Department of Distance Education

Punjabi University, Patiala

Class : M.A. I (Economics)   **Semester : 2**
Paper : IV (Economics of Growth and Development)
Unit : II   **Medium : English**

Lesson No.

Department website : www.pbidde.org

Human Capital - I: Education and R&D
(Theory, Investment, Returns and Policy Issues)

Structure
1. Objectives
2. Introduction
3. Human Capital in Neo-Classical Framework
4. Human capital in New Growth Theories
   4.1 Growth of Physical Capital

   4.2 Growth of Human Capital
        4.2.1 Internal Effects of Human Capital
        4.2.2 External Effects of Human Capital
5. Private and Social Returns to Education and Policy Implications
6. Research and Development
7. Summary
8. Glossary

9. Suggested books and References

10. Suggested questions

1. Objectives: This lesson aims to –
(i) examine the role of human capital in the neo-classical as well as new growth theories
(ii) analyse the internal as well as external effects of human capital on growth of the economy
(iii) know about the private and social costs and benefits of education
(iv) understand the policy stances in case of primary, secondary and higher education
(v) examine the role of research and development in growth of economy
(vi) understand the policy initiatives for development of research and development in any economy.

2. Introduction: In the Neo-classical growth theories, in the long-run the economy reaches in the steady state that is independent of the initial conditions. These theories believe that all growth is due to advances in technology but technological progress is taken as an exogenous factor. But the experiences of several counties show that with same saving ratios, population growth and same level of technical progress, different growth rates have been registered. Actually, when we say that capital accumulation is an important determinant of growth, in the neo-classical sense, we merely talk about physical capital. The new growth theories emphasize that one of the most important determinant of differences in growth of income and output is the difference in acquisition of skills, say the

human capital. This human capital includes both schooling and on-the-job training. Human capital refers to all the attributes of workers that potentially increase their productivity in all or some productive tasks. Human capital theory was developed primarily by Becker (1964) and Mincer (1974) but now we can find a number of studies which establish that the accumulation of human capital is as important as that of the physical capital. According to new growth theory, long-run economic growth is affected by deliberate economic behaviour and human actions such as innovation and education. However, research on the benefits of education consistently produces substantial returns to education. Similar examples can be given for expenditures on R&D. However, while the benefits from schooling and R&D can be large, it typically takes a long time before these gains arrive, and also the costs associated with the investments need to be taken into account. For this purpose, we need to introduce a link between inputs such as knowledge, R&D and human capital and the output variable of interest (e.g. economic growth). We also need to evaluate the impact of policy on the decisions to invest in education and R&D. In this perspective, first of all, we'll discuss the theories of human capital and growth.

3. Human Capital in Neo-Classical Framework: Shultz in 1961 and Becker in 1964 put forth the idea that education enhances a person's skill level and thereby his/her human capital. Attaining higher level of education increases the productivity which shifts the production possibility curve in upward direction showing higher capacity to produce with same amount of resources. But as we pointed out earlier, the knowledge part is supposed to be given i.e. determined exogenously and with a given and constant level of knowledge, the diminishing returns are applicable leading the economy to steady state in the long run. Lucas (1990) very rightly pointed out that differences in returns to physical capital become much smaller when account is taken of differences in human capital across countries: physical capital tends to move towards countries with more human capital. In the 1990s the standard neoclassical growth model has been revised by introducing human capital. In Neo-classical framework, assuming Cobb-Douglas production function, production in a single sector economy can be represented by the following equation:

$$Y = f(K, H, EL)$$

Here, K is the stock of physical capital, L – the units of labour, H – human capital stock of the work force, and E is the efficiency parameter. With constant returns to scale, the production function assumes the form:

$$Y = K^{\propto} H^{\beta} (EL)^{1-\alpha-\beta}$$

Here, $\propto$ is the production elasticity of physical capital, β is the production elasticity of human capital ($\propto + \beta < 1$) and $1 - \alpha - \beta$ is the production elasticity of labour. The inclusion of human capital in the production function has several implications. It indicates that an increase in the average years of schooling increases the human capital stock 'H' as well as the efficiency level of the workers. But in a neo-classical framework under constant returns to scale, we can say that an increase in human capital stock

increases the total production as well as marginal productivity of the factors. Since, the payment to the factors are function of their marginal productivities, an increase in human capital also increases the returns to the factors.

$$r = \frac{\partial Y}{\partial K} = \propto K^{\alpha-1} H^{\beta} (EL)^{1-\alpha-\beta}$$

and,

$$w_{EL} = \frac{\partial Y}{\partial EL} = (1 - \alpha - \beta) K^{\alpha} H^{\beta} (EL)^{-\alpha-\beta}$$

where r is the return to physical capital, and w_{EL} is the wage rate. So the increase in human capital stock not only increases the wage rate, but also the interest rate and therefore, to adjustments in the physical capital stock (which will decline with an increase in rate of interest). In this way, the changes in the stock of human capital only have a temporary impact on the growth rate; the balanced growth rate is again determined by the exogenous rate of labour-augmenting technical progress and population growth.

4. Human capital in New Growth Theories: In New Growth theories, the human capital is taken as an endogenous factor. In contrast to the neo-classical theory, this theory assumes that the economy can experience increasing returns to scale even in the long run which is possible through technological spillovers and accumulation of human capital. In this context Lucas (1988) and Barro and Sala-i-Martin (1995) proposed the following production function:

$$Y = K^{\alpha} (uH)^{1-\alpha}$$

Where, 'H' is human capital and 'u' is time devoted to production. Lucas emphasised that the time spent on accumulation of human capital through years of schooling and learning by doing. Assuming that existing level of human capital has diminishing returns and under a dynamic economy, it also depreciates at a constant rate, and taking 1-u as the time devoted to accumulation of human capital, the growth of human capital (γ_H) can be represented by –

$$\gamma_H = B(1 - u) - \delta$$

Here, 'B' the transformation rate of learning in to human capital and δ is the depreciation rate of human capital. The core of the model responsible for endogenous growth is that there are constant returns to scale with regard to reproducible factor inputs physical capital and human capital. As economic growth is determined from structural parameters, growth is called endogenous. Notice that human capital is a purely private commodity in this analysis. In a simple framework, increases in human capital translate into higher wages, and there are no external effects. So, there is no government interference. In a growing economy, the savings increase as the income increases. Part of the savings is spent in accumulation of physical capital and its part is spent on accumulation of human capital. These two components of savings can be expressed in the following manner. Firstly, taking 's' as the part of savings being used for accumulation of physical capital i.e.

$$k(t+1) - k(t) = sy(t)$$

Secondly, the proportion of savings spent on accumulation of human capital can be expressed as

$$h(t+1) - h(t) = qy(t)$$

Thus, *sy(t)* and *qy(t),* respectively show the total amount of resources spent on accumulation of physical and human capital. For self-sustained growth of the economy, 'y', 'k' and 'h' should be growing at the same rate. The rate of economic growth actually depends upon the growth of investment in physical as well as human capital. Therefore, it is important to find the ratio of investment in human capital to that of physical capital. For this, we would have to find the growth of these two types of capital in an economy.

