have an autonomous component X and a component that depends negatively on current nominal income, expressed as last periods value plus the charge $\Delta Y$, Net Capital inflows $\Delta F$ are exogenous with nominal output endogenous and the balance of payments identity added, the Polak model consists of two equations in the two unknowns $\Delta R$ and $\Delta Y$.

The interaction between the money market equilibrium condition (4) and the balance of payments identity (6) in determining nominal income and the balance of payments is illustrated in Figure 1. equation (4) is

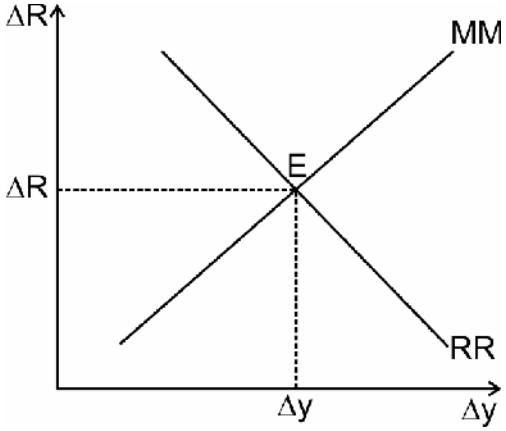



**Figure14. 1 : The Polak Model**

Depicted as the positively sloped MM locus, while (6) is the negatively sloped locus RR. The equilibrium values of the balance of payments and the change in nominal income are determined by the point of intersection, E. It is easy to show that in this model an increase in the rate of expansion of credit will cause the balance of payments to deteriorate and nominal income to rise, while an increase in exogenous receipts of foreign exchange will improve the balance of payment and raise nominal income. This “Polak” form of the financial programming model can be given a “classical” closure - that is, it can be solved for the domestic price level, taking real output as exogenous or a “Keynesian” closure, in which it is solved for changes in real output, taking the price level as given.

‘A Marriage’ between the Fund and Bank Models; some comments

Despite its underlying appeal, Polak (1990) has criticized this so called ‘marriage’ between the Fund and Bank models’ by drawing attention to what he feels are the three principal weaknesses of the merged model;

1) The effort to cross these two models incapacitates each from doing its own job; 2) The simplicity that accounted for part of the attraction of the two models is lost in the merger; and 3) The

merger provides scant rewards in terms of new insights with respects to growth oriented adjustment.

Considering the fact that we shall be using the essence of this approach to provide on alternative design structure for intengrating growth into the basic monetary model, it would be pertinent to examine the validity of these criticisms. As shown in the earlier section, the effort to cross these two models has implied that both of them have gained from the merger. The fund model treats real output as exogenous and the Bank model supplies the missing output equation, that thus providing a way to close the Fund model with endogenous output. In a similar manner, the Bank model treats the flow of foreign savings as exogenous and the merger renders it partly endogenous because it supplies the missing reserve accretion equation, that is, thus providing a way to close the Bank model with endogenous foreign savings.

The basic drawback of the Bank approach is that if a country has spare reserves to finance a current account deficit, then this element of foreign saving enters the Bank model directly. However, this is a theoretically incorrect approach because spare reserves per se do not automatically translate into increased foreign savings, unless these reserves are used for import liberalization as Polak himself has noted. However, with imports being given endogenously, and with the marginal propensity to import being treated as a fixed (technical) parameter, imports cannot be increased merely by the presence of spare reserves, unless these reserves are run down by domestic credit expansion which would, by increasing the demand for investment and output, lead to an increase in imports, thereby worsening the current account deficit. As increasing credit expansion is compatible with fallowing reserves which, in turn, is compatible with rising investment, the type of model closure used to merge the two approaches is not merely the only one of interest, but the only one feasible. Therefore, the first comment regarding the incapacitation of both models as a result of the merger is seen to be invalid.

As far as the third comment is concerned that

$$\frac{\partial \Delta y}{\partial \Delta DC} > 0, \text{and } \frac{\partial \Delta y}{\partial \Delta E} < 0 ,$$

Which corroborates the empirical evidence (see IMF 1985) regarding the expansionary (contractionary) effects of credit expansion (devaluation) on output. Therefore, by endogenizing output and showing how policy variables associated with Fund programmes can affect growth rate, the merger does provide substantive insights with regard to designing growth oriented adjustment programmes.

The only substantive criticism by for is the second comment that the simplicity which accounted for part of the underlying appeal of the two models is lost in the merger. Ironically, however, it is not the complexity of the merged Bank Fund. Model that is its principal weakness - although this criticism is certainly valid - but rather some of the simplifying assumptions it invokes in order to effect the merger.

The most damaging of these simplifications occurs in the final merger, which, under the assumption that $\Delta P \Delta Y$ can be ignored, technically yields the following 'simplified' version of the Bank equation :

$$\Delta y = KI$$

Implying that increases in the price level have no impact on real investment and, consequently, it is nominal investment which determines real output. This represents the most serious flaw in the overall design of the merged model.

As will be shown in the next section, the problem lies not so much with the omission of the interaction term but rather in the specification of the Bank equation itself which, under the present circumstances, needs to be modified in order to accommodate the adverse impact of inflation on the growth rate even after the inevitable simplification has been carried out.

## 14.4 FINANCIAL CRISES

### 14.4.1 Types of Crises:

A number of broad type of financial crises can be distinguished in the literature. A currency crisis may be said to occur when a speculative attack on the domestic currency results in its devaluation, or forces the central bank to defend the currency by expending large volumes of foreign exchange reserves or by sharply raising interest rates. A banking crisis refers to a situation in which actual or potential bank runs or failures either induce bank to suspend the internal convertibility of their liabilities or compel the government to intervene by extending assistance on a large scale. Broadly speaking, when defining a banking crisis, a distinction is made between financial distress and financial panic which refer, respectively, to situations of insolvency and illiquidity in the banking system. The former occurs when 'a significant fraction of the banking sector is insolvent but remains open' while the latter occurs when 'bank debt holders suddenly demand that banks' convert their debt claims into cash to such an extent that banks are forced to suspend to convertibility of their debts into cash. Finally, a debt crisis is a situation in which a country cannot service its foreign debt, whether sovereign or private.

Crises of all types have often had common origins in the build-up of unsustainable economic imbalances and misalignments in asset prices or exchange rates, often in a context of financial sector distortions and structural rigidities. Thus, a crisis may be triggered off by a loss of confidence in the currency or banking system, often prompted by developments such as a sudden correction in asset prices or by a disruption to credit or external financing flows, that expose the underlying economic and financial fragilities. The ensuing crises may then involve sharp declines in asset prices, and failures of financial institutions and non financial corporations. However, not all corrections of imbalances entail a crisis whether or not they do depend apart from the magnitude of the imbalances themselves, on the credibility of policies designed to correct the imbalances as well as on the inherent robustness of the country's financial system to economic shocks.

At times, elements of currency, banking and debt crises may be present simultaneously as in the recent East Asian crisis. This close association was initially noticed by Diaz-Alejandro (1985) who suggested that the bailout of the banking sector in Chile was the cause of inconsistency between its monetary and exchange rate policy and a possible explanation for the speculative attack on the Chilean peso in the early - 1980s. Since then several other channels have been proposed to explain the close timing of these events. In particular, banking problems, by resulting in on excess supply of money or by causing a sudden downward shift in the money demand function, may cause or precipitate a currency crisis. On the other hand, a currency crisis may cause a banking crisis either by deteriorating the net worth of the banking system or by inducing a withdrawal of deposits. However, as is very often the case, the fact that one type of crisis precedes another does not necessarily imply causality. This is because banking sector difficulties may not always be very apparent, especially in poorly supervised and inadequately regulated system or in situation where lending booms and asset price inflation mask banking problems, until a correction in asset prices exposes the fragility of the financial system. The same is true for corporate sector indebtedness. In these situations, the actual weakness of the banking system or the corporate sector may be fully revealed only after a run on the currency has precipitated speculative attacks that exacerbate banking and debt problems. This has been a feature of the recent East Asian crisis, as illustrated most clearly in the case of Indonesia.

### 14.4.2 Origins of Crises:

The factors that underlie the emergence of imbalances that render an economy vulnerable to financial crises may be grouped under the following headings : unsustainable macroeconomic

policies, global financial conditions, exchange rate misalignments, weakness in financial structure, and political instability. Macroeconomic insfeasibility has been one of the most important underlying factors in many financial crises. Very often, overly expansionary monetary and fiscal policies have spurred lending booms, excessive debt accumulation, and over investment in real assets, which have driven up equity and real estate prices to unsustainable levels. The eventual tightening of policies to contain inflation and promote external sector adjustment, and the inevitable correction of asset prices that have followed as a result, have then led to economic recession, debt - servicing difficulties, declining collateral values and net worth, and rising levels of non performing loans that have threatened the solvency of the banking system (see Eichengreen and Rose 1998)

In addition to domestic macro-economic policies, external conditions have also played a role in precipitating financial crises especially in the emerging market economies. Most notable have been the sudden large shifts in the term of trade and in world interest rates. Movements in interest rates in the major industrial countries have become increasingly important to emerging market economies worldwide, reflecting the increasing integration of world capital markets and the globalization of investment. An obrupt rise in world interest rates can drastically reduce the flow of foreign financing to the emerging markets, raising the cost to domestic banks of funding themselves offshore, thereby increasing adverse selection and moral hazard problems and the fragility of the financial system. The composition of capital inflows has also been considered an important factor in a number of currency crises in emerging markets. In both the recent crisis in Thailand and in the 1944-5 Mexican crisis, the reliance on short-term borrowing to finance large current account deficits was a crucial ingredient in precipitating the crisis.

Another lesson of the recent crises is that currency mismatches in private sector balance sheets (of either financial institutions or corporation) may be more of a problem in countries with inflexible exchange rates, since an exchange rates peg may encourage borrowers to ignore exchange rate risks which are implicit in any crawling pegregime, especially if the rate depreciation of the so called shadow floating exchange rate is higher than the crauil rate. Experience suggests that in such cases of unsustainable crauil rates, countries with foreign currency denominated debt or foreign debt intermediated through domestic financial institutions are particularly vulnerable to financial crises.

Financial sector distortions, in conjunction with macroeconomic volatility, from another group of factors behind many banking crises, often these distortions arise in times of rapid

financial liberalization and innovation in countries with weak supervisory and regulatory policies or where the government intervenes directly in the allocation or pricing of credit. Insufficiently stringent regulatory regimes in more liberalized financial environments have created moral hazard by encouraging financial institutions with low capital ratios to assume imprudent risks thereby worsening the quality of their asset portfolios. And deficiencies in accounting, disclosure, and legal frameworks add to the problem by allowing financial institution to disguise the extent of their difficulties. Finally, to all this must be added the frequent failure of governments, primarily because of political instability, to take prompt corrective action when problems initially emerge, with the result that when they finally do so because of external compulsions, it becomes a case of too little, too late.

### 14.4.3 Identifying Crises:

A currency crisis could be identified simply as a substantial nominal currency devaluation. This criterion, however, would exclude instances where a currency came under severe pressure but the authorities successfully defended it by intervening heavily in foreign exchange market, or by raising interest rates sharply, or by other means. Thus, an alternative approach is to construct an index of speculative pressure that takes into account not only exchange rate changes, but also movements in international reserves or interest rates that absorb such pressure and thus serve to moderate exchange rate variations.

Banking crises are more difficult to identify empirically, partly because of the nature of the problem and partly because of the lack of relevant data. Although data on bank deposits are readily available for most countries, and thus could easily be used to identify crises associated with runs on banks, most major banking crises in recent years have not originated from the liabilities side of banks balance sheets, that is they have not been associated with runs on deposits. However, whenever such runs on deposits have occurred - such as during the recent financial crisis in Indonesia - they have tended to fallow the disclosure of difficulties on the assets side or widespread uncertainty abut the future value of the currency. Similarly, a failure to roll over inter bank deposits, as in Korea recently, can have results similar to those of a run on banks. Thus, in general, runs on banks have been the result rather than the cause of banking problems.

Banking crises, which generally stem from the assets side of banks balance sheets, are basically due to a protracted deterioration in asset quality. This suggests that variables such as the share of non-performing assets (NPAs) in banks portfolios, large fluctuations in real estate and stock market prices, and

indicators of business failures could be used to identify banking crises episodes. The difficulty is that data on such variables are rarely available in many developing countries or are incomplete. In cases where central banks have detailed information on NPAs, it is usually laxity in the follow up action in response to such data that allows the situation to deteriorate to a point of crisis. Given such limitations, banking crises have usually been dated by researchers on the basis of a combination of events such as forced closure, merger, or government takeover of financial institutions, runs on banks, or the extension of government assistance to financial institutions.

Using some of the above criteria, 116 currency crises and 42 banking crises were identified by the IMF )1948) for a group of 31 developing countries for the period 1975-97. in 42 of the currency crises, the exchange rate component of the index accounted for more than 75 percent of its overall value and these episodes were termed as ‘currency crashes’ cases in which more than one country was affected by a crisis, either because of a common shock or because of contagion effects, were counted as more than one crisis. For instance, the recent East Asian Financial crises comprised five currency crises several interesting points emerge from the data.

On the basis of the operational criteria used by the IMF, currency crises were seen to be relatively more prevalent during 1975-86. the number of currency crises was particularly high in the mid - 1970s (a period of large external shocks to many countries) and in the early to mid - 1980s (Latin American debt crises) Banking crises, in contrast were somewhat more prevalent during 1987-97, probably reflecting the increasing incidence of financial sector liberalization that occurred in many emerging market economies during this period.

Given that the two types of crises often have common origins, or that one type of crisis often induces the other, it is not surprising that countries appear to have banking and currency crises at around the same time. In these instances, banking crises have preceded currency crises more often than the other way around (see Kaminsky and Reinhart 1996). However, since the late 1980s, currency and banking crises seem to have become more contemporaneous. This evidence, while suggestive, should be interpreted with caution in view of some of the aforementioned difficulties in dating the genesis of banking crises.

Financial crises can be very costly both in terms of the direct fiscal and quasi-fiscal costs of restructuring the financial sector, as well as in terms of the indirect effects on economic activity an account of the inability of financial markets to function effectively.

Resolution costs for banking crises have in certain cases reached to over 40 percent of GDP (for example, in Chile and Argentina in the early 1980s) while NPAs have exceeded 30 percent of total bank assets (for example, in Malaysia during 1988 and Sri Lanka during the early 1990s). in addition to their fiscal and quasi-fiscal costs, banking and currency crises have also led to misallocation and under utilization of resources thereby resulting in real output losses.

As seen in table 14.1 for currency crises, on average, output growth returned to trend in about one and a half years, while the cumulative loss in output growth per crisis was slightly under 5 percentage points (relative to trend). For currency crashes, the average recovery time and cumulative loss of output growth increased to almost 2 years and 8 percentage points, respectively. Banking crises, not surprisingly, were more pronounced and more costly than currency crises : on average, it took almost 3 years for output growth to return to trend, and the cumulative loss in output growth was 12 percentage points. When banking crises occurred within a year of currency crises, the losses were substantially larger, amounting to almost 13.5 percentage point on average. The study also reveals on interesting feature; for both currency and banking crises, the recovery time is shorter in the emerging market economies than in the industrial countries, although the cumulative output losses are much higher.