4.1 Growth of Physical Capital: The growth of physical capital can be derived from equation (2) by putting the value of *y(t)* and also dividing both sides by *k(t)*. The resultant equation can be written as

$$\frac{k(t+1) - k(t)}{k(t)} = \frac{s.[k(t)]^{\alpha}[h(t)]^{1-\alpha}}{k(t)} = s.\left[\frac{h(t)}{k(t)}\right]^{1-\alpha}$$

If we take h(t)/k(t)=r, this equation can be written as

$$\frac{k(t+1) - k(t)}{k(t)} = sr^{1-\alpha}$$

Similarly, we can also derive the equation for growth of human capital

4.2 Growth of Human Capital:

$$\frac{h(t+1) - h(t)}{h(t)} = qr^{-\alpha}$$

Since, in the long period the growth of human capital as well as the physical capital are equal, therefore,

$$sr^{1-\alpha} = qr^{-\alpha} \quad \text{or} \quad r = \frac{q}{s}$$

This 'r' can be used as long term growth rate and since in the long period, the growth of income, physical capital and human capital are the same, therefore,

$$\frac{y(t+1) - y(t)}{y(t)} = sr^{1-\alpha} = qr^{-\alpha}$$

$$= s^{\alpha} q^{1-\alpha}$$

Thus, the long term growth of the economy depends upon the rate of physical capital formation as well as human capital formation. It is the human capital which compensates the fall in growth of output due to applicability of diminishing returns on physical capital. The human capital investment, rather ensures increasing returns by its internal as well as external positive and output stimulating effects. The internal and external effects of human capital formation in any economy can be discussed as below.

4.2.1 Internal Effects of Human Capital: According to Lucas, the total time of a human being, particularly a worker, can be divided in to two components – the time spent in production and the time spent in accumulation of human capital. If we denote the proportion of time spent in production as *μ(h)*, then the time spent in accumulation of human capital will be 1- *μ(h)*. In any economy, the size of the labour force as well as its productivity per hour significantly influences the level of output. Therefore, instead of having merely the size of the work force, Lucas has put forth the idea of 'effective labour force' which is shown as the product of size of the labour force and the time spent on producing goods and services for a given level of human capital.

$$N_e = \int_0^\infty \mu(h).N(h).dh$$

Here, *N(h)* is the size of labour force and N_e is the effective labour force. Thus, we can express production as function of physical capital stock and effective labour force.

$$Y = f(K, N_e)$$

The level of human capital would not only have a macro economic impact upon the aggregate output of the economy but will also have accrue certain private benefits to the holders of the human capital as in a competitive market economy, wages are paid according to the marginal product of workers. Since, the workers with higher level of human capital are more productive, they would have higher earnings. Total wages in an economy for a given level of human capital can be calculated as follows:

$$\text{Total wages} = f'(K, N_e).h.\mu(h)$$

Where, $f'(K, N_e)$ shows the marginal productivity of labour.

4.2.2 External Effects of Human Capital: Increase in level of human capital formation in any economy, undoubtedly increase the level of productivity of a single worker but also have an overall effect upon the average productivity of the economy as a whole. Even a single worker with higher human capital in a production unit has huge ripple effects in the production unit. Same is true for the economy as a whole. But in a perfectly competitive economy, it is generally assumed that human capital of an individual would not affect the average level of human capital, yet its opposite is not true as average level of human capital in any economy determines the minimum target to be achieved by average workers to ensure their employability. Hence human capital investments and its attainments by the private individuals are largely determined by the average level of human capital for the country as a whole. This is termed as external effect of human capital. In order to know about the external effect of human capital, it is important to know about the average level of human capital which can be calculated as below.

$$h_a = \frac{\int_0^\infty h.N(h).dh}{\int_0^\infty N(h).dh}$$

Here, h_a is average level of human capital in a country. Now, we can easily adjust our production function by incorporating internal as well as external effect of human capital. First of all let us have production as function of capital and effective labour force.

$$Q = A.K^\beta(t)N_e^{1-\beta}(t)$$

Putting the value of N_e, we get

$$Q = A.K^\beta(t)[\mu(t).h(t).N(t)]^{1-\beta}$$

Incorporating the external effects of human capital or the average level of human capital for the society as a whole

$$Q = A.K^\beta(t)[\mu(t).h(t).N(t)]^{1-\beta}[h_a(t)]^\tau$$

Any increase in time for accumulation of human capital i.e. 1- $\mu(t)$ will raise the individual as well as the average level of human capital of the society which will have positive effects upon the level of output at an increasing rate. Thus, the economy would grow at a faster rate due to applicability of increasing returns to scale in the production sector. The change in human capital, which is the main force behind the faster growth of the economy, can be measured as

$$\hat{h}(t) = [h_a(t)]^\tau.G[1-\mu(t)],$$

here, G is the growth of human capital and it is always positive i.e. G > 0 but the existing level of human capital or say the knowledge, which the society has attained so far, will have diminishing returns to output, therefore, $\tau < 0$. In order to simplify the analysis if we simply assume $\tau = 1$, then the equation showing the change in human capital at any point of time can be shown as

$$\hat{h}(t) = [h_a(t)].G[1-\mu(t)]$$

We can discuss here two extreme cases, one is when whole of the time is spent in accumulation of human capital i.e. $\mu(t) = 0$, and the second when whole of the time is spent in production only i.e. 1- $\mu(t) = 0$. In the first case, the change in human capital will be

$$\frac{\hat{h}(t)}{[h(t)]^\tau} = G$$

i.e. the economy can achieve the highest growth of human capital equal to G while in second case there will be no change in growth of human capital and it would be zero. In

absence of any change in human capital, since diminishing returns to scale are applicable to the existing level of human capital, the economy will also grow at diminishing rate for any change in its inputs. But in real life, the value of $\mu(t)$ or that of 1- $\mu(t)$ varies between 0 and 1 which shows that the economy moves on a continuous growth path and the rate of growth of the economy will be higher for higher growth of human capital. This fact points towards the fact that the applicability of diminishing returns can be postponed by increasing investment in human capital. Thus, the economies with higher rate of growth of human capital experience a higher growth of income and the economies with lower investment in human capital will experience the lower growth of income, leading to divergence between the two types of the countries. Through this fact, Lucas pointed out that the the gap between growth of rich and poor economies can be explained by the gap in investments in human capital in these economies.

Finally, Lucas also differentiated between the optimum growth path and the equilibrium growth path. By optimal path he means that the society wants to maximise its utility function by achieving the optimal level of per capita income with an optimal combination of $K(t)$, $N(t)$, $c(t)$, $h(t)$ and $h_a(t)$. On the other hand, equilibrium path means that there is simultaneous equilibrium in all firms and households as well the economy as a whole. Assuming that $h_a(t)$ is exogenously determined and it is expected that each individual will follow the same path so that the actual behaviour coincides with the expected behaviour and there is no gap between demand and supply (i.e. AD = AS), for given physical and human capital stock. The solution will be achieved in both the cases if $h(t) = h_a(t)$. Any divergence between the two will mean the divergence from the equilibrium as well as the optimal growth path.

However, Lucas' Model also gave little attention to the structural rigidities of the developing economies. There is no doubt that these economies have low level of productivity due to low level of human capital but at the same time these economies are also suffering from the problem of misallocation and misutilisation of the available resources. These economies not only face the problem of skilled workers but also underutilisation of existing human capital. Due to lack of opportunities and low returns to human capital, there is little incentive to invest in human capital by private individuals. On the other hand, the skilled workers have a greater tendency to migrate to other countries in search of better opportunities. These are the workers which the economy needs the most and their emigrations means the drain of most essential and productive resources. This loss of intellectual capital has a huge and long run adverse effect upon the economic growth of the poor countries. These aspects are ignored by Lucas' model of endogenous growth, yet there is no doubt that the developing economies can learn a lot from this model that it is the higher level of human capital that can ensure higher productivity of other resources. So, it must be attained as well as retained.

5. Private and Social Returns to Education and Policy Implications: There is no doubt that with increase in education, the income of the individuals increases during his/her life time but the investment in education depends upon the comparative analysis of direct and

indirect costs and benefits. When people spend time in acquiring education, they not only bear the direct cost in terms of fees, cost of books and stationery etc. but they also postpone their participation in the labour market and thus forgo the income which they would have been earning had they not being involved in acquiring formal education. This income foregone can be counted as indirect cost of attaining education. On the other hand increase in income at higher level of education is considered to be the benefit of education. A person will have a rational decision of investing in education if the cost of attaining education is lower than its benefits. This can be understood with help of figure 1.

**Figure 1: Costs and Benefits of Education**
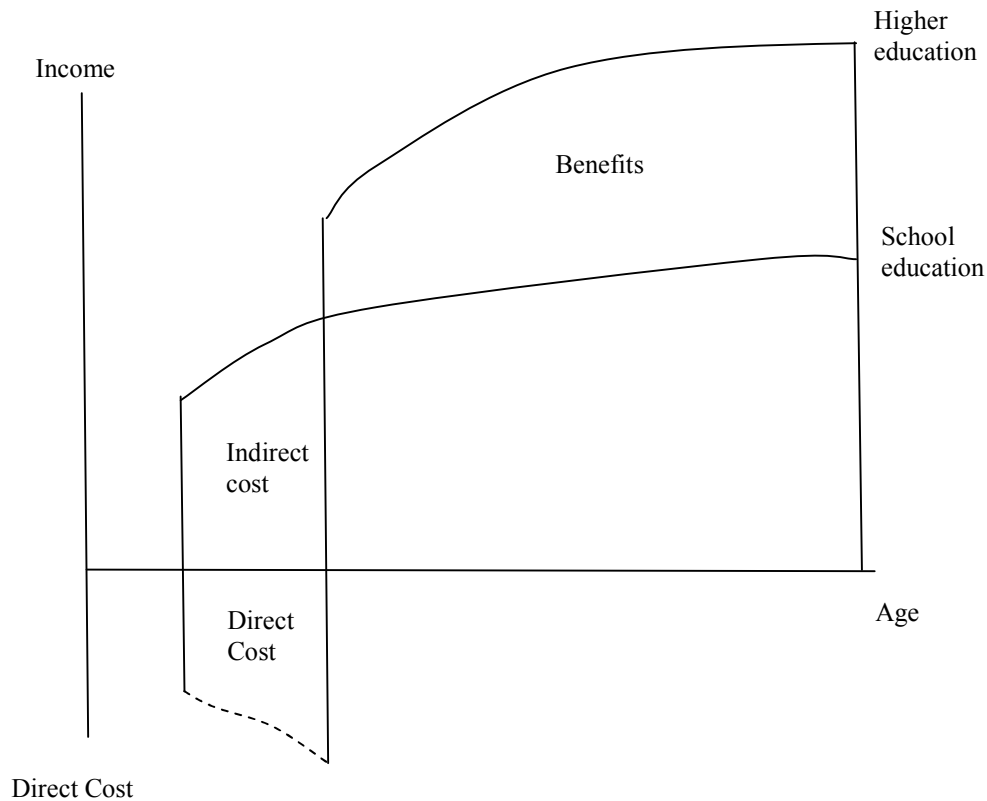
In this figure, we can see that as the person tries to attain the higher education, he/she has also entered the adulthood and so eligible for entering in the labour market. At this age, an individual has two alternative choices, one is to spend time in attaining higher education and other to enter the labour market. If the decision is in favour of attaining education, then he/she have to sacrifice the income from their labour services, hence, it is shown as the indirect cost of attaining education apart from the direct cost of education shown in the lower quadrants of Y-axis. On the other hand, the earnings for the persons with higher education increase at a much higher rate. We can see here that the total cost (total of direct and indirect cost) of attaining higher education is lower than the total

benefits of higher education over the life time of an individual. But we know that while the individuals have to bear the cost in the current time period, the benefits will accrue to them after the education is over and the person is able to get employment on basis of that education. Hence, for a true cost-benefit analysis, it is necessary that we should have a discounted value of these benefits. This can be given as under:

$$PVB_{edu} = \sum (E_t - N_t)/(1+i)^t$$

Here, PVB_{edu} is the present value of the benefits of education which is the discounted difference of earnings with higher education (E_t) and without higher education (N_t) at any point of time (t) and 'i' is the time rate of discount. The higher is the time rate of discount, lower will be the present value of the benefits of higher education. However, education is a private as well as social good. Therefore, in order of have a rational policy choice regarding public and private expenditure on education, it is better to analyse its private as well as social costs and benefits over a longer time horizon. On the cost side, the societies incur a lot on providing infrastructure, subsidies on education and research etc. They also have to divert the resources from some sectors in order to increase the funding for education, the sectors foregone in favour of education is considered as the indirect cost of education. On the other hand apart from the direct social benefits in terms of more productive workforce, the societies also enjoy some indirect benefits in terms of more law abiding, peace loving and disciplined citizens. With spread of education in general and primary education in particular, many government programmes on health, sanitation, population control etc. become easier to be implemented. The educated people are quick to adapt themselves to any change. However, it is generally propounded that the spread of primary education has more of the social benefits while the higher education has more of the private benefits. There are actually fewer sectors of the economy which benefit from the higher education, therefore, the social benefit curve becomes flatter for higher education while that of the private benefits increases sharply while in case of costs, the opposite is the case as the private cost is low due to availability of public research and education infrastructure but social costs are very high as the countries have to spend a lot for providing costly books, laboratories and other research facilities. This can be observed from figure 2.