**Table 14.1 : Characteristics of Financial crises in Emerging Market Economics :**

| | Number of crises | Average Recovery Time[a] (in years) | Cumulative loss of output per crisis (in percentage points) |
|---|---|---|---|
| Currency crises | 116 | 1.5 | 4.8 |
| Currency crashes[C] | 42 | 1.9 | 7.9 |
| Banking crises | 42 | 2.8 | 12.1 |
| Currency and Banking crises[d] | 26 | 2.6 | 13.6 |

### 14.4.4 Signals of Crises:

In view of the costly adjustment that economies undergo in the wake of financial crises there has been considerable interest in identifying configuration of economic variables that can serve as early warning signals of crises. Such indicators of vulnerability could be used to identify situations in which an economy faces the risk of a financial crisis being triggered by changes in either world

economic conditions, spillovers from crises in other countries, or other forces that are liable to cause a sudden shift in market sentiment if imbalances go unaddressed. A commonly used approach to constructing on 'early warning system' is to identify a set of variables whose behaviour prior to episodes of financial crises is systematically different from that during normal or tranquil periods. By closely monitoring these variables, it may be possible to detect behaviour patterns similar to those that, in the past, have preceded crises.

There are potentially a large number of variables that could serve as indicators of the vulnerability to currency and banking crises. The choice is ultimately determined by one's understanding of the causes and proximate determinants of crises considering that, by and large currency crises have been caused by external sector problems, variables such as the real exchange rate, the imports-to-reserve ratio, the current account balance, changes in the terms of trade, the differential between foreign and domestic interest rates, and changes in the level and maturity structure of foreign capital inflows tend to feature prominently in the set of leading indicators (see Kaminsky, Lizondo, and Reinhart 1977) on the other hand, considering that banking crises have been largely due to weaknesses in the financial sector, exacerbated by the failures of prudential regulations to keep pace with advances in financial liberalization, variables such as domestic audit expansion, measures of financial liberalization (such as the M3 to M1 ratio), the level of shortterm foreign indebtedness of the banking system, the structure of domestic real interest rates, changes in the money multiplier and equity prices, and the quality of bank assets as measured by the extent of NPAs have been extensively used as indicators (see Demiraguickunt and Detragiache 1997).

However, unlike currency crises, where sharp changes in high frequency variables such as international reserves, interest rates, and the exchange rate itself make the dating of crises relatively straight forward, the lack of such high frequency data that could be used to consistently mark the onset of banking problems makes the construction of leading indicators of banking crises more difficult as mentioned earlier, the dating of banking crises is much more approximate than that of currency crises because it depends on the occurrence of actual 'events' such as the closure or government takeover of financial institutions, bank runs, and the like. Therefore there is always the risk of dating crises either 'too late' - since financial problems usually begin well before bank closures or runs actually occur or 'too early'. Since the peak of a crisis is generally reached much later in many cases nevertheless, even with approximate dates for the onset of banking crises, an analysis of the behaviour of pertinent variables around the time of

crises has been found to be useful in constructing an early warning system of vulnerability indicators.

Finally, considering the empirical evidence that currency and banking crises seem to have become more contemporaneous of late, as well as the fact that the causes of banking crises are very often similar to those of currency crises in particular, loose. Monetary conditions, overheating of the economy, the bursting of asset price bubbles, and increasing financial integration it is possible to assume that many of the leading indicators of both types of crises could be broadly similar. Such an approach was used by the IMF (1998) to analyse the behaviour of a number of macroeconomic variables around the time of currency crises during the period 1975-97, for a group of 50 advanced and emerging market countries. In many instance, the behaiour of several key monetary, financial, and trade related variables was found to be different in the months leading to a crisis from their corresponding behaviour during the tranquil periods. Unfortunately, several such differences were only suggestive and the concerned variables could not be used with any confidence as an early warning system of crises for three reasons. First, the statistical significance of the differences identified was not established. Second, a number of variables were unable to signal vulnerability until a crisis was just about to occur. Finally information about the behaviour of many variables was available with a lag too long to make them useful as an indicator.

When these requirements were taken into account, only a handful of variables could be considered to consistently provide information about such vulnerability in the sense that, apart from correctly signaling crises a significant number of times (without sounding frequent false alarms), they provided such signals early enough for appropriate countermeasures to be taken. These variables were the real exchange rate, credit growth and the M3 to reserves ratio. Together they provide some useful information about the risks of a possible currency crisis. Specifically, if these variables have been consistently above their average levels during normal times, then a country would seem to be potentially vulnerable to a crisis in the event of any exogenous shock that adversely affects investor confidence.

The overvaluation of the real exchange rate was one of the earliest and most persistent signals of vulnerability. As early as 13 months before a crisis, a real appreciation of the domestic currency relative to it’s previous two year average tended to signal a currency crisis. Moreover, this signal persisted throughout the buildup to the crisis. Other variables that displayed these properties were the growth of domestic credit and the M3 to reserves ratio : Low domestic real interest rates, reflecting easy monetary

conditions, and equity price declines significantly signaled currency crisis only for the industrial countries. On the other hand a terms of trade deterioration of around eight months prior to the crisis provided a strong signal only for the emerging market countries. Finally, and interestingly enough, the world interest rate was not a significant indicator for both groups of countries except of times that were very close (about 3 months prior) to a crises.

Table 14.2 : Significance of Early Warning Indicators of Vulnerability To Currency Crises in Emerging Market Economies (1975 - 97)

| Indicator | Months prior to a crisis | | |
|---|---|---|---|
| | 13 | 8 | 3 |
| Real exchange rate appreciation | ▪ | ▪ | ▪ |
| Domestic credit expansion | | ▪ | ▪ |
| M3-to-reserves expansion | ▪ | ▪ | ▪ |
| Terms of trade deterioration | | ▪ | |
| World real interest rate increase | | | ▪ |

However, it needs to be noted that in instances where a crisis in one country spills over or spreads contagiously to other countries owing, say, to trade or financial linkages these variables may not provide the best indicators for the non-originating countries. In such cases, a crisis in a closely linked economy, or in an economy perceived to have broadly similar characteristics, may be the most informative signal of all Nevertheless, the obovet variables can serve as indicators of the vulnerability to spillovers. In the recent Asian currency crisis, although contagion effects were certainly evident in spreading the crisis, the affected economies, by and large, also displayed signs of macroeconomic vulnerability when measured against the yardstick of the above indicators.

The appreciation of the real exchange rate, the growth of real domestic credit, and the growth of unbacked domestic banking sector liabilities (the ratio of M3 to international reserves) were used to form on index of macroeconomic vulnerability to a currency crisis, which was calculated for six Asian and four Latin American countries. The index indicated that, beginning in early 1997, vulnerability increased in almost all of the East Asian economies most affected by the recent turmoil : Thailand, Malaysia, and to a lesser extent, Indonesia and Korea were all vulnerable according to the index. A sustained buildup in macroeconomic imbalances was often followed by a sudden jump in the index of foreign exchange market pressure that was used to identify the eruption of a potential currency crisis. This was most evident in the cases of Thailand and Malaysia. Such a build up was also present in the 1994-5 Mexican crisis. In the major merging market countries that successfully

resisted contagion and spillover effects from the East Asian crisis, there were no such signs of vulnerability. For instance, Argentina, Brazil, Chile, Mexico, and Singapore showed little sings of vulnerability. The rather short lead up to the crisis in Thailand shown by the index and the absence of vulnerability in some of the non - Asian emerging market economies that did actually experience contagion effects, suggests that other indicators such as the 'shadow floating exchange rate, also need to be monitored.

However, it is unlikely that any single index is ever going to capture the complexity of developments leading up to a crisis, which usually includes significant element of vulnerability coupled with economic disturbances, political events, or changes in investor sentiment associated with contagion effects. Indicators of vulnerability need to be supplemented with country - specific information in order to arrive at a judgment concerning a country's true vulnerability to a currency crisis. As noted earlier, the usefulness of the index as an early warning system depends also on the availability of timely information. If the relevant information is not available on a timely basis, the index merely serves to summarize certain elements of vulnerability after the event and is useful only as an analytical tool to study historical crises.

## 14.5 ANALYTICAL AFTERTHOUGHTS ON THE ASIAN CRISIS:

The East Asian economic meltdown not only vindicated currency crisis theory but also demonstrated in a devastatingly through manner the importance of the subject; that in a world of high capital mobility, the threat of speculative attack becomes a central issue indeed, for some countries the only issue - of macroeconomic policy. However, viewing the Asian crisis through the lens of conventional currency crisis theory which basically focuses attention on the relationship between fiscal monetary, and exchange rate policies seems to reveal the inadequacy of existing crisis models.

In order to explain this lacuna we have to, following krugman (1990a), think in term of the distinction between first generation and 'second - generation' crises models. The canonical first - generation crisis models, exemplified by krugman (1979), in effect explain crises as the product of budget deficits; it is the ultimately uncontrollable need of the government for seignorage to cover its deficit that ensures the eventual collapse of a fixed exchange rate, and the efforts of investors to avoid suffering capital losses when that collapse occurs provoke a speculative attack on the currency when foreign exchange reserves fall below a critical level. The self fulfilling multiple equilibria second-generation models, exemplified

by obstfeld (1994), instead explain crises as a result of a conflict between a fixed exchange rate and the desire to pursue a more expansionary monetary policy ; when investors begin to suspect that the government will choose to let the partly go, the resulting pressure on interest rates can itself push the government over the edge, that is the crisis is a result of self-fulfilling outcomes. Both class of models have had considerable relevance to currency crises in the 1990s for example, the sterling crisis of 1992 was driven by the (correct) perception that the UK government would, under pressure, choose domestic employment over exchange rate stability; while the Mexican crisis of 1994 was caused by the banking sector bailout which, by increasing the supply of money, hastened reserve depletion thereby precipitating a speculative attack.

Despite the usefulness of these models in making sense of many historical crises, it has become clear that they miss certain important aspects of the Asian crisis which seem to have differed from the standard story in several distinctive ways. To begin with, the Asian crisis arrived suddenly with little warning. By normal criteria, most of the macroeconomic indicators were in good shape, apart from the fact that current account deficits were larger in Thailand and Malaysia. On the eve of the crisis, all the governments were more or less in fiscal balance; nor were they engaged in irresponsible credit creation or runaway monetary expansion. Indeed, right up to the summer of 1997, many observers echoed the conclusions of the now-discredited World Bank (1930) report, The East Asian Miracle that good macroeconomic and exchange-rate management was the key ingredient in the Asian recipe for success. Thus none of the fundamentals that drive the first generation crisis models seem to have been present in any of the afflicted Asian countries. Second, although there was some slowdown in growth in 1996, the Asian Victims did not have any substantial unemployment problems when the crisis began, implying that there was no strong case for any of these countries to carry out a devaluation for competitive or macroeconomic reasons. In other words, there did not seem to be the kind of incentive to abandon the fixed exchange rate in order to pursue expansionary monetary policy that is generally held to be the cause of the 1992 ERM crisis in Europe. Thus, the Asian crisis was not brought about by macroeconomic temptation as is usually the case in the second generation models. Clearly something else was at work; implying the need for a ‘third generation’ crisis model which approximates as closely as possible the stylized fats of actual experience. In such a context, it has been suggested that there are three essential aspects that such a model should attempt to capture.

### 14.5.1 Contagion :

The first and most stunning aspect of the recent financial crisis has been the extent to and the speed by which instability in foreign exchange markets was transmitted across countries. The initial speculative attack on one currency (the Thai baht in June 1997) led to a contagious attack on another in a matter of weeks (The Philippines' peso in July 1997), spilled over to the currencies of two more countries (The Malaysian ringgit and Indonesian rupiah by September 1997), before spreading all the way over to the currency of another country with apparently sound fundamentals (the Korean won in December 1997). The fact that even South Korea succumbed to the contagion seems to suggest, as Krugman (1998) put it, that,... bahtulism apparently mutated into an even more virulent strain by the time it reached northeast Asia. The virus then attacked, albeit with less force and persistence, Hong Kong, Singapore, and Taiwan within the region, and finally before dying, in a highly attenuated form, it managed to affect a number of emerging market economies in other regions as well.

In such a context, it would be useful to distinguish three sets of reasons as to why currency crises tend to be clustered in time. One is that crises may stem from a common cause for instance major economic shifts in industrial countries that trigger crises in emerging markets in what has been referred to as 'monsoonal effects'. The sharp increases in US interest rate in the early 1980s was an important factor in the Latin American debt crisis. Similarly, the large appreciation of the US dollar, especially versus the Japanese yen, between mid 1995 and 1997 contributed to the weakening of the external sector in several South East Asian countries. But while external event may contribute to or precipitate a crisis, a country's vulnerability to a crisis depends on domestic economic conditions and policies, such as overborrowing for unproductive purposes, a fragile financial sector, or an inflexible exchange rate system. A second reason why crises may be clustered is that a crisis in one country may affect the macro economic fundamentals in another country, either because of trade and capital market linkages (for example, a devaluation in one country adversely affects the international competitiveness of other countries) or because of interdependences in creditors' portfolios (for example, illiquidity in one market forces financial inter-mediaries to liquidate assets in other markets). Such 'spillovers' resulting from interdependences have been cited as contributing in important ways to the spread of the East Asian crisis. Finally, a third reason for clustering is that a crisis in one country may lead creditors to re-evaluate the fundamentals of other countries even if these have not changed objectively, or may lead creditors to reduce the riskiness of their portfolios and flee to quality. This is often associated with 'herding' by investors, resulting from bandwagon

effects driven by asymmetric information or from incentives faced by fund managers. It is this third effect, specifically, that is sometimes referred to as contagion (or 'pure' contagion) which arises only if financial markets exhibit multiple equilibria and self fulfilling speculative attacks (see Masson 1998).

The evaluation of the East Asian crisis suggests that spillover and contagion effects played a role, although formal empirical evidence at this juncture is sketchy. However, it needs to be noted that such evidence on contagion can never be definitive because it is impossible to be certain that the estimated model incorporates the true fundamentals, or does so correctly. For instance, given the known difficulties involved in modeling the banking sector weaknesses, it may be impossible to graft this aspect formally into any of the various versions of the financial programming models developed in this study. As such, systematic empirical modeling of contagious financial crises can still be considered to be in its infancy, especially for emerging markets, although, as mentioned earlier, there have been important developments in the construction of early warning signals of the vulnerability to currency crisis.

### 14.5.2 Balance sheet and Transfer Problems :

Descriptive accounts, both, of the fundamentals of the crisis countries and of the policy discussions that led the crisis to be handled in the way it was, place extensive emphasis on the problems of firms, balance sheets on the one side, the deterioration of these balance sheets played a key role in the crisis itself notably, the explosion in the domestic currency value of the dollar debt had a disastrous effect on Indonesian firms, and the fear of corresponding balance sheet effects on other countries was the principal reason as to why the IMF was so overly concerned to avoid any further deterioration of the affected currencies. On the other side, the prospects for any rapid recovery are especially difficult because of the weakened financial conditions of firms, whose capital in many cases was all but wiped out by the combination of declining sales, high interest rates, and a depreciated currency.

Despite the fact that the role of balance sheet problems in constraining firms has been the subject of some recent work in the macroeconomics literature, this issue has been neglected in the currency crisis literature. However, in the last year or so, a number of economists seem to have converged towards a view about the Asian crisis that might be described as 'open-economy Bernanke-Gertler' (see Krugman 1999b; Aghion, Bachetta, and Banerjee 1999). The key idea here is this; suppose that as argued by Bernanke and Gertler (1989) investment is often wealth constrained

that is, because firms face limits on their leverage, the level of investment is strongly affected by the net worth of their owners. And also suppose that for some reason, many firms have substantial debt denominated in foreign currency. Then two distinct possibilities can emerge. First, a loss of confidence by foreign investors can be self justifying, because capital flight leads to a plunge in the currency and the balance sheet effects of this plunge leads to a collapse in domestic investment. Second, the normal response to recession printing more money becomes counter productive because loose money reinforces the currency depreciation, and thereby worsens the balance sheet crunch and hence the Asian crisis seemingly irrelevant events triggered of self fulfilling collapses in confidence, and conventional macroeconomic remedies were of no avail.

And finally, if there is a single statistic that captures the violence of the shock caused by the Asian financial crisis, it is the dramatic reversal in the current account of some of the afflicted economies. In the case of Thailand, for instance, the country was forced, on account of the reversal of capital flows, to convert a current account deficit of about 10 percent of GDP in 1996 to a current account surplus of about 8 percent by 1998, that is an 18 percentage point reduction in its current account over a two year period. This desperate need to effect such a huge reduction in the current account represents what may be history's most spectacular example of the classic transfer problem debated by Keynes and ohlin in the 1920s.

Yet despite the evident centrality of the transfer problem which was effected partly through massive real depreciation and partly through severe recession that produced a drastic compression of imports to what actually happened in Asia, this issue has been conspicuously missing from all formal currency crisis models, perhaps because the modelers have been more concerned with the behaviour of investors rather than with the real economy per se, all of the major models in this context have been one good models in which domestic goods can be freely converted into foreign goods and vice versa without any movement in either the terms of trade or the real exchange rate. However, many economists are of the opinion that this is an unacceptable strategic simplification because the difficulty of affecting a transfer, and the need to achieve a current account counterpart of a reversal of capital flows, either via real depreciation or via recession, is the actual heart of the recent financial crisis.

## 14.6 SUMMARY

- Twin deficit means fiscal and trade deficit.
- Borrowing countries need to implement the macroeconomic stablisation programme after they borrow from IMF.
- This programme is IMF Financial Programming Model.
- Various types with various origins of financial crises are there.
- The Asian crisis was mainly currency crisis.

## 14.7 QUESTIONS

1) Discuss the twin deficit approach to growth.
2) Explain the IMF financial Programming Model.
3) Discuss the concept of Financial Crisis.
4) Write a note on analytical afterthoughts on the Asian Crisis.

# 15

# FINANCIAL CRISIS

**Unit Structure:**

15.0 Objectives
15.1 Introduction
15.2 A Theoretical model of financial Crisis
15.3 Policy Co - ordination and sustainability
15.3.1 ISO-Reserves Line
15.3.2 ISO Parity Curve
15.4 Policy Implications
15.5 On the design of Policy
15.5.1 The Estimated Model
15.5.2 Overambitious Targets and Unsustainable Policies
15.5.3 Domestic Adjustment and Sustainability
15.6 Summary
15.7 Questions

## 15.0 OBJECTIVES

After having studied this unit, you should be able -

- To know a theoretical model of financial crisis.
- To understand the policy co-ordination and sustainability.
- To understand the policy implications.

## 15.1 INTRODUCTION

This unit discusses the theoretical model of financial crisis, Policy Co-ordination and sustainability and the Policy implications.

## 15.2 A THEORETICAL MODEL OF FINANCIAL CRISIS

**The modified Mundell-fleming Model of financial crisis:**

Based upon the above insights, krugman (1999a) spelt out the frame work for constructing a highly simplified crisis model which also has apparent implication for policy. It is based on the mundell – Fleming framework which in its simplest version has

three equations. First is an aggregate demand equation relating domestic spending to real income (y) and the interest rate (i), together with net exports (NX) that depend on the real exchange rate (EPf/P), that is :

$$Y = D(y, i) + NX(y, EP_f/P) \quad (1)$$

Second, is a demand function fro real balances (M/P) given by:

$$M/P = L(y, i) \quad (2)$$

And, finally, in the simplest version, investors are assumed to be risk-neutral and have static expectations about the exchange rate, implying an interest. Arbitrage equation given by

$$I = if' \quad (3)$$

where $if$ if is the world interest rate.

Admittedly the model is simple: in particular, nobody believes in static expectations about the nominal exchange rate(E). However under a crawling peg exchange rate arrangement, with a pre-announced fixed rate of carvel, we can assume that once the exchange rate has been revised upwards given the existing rule, there would be no further expectations of an exchange rate depreciation during the current period thereby yielding.

The above framework can be regarded as simultaneously determining both out put (y) and the nominal exchange rate (E). The vertical line AA in fig. 15.1 should all the points at which, given (equ. 3) the domestic and foreign interest rates are equal. Meanwhile; the line GG shows how output is determined given the exchange rate; it is upward sloping because depreciation increases net exports and therefore stimulates the economy. The intersection of there two lines at E. indicates the equilibrium, level of output and the exchange rate.

To convert this into a model that can yield crises, all we need to do is to add a strong open-economy bernanke-Gertler effect Assume, then that many firms are highly leveraged, that a substantial portion of their debt is denominated in foreign currency, and that under some circumstances their investment will be constrained by their balance sheets. Then the aggregate demand equation will have to incorporate a direct dependence of domestic demand on the real exchange rate. Consequently, (eq.1) will need to be modified as follows;

$$Y = D(y, i, EP_f/P) + NX(y, EP_f/P) \quad (4)$$

Now at very favourable real exchange rates, few firms could be balance sheet constrained; so at low $EP_f/P$, the direct positive effect of the real exchange rate on aggregate demand would prevail, and hence the curve GG would be upward sloping. At very unfavourable real exchange rates, firms with foreign – currency debt would be practically bankrupt and unable to invest at all and

therefore once again the direct positive exchange rate effect on aggregate demand would dominate. As such, the curve GG would also be upward sloping in this region. However, over an intermediate range, the indirect negative balance sheet effect of the real exchange rate on investment might overwhelm its direct positive effect on export competitiveness, so that over that range, depreciation of the currency would be contractionary rather than expansionary.

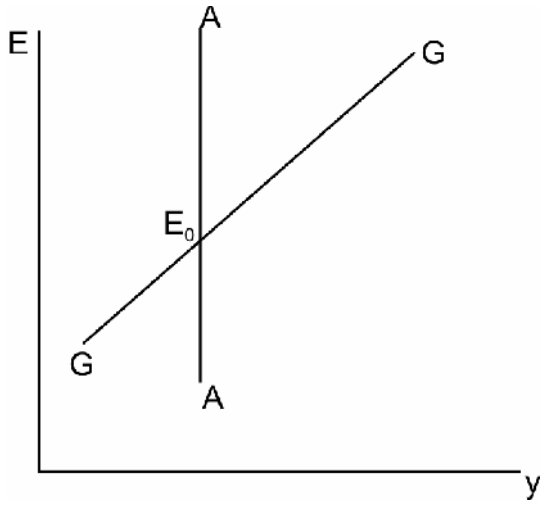



**Fig 15.1 : The Mundell- Fleming model**

In short, as pointed out by Aghion et al (1999), we, might expect the GG curve to have a backward-bending segment, as in Figure 15.1. Hence, there could be multiple stable equilibria, one with a 'normal' exchange rate, such as Eo, and one with a hyper depreciated exchange rate, such as E2. We immediately have here a simplified version of an Asianstyle financial crisis. Any event a financial crisis in another country, political instability, economic sanctions, or deliberate market manipulation by big speculators – cause a sudden lare currency depreciation; this depreciation creates have with balance sheets, and the economy plunges from normal equilibrium (Eo) into the crisis equilibrium (E2).

In the above framework, it is clear that the application of fiscal austerity does not help prevent or cure an Asian-style financial crisis. As indicated in fig. 15.2 fiscal contraktion shifts the GG curve to the left towards its new position G G and, consequently both the normal and crisis equilibrium points shift up towards Eo' and E2' implying a still greater depreciation in the currency. If these austerity measures are pushed hard enough, then the GG curve could shift so far to the left that the normal equilibrium could be eliminated altogether, leaning behind only an even more hyper depreciated crisis equilibrium.

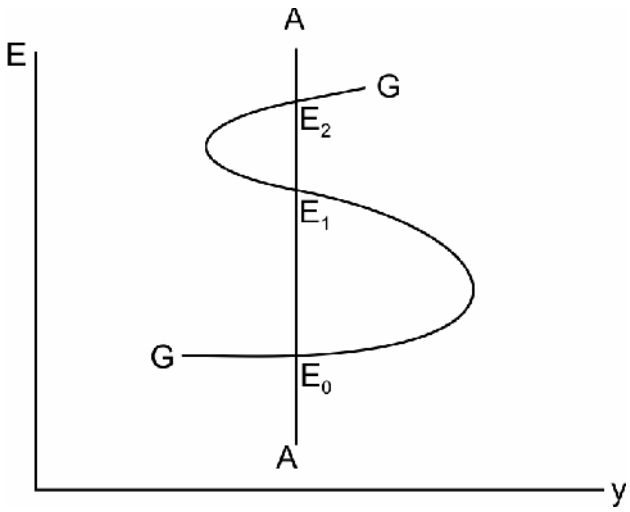



**Figure 15.2 : The modified mundell- Fleming Model**

Fiscal expansion, on the other hand. Just above might work : it shifts GG to the right, and undertaken on a sufficient scale can rule out the crisis equilibrium. The question is whether countries are able to undertake the requisite expansion Deficit spending after all did strengthen the yen, just as mundell-Fleming would predict, but is may not be a sustainable option for smaller countries that are debtors rather than creditors. Be that as it may, it is by now well know that although during the early stages of the Asian crisis the IMF imposed fiscal austerity, currently the recovery is being party driven by deficit spending.

Finally, the above model provides a rationale for the principal and much-disputed, tool in IMF stabilization programmes which is a temporary sharp tightening of monetary policy to support the exchange rate, followed by a gradual loosening on gradual loosening once confidence has been restored. Consider figure 15.3 and imagine that for some reason markets appear to have become convinced that the economy is heading for the crisis equilibrium (E2) – a belief that, if unchecked, will become self-fulfilling. One way to pre-empt this is to drastically tighten monetary policy, shifting the AA curve so far to the left that it becomes AA – that is far away to rule out the crisis equilibrium. Once investors are convinced that the exchange rate is not going to depreciate massively, this monetary contraction can be relaxed.

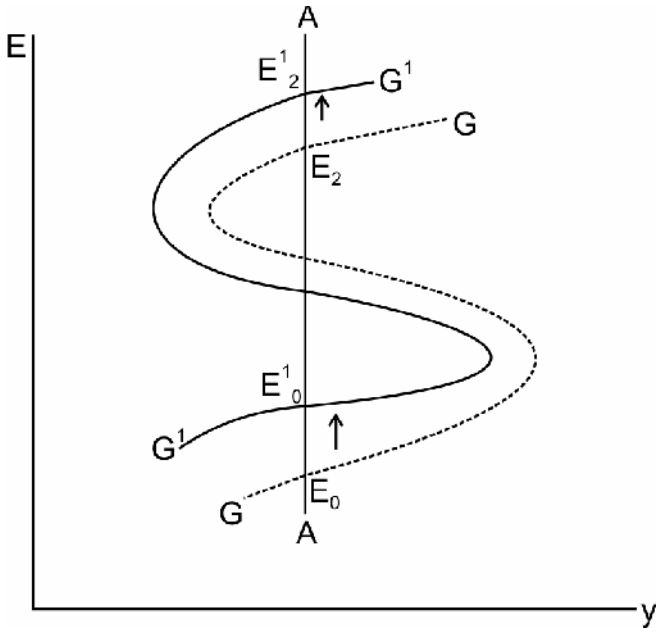



**Figure 15.3 : Fiscal Policy, in the modified model**

The problem, however, is that all along the way, from Eo to E'o the economy faces a sharp and sustained contraction in real output which is a feature that is being currently witnessed in these. Asian countries. However, it could be argued, based on the korean experience, that this strategy-of imposing a temporary monetary contraction and hence a severe, but hopefully short-lived, real contraction-works in the end. By and large, the analysis indicates that a moderate fiscal (monetary) expansion (contraction) could serve to defuse the situation by dissipating the crisis equilibrium although, going by the government budget constraint specified in Figure 15.4 it is immediately obvious, that this would in the absence of official capital floures-entail an increase in private sector borrowing. This, in turn, could, given the private sector budget constraint specified in Figure 15.4 could out private investment even in the normal equilibrium putting the economy into a different but equally unpleasant form of a low level trap. All this implies that, because the stabilization policy options confronting the authorities during a crisis could be derived from a model which integrates the best ingredients of both, the above analytical framework as well as the early warning signals of vulnerability discussed earlier.

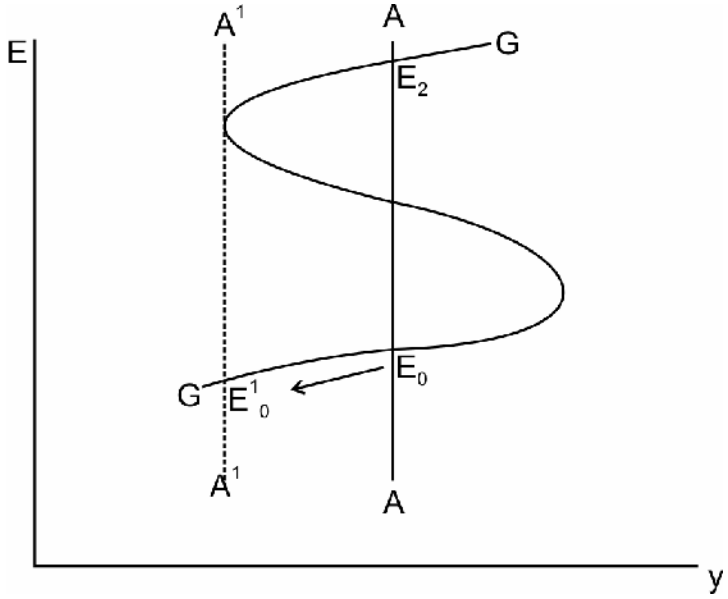



**Fig 15.4: Monetary policy in the modified model**

## 15.3 POLICY COORDINATION AND SUSTADNABILITY AN ANALYTICAL FRAMEWORK:

In such a context, we now set out a framework (see Rao 1997a, b; Rao and Singh 1998) for a currency crisis model which, as will be shown later using, empirical illustrations for the Indian economy, has considerable policy implications.