Figure 2: Private and Social Costs and Benefits of Education
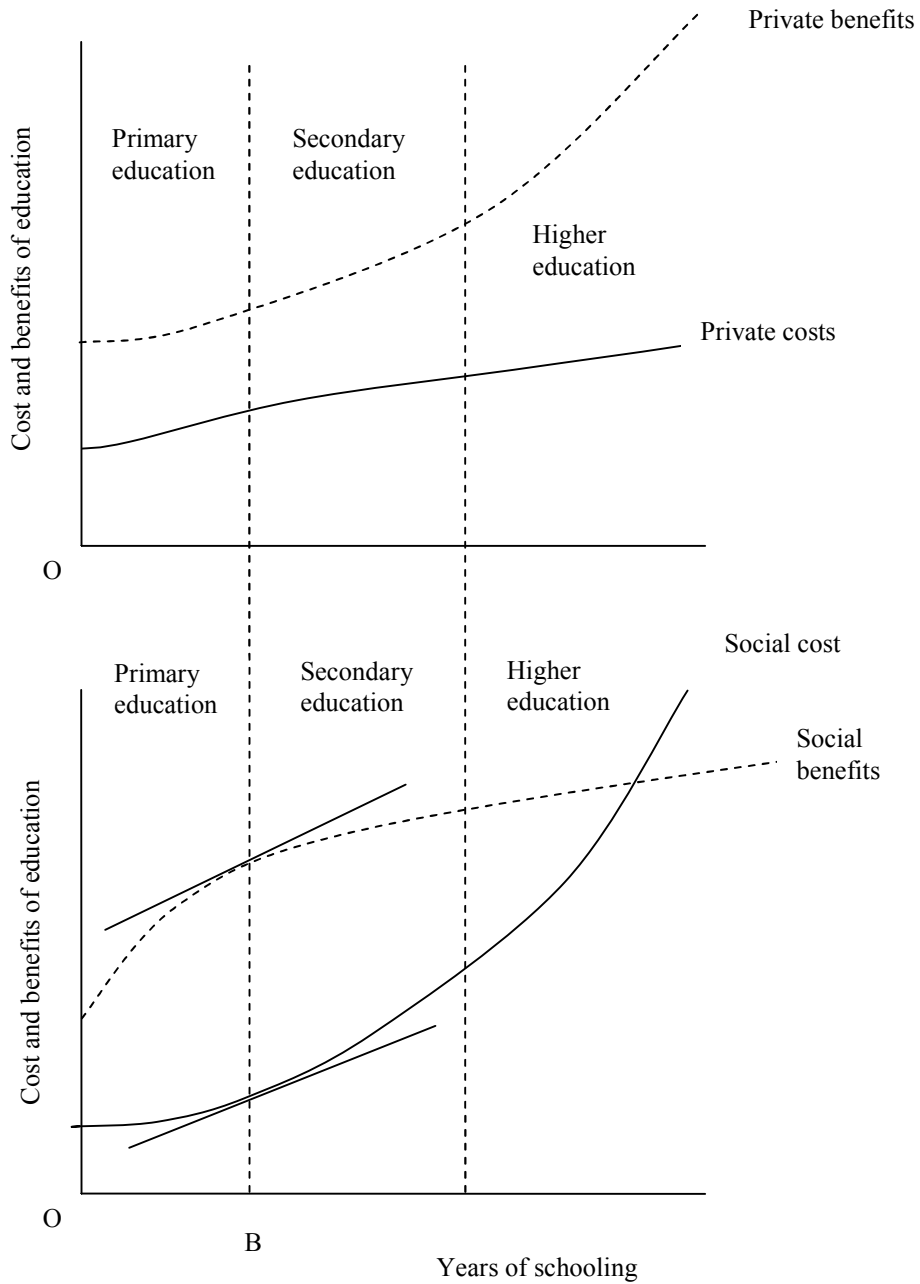

In figure 2, we can see that after primary education, the difference in private benefits and costs goes on increasing while that of the social benefits and the costs goes on diminishing.

We can also see that as the society moves from primary to secondary and higher education, the private benefits surpass the social benefits. On basis of this argument, it is sometimes suggested that the higher education should be privatised and the burden on government exchequer for providing higher education should be minimised. They say that the individuals stand to gain more than the society for increasing years of the education, therefore, the cost should also be borne by the individuals than the society. In figure 2, we can see that the net social benefits are the highest upto primary level of education i.e. up to point 'B' in the figure. We can see that the tangent lines to the social cost curves and the social benefits curve are parallel to each other at point 'B'. This indicates that the difference between the two is the maximum at this point. This suggests that the education up to this level must be provided by the government either free of cost or at least at the subsidised rate. Beyond the point 'B', the private benefits are increasing at a faster rate while the social benefits are growing at a much slower rate, therefore, it is suggested that it will be rational policy decision if beyond this point, we shift the cost of education to the private sector. It is being warned that over-subsidisation of education actually leads to inefficient allocation of scarce resources. The advocates of privatisation of education point out that the irrational pricing of education leads to inefficient allocation of resources through unnecessary spread of education. When more of the people are attaining higher education than the markets can absorb, it results in either unemployment of the educated youth or their employment in the jobs which require low qualifications. This leads to waste of the private as well as social resources spent on attaining higher education apart from under-utilisation of human skills. Many a times, the educated youth, especially those trained in professional education emigrate to other countries in search of better job prospects. This again leads to wastage of resources as the country would have spent a huge amount in training those professionals whose benefits will now be enjoyed by other countries where they have migrated. Therefore, it is suggested that the resources of education should be spent more on improving the primary education both quantitatively as well as qualitatively. Moreover, these resources can also be spent on rural development and enhancing the capabilities of rural masses.

6. Research and Development: We all know that for long run growth we need investments in physical as well as human capital. Investment in either of the two alone is not sufficient as with same type of physical capital and same level of human capital, diminishing returns soon set in and the economy may reach the stationary state. For ensuring increasing returns to scale in the long run as well, there is a need of continuous improvement of physical as well as human capital which needs a steady involvement of the society in research and innovations. The investment in research not only increases the capabilities of the existing resources but new methods, new products also come in to being which increase the level of consumption and quality of life of the masses. The expenditure on R&D leads to technical change and so the productive capacity of the economy which shifts the production possibility curve in the outward direction as shown in figure 3.

Figure 3: R&D, Technical Change and Production Possibility Frontier
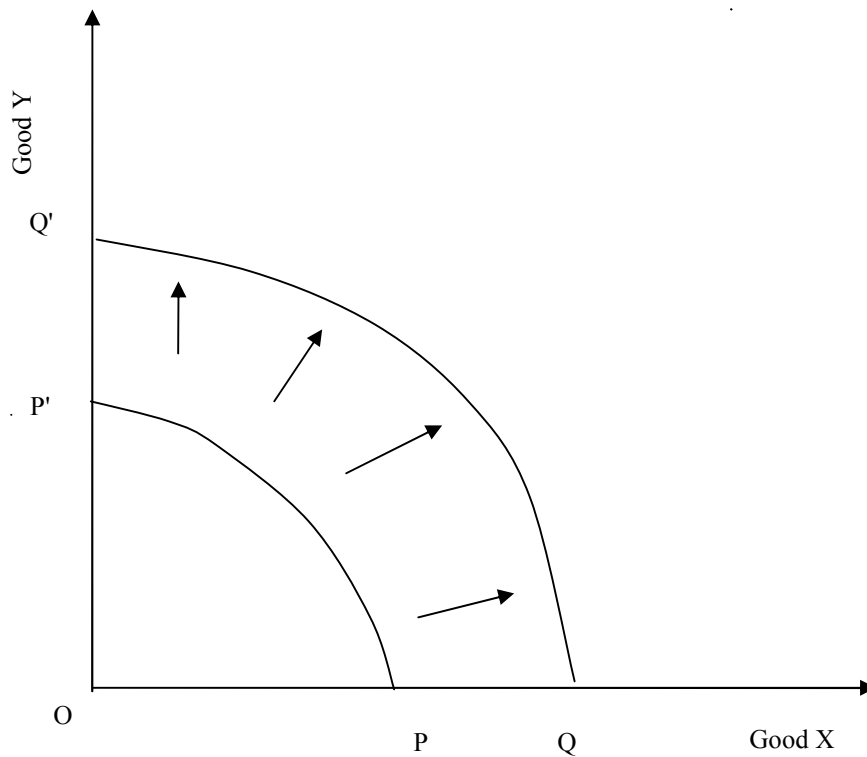
In this figure PP′ is the initial production possibility frontier showing full utilisation of all the available resources. With R&D, new techniques of production are utilised pushing the new production possibility curve at QQ′ which shows greater capacity of the economy to produce both X and Y commodity with same amount of resources.

The private investors invest in R&D if they think that the resulting net returns would be sufficient to induce them invest in these ventures. In neo-classical framework, it is generally propounded that the investment in research actually leads to improvement in variety of intermediate goods. Therefore, we can introduce the capital stock as an intermediate good with available varieties of the same. Hence, the production function assumes the following form:

$$Y = TL^{1-\alpha} N^{1-\alpha} X^{\alpha} \qquad\qquad \ldots\ldots \quad (1)$$

Here, T stands for technical factor, L – units of labour, N – number of varieties available of the intermediate goods, X – quantity of intermediate goods and $0 < \alpha < 1$. With research and development, new varieties of the intermediate goods are found and 'N' increases which has a positive effect on level of output 'Y'. From this equation, we can also derive the growth of output which will be a function of change in technical component, growth of number of varieties and quantity of intermediate goods along with the growth of labour

force.

$$\frac{\hat{Y}}{Y} = \frac{\hat{T}}{T} + (1-\alpha).\frac{\hat{N}}{N} + (1-\alpha).\frac{\hat{L}}{L} + \alpha.\frac{\hat{X}}{X} \qquad\qquad ..... \ (2)$$

For economic growth the Total Factor Productivity or TFP must be increasing. If we denote the growth of TFP as $\hat{g}$ then it can be derived by subtracting the growth of labour and capital goods from growth of income. This can be shown as under:

$$\hat{g} = \frac{\hat{Y}}{Y} - (1-\alpha).\frac{\hat{L}}{L} - \alpha\frac{\hat{X}}{X} = \frac{\hat{T}}{T} + (1-\alpha)\frac{\hat{N}}{N} \qquad\qquad ..... \ (3)$$

Thus, we can say that the growth of TFP depends upon the change in technology as well as change in number of varieties of the intermediate goods, which in turn depends upon the expenditure on R&D.

$$\hat{N} = (1/\eta)(R \& D) \qquad\qquad ...... \ (4)$$

Here, η is per unit expenditure of R&D required to create a new variety (N) of the intermediate goods. From equation 4, we can also derive, the growth of number of varieties of the intermediate goods ($\hat{N}/N$).

$$\hat{N}/N = (1/N\eta)(R \& D) = (R \& D)/(market\,value\,of\,past\,R \& D) \qquad\qquad ...... \ (5)$$

Here, $N\eta$ is the product of existing number of varieties of the intermediate goods and the per unit expenditure of R&D on production of these varieties. Thus, it denotes the past expenditure of R&D on creation of these varieties. Now by putting this value of $\hat{N}/N$ in equation (3), we can get the function of growth of TFP.

$$\hat{g} = \frac{\hat{T}}{T} + (1-\alpha)(R \& D)/(market\,value\,of\,past\,R \& D) \qquad\qquad ...... \ (6)$$

Thus, we can see that the growth of TFP is positively related with the current expenditure on R&D and inversely related with market value of past R&D. This is due to the reason that if the firms have already spent a lot on R&D on creating new varieties, they have already explored the possible varieties and therefore, there are fewer chances of further growth. But in developing economies, most of these possibilities are yet unexplored and hence, initial doses of expenditure on R&D results into much faster growth rates. However, the expenditure on R&D is also fraught with high risks and uncertainties, therefore, the private investors are less interested in making such investments. But we know that in macro perspective, the increase in variety of intermediate products increases the aggregate output, hence, we can say that even though the private benefits of R&D are lower, their social benefits are much higher. Therefore, it is generally suggested that under riskier and uncertain circumstances, the government should step in to build the R&D infrastructure or it may also direct involve itself in creating new varieties through R&D or alternatively, it may also provide incentives/subsidies to the private sector for encouraging private investment in R&D.