Consider an open economy in which residents consume a single tradeable good, whose foreign currency price (that is the nominal exchange rate) is revised upwards at a constant pre-announced rate every period. The domestic inflation rate is a weighted average of excess money growth relative to real output growth (the quantity theory assumption) and nominal exchange rate variations (the purchasing power parity theory assumption) with complete financial openness, the domestic nominal interest rate is governed by the uncovered interest rate parity equation. There are private banks, so that the money stock is equal to the sum of domestic credit issued by the central bank (which is assumed to expand at a constant rate) and the domestic currency value of foreign reserves held by the central bank finally, it is assumed that the output growth rate is positively influenced by the supply or real domestic credit (the credit availability effect) and real interest rates (the financial repression hypothesis), although very high real rates can reduce investment and thereby lower the growth rate (the crowding – out effect) The model is thus defined by the following equations:

$$M - P = C + \alpha y - \beta i, \quad (5)$$

$$M = \sigma_1 D + \sigma_2 R, \quad (6)$$

$$\Delta D = \mu, \quad (7)$$

$$\Delta e = \varepsilon \quad (8)$$

$$\Delta p = \delta_1 (\Delta M - \Delta y) + (1 - \delta_1) \Delta e, \quad (9)$$

$$I = i_f + \Delta e, \quad (10)$$

$$\Delta y = \theta_0 + \theta_1 (\Delta D - \Delta p) + \theta_2 (i - \Delta p) - \theta 3 (i - \Delta p)^2 , \quad (11)$$

Where M is the nominal money stock, D is domestic credit, R is the domestic currency value of foreign exchange reserves, e is the nominal exchange rate, p is the price level, y is the real output, i is the domestic nominal interest rate and if is the (exogenous) foreign interest rate. All variables, except interest rates, are measured in logarithms, and therefore the change in the logarithmic values of these variables would denote their growth rates. As such, in all the ensuing derivations we shall set

$\Delta p = \pi$ (inflation rate) and

$\Delta y = g$ (real growth rate)

(Eqn 5) relates real money demand positively to real income and negatively to the nominal interest rate (Eqn 6) is a log-linear approximation of the identity defining the money stock as the sum of domestic credit and reserves (Eqn 7) specifies that domestic credit grows at a constant rate (Eqn 8) specifies a crawling peg exchange rate arrangement with the nominal exchange rate being depreciate at a constant rate each period (Eqn 9) indicates that the inflation rate is a weighted average of the relative excess liquidity and the depreciation rate (Eqn 10) provides the interest rate parity equation where the expected rate of depreciation is replaced by the actual rate under the assumption of a constant pre-announced rate of crawl. Finally (Eqn 11) indicates that the real growth rate of output is positively related to the growth rate of real domestic credit as well as the real rate of interest. The negative parabolic term is introduced on the assumption that once the real rate crosses a critical threshold, the resulting crowing out of investment would adversely affect growth.

### 15.3.1 ISO - Reserves Line

Combining together the time derivatives of (equ 5), (eqn 6) and (eqn 10) yields:

$$\sigma_1 \Delta D + \sigma_2 \Delta R - \pi = \alpha g - \beta \Delta if \quad (12)$$

Substituing (eqn 9) into (eqn 12) above, and using (eqn 7) and (eqn 8), yields:

$$\sigma_1\mu + \sigma_2 \Delta R - \delta_1 (\delta_1 \mu + \delta_2 \Delta R - g) -$$

$$(1 - \delta_1) \varepsilon = \alpha g - \beta \Delta \text{ if} \qquad (13)$$

Assuming $\alpha = 1$ yields the following equation for the rate of change of reserves :

$$\Delta R = [(g + \varepsilon - \sigma_1\mu) / \sigma_2] - [\beta\Delta_{if} / \sigma_2 (1 - \delta_1)] \qquad (14)$$

(Eqn 14) indicates that if domestic credit expansion exceeds the sum of the real growth rate and the crawl rate, reserves are depleted each period. Thus, any finite stock of reserves will be exhausted in a finite period of time.

The equation also indicates that even if there is consistency between monetary and exchange rate policy in as much as $\varepsilon=\sigma_1\mu$-9 so that the first term in parentheses on the right hand side of (eqn 14) vanishes, reserves can still be depleted if foreign it interest rates are rising, that is $\Delta$ if > 0, because this would lead to reserve depletion via capital outflows. To offset this, even higher craul rates would automatically be required. Thus, exchange rate management, a part from necessarily being consistent with monetary policy, would also be subject to foreign influences.

We now assume an import demand function of the following form.

$$Z - e\ 1_n A + my - b\ (e - p), \qquad (15)$$

where Z is the logarithm of imports measured in domestic currency units. Thus, (eqn 12.15) implies that imports measured in foreign currency unit, that is z – e, are positively related to real output and negatively to the real exchange rate; with m and b measuring their corresponding elasticities. If we assume that reserves should be some fixed fraction of imports, then regardless of the value of this constant, we have:

$$\Delta R = \Delta z = mg + (1\text{-}b)\ \varepsilon + b\pi, \qquad (16)$$

Linking up (eqn 14) and (eqn 16) and assuming that $\pi = \pi^*$ (the desired inflation rate) and $g = g^*$ (the desired growth rate) yields :

$$mg^* + (1-b)\varepsilon + b\pi^* = \left[\left(g^* + \varepsilon - \sigma_1\mu\right) / \sigma_2\right] - \left[\beta\Delta_{if} / \sigma_2 (1-\delta_1)\right] \qquad (17)$$

Rearranging terms, we obtain the ISO reserves line given by:

$$\sigma_1\mu + [(1-\sigma_{2m})\, g^* - \sigma_2\, b\, \pi^* - \beta\, \Delta \text{if} / \sigma_2 (1-\delta_1)] + [1-\sigma_2 (1-b)\, \varepsilon, \qquad (18)$$

which is an upward sloping line in $\varepsilon$ - $\mu$ space – depicted by the RR line in figure 15.5 implying that higher domestic credit

growth rates (which deplete reserves) would have to be compensated by rising crawl rates (which attract reserves) in order to maintain reserve) at their desired level. All points lying above (below) the RR line indicate that the actual reserves accretion rate is lower (higher) than the desired target rate.

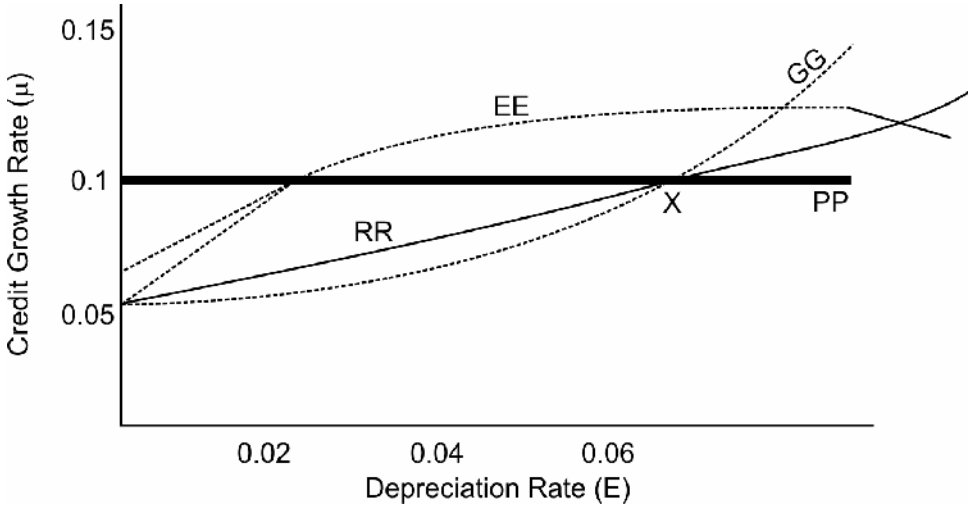



**Figure 15.5 : An Analytical Model of currency crisis**

**ISO – Lnflation Line**

Substituting the time derivative of (eqn 6) into (eqn 9) yields :

$$\pi = \delta_1 \left( \sigma_1 \mu + \sigma_2 \, \Delta R - g \right) + \left( 1 - \delta_1 \right) \varepsilon, \qquad (19)$$

which is obtained by invoking (eqn 12.7) and (eqn 12.8). Substituting (eqn 16) into (eqn 19) above and setting, as before, $\pi = \pi^*$ and $g = g^*$ yields the iso-inflation line given by :

$$\delta_1 \sigma_1 \mu = \left[ \delta_1 \left( 1 - \sigma_2 m \right) g^* + \left( 1 - \delta_1 \sigma_1 b \right) \pi^* \right]$$

$$- \left[ \delta_1 \sigma_2 \left( 1 - b \right) + \left( 1 - \delta_1 \right) \right] \varepsilon \qquad (20)$$

which is a downward sloping line in $\varepsilon$ - $\mu$ space – the PP line in Figure 15.5 implying that high credit expansion rates (which increase inflation) would have to be offset by low crewel rates (which decrease inflation) to keep the inflation rate at its desired level. All point above (below) the PP line indicate that the actual inflation rate is higher (lower) than the target rate.

**ISO – Growth Curve**

Substituting (eqn 10) into (eqn 11), setting $\pi = \pi^*$ as well as $g = g^*$ and using (eqn 7) and (eqn 8) yields the following ISO – growth curve.

$$\theta_1 \mu = \left[ g^* - \theta_0 + \theta_1 \pi^* - \theta_2 \left( if - \pi^* \right) + \theta_3 \left( i_f \pi^* \right)^2 \right]$$
$$- \left[ \theta_2 \, 2\theta_3 \left( i_f - \pi^* \right) \right] \varepsilon + \theta_3 \varepsilon^2 \qquad (21)$$

Which is a parabola in $\varepsilon$ - $\mu$ space shown by the convex curve GG in figure 15.5. As the real interest rate is given by r (= i - $\pi$ = if + $\varepsilon$ - $\pi$), it implies that risign values of $\varepsilon$, by increasing the real rate, would be initially growth inducing, there by requiring lesser rtes of credit expansion. However, once $\varepsilon$ exceeds a critical level, the ensuing high value of r would retard growth and therefore require a higher rate a credit expansion to offset this effect.

**15.3.2 ISO Parity Curve**

Assume that there exists a pre-announced minimum level of reserves (Rmin) which, if reached, would compel the central bank to abandon the crawling peg exchange rate system, withdraw-from the foreign exchange market, and allow the exchange rate to float freely. In such a situation, the speculative attack would occur at the point where the 'shadow floating exchange rate', which reflects market fundamentals, is equal to the prevailing crawling rate.

The shadow floating rate is the exchange rate that would prevail with the current credit stock if reserves had fallen to the minimum level and the exchange rate was allowed to float freely. As long as the official parity is more depreciated than the shadow floating rate, the existing exchange rate regime is viable; beyond that point, it is not sustainable, The reason is that if the shadow, floating rate is below the official parity, speculators would not profit by driving down the central bank's stock of reserves to its pre-announced lower bound (R min) thereby forcing an adoption of a floating rate regime, since they would experience an instantaneous capital loss on their purchases of foreign reserves. On the other hand, if the shadow floating rate is above the official parity, speculators would experience an instantaneous capital gain. As neither anticipated capital gains or losses are compatible with a perfect foresight equilibrium, speculators will compete with each other to eliminate such opportunities. This type of behavioir incorporates the arbitrage condition that the pre-attack official parity must equal the post-attack floating rate at the exact time of the speculative attack.

The first step, therefore, is to find the expression for the money stock when reserves reach their lower bound (R min). This is given by :

$$M = \sigma_1 D + \sigma_2 R_{min} \qquad (22)$$

Setting, as before, $\alpha = 1$, taking the rates of change of (eqn 5) and (eqn 22), using (eqn 7) and (eqn 9) and rewriting the resultant expression in terms of $\Delta e$, which has now to be interpreted as the rate of change of the shadow floating exchange rate, yields :

$$\triangle e = \sigma_1 \mu - g + \beta \triangle i_f \,/\, (1 - \partial_1) \qquad (23)$$

(Eqn 23) indicates that the lower the credit growth rate, $\mu$; or the higher the real growth rate, g; the slower will be the rate of depreciation of the shadow floating exchange rate. As far as the parameters are concerned, it is seen that the larger the proportion of domestic credit in the money stock, $\sigma_1$; or the greater the sensitivity of the inflation rate to excess money growth, $\sigma_1$; or the larger the value of $\beta$, the faster will be the rate of depreciation of the shadow floating rate.

Now, if the crawling exchange rate is above the shadow floating exchange rate to begin with (which is a necessary condition for the crawling-peg-regime to be viable in the first place), and the rate of crawl is exactly equal to the rate at which the shadow floating rate is depreciating, then the shadow floating rate can never overtake the prevailing crawling rate. And as long as the crawling exchange rate is more depreciated than the shadow floating exchange rate, the crawling peg regime is viable.

The optimal crawl rate is therefore obtained by equating the right hand side of (eqn 8) which is the rate of change of the official parity – with the right – hand-side of (eqn 23). which is the rate of change of the shadow floating rate doing so yields;

$$\varepsilon = \sigma_1 \mu - g + \beta \triangle i_f \,/\, (1 - \delta_1), \qquad (24)$$

Which, from (eqn 14) is seen to be exactly equal to that rate of craved for which $\Delta R = 0$. Thus, the results indicate that if the craved rate is less than the rate at which the shadow floating rate is depreciating, sustained speculative attacks would result in reserve depletion. Substituting. (eqn 11) into (eqn 24) above and setting. $\pi = \pi^*$ as before, yields the so-called iso-parity curve which is given by:

$$(\sigma_1 - \theta_1)\mu = [\theta_0 - \theta_1 \pi^* + \theta_2 (\text{if} - \pi^*)$$
$$- (\theta_3 (\text{if} - \pi^*)^2 - \beta \triangle \text{if} /$$
$$(\ell - \delta_1)] + [(1 + \theta_2) + 2\theta_3 (\text{if} - \pi^*)]\varepsilon - \theta_3 \varepsilon^2 \qquad (25)$$

Which is parabola in ε - μ space depicted by the concave curve EE in figure 16.5. This ISO-parity curve defines all combinations of μ and ε for which the rate of depreciation of the shadow floating exchange rate. It needs to be noted that all points lying above the EE curve indicate combination of μ of and ε at which the shadow floating exchange rate would be depreciating faster than the official parity (whose depreciation rate is given by ε). This implies that the shadow exchange rate would eventually overtake the official parity, culminating in a speculative attack which would ultimately compel the abandonment of the official parity. Consequently, all points lying in this region indicate an unsustainable mismatch between monetary and exchange rate policy.

## 15.4 POLICY IMPLICATIONS

Thus, all policy mixes can be categorized as follows consistent and sustainable, inconsistent but sustainable, consistent but unsustainable, and in consist and unsustainable.