7. Summary: In the evolution of growth theories, we can observe that the classical economists talk about given state of technology, knowledge and skills and say that the economy reaches the stationary state in the long run on the other hand, the neo-classical economists admit the role of technology in economic growth but they take it as an exogenous factor and although they say that as the technology improves, the economy moves on a higher production possibility frontier but for a given technology, the diminishing returns to scale are applicable. It is only the new-growth theories which have recognised the externalities of the human capital in terms of technological spillovers, knowledge spillovers and the role of R&D. These theories have pointed out that it is through enhancing the level of human capital that the economies can ensure increasing returns to scale even in the long run. Lucas says that although, the difference in units of capital is an important determinant of differences in growth of income across the economies but as we include the human capital in production function, these differences are reduced significantly. We have also found that the primary education has more of the social benefits while the higher education has more of the private benefits. Therefore, on basis of this argument, it is generally suggested that the primary education should be the responsibility of the state while it would be a rational policy decision if the higher education is privatised. Further, we have also discussed that the existing stock of physical as well as human capital is not sufficient for attaining self-sustained growth of income and output in the long-run. This can be ensured only if these two components of capital are being improved continuously. This is possible only through increasing expenditure on R&D. We have seen that the expenditure on R&D at any point of time positively affects the growth of total factor productivity and hence, the growth of income in any society. However, it is being realised that in most of the developing economies, the investment in R&D is generally risky as the returns are not sure, therefore, the private investors shy away from making these investments. However, the social returns of research and development are much higher than the private returns in a developing economy, hence it is suggested that the in such societies, the government should step forward to invest in required R&D or create a well defined structure of the incentives so that the private sector is encouraged to make such investments.

8. Glossary

Convergence and Divergence: The convergence or 'catch up' is the situation when the poor as well as the richer economies come closer to each other in terms of per capita income. On the other hand, if the gap between the per capita income of the rich and poor economies widens over period of time, then this phenomenon will be termed as divergence.

Empirical Evidence: Empirical evidence means to support a given set or inference of theory with real life observation or data. Any evidence based on verifiable observations or experience rather than pure logic or theory can be termed as empirical evidence.

Endogenous and Exogenous Variables: An endogenous variable is a dependent variable generated within a model whose value is changed by one of the given functional

relationships within the model. On the other hand, the independent variable that affect the dependent variable without being affected by it and whose variation is not determined by any of the factors given within the model. This type of variable is used for given external conditions not controlled by the economy or the economic model itself.

Human Capital: Human capital refers to the skills, knowledge and experience possessed by an individual, labour force or population in general. The level of human capital largely determines the level of productivity in the economy.

Spillover Effects: By spillover, we mean to say a process of overflowing or spreading the effects of something in one area into another area. Here, in this lesson, we have taken the spillover effects of technology in the sense that the benefits of technology used or created in one sector of the economy are soon spread to other sectors of the economy leading to multiplier effects on productivity.

9. Suggested Books and References

Barro, R.J. and X. Sala-i-Martin (1995), *Economic Growth*, New York: McGraw-Hill.

Becker, G.S. (1964), *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, NBER, Cambridge, MA.

Lucas, R.E. (1988), "On the mechanics of economic development", *Journal of Monetary Economics*, Vol. 22, No. 1, pp. 3–42.

————— (1990), "Why doesn't capital flow from rich to poor countries?", *American Economic Review Papers and Proceedings*, Vol. 80, No. 2, pp. 92–96.

Mankiw, N.G., D. Romer and D.N. Weil (1992), "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, Vol. 107, No. 2, pp. 407–437.

Ray, D. (1998). *Development Economics.* Oxford University Press, New Delhi.

Romer, D. (2001), *Advanced Macroeconomics (2nd ed.)*, New York: McGraw-Hill.

Romer, P.M. (1986), "Increasing returns and long-run growth", *Journal of Political Economy*, Vol. 94, No. 5, pp. 1002–1037.

————— (1990), "Endogenous technological change", *Journal of Political Economy*, Vol. 98, No. 5, pp. S71–S102.

Schultz, T.W. (1961), "Investment in human capital", *American Economic Review*, Vol. 51, No.1, pp. 1–17.

Todaro, M. P. And Smith, S. C. (2011). *Economic Development.* 8th Edition. Pearson. New Delhi.

10. Suggested questions

Q 1. Discuss briefly the role of human capital in economic development.

Q 2. Give the basic idea of endogenous growth models.

Q 3. What do mean by external effect of human capital?

Q 4. Differentiate between optimal growth and equilibrium growth path as suggested by Lucas.

Q 5. Give a cost-benefit analysis of education.

Q 6. Discuss the role of R&D in long term economic growth.

Q 7. Give policy suggestions for investment in education and R&D.

Lesson No 2.2   **Author: Dr. Anupama**

Human Capital – II: Health and Nutrition

Structure
1. Objectives
2. Introduction
3. Measuring Health
   3.1. Mortality Indicators
   3.2. Morbidity Indicators
   3.3. Composite Indicators
   3.4 Nutrition Based Indicators
   3.5 Some other commonly used indicators
4. Why Health Matters?
   4.1 Gains in Worker Productivity
   4.2 Health and Education
   4.3 Health and Eradication of Poverty
   4.4 Demographic Achievements
5. Health Policy
6. Summary
7. Glossary

8. Suggested books and References
9. Suggested questions

1. Objectives: This lesson aims to –
(i) Examine the role of health in economic growth
(ii) Know about the methods to measure the health status of any country
(iii) Know about the direct and indirect effects of health status
(iv) Understand the policy stances to raise the health status of a country.

2. Introduction: Saved off and Schultz (2000) say, "Common sense tells us that healthier people are more productive and that wealthier people can obtain goods and services and shape their environments in ways that make them healthier. The magnitude of these causal relationships, however, is largely unknown. Quantifying how health affects productivity and how wealth affects health is a starting point for understanding the role of health as a form of "human capital"–that is, as a characteristic of individuals, like education, that can be increased and improved through investing time and resources. In turn, understanding this aspect of human capital is critical for evaluating the role public policy plays in encouraging households, private firms, and public sector institutions to make efficient

investments in health". In case of health as in case of education, it must be kept in mind that it is also an important part of t human capital which is acquired human capacities that are influenced through conscious choices and behaviors. The key aspects of human capital are that the investments involved affect human capacities over a reasonably long time frame and that people make choices regarding these investments in relation to the expected future returns of those "investments," among other things. Therefore, in case of health, it is important to know the effect of change in health status on the productive capacities of the individuals.