If, now, for a desired $\pi_0$ and $g_0$ the iso-reserves line (RR), the iso inflation line (PP), and the iso-growth curve (GG) all interest at a unique point below the iso-parity curve (EE)such as the point X in figure ….. it would imply that the policy mix is both consistent as well as sustainable. If, however, one targets over-ambitious inflation and growth rates, such as $\pi < \pi_0$ and , $g_1 > g_0$, then the resulting policy stance would imply a higher $\mu$ (to increase g) and a lower 8 (to reduce $\pi$) and this pair could very well lie above the iso-parity curve implying unsustainability.

Equally true, an increase in the world interest rate (if) could push down the iso-parity curve to such an extent that, although there is no change in the original targets $(\pi_0, g_0)$, the intersection of the PP and GG lines could still take place above the EE line rendering the policy stance un-sustainable. This would imply that the desired growth (inflation) target would have to be revised downwards (upwards) until the iso-inflation and iso-growth lines once again interest below the iso parity line.

All this suggests that there is much more to policy coordination under complete financial openness than what is implied in the literature, because it entails a very high level of policy

discipline and flexibility in order to rapidly adapt to changing circumstances.

## 15.5 ON THE DESIGN OF POLICY : SOME ILLUSTRATIVE EXAMPLES

### 15.5.1 The Estimated Model:

In order to empirically apply the above theory in the Indian context, as we have been doing all along, we need numerical estimates of the parameters of the model. These were obtained by using annual time. Series data on the Indian economy over the 12 year period 1986-7 to 1997-8 and then applying the Kalman filtering and smoothing recursion algorithms (see Rao 1997b) to this data set. The five estimated behavioural equations of the model were :

M - P = - 0.3565 + Y - 0.5335 i, (12.26)

M = 0.9547 D + 0.0641 R (12.27)

$\pi = 0.8453(\Delta m - g) + 0.1547 \Delta e$, (12.28)

$g = 0.0378 + 0.4286\ (\mu - \pi) + 0.6576(i - \pi) - 8.3594(i - \pi)^2$ (12.29)

z = - 15.6171 + 2.1328 y - 0.2455 e + 1.2455 p (12.30)

Equation 12.26) was estimated by initially constraining $\alpha$ to be equal to unity following the analytical derivation (equation 12.28) implies that about 85 percent of the inflation rate is explained by excess money growth and the remaining 15 percent by exchange rate variations (equation 12.29) besides validating the credit availability hypothesis, indicates that, while growth rates would be stimulated at low real rates of interest once real rates cross a threshold limit of about 4 percent, growth rates would be adversely affected. (equation 12.30) indicates that the elasticity of nominal imports with respect to the nominal exchange rate is about 0.25.

### 15.5.2 Overambitious Targets and Unsustainable Policies:

Using the parameter values listed above in (equation 12.26) - (equation 12.30), we initially set the desired growth rate at 7 percent, the desired inflation rate at 2 percent, and the foreign interest rate at 9.3 percent. We thus obtained the following four curves in $\varepsilon - \mu$ space :

ISO - reserves : $\mu$ = 0.0616 + 1.06398 (12.31)

ISO - inflation : $\mu$ = 0.0864 - 0.1752 E (12.32)

ISO - growth : $\mu$ = 0.0871 + 1.3133 E + 19.5040 $E^2$, (12.33)

ISO - Parity : $\mu$ = 0.0621 + 0.8309 E - 15.8894 $E^2$ (12.34)

The intersection of the ISO - inflation and ISO-growth lines yields a domestic credit expansion rate of 8.7 percent (that is $\mu$=0.087) and a fixed exchange rate (that is $\varepsilon = 0.00$). However, these policy settings are unsustainable because the intersection of the PP and GG lines occurs above the iso-parity curve EE. The reason for this is with such a high world interest rate, both (equation 14) and (equation 24) indicate that the rate of credit expansion (exchange rate depreciation) should be decreased (increased) to ensure sustainability. However, the overambitious target levels do not permit these adjustments because that would imply a lower (higher) growth (inflation) rate. This implies that there is every likelihood of succumbing to a speculative attack because the optimal crauul rate necessary to realize the desired inflation rate which in this case is zero percent is lower than the rate at which the shadow exchange rate is depreciating making it profitable for speculators to attack the currency eventually. Therefore, the results highlight the fact that an overheating economy because it involves destabilizing increases in the shadow floating exchange rate, becomes increasingly vulnerable to external shocks. This could be one possible reason as to why, despite such strong economic fundamentals, some of the afflicted East Asian countries succumbed to the contagious speculative attack, because, by and large, the macroeconomic policies they pursued to achieve such overambitious targets were in fact, unsustainable.

### 15.5.3 Domestic Adjustment and Sustainability:

Repeated iterations indicated that a 4.3 percent growth rate and a 5 percent inflation rate would entail a sustainable policy mix at this high world interest rate of 9.3 percent. Using these altered settings, we obtained the following four curves in $\varepsilon - \mu$ space (see Figure 15.5)

ISO - reserves : $\mu$ = 0.0347 + 1.0639 E (12.35)

ISO - inflation : $\mu$ = 0.0967 - 0.1752 E (12.36)

ISO - growth : $\mu$ = 0.0322 + 0.1430 E + 19.5040 $E^2$, (12.37)

ISO - Parity : $\mu$ = 0.0555 + 1.7842 E - 15.8894 $E^2$ (12.38)

For these revised target values, the iso-infaltion, iso-growth and iso-reserves lines all intersect at a unique point which is denoted by the point x in Figure 15.5 that is below the iso-parity curve - indicating a domestic credit growth rate of 8.8 percent and on exchange rate depreciation of 5 percent. As the PP and GG lines intersect on the RR line, it implies that the instrument pair $\left(\mu^* = 0.088, \varepsilon^* = 0.05\right)$ not only attains the corresponding target pair ($g^{\&}$ = 0.043, $\pi^*$ = 0.05), but is also simultaneously compatible with the reserve accretion target. Thus, it is seen that consistency and sustainability have both been achieved, albeit at a considerable

cost as it entails scaling down the growth rate besides allowing inflation to build up. This is broadly indicative of the extent of domestic adjustment that is necessary to accommodate external shocks of such a magnitude. Thus, it is noticed that the resulting strategy of imposing a drastic real monetary contraction and hence a severe real contraction prevents the speculative attack and, analogous to the model thereby rules out the crisis equilibrium.

## 15.6 SUMMARY

- A theoretical model of financial crisis is nothing but the modified Mundell - Fleming model of Financial Crisis.
- For sustainability in the economy, there should be policy co-ordination.

## 15.7 QUESTIONS

1) Explain the theoretical model of Financial Crisis.
2) Discuss the link between policy co-ordination and sustainability.
3) Write a note on Policy implications.
4) Explain the estimated model.

# 16

# MODULE - 6

# PUBLIC FINANCE WITH MANY JURISDICTIONS FISCAL FEDRALISM

**Unit Structure**

16.0 Objectives
16.1 Decentralization Theorem
16.2 Tax assignment problem: meaning & models/ approaches
16.3 Public choice and fiscal federalism
16.4 Fiscal federalism as a Principal-agent problem
16.5 Questions

## 16.0 OBJECTIVES

- To introspect the theorem of decentralization.
- To bring out the approaches related to allocation of tax jurisdictions between national and sub-national governments in a federal set up.
- To discuss the implications of fiscal federalism

## 16.1 DECENTRALIZATION THEOREM

The 'Decentralization theorem' is central to the discussion of fiscal federalism. Whether fiscal responsibilities should be assigned to a centralized authority or be decentralized has been a long debated issue in public economics. The decentralization theorems given by Wallace Oates (1972) states that in the absence of cost savings from centralization and inter jurisdictional externalities, responsibilities should be decentralized. This argument implicitly assumes that the center is unresponsive to preference heterogeneity and thereby is only able to implement uniform policies. More specifically individual local governments are presumably much closer to the people; they posses knowledge of both local preferences and cost conditions that a central agency is unlikely to have" (Oates, 1999, p.1123). If the geographical scope of a jurisdiction falls short of the spatial pattern of spending benefits, the optimal assignment of policy tasks is deduced by

trading off the welfare costs of policy uniformity against the welfare gains from internalizing spillovers in policy-making. (Alesina and Barro (2002).

The propositions and implications of the theorem are:

1. Centralization yields higher welfare when spillovers/externalities are sufficiently high.
2. Centralization yields higher welfare when spatial spread is low and people's tastes and preferences are aligned or more similar across jurisdictions.
3. Decentralization yields higher welfare when spillovers are sufficiently low.
4. Decentralization yields higher welfare when spatial spread is wide and people's tastes and preferences are heterogeneous across jurisdictions.

Decentralization results in welfare gains due to improved allocation of resources in the public sector. Individual/local jurisdictions/governments can adjust levels and composition of public goods and services as per the tastes, requirements and features of respective communities.

The decentralization theorem given by Wallace Oates also brings out the precise nature of these gains and the determinants of their magnitude.

The discussion is divided into 2 parts: divergences in demand for local public goods and the issue of cost differentials across jurisdictions.

**Welfare loss from centralization and welfare gains from fiscal decentralization:**

Figure 16.1 given ahead depicts the demand curves for a local public good of the representative residents of jurisdictions one and two. The assumption here is that the local public good can be provided at a constant cost per unit per resident of *MC.* The good is thus taken to be subject to congestion in the same way as a private good. We see in the figure that the optimal outputs of the local public good are $E_1$ in jurisdiction one and $E_2$ in jurisdiction two.

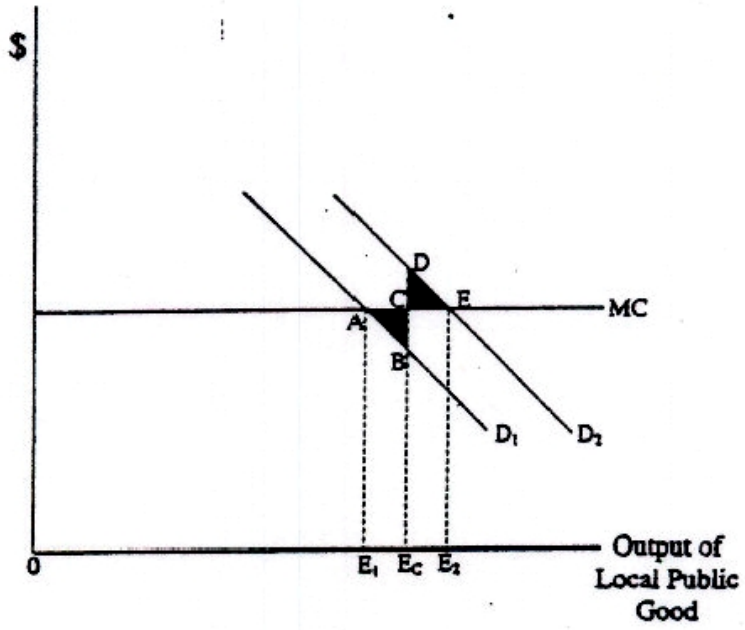



**Figure 16.1**

Suppose, however, that instead of this decentralized outcome, the central government determines that a uniform level of output of $E_c$ is to be provided in all jurisdictions. It is straightforward to measure the loss in social welfare from centralized provision: the loss is triangle *DCE* for each resident of jurisdiction two and triangle *ABC* for each resident of jurisdiction one. It is thus clear that in this setting, a uniform, centrally determined level of local public outputs will result in a lower level of social welfare than an outcome in which each local jurisdiction provides its own optimal output. This result incidentally, is the so-called Decentralization Theorem [OATES,1972, ch. 2].

In addition, we can see in Figure 16.1 what determines the magnitude of the gain in social welfare from fiscal decentralization. First it is clear that the greater the divergence between$D_1$ and $D_2$, the larger will be the triangles *ABC* and *CDE.*

Fiscal decentralization is thus more important where the demand for local public goods has greater variation across jurisdictions. In countries where the population is quite homogeneous in terms of the demand for local public goods, the potential gains from fiscal decentralization are correspondingly smaller; we would thus expect the political and social forces pushing for devolution to be somewhat weaker than in countries characterized by greater divergences in demands for local outputs.

The extent of the welfare gains from fiscal decentralization depends importantly on the grouping of populations according to

individual demands for local public goods. Such segregation can be facilitated to some extent by the mobility of households in response to local fiscal differentials. In the limiting case, the famous Tiebout model [1956] envisions an outcome in which perfectly mobile households locate in jurisdictions that satisfy precisely their demands for local public goods. A Tiebout equilibrium is thus one in which localities are perfectly homogeneous in terms of demands for the local public good. Figure 1 depicts such an outcome. This kind of sorting process maximizes the capacity of fiscal decentralization to promote social welfare.

Figure 1 also provides an insight into the determinants of the welfare gains from fiscal decentralization. The size of the welfare-gain-triangles *ABC* and *CDE* depend on the slope of the demand curves. More specifically, the steeper are the demand curves, the larger are the triangles and hence the greater is the gain in social welfare from differentiation in local outputs. This follows because where demand is less price responsive, i.e. relatively inelastic demand, individual valuations of marginal units change relatively rapidly as we move away from the optimum.

**Inter jurisdictional cost differences and the welfare gains from fiscal decentralization:**

Cost differentials across jurisdictions (as well as differences in demands) can be a source of welfare gains from fiscal decentralization. Figure 2 given below, depicts such a case. Suppose that everyone has the same demand for local public goods, namely demand curve *DD,* but that the marginal cost of providing a unit per person differs between the two jurisdictions *($MC_1$,* in jurisdiction one and $MC_2$ in jurisdiction two). In Figure 2 the Pareto efficient outcomes are $E_1$ and $E_2$ respectively. In this case, centralized provision of a uniform level of output, *E* results in welfare losses per resident of triangle *ABC* in jurisdiction one and triangle *CDE* in jurisdiction two.

We see two results from the diagram. First, it is obvious that the size of the welfare loss triangles vary directly with the magnitude of the inter jurisdictional cost differential. The greater the distance between $MC_1$ and $MC_2$ the optimal outputs in the two jurisdictions will farther diverge from one another and the larger will be the social loss in welfare from a centrally determined, uniform level of output.

**On the Welfare Gains from Fiscal Decentralization**



FIGURE 2

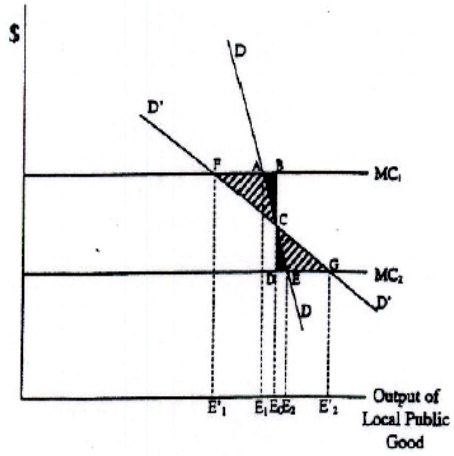



**Figure 16.2**

Second, in contrast to the variation in the demand case, we find that, for the case of cost differences, the welfare gain from fiscal decentralization varies *inversely* with the absolute value of the slope of the demand curve. In this case, the less steep are the demand curves *(i.e.,* the more price-responsive is the demand for local outputs), i.e. with relatively elastic demand, the more divergent will be the efficient outputs in the two jurisdictions and the greater the loss in social welfare associated with a centrally prescribed and uniform level of local outputs. In Figure 2, we see that for the more price-elastic demand curve *D'D',* the welfare-gain triangles from fiscal decentralization increase to *CBF* in jurisdiction one and *CDG* in jurisdiction two. So we find that the effect of the price responsiveness of demand on the potential welfare gains from fiscal decentralization depends on whether the divergence in Pareto-efficient local outputs has its source in inter jurisdictional variation in demand or variation in costs.

Inter jurisdictional cost differentials can result from two different sources. First, it may simply require more of inputs to provide a given level of output in one place than another. For example, keeping the roads clear of snow in the winter will require more effort in an area that gets lots of snow than in one with a milder winter season. Thus, the difference between $MC_1$ and $MC_2$ *in figure 2* may simply result from differences in the production functions such that one jurisdiction requires more inputs per unit of output than another.

The second source of cost differentials is that the public goods may exhibit quite different congestion characteristics. In the Tiebout model, for example, an extreme assumption is made: local public goods are postulated to be fully congestible (like private goods) in the sense that a doubling of the number of people who consume the good requires a doubling of the quantities of inputs in order to keep public output per resident the same.

## 16.2 TAX ASSIGNMENT PROBLEM: MEANING & MODELS/ APPROACHES

**Introduction:**

Tax assignment problem refers to the task of deciding who should tax what and how in a federal structure. The tax assignment problem thus revolves around 4 questions:

- which level of government chooses the taxes that a given level imposes
- who defines tax bases
- which government sets tax rates
- Which government administers the various taxes.

The conventional model of tax assignment in a multi-tier governmental structure under which all productive revenue sources are assigned to the central government. Since this favours with the needs and wishes of most central governments, it is the pattern found in most countries. But in the context of increasing decentralization of important expenditures in many of these countries, this centralized assignment of revenues has put an increasing strain on intergovernmental fiscal transfers and, in some instances, facilitated irresponsible behavior by some sub national governments. One obvious way to relieve at least some of these problems would be to strengthen sub national tax regimes.

Two new approaches to accomplish such strengthening in developing countries are proposed. The first, and most important, approach is to establish sub national value-added taxes that are not subject to the well-known problems that have long been thought to preclude such taxes. Although not yet implemented in any developing country, the way in which this can be done is now well worked out in principle and has actually been implemented in practice to some extent in one developed country. Moves in this direction seem both inevitable and desirable in the fairly near future, particularly in large countries with important regional governments such as Brazil and India.

The second approach suggested here is less well worked out as yet, but seems sufficiently promising to be worth careful consideration. The fundamental idea is to replace all or some of the various unsatisfactory state and local taxes on business that exist in most countries by a "business value tax" – in essence, a relatively low rate flat tax levied on an income-type value-added base.

## 16.2.1 Models/Approaches of Tax Assignment

1. The conventional/ public economics model
   Richard Musgrave (1983) characterized the question of tax assignment in a multi-level government as "Who should tax, where, and what?" Recently by Wallace Oates answers these questions in 3 ways:

1) Lower levels of government... should, as much as possible, rely on benefit taxation of mobile economic units, including households and mobile factors of production.

2) To the extent that non-benefit taxes need to be employed on mobile economic units, perhaps for redistributive purposes, this should be done at higher levels of...government.

3) To the extent that local governments make use of non-benefit taxes, they should employ them on tax bases that are relatively immobile across local jurisdictions.

In practice, what these principles have implied is that local governments should rely primarily on user charges and taxes on real property. Only central (national) governments could, or should impose a corporate income tax (CIT) or levy a progressive personal income tax (PIT). As for intermediate (regional) governments such as states or provinces, the apparent professional consensus, if there was one, was, until recently, probably that the only really acceptable general tax was a simple single-stage and preferably retail sales tax (RST) levied directly on final (resident) consumers along with, perhaps, a few relatively uniform excise taxes.

The conventional model of tax assignment in public economics, has invariably led to the sub national governments with less in "own revenues" than the expenditures for which they are responsible. In practice, the resulting "vertical fiscal imbalance" is almost invariably resolved by transfers. It does not explain reality. It does not provide a good guide as to how to change reality. Applications of the basic guiding prove both economically undesirable and politically unsustainable in the circumstances of most countries.

### 16.2.2 Richard Musgrave's approach based on functions & benefit tax

Richard Musgrave (1959) usefully distinguishes three fiscal functions of government, which he conceives of being performed by three conceptually distinct "branches" of government: resource allocation, income redistribution, and macroeconomic stabilization. Taxes of the allocation branch should, to the extent possible, reflect benefits of public services. By comparison, progressive individual income taxes and corporate income taxes would be used to implement income redistribution and, through their countercyclical effects on revenues and disposable income, endogenous macro-economic stabilization. Taxes might also be varied exogenously to implement stabilization policy.

Given the importance of benefit taxation in the theory of tax assignment, the optimal assignment of taxes of Musgrave's allocation branch depends on the assignment of expenditure functions. Taxes intended to reflect benefits of public services (e.g., education or construction and maintenance of roads and highways) or to charge for other costs imposed on society (e.g., for medical care for smokers and those who consume alcoholic beverages) should be assigned to the level of government incurring the costs. That governments at all levels should charge those who create costs is required for fairness and economic efficiency, as well as for the financial viability of governments. Thus, while state or local financing is generally appropriate in the case of local roads and state highways, federal financing is more likely appropriate for an interstate highway system. Taxes that closely reflect benefits of public services generally would not be adequate to finance governments at any level; substantial amounts of expenditures must be financed with taxes that are only loosely related to benefits, if at all.

### 16.2.3 The Public Choice Approach

Some, but by no means all, of these defects are remedied by what may be called the "public choice" approach to tax assignment. This approach too has various manifestations. Starting from the basic statement of "the principle of fiscal equivalence" in Olson (1969), this approach has been perhaps most fully developed by Hettich and Winer (1984), although perhaps the best-known manifestation of this general line of thinking may be found in Brennan and Buchanan (1981). In a sense, this approach also lies at the root of such well-known "folk" sayings in the field as "every tub on its own bottom" and "match revenue and expenditure responsibilities."

Hettich and Winer (1984), for example, make the obvious, but important, point that governments decide which taxes to impose in terms of a political rather than an economic calculus and that such critical aspects of the latter as, for example, competition for tax base affect political decision-making with respect to taxation only to the extent they are perceived to affect (for example) the probability of being re-elected. In the tax assignment context, what this implies is that the taxes assigned to lower governments, something which in most countries is essentially determined at the discretion of higher-level governments, will fall into one of three categories -- (1) those that are too small to bother with (the minor nuisance taxes found at the local level in so much of the world); (2) those that can plausibly be labelled "local," are difficult or costly for central governments to administer, and potentially politically troublesome (the property tax); and (3) local business taxes of one sort or another. This approach yields a fairly accurate description of what one sees around the world in the way of local taxation, but it does not, of course, provide any normative guidance as to what "should" be done.

In contrast, Brennan and Buchanan (1981) are crystal clear as to what should be done. In their model, in contrast to the prescriptions of the conventional model -- which, they correctly observe, can be interpreted as a revenue-maximizing model – sub national taxes should be imposed on mobile factors in contrast to the immobile factors of the conventional model -- that is, competition is as healthy and beneficial between governments as between private economic agents.

### 16.2.4 A Policy Perspective

This simple approach suggests two useful guidelines for rethinking tax assignment problems. First, as is of course obvious, the importance of the problem depends very much upon the assignment of spending responsibilities. If, for example, local governments are responsible only for sweeping the streets and picking up the garbage, user fees and some sort of low-rate general local tax such as a uniform tax on real property will do the job. In effect, the conventional prescriptions of the public economics model produce roughly the right results in this case. On the other hand, if sub national governments (local, or perhaps more commonly, regional) are responsible for expensive services such as health or (in developing countries) especially education, the pressures on sub national revenues will be much greater and the conventional prescription seems less likely to produce sustainable results.

Second, it is critical to be clear that meaningful tax assignment refers to the assignment of the ability (and responsibility) to determine own revenues in some meaningful way.

Sub national governments may be fully financed from what they (and others) may consider their "own" taxes: but if they cannot determine which taxes they levy, what the tax base is, what tax rate is imposed, or how intensively the tax is enforced, they actually have no control at all over revenues and hence have really been "assigned" no revenue power.
Conclusion

The solutions suggested here to at least some of the existing problems of sub national taxation in developing countries are essentially based on three simple principles: (1) More attention should be paid to matching expenditure and revenue needs. (2) More effort should be made to ensure that all governments bear significant responsibility at the margin for financing the expenditures for which they are politically responsible. (3) Sub national taxes should not unduly distort the allocation of resources.

## 16.3 PUBLIC CHOICE AND FISCAL FEDERALISM

A central tenet of the public-choice approach is the view that public decision-makers are utility maximizers with their own objective functions. And this has produced, in certain instances, a rather different view of the normative properties of fiscal decentralization.

One prominent theme in this literature has its source in the Niskanen contention that public agents can be usefully characterized as seeking to maximize the size of their budgets.

Budget maximization is taken here to serve as a proxy for a variety of objectives including enhancement of power and influence, large staffs, and higher salaries. Brennan and

Buchanan (1980) extended this view to the proposition that the public sector can itself be envisioned as a monolithic agent, a “Leviathan,” that seeks its own aggrandizement through maximizing the revenues that it extracts from the economy. What is relevant here is the implication of the Leviathan view for fiscal federalism. Brennan and Buchanan see fiscal decentralization as a mechanism for constraining the expansionary tendencies of government.

Competition among decentralized governments, much like competition in the private sector, can limit the capacities of a monopolist, in this case a monopolistic central government, to increase its control over the economy’s resources. In their view, competition among governments in the context of a decentralized fiscal system with mobile households and firms “can offer partial or possibly complete substitutes for explicit fiscal constraints on the

taxing power" (1980, p. 184). The Brennan-Buchanan approach thus offers a very different view of the role of decentralization in the fiscal system. The favorable normative implications of the Leviathan viewof fiscal competition contrast sharply with some later treatments of this issue. There is now a large literature that contends that active fiscal competition among jurisdictions can itself result in distorted levels and patterns of economic activity. This topic is currently of major importance in light of the ongoing economic and political integration of Europe with numerous proposals for the harmonization of member-country fiscal and regulatory policies. Ongoing theoretical and empirical work is providing an illuminating range of insights into both the efficiency enhancing and distorting effects of such competition. I will return to the issue of fiscal competition in a somewhat expanded context in the concluding section.

But two further comments are in order. First, as to the Leviathan view of the role of fiscal decentralization, the early evidence was at best mixed; my own work found little support for the proposition that fiscal decentralization provided an effective constraint on the growth of government (Oates, 1985, 1989). But more recent work has provided some important clarification on this matter. In particular, Rodden (2003) and others, developing a central theme in the SGT and in this paper, point out that it is not fiscal decentralization per se that matters, but *what form it takes*. Rodden finds that, where decentralization involves reliance on own taxation at provincial and local levels, it is indeed associated with smaller government. But where decentralized government is financed primarily with transfers from above, the opportunities for "raiding the fiscal commons" can result in perverse programs that actually increase the size of the overall public budget.

Second, also one need to keep in mind that focus here on the Leviathan issue should not obscure the much broader range of contributions of the public-choice literature (the political-economy approach) to fiscal federalism. This literature has provided an invaluable focus on the structure of fiscal institutions and the incentives they create for public decision-makers.

On the Emerging Second-Generation Theory (SGT) of Fiscal Federalism:

The "new" literature on fiscal federalism is wide ranging both in terms of its sources and directions. It reaches not only across fields in economics but across disciplines with important contributions from political scientists and others. This makes it difficult to characterize in a simple and systematic way, but let me try. The SGT draws heavily on two basic sources:

(1) Work in public choice and political economy that focuses on political processes and the behavior of political agents. Unlike the FGT which largely assumed that public officials seek the common good, this work takes as its point of departure the assumption that participants in political processes (*both* voters and officials) have their own objective functions that they seek to maximize in a political setting that provides the constraints on their behavior. Officials don't simply act on behalf of the welfare of their constituents.

This body of work thus involves modeling of political institutions with explicit attention to the incentives they embody. Inman and Rubinfeld (1997a) have characterized the FGT as "economic federalism," which they contrast with more recent models that explicitly account for political processes and their impact on outcomes.

(2) The expansive literature on problems of information. The outcomes from collective choice institutions depend in fundamental ways on the information that the various agents possess. In particular, in settings of asymmetric information, where some participants have knowledge of such things as preferences, cost functions, or effort, knowledge that is not available to other participants, the literature has shown us that optimal "procedures" or institutions are likely to be quite different from those in a setting of perfect information. The SGT is thus drawing heavily on much of the work in industrial organization and microeconomic theory that has explored these information issues.

What emerges from these two sources is a new literature on fiscal federalism that examines the workings of different political and fiscal institutions in a setting of imperfect information and control with a basic focus on the incentives that these institutions embody and the resulting behavior they induce from utility-maximizing participants. In this setting, the basic issue of whether to centralize or decentralize certain public activities appears in a somewhat (but not altogether) different light. Under the FGT, the tradeoffs involved in this decision were, on the one hand, the inefficiencies under centralized provision of public services stemming from more uniform outputs that fail to reflect divergences in local tastes and conditions versus, on the other hand, inefficiencies in local provision resulting from the failure to internalize inter jurisdictional externalities. Where spillover effects across local jurisdictions are relatively small and where the variation in efficient levels of local outputs is relatively large, the case for decentralized provision is obviously stronger (and vice-versa).

## 16.4 FISCAL FEDERALISM AS A PRINCIPAL-AGENT PROBLEM

On the Tradeoff between Centralization and Decentralization in SGT Models:

In the wide range of models encompassing the SGT, the centralization versus decentralization trade off takes a variety of different forms. In several SGT models (e.g., Seabright, 1996), we find a tradeoff in terms of local "accountability" (sensitivity of outcomes to local preferences) versus a coordination of policies under centralization that serves to internalize inter jurisidictional interdependencies. In two important papers, Lockwood (2002) and Besley and Coate (2003) provide frameworks in which the centralized outcome is a vector of local outputs determined by a central legislature composed of locally elected representatives. These papers thus explicitly depart from the earlier assumption under the Decentralization Theorem of uniform levels of output under a centralized regime. Depending on how the central legislature functions, the centrally determined outcome exhibits various sorts of misallocations.

These inefficiencies must be weighed against the losses under a decentralized outcome where localities ignore the spillovers associated with their decisions. It is interesting that although the models under the SGT umbrella differ in fundamental ways from the FGT, many of them produce a tradeoff between centralization and decentralization that is in a somewhat similar spirit to their earlier counterparts. As Besley and Coate (2003) point out, "All of this notwithstanding, the key insight remains that heterogeneity and spillovers are correctly at the heart of the debate about the gains from centralization". But this tradeoff can been seen in intriguing new ways.