3. Measuring Health Status: Before establishing the relationship between health and economic growth, it is important to know the indicators of health. There is no single metric to measure the health status and different economists have different views on each of the indicators just to add to the confusion about an accurate and adequate measure of health status in the micro as well as macro sense. Some say that it can be life expectancy at birth while other advocate the disease free and disability free life years. Apart from selecting the single or a composite indicator, the problem is also there in assigning the weights to each of the indicators. This problem arises specially in case of assigning weights to the inherited, environmental, and behavioral factors that affect health outcomes. All these factors actually have different impact on labour productivity and therefore are needed to be decomposed for appropriate policy strategy. Broadly in economic literature, we may find three sets of indicators measuring the health status of the economy. These are:

3.1. Mortality Indicators: By mortality we generally mean to say the number of deaths in a given area or period, or from a particular cause. The mortality rate is generally calculated as number of deaths per thousand of population in a particular area/country. Higher mortality rates in a country show low health status of the families.

3.1.1 Childhood Mortality Indicators: There are a number of indicators showing the mortality rate of children at different ages of their childhood. The most used mortality indicators are childhood mortality indicators (perinatal, neonatal, infant and under five mortality), maternal mortality indicators and life expectancy at birth.

a. Perinatal Mortality Rate: The peri-natal mortality rate is defined as the annual number of deaths from 28wks gestation to 7 days after birth per 1000 live births.

b. Neonatal Mortality Rate: Neo-natal mortality rate shows the annual number of deaths at 0 to 28 days of age per 1000 live births.

c. Infant Mortality Rate: The infant mortality rate is defined as the annual number of children less than 1 year old who die per 1000 live births.

d. Under Five Mortality Rate: The Under Five Mortality Rate or U5 mortality is defined

as the annual number of children less than 5 year old who die per 1000 live births

Infant mortality rate is the single most sensitive index of health and standard of living. Infant mortality decreases sharply when living standard and nutritional status improves. Due to its multifaceted character, infant mortality rate is a frequently used indicator when comparing health situation on a global scale. Other childhood mortality indicators are also equally useful in different contexts. The higher the childhood mortality is in a country, the lower will be quality of life. It will have adverse impact on productive capacities and cognitive abilities of the human factors which in turn adversely affects the income level of the country.

3.1.2. Maternal Mortality Ratio and Maternal Mortality Rate: In a prosperous economy, with better nutrition and access to better medical facilities, the mortality among the females, especially at the time of birth of babies is low. Here, we have two concepts related with maternal mortality. One is maternal mortality ratio which is defined as annual number of deaths of women from pregnancy related causes per 100,000 live births. On the other hand, another concept related with maternal mortality is Maternal Mortality Rate which is defined as the annual number of pregnant women that die from pregnancy related causes per 100,000 women in the reproductive age group. The lower is the maternal mortality rate and maternal mortality ratio, the healthier will be female population of a country which refers to half the population as well as half the labour force. Healthier females have other positive externalities as well in terms of healthier children, higher labour force participation and even qualitatively better education standards of the children.

3.2. Morbidity Indicators: Morbidity indicators are useful measures of disease burden and occurrence especially for chronic and non-fatal diseases. There are three types of morbidity indicators viz. Incidence rates, Prevalence rates, Case fatality rates. Here the Prevalence Rate is defined as the total number of current cases suffering from particular diseases of certain population at a specific point of time; the Incidence refers to the number or rate of new cases of a particular condition/disease during a specific time; and the Case Fatality Rate refers to the number of deaths among observed new cases.

3.3. Composite Indicators: The World Health Organisation (WHO) defines health as a complete state of physical, mental and social well-being, and not merely the absence of disease or infirmity. Therefore, it is proposed that instead of relying on a single variable of life expectancy or mortality etc, it is better to rely upon a composite indicator which includes the quality of life which is free of diseases. For this purpose mainly two types of indicators are used. These are – the Disability Adjusted Life Years (DALY) and Health Adjusted Life Expectancy (HALE).

3.3.1 Disability-Adjusted Life Years (DALY): The disability-adjusted life year (DALY) is a measure of overall disease burden, expressed as the number of years lost due to ill-health, disability or early death. One DALY is equivalent to one lost year of "healthy" life. The sum of these DALYs across the population, or the burden of disease, can be thought of as a measurement of the gap between current health status and an ideal health situation where the entire population lives to an advanced age, free of disease and disability (www.who.org as accessed on 29th December, 2015). Thus, DALYs for a disease or health condition are calculated as the sum of the Years of Life Lost (YLL) due to premature mortality in the population and the Years Lost due to Disability (YLD) for people living with the health condition or its consequences, its calculation method is given below:

$$DALY = YLL + YLD$$

The YLL is calculated as the number of deaths multiplied by the standard life expectancy at the age at which death occurs:

$$YLL = N \times L$$

Where, N = number of deaths AND L = standard life expectancy at age of death in years.

The second component of DALY i.e. YLD is calculated as the number of incident cases in a particular period multiplied by the average duration of the disease and a weight factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (dead). The basic formula for YLD is given below:

$$YLD = I \times DW \times L$$

Where, I = number of incident cases; DW = disability weight; and L = average duration of the case until remission or death (years)

YLD can also be calculated on basis of Prevalence, instead of incidence as given above.

$$YLD = P \times DW$$

Where, P = number of prevalent cases; DW = disability weight

3.3.2 Health-Adjusted Life Expectancy (HALE): Health-adjusted Life Expectancy or HALE is defined as the total years of life a newborn can expect to live in full health. In contrast to conventional life expectancy, which considers all years as equal, to calculate HALE, years of life are weighted by health status. Thus, it is a measure of burden of disease based on life expectancy at birth, but including an adjustment for time spent in poor health. In other words, it is the number of years in full health that a person can expect to live, based on current rates of ill health and mortality. It can be calculated in the following manner:

$$HALE = \text{overall life expectancy - number of unhealthy years}$$

3.3.3 The Burden of Ill Health: The difference between life expectancy and HALE estimates represents the burden of ill health. The magnitude of this gap differs by sex and it has been found in most of the developing countries that the burden of ill health is higher among women. Women's longevity is one factor behind this disparity. Since the prevalence of chronic conditions increases with age and women live longer, they spend a longer period with chronic
conditions. Also, at age 65 and over, women tend to be in notably poorer health than men the same age.