Let us note a few preliminary observations. Some of the work in the SGT spirit flows from a direct application of industrial-organization models to a public-sector setting. For example, one of the basic, workhorse models in the I-O literature is the standard principal agent model. Here, in a setting of asymmetric information with imperfect monitoring, the principle seeks to provide an incentive that will induce the agent to respond with a level of the relevant activity (or effort) that maximizes the utility of each party in a context where outcomes have a stochastic component. The solution takes the form of a contract based on observed behavior (or output). In an I-O setting, the model has been employed to study the organization of the firm, where the owner or manager is typically treated as the principal and the worker as the agent. It is not immediately clear how this framework translates into an

electoral setting with multiple levels of government (Wildasin, 2004). In such a public-sector context, who exactly is the “principal” and who are the “agents”?

The SGT literature has handled this issue in two quite different ways. The first approach is essentially to ignore the electoral dimension of public sector structure and to treat the vertical structure of the public sector much like that of the firm. The central government in this formulation, acting as the principal, seeks to structure intergovernmental fiscal relations in such a way as to get regional or local governments, the agents in the problem, to behave in ways that promote the objectives of central officials in a setting where the center has only imperfect information and control over the fiscal activities of decentralized public agents. In one such application, Levaggi (2002) makes use of a principal-agent approach to show that, under certain conditions, a central government, rather then providing a lump-sum budget to local governments, may do better by limiting the flexibility of local government service provision through the use of a “double budget constraint” which prescribes not only the overall budget but also the precise distribution of the expenditure of the funds on specified functions (or programs).

This particular approach to applying the principal-agent model to fiscal federalism is limited in its application. It basically describes a setting that Inman (2003) and others call “administrative federalism,” where regional or local governments are largely agencies that respond to central directives. There are surely cases where such an application makes sense, but much of the electoral and fiscal autonomy that we usually associate with decentralized finance under a federal system seems outside the scope of this kind of model.

A second approach that explicitly accounts for the fiscal autonomy of different levels of government in the context of electoral processes adopts a different definition of principals and agents. In this framework, the electorate itself is taken to be the principal(s) and elected officials to be the agents. To take one example, Tommasi (2003) has recently formulated a model in which the principal is not a single individual, but rather the electorate as a whole. He employs a variant of the “common-agency” model to capture the problem of the control of public officials by the citizens themselves. In this framework, a centralized system takes the form of a single agent (elected public official) who serves the whole population, while decentralization consists of one agent in each jurisdiction. The problem then becomes the design of the “optimal contract” involving a reward scheme to the agent (public official).

The analysis leads to an interesting comparison of outcomes under the alternative regimes in which we find that the case for centralization becomes stronger the larger are the externalities associated with local public outputs. This, of course, is similar in spirit to the FGT. But what is new here is the finding that decentralization may be preferable even in cases of perfect homogeneity of preferences across local jurisdictions. The case for fiscal decentralization depends not only on differences in tastes, but on the potential for better local control or "accountability" under decentralized provision.

While these models offer an enriched perspective on the choice between centralization and decentralization, they often have little to say about the structure of fiscal institutions.

## 16.5 QUESTIONS

1. Explain the theorem of decentralization.
2. Discuss the problem of tax assignment.
3. Write a note on Public Choice and Fiscal Federalism.
4. Examine fiscal federalism as a Principal-Agent problem.

# 17

# INTERGOVERNMENTAL GRANTS & REVENUE SHARING

**Unit Structure**

17.0 Objectives
17.1 Intergovernmental Grants and Revenue Sharing
17.2 The Flypaper Effect
17.3 Federal System in India
17.4 Questions
17.5 Reference

## 17.0 OBJECTIVES

- To study the challenges of resource transfers in federal set up.
- To analyze the issues related to grants provided by centre to states/regional jurisdictions
- To understand the various aspects/components of India's federal structure

## 17.1 INTERGOVERNMENTAL GRANTS AND REVENUE SHARING :

Intergovernmental grants constitute a distinctive and important policy instrument in fiscal federalism that can serve a number of different functions. The literature emphasizes three potential roles for such grants: the internalization of spillover benefits to other jurisdictions, fiscal equalization across jurisdictions, and an improved overall tax system. Grants can take either of two general forms. They can be "conditional "grants" that place any of various kinds of restrictions on their use by the recipient. Or they can be "unconditional," that is, lump-sum transfers to be used in any may the recipient wishes. The theory prescribes that conditional grants in the form of matching grants (under which the grantor finances a specified share of the recipient's expenditure) be employed where the provision of local services generates benefits for residents of other iurisdictions. The rationale here is simply the usual Pigouvian one for subsidies that induce individuals (in this case policy-makers or the electorate) to incorporate spillover benefits into their decision-making calculus.

The magnitude of the matchinc shares in such instances, should reflect the extent of the spillovers.

In contrast, unconditional "grants are typically the appropriate vehicle for purposes of fiscal equalization. The purpose of these grants is to channel funds from relatively wealthy jurisdictions to poorer ones. Such transfers are often based on an equalization formula that measures the "fiscal need" and "fiscal capacity" of each province, state, or locality. These formulae result in a disproportionate share of the transfers going to those jurisdictions with the greatest fiscal need and the least fiscal capacity. Although widely used, equalizing intergovernmental grants are by no means a necessary feature of fiscal federalism (Dan Usher 1995; Robin Boadmay 1996). Economists normally think of redistributive measures from rich to poor as those that transfer income from high to low-income individuals. Intergovernmental equalizing transfers require a somewhat different justification based on social values. In practice, such equalizing grants play a major role in many countries: in the fiscal systems of Australia, Canada, and Germany, for example, there are substantial transfers of income from wealthy provinces or states to poorer ones. In the United States, in contrast, equalizing grants from the federal to state governments have never amounted to much. Intergovernmental grants in the U.S. typically address specific functions or programs, but usually do not accomplish much in the way of fiscal equalization. At the levels of the states, however, there are many such programs under which states provide equalizing grants to local jurisdictions, notably school districts.

Fiscal equalization is a contentious issue from an efficiency perspective. Some observers see such grants as playing an important role in allowing poorer jurisdictions to compete effectively with fiscally stronger ones. This view holds that, in the absence of such grants, fiscally favored jurisdictions can exploit their position to promote continued economic growth, some of which comes at the expense of poorer ones. Fiscal equalization, from this perspective, helps to create a more level playing field for inter jurisdictional competition. But the case is not entirely persuasive. Others have argued that fiscal equalization can stand in the way of needed regional adjustments that promote development in poorer regions. McKinnon (1997a), for example, contends that in the United States, the economic resurgence of the South following World War II resulted from relatively low levels of wages and other costs. It was this attraction of low wages and costs that ultimately induced economic movement to the South, bringing with it a new prosperity. Fiscal equalization, from this perspective, may actually hold back the development of poorer areas by impeding the needed interregional flow of resources (both emigration and immigration) in response to cost differentials.

But the primary justification for fiscal equalization must be on equity grounds. And it is as a redistributive issue that it continues to occupy a central place on the political stage. In some cases, as in Canada, it may provide the glue necessary to hold the federation together. In other instances, like Italy, it may become a divisive force, where regions, weary of large and longstanding transfers of funds to poorer areas, actually seek dissolution of the union. Fiscal equalization is a complex economic and political issue.

The third potential role for intergovernmental grants is to sustain a more equitable and efficient overall tax system. For reasons we have discussed, centrally administered, non benefit taxes with a single rate applying to the national tax base will not generate the sorts of locational inefficiencies associated with varying rates across decentralized jurisdictions. Moreover, central taxes can be more progressive, again without establishing fiscal incentives for relocation. There is, in fact, considerable evidence to indicate that state and local systems of taxes are typically more regressive than central taxation (e.g., Howard Chernick 1992). There is thus some force in an argument for "revenue sharing" under which the central government effectively serves as a tax-collecting agent for decentralized levels of government. The central government then transfers funds, in a presumably unconditional form, to provinces, states, and/or localities. It is certainly possible, where the polity wishes, to build equalizing elements into these transfers. While there is here a real case for the use of intergovernmental grants, a most important qualification is that such a system of grants must not be too large in the sense of undermining fiscal discipline at lower levels of government.

The prescriptive theory of intergovernmental grants thus leads to a vision of a system in which there exists a set of open-ended matching grants, where the matching rates reflect the extent of benefit spillovers across jurisdictional boundaries, and a set of unconditional grants for revenue sharing and, perhaps, equalization purposes. Such a conception has, however, only modest explanatory power. We do, in fact, find federal matching programs that have supported a number of state and local activities with spillover effects, including, for example, grants for interstate highway construction. However, on closer examination, important anomalies appear. These grants are often closed, rather than open, ended. They thus do not provide incentives for expansion at the margin. Moreover, the federal matching shares are typically much larger than justifiable by any plausible level of spillover benefits. More generally, in a careful study of the intergovernmental grant system, Inman (1988) concludes that the economic theory of intergovernmental grants does not provide a very satisfactory explanation of the structure of U.S. grant programs; he finds that a political model can do a much better job of explaining U.S. grant

programs. Some years ago, David Bradford and I (1971a,b) tried to lay the foundations for a positive theory of the response to intergovernmental grants by setting forth a framework in which the budgetary decisions of the recipients of such grants are treated explicitly in a collective- choice setting. In short, we treated these grants, not as grants to an individual decision-maker, but rather as grants to polities that make budgetary decisions by some collective algorithm (such as simple majority rule). This exercise produced some intriguing equivalence theorems. For example, it is straight forward to show that a lump-sum grant to a group of people is fully equivalent in all its effects, both allocative and distributive, to a set of grants directly to the individuals in the group. Moreover, this result applies to an important class of collective-choice procedures, encompassing several of the major models employed in the public-finance literature. These theorems, known as the "veil hypothesis," thus imply that a grant to a community is fully equivalent to a central tax rebate to the individuals in the community; intergovernmental grants, according to this view, are simply a "veil" for a federal tax cut.

The difficulty is that this hypothesis has not fared well in empirical testing. It implies that the budgetary response to an intergovernmental transfer should be (roughly) the same as the response to an equal increase in private income in the community. But empirical studies of the response to grants have rejected this equivalence time and again. Such studies invariably find that state and local government spending is much more responsive to increases in intergovernmental receipts than it is to increases in the community’s private income. And this has come to be known as the "flypaper effect-money sticks where it hits. While this finding may not be all that surprising, it is not so easy to reconcile with models of rational choice, for it suggests that the same budget constraint gives rise to different choices depending on what form the increment to the budget takes. There is now a large literature that tries in a variety of ways (some quite ingenious) to explain the flypaper effect." James Hines and Richard Thaler (1995) have suggested recently that this is just one of a more general class of cases where having money on hand (e.g., from grants) has a much different effect on spending behavior than where the money must be raised (e.g., by taxation).

Much of the early empirical work on the expenditure response to intergovernmental grants studied the period from the 1950’s through the 19707s, when these grants exhibited a continuing path of expansion. As a result, much of the interest focused on the budgetary response to increases in grants. However, in more recent times, efforts at fiscal retrenchment and devolution have led to large cuts in a wide range of federal grant programs. And this has raised the interesting and important

question of whether the response to cuts in grants is similar in sign and magnitude to the response to increases in these grants. Gramlich (1987), for example, observed that during this period of retrenchment, state and local governments responded to the cutbacks in grants by picking up much of the slack: they increased their own taxes and replaced in large part the lost grant funds so as to maintain levels of existing programs. If Gramlich is right, then we should observe a basic asymmetry in response: the spending of recipients should be more responsive to increases in grant monies than to decreases in these revenues. This issue is of some importance if we are to understand the budgetary implications of the ongoing process of fiscal decentralization. In the first study of this issue, William Stine (1994), examining the response of county governments in Pennsylvania, found just the opposite of Gramlich's prediction: his estimates imply that these county governments not only failed to replace lost grant revenues, but that they reduced their spending from own-revenues on these programs as well, giving rise to a "super-flypaper effect." Our findings are thus consistent with the proposition that the flypaper effect operates symmetrically in both directions. But much clearly remains to be done on this issue.

## 17.2 THE FLYPAPER EFFECT

**Meaning:**

The flypaper effect is a concept from the field of public finance that suggests that a government grant to a recipient municipality increases the level of local public spending more than an increase in local income of an equivalent size. When a dollar of exogenous grants to a community leads to significantly greater public spending than an equivalent dollar of citizen income: money sticks where it hits, like a fly to flypaper. Grants to the government will stay in the hands of the government and income to individuals will stay with these individuals.

**Background:**

The concept was first described in a metaphorical way by Arthur Okun in response to the research of his colleague Edward Gramlich, which was published in 1979 as *The Stimulative Effect of Government Grants*. Gramlich, together with Courant and Rubinfeld, sought an explanation for the phenomenon that nonmatching grants stimulate much more local spending per dollar of grant than does income going to private citizens within the community. The flypaper effect in this paper is defined as: "bureaucrats and politicians find it easier to avoid cutting taxes when the government receives revenue-sharing monies than they

do to raise taxes when some exogenous event raises the income of the community."

In this case, the finding was that a grant from federal government to local government would raise spending of that local government by a greater amount than an equivalent increase in local income. Local public resources come from both fiscal transfers from the central government in the form of grants and from the income of individuals. Henderson and Gramlich specified the demand equations of individuals by maximizing their utility subject to that individual's income constraint, which is specified as the sum of personal income and the individual's share of his government's unconstrained fiscal transfers. This specification would mean that the individual income and the individual's share of the fiscal transfers would have an identical impact on spending. The Flypaper effect however suggests that this is not the case. This can be described as an anomaly since it is difficult to rationalize: one would expect that a government grant and an equivalent increase in local income to have the same effect.

**The flypaper model:**

The benchmark for both the policy and political economy literatures is how a politically decisive citizen would like to see government resources allocated, specified by the maximization of that representative citizen's welfare over private (x) and public (g) goods, indexed by U(x, g), subject to a current period budget constraint specified as:

$$Y = \{I + h \bullet z\} = x + pg \bullet g$$

where I is the citizen's private income (or tax base), h is the citizen's share of *unconstrained or lumpsum* intergovernmental transfers per capita (z) specified as h = I/I with I equal to the average income (or tax base) in the citizen's political jurisdiction, and pg is the "tax price" for government services (g) equal to c* (1 - m)* h where c is the per unit production cost of g and m is the matching rate for open-ended matching federal aid. Private goods cost $1. Y is called the citizen's "full income." The citizen's preferred allocations will be x = x(1, pg, Y) and g = g(1, pg, Y), where:

$\Delta g_I = (\delta g/\delta Y)*(\delta Y/\delta I)* \Delta I = (\delta g/\delta Y)* (\Delta I = \$1)$,
for an extra dollar of personal income and:

$\Delta gz = (\delta g/\delta Y)* (\delta Y/\delta z)* \Delta z = (\delta g/\delta Y)* h* (\Delta z = \$1)$
for an extra dollar of aid, implying that estimated marginal effects of aid to income should be related as $\Delta gz/\Delta gI = h$.

In most political jurisdictions the representative citizen has a tax base (often specified as the median tax base) less than the average tax base; thus, $h = I/I < 1$ in most cases. If our representative citizen has had her way, then we should expect $\Delta gz/\Delta gI = h < 1$.

The overwhelming empirical evidence summarized by Gramlich (1977), Inman (1979), Fisher (1982), and Hines and Thaler (1995) shows just the opposite, however; $\Delta gI$ ranges from \$.02 to \$.05 while the companion estimates of $\Delta gz$ typically fall between \$.30 to \$1.00. Income to the citizen stays with the citizen; grants to the government stay with the government. Money sticks where it hits.

**Explanations:**

1. A common explanation of the flypaper effect has a focus on the role of public officials. It was first argued by Niskanen that public officials tend to maximize their budgets by the budget-maximizing model. As public officials are budget maximizers, the bureaucrats have no incentive to inform citizens about the true level of grant funding that a community receives. If public officials conceal this information, they may trick citizens into voting for a higher level of funding than would have otherwise been the case. In this respect, the flypaper effect occurs because citizens are unaware of the true budget constraint.

2. The flypaper effect results when a dollar of exogenous grants-in-aid leads to significantly greater public spending than an equivalent dollar of citizen income: Money sticks where it hits. Viewing governments as agents for a representative citizen voter, this empirical result is an anomaly.

3. Four alternative explanations have been offered. First, it's a data problem; matching grants have been mis-classified as exogenous aid. Second, it's an econometric problem; exogenous aid is correlated with omitted variables leading to a downward bias in estimates of income's effects and an upward bias in estimates of aid's effects. Third, it's a specification problem: the representative citizen either fails to observe lump-sum aid, or sees aid but misperceives its impact as an average price effect, or finally, sees and understands aid's budgetary effects but allocates "public" and "private" monies through separate "mental accounts."

**Empirical evidence:**

The empirical evidence suggests none of these explanations is sufficient. A fourth explanation seems most promising: It's politics. Rather than an anomaly, the flypaper effect is best seen as

an outcome of political institutions and the associated incentives of elected officials. The empirical analyses of Henderson and Gramlich revealed something unexpected, however. An extra dollar of personal income increased government spending on the order of $.02 to $.05 but an equivalent extra dollar of grants-in-aid increased government spending by $.30 to often as much as a full dollar. When Gramlich first presented his results, his colleague Arthur Okun called this larger effect of lump-sum aid on government spending a “flypaper effect” noting that “money seems to stick where it hits.” The label stuck too, as has the puzzle of why intergovernmental transfers are so stimulative. Over 3,500 research papers have now been written documenting and seeking to explain *the flypaper effect*.

**Implications/importance:**

First, as a matter of policy, understanding *how* recipient governments spend intergovernmental transfers is essential for the design of efficient fiscal policy in federal economies. Second, as a matter of science, understanding *why* governments spend citizens’ incomes as they do provides valuable insights as to how citizen preferences are represented in government policies. The taxation of citizen incomes and the allocation of grants-in-aid provide two “tracers” as to the inner workings of political decision making, one (taxes) that is directly observed and controlled by citizens and the other (grants) perhaps only imperfectly so.

## 17.3 FEDERAL SYSTEM IN INDIA

**Introduction:**

India is a Federal state. Power is divided between the Union and the states. The functions have been classified as exclusively for the centre, exclusively for the state and both for the centre and the states.

The constitution provided a way to distribute the revenues from captain taxes collected by the centre among the states. To ensure a fair and judicious balance between the respective shares of the centre and states, Article 280 of our constitution empowers the President of India to set up a Finance Commission every five years.

**Taxation powers:**

In India, taxes are divided into six groups:

**(I) Taxes Levied, Collected and Retained by the Centre:**

These taxes are belonging to the centre exclusively. In other words, no part of the proceeds of these taxes can be assigned to the states. The following taxes fall under this category:

(i) Corporation Tax (Corporate tax)
(ii) Customs Duties.
(iii) Surcharge on Income Tax.
(iv) Taxes on capital value of assets of individual and companies.
(v) Fees on matters of the Union list.

**(II) Taxes Levied by the Centre but Collected and Appropriated by the States:**

The following taxes are included in this category:

(i) Stamp duties on bills of exchange, cheques, promissory notes and others.
(ii) Excise duties on medicinal and toilet preparation containing alcohol.

There taxes which form part of the union list are levied by the centre but (a) collected by the states within which such duties are levied; and (b) collected by the centre when such duties are levied within any Union Territory.

**(III) Taxes Levied and Collected by the Centre but Assigned to the States:**

This category includes the following duties and taxes:

(i) Duties on succession to property (other than agricultural land).

(ii) Estate duty on property (other than agricultural land).

(iii) Terminal taxes on goods and passengers carried by railways, sea and airways.

(iv) Taxes on railway fares a freights.

(v) Taxes on transaction in stock exchanges and future markets (other than stamp duties).

(vi) Taxes on the sale or purchase of newspapers and taxes on advertisements published in them.

(vii) Taxes on the sale or purchase of goods in course of inter-state trade or commerce (other than newspaper).

(viii) Taxes on the consignment of goods in the course of inter-state trade or commerce.

The net proceeds of these duties and taxes are assigned to states in accordance with the principles laid down by the parliament.

**(IV) Taxes Levied and Collected by the Centre and Compulsorily Distributed between the Centre and the States:**

Taxes on income (other than agricultural income and corporation tax) shall be levied and collected by the centre but compulsorily distributed between the centre and the states in such manner as prescribed by the president on the recommendations of the Finance Commission. The obligatory sharing of income tax is provided by Article 270 of the Constitution.

**(V) Taxes Levied and Collected by the Centre and may be distributed between the Centre and the States:**

Under this category falls the excise duties included in the Union list except those on medicinal and toilet preparations. These are levied and collected by the centre. The net proceeds of such duties can be paid to states out of the consolidated Fund of India only if the parliament so provides.
Further, the principles of distribution shall also be laid down by the parliament. It is to be noted that sharing of the proceeds, of income tax is obligatory, while that of excise duties is permissible.

**(VI) Taxes Levied and Collected and Retained by States:**

The following taxes and duties exclusively belong to states. They are mentioned in the State list. Every state is entitled to levy, collect and appropriate these taxes. The taxes are

(i) Duty on succession to agricultural land.
(ii) Estate duty on agricultural land.
(iii) Land revenue.
(iv) Tax on agricultural income.
(v) Tax on land and buildings
(vi) Capitation taxes.
(vii) Tax on mineral rights.
(viii) Tax on the consumption or sale of electricity.
(ix) Tax on vehicles.
(x) Tax on the sales and purchase of goods (other than newspaper) for e.g. Sales tax.
(xi) Tolls
(xii) Tax on professions, trades and employment.

**Functions/Responsibilities:**

**1. Centre:**

The functions of the central government can be classified as those required to maintain macroeconomic stability, international trade and relations and those having implications for more than one state, for reasons of economies of scale and cost efficient provision of public services. Issuing currency and coinage, dealing in foreign exchange, foreign loans, the operation of the 10 central bank of the country (Reserve Bank of India or RBI), international trade, banking, insurance and operation of stock exchanges are some of the major functions assigned to the central government to maintain macroeconomic stability. Functions like the operation of railways, posts and telegraphs, national highways, shipping and navigation on inland waterways, air transport, atomic energy, space, regulation and development of oilfields and major minerals, inter-State trade and commerce and regulation and development of inter-state rivers are the major functions assigned to the center for reasons scale economies and spillovers in respect of services with benefits spanning more than one State.

**2.** States: The major subjects assigned to the states comprise public order, police, public health, agriculture, irrigation, land rights, fisheries and industries and minor minerals. As mentioned earlier, the states do have jurisdiction over concurrent items and can take initiatives with regard to these subjects.

However, in the event of conflict between the center and the states, the former has overriding powers. Subjects like public health, agriculture and irrigation involve considerable governmental intervention and expenditures. Even in regard to the subjects in the concurrent list like education and transport , social security and social insurance, in a democratic polity, being proximate to the people, the states would be compelled to assume a significant role.

**Rationale:**

The assignment of tax powers, however, is based on the principle of “separation”, and the tax handles are exclusively assigned either to the center or to the states. Most of the broad-based and productive tax handles have been assigned to the center perhaps, for reasons of stabilization and redistribution stated earlier. These include taxes on income and wealth from nonagricultural sources, corporation tax, taxes on production (excluding those on alcoholic liquors, opium, hemp and other narcotics) and customs duty. A number of tax handles are assigned to the states as well. These include taxes on agricultural income and wealth, taxes on the transfer of property (stamp duties and registration fees), taxes on motor vehicles, taxes on the transportation of goods and passengers, sales tax on goods,

excises on alcoholic beverages, entertainment tax, taxes on professions, trades, callings and employment, property tax and taxes on the entry of goods into a local area for consumption, use or sale (octroi). However, from the viewpoint of revenue productivity, only the tax on the sale and purchase of goods is important. The center has also been assigned all residual powers which implies that the taxes not mentioned in any of the lists automatically fall into its domain. The Constitution recognizes that assignment of tax powers and expenditure functions would create imbalances between expenditure needs and abilities to raise revenue. Therefore, the constitution provides for the assignment of revenues (as contrasted to assignment of tax powers), sharing of the proceeds of certain centrally levied taxes with the states, and making grants to the states from the Consolidated Fund of India.

**Role of finance Commission:**

The shares of the center and the states and their allocation among different states of both the taxes are to be determined by the Finance Commission appointed by the President of India every five years or earlier as needed. As this arrangement is alleged to have created serious disincentives to the central government in raising revenues from the two taxes, and based on the recommendation of the Tenth Finance Commission, the central government has proposed replacing of the sharing of the two taxes with sharing of revenues from all central taxes.

In addition to tax devolution, the Finance Commission is also required to recommend grants to the states in need of assistance under article 275. The quantum of shared taxes, their distribution among the states and the amount of grants to be given to the states is determined by an independent quasi-judicial body, the Finance Commission, appointed by the President of India every five years (or earlier).

**Trends:**

The states on an average raise about 35 per cent of total revenues, but incur about 57 per cent of total expenditures. The revenues derived from exclusive central taxes constitute about 24 per cent; those from exclusive state taxes 31 per cent, shareable sources contribute about 27 per cent and the remaining 17 per cent consists of non-tax revenues. The major taxes levied exclusively by the center consist of customs duty (15 per cent of total tax revenue) and corporation tax (8 per cent). Among the state taxes, the revenue from sales tax constitutes almost 16 per cent. Other state taxes individually contribute less than 6 per cent of total tax revenue. It is seen that the expenditure shares of central and state governments indicate a fairly high degree of decentralization.

However, states' control over expenditure policies is much less. About 15 per cent of states' expenditures was on central sector and centrally sponsored schemes which are specific purpose transfer schemes administered by various central ministries. States' expenditures on these schemes has actually shown an increasing trend.

The pattern of expenditures shows that the central government plays a major role in providing defense, meeting interest payments and industrial promotion. The states on the other hand, have a predominant share of total expenditures on internal security, law and order, social services like education, health, family welfare, housing and social security and on economic services like agriculture, animal husbandry, forestry, fisheries, irrigation and power and public works. The states' share in expenditure on administrative services is about two-thirds; on social services they spend over 85 per cent and on economic services their expenditure share is almost 60 per cent. It is also seen that over the last decade, while the share of the states in raising revenues has remained constant, their expenditure share has shown a steady increase particularly since 1991 by about five percentage points. This has occurred because fiscal reforms initiated in 1991 have led to a deceleration in the growth of central government expenditures, but not so much in central transfers to states. Consequently, state expenditures have continued to increase even as central government expenditures decelerated. Thus, the states' expenditure share has increased both in current and capital expenditure.

The functions of the finance commission can be summarized as follows:

1. Distribution of net proceeds of taxes between Centre and the States, to be divided as per their respective contributions to the taxes.
2. Determine factors governing Grants-in Aid to the states and the magnitude of the same.
3. To make recommendations to president as to the measures needed to augment the Consolidated Fund of a State to supplement the resources of the panchayats and municipalities in the state on the basis of the recommendations made by the Finance Commission of the state.
4. any other matter related to it by the president in the interest of sound finance

## Local governments:

Local governments got a fillip after the 73rd [1] and 74th [2]Constitution Amendment Acts. Later in 1992, the 73rd and 74th constitutional amendments were passed by the Parliament.

- The **73rd Amendment** is about **Rural Local Governments** (which are **also known as Panchayati Raj Institutions or PRIs**) and
- The **74th amendment** made the provisions relating to **Urban Local Governments** (**Nagarpalikas**).

All municipal acts in India provide for functions, powers and responsibilities to be carried out by the municipal government. These are divided into two categories - obligatory or discretionary.

### Obligatory functions

- supply of pure and wholesome water
- construction and maintenance of public streets
- lighting and watering of public streets
- cleaning of public streets, places and sewers
- regulation of offensive, dangerous or obnoxious trades and callings or practices
- maintenance or support of public hospitals
- establishment and maintenance of primary schools
- registration of births and deaths
- removing obstructions and projections in public streets, bridges and other places
- naming streets and numbering houses

### Discretionary functions

- laying out of areas
- securing or removal of dangerous buildings or places
- construction and maintenance of public parks, gardens, libraries, museums, rest houses, leper homes, orphanages and rescue homes for women
- public buildings
- planting and maintenance of roadside and other trees
- housing for low income groups
- conducting surveys
- organizing public receptions, public exhibitions, public entertainment
- provision of transport facilities with the municipality
- promotion of welfare of municipal employees

**Conclusion:**

Some of the functions of the urban bodies overlap with the work of state agencies. The functions of the municipality, including those listed in the Twelfth Schedule are left to the discretion of the state government. Local bodies have to be bestowed with adequate powers, authority and responsibility to perform the functions entrusted to them by the Act.

## 17.4 QUESTIONS

1. Write a note on intergovernmental grants and revenue sharing.
2. Evaluate the concept of Flypaper effect.
3. Describe the federal system/ structure in India.

## 17.5 REFERENCES


- Wallace E. Oates (1997), 'On the welfare gains from fiscal decentralization', Economia, University of Maryland, Washington DC.
- Richard M. Bird April 1999, 'Rethinking Tax Assignment: The Need for Better Subnational Taxes'.
- Charles E. McLure, Jr., The Tax Assignment Problem: Ruminations on How Theory and Practice Depend on History National Tax Journal Vol. 54 no. 2 (June 2001) pp. 339-364.
- Wallace E. Oates (Sept. 1999), 'An Essay on Fiscal Federalism', *Journal of Economic Literature*, Vol. 37, No. 3, pp. 1120-1149.
- Smriti Chand, 'Taxes in India',

  **http://www.yourarticlelibrary.com/tax/in-india-taxes-are-divided into-the-following-six-groups/23401/**
- M. Govinda Rao and Nirvikar Singh (Dec. 1998), 'The assignment of taxes and expenditures in India'.
- https://en.wikipedia.org/wiki/Local_self-government_in_India