3.4 Nutrition Based Indicators: Generally, it is the energy intake or the calorie intake which is most popularly used as an indicator of nutritional standards. Calorie availability is computed by converting food quantities (purchases and consumption from own production) into nutrient intakes, using standard food composition tables. But, it is generally suggested that calorie intake is actually an 'input' indicator while health is better measured in terms of 'output' indicators such as height, weight and Body Mass Index (BMI). The nutritionist rely more upon BMI as it gives the ratio of weight to height. On an average, a prime-age male is supposed to have a BMI of about 25; BMIs are considerably lower in poor countries and it is also being found that extremely low (below 18) and high (above 30) values have been associated with higher adult mortality (Waaler 1984 and Fogel 1994).

3.5 Some other commonly used indicators

In addition to the indicators explained earlier, there are plenty of other indicators that are used when working with different global health related questions. Some examples of other useful indicators:

a. Demographic Indicators: There are certain demographic indicators which show the health status as well. Some of these indicators are − population growth rate (percentage increase in population in a year), Total fertility rate (births per 1000 women at 15-49 years of age in a year), Crude birth rate and death rate (births and deaths per 1000 inhabitants in a year). At higher health status, the total fertility rates, the birth rates and the death rates are lower and so the population growth rate is also lower.

b. Indicators of Nutritional Status: The indicators of nutritional status include, the proportion of babies with low birth weight (<2.5 kg at birth) of all births, Proportion of under 5-year-old malnourished children, children and adults with severe anemia and the proportion of the obese population.

c. Social Indicators: Sometimes, the social indicators can also be used for a proxy of the health status of the society. One such indicator is percentage of people who have access to clean water and sanitation facility.

d. Health Systems Indicators: These indicators actually show the availability and access to health infrastructure in a society. These indicators include – Proportion of attended deliveries, Vaccination coverage, Percentage of people with access to health care, Contraceptive use, Number of health care personnel per population, Population covered by a hospital, doctor, ANM etc.

e. Economic Indicators: This indicator includes the health expenditure by public as well as the private sector. This can be Health spending per capita by the public sector or for the society as a whole. Besides, we can also calculate the out-of-pocket expenditure on health by the households or the individuals.

f. Anthropometrical indicators: There are some anthropometrical indicators which can be used for measuring health as well as the nutritional status of the individuals. Some of the commonly used indicators are – Weight for height, Height for age, BMI etc.

4. Why Health Matters?: Good health is not only an important factor for the well-being of the individuals but it also contributes to economic growth in a number of ways. It reduces production
losses caused by worker illness; it permits the use of natural resources that had been totally or nearly inaccessible because of disease; it increases the enrollment of children in school and makes them better able to learn; and it frees for alternative uses resources that would otherwise have to be spent on treating illness (WDR, 1993). The economic gains from improvement in health facilities are discussed below:

    4.1 Gains in Worker Productivity: Health is an important form of human capital. It can enhance workers' productivity by increasing their physical as well as mental capacities. There are many studies across countries that have established a positive relationship between health and productivity for both unskilled and skilled workers. A link also exists between health and income at the macroeconomic level (Bloom and Canning, 2005). Strong positive correlation has been found between measures of aggregate health, such as life expectancy or child mortality, and per capita income (Preston 1975; World Bank 1993). Higher incomes promote access to many of the goods and services believed to produce health and longevity, such as a nutritious diet, safe water and sanitation, and good health care.

Figure 1: Nutrition and Capacity to Work
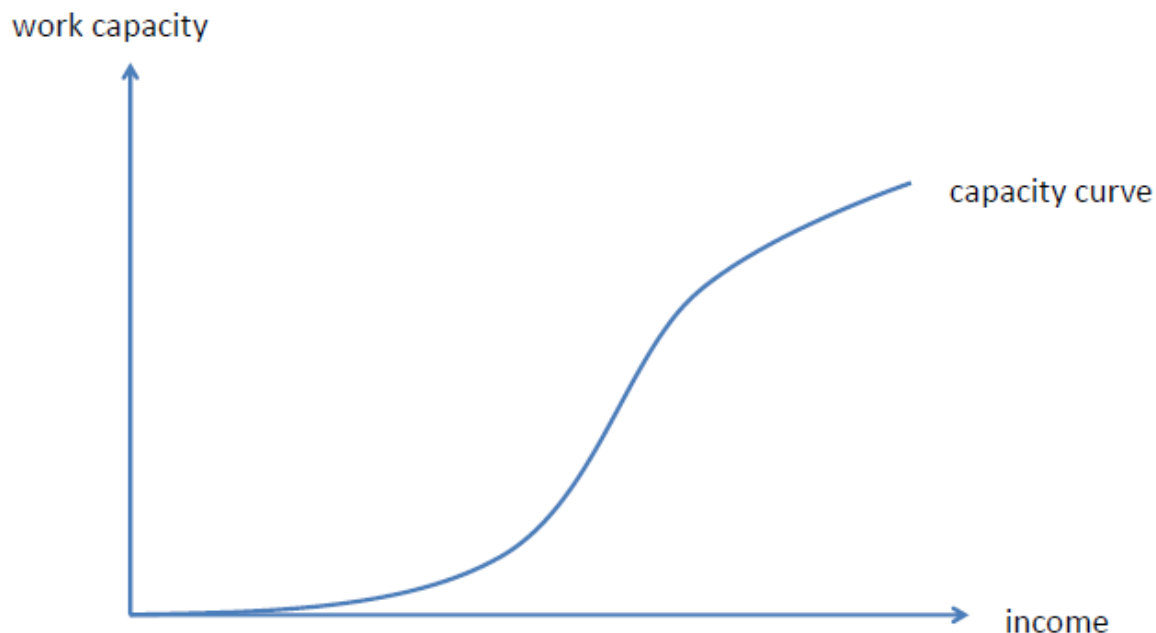
In figure 1, the nutrition level is actually measured by the income level. The shape of this curve depicts the different impacts of nutrition on work capacity of the individuals at different income levels. Initially as the income levels increases, due to inadequate nutritional standards at the low income levels, the work capacity rises slowly, then as the adequate nutrition is achieved, the work capacity rises sharply and then for higher income levels or the nutritional standards, the work capacity curve stagnates. Thus, we can say that as the nutritional standards are raised the work capacity of the individuals and overall productivity for the economy is raised which leads the economy on faster economic growth. The shape of this curve clearly suggests that for raising the nutritional standards and hence the productivity level of the workers to a minimum required level, the income must be raised up to that level which is supposed to be the critical minimum of the income for any society.

On employer's side, the most obvious gain of good health are fewer work days lost to illness, increased productivity, greater opportunities to obtain better-paying jobs, and longer working lives. Healthier workers earn more because they are more productive and can get better-paying jobs. In case of illness, an individual's lost output and earnings often deplete and savings fall while a healthy life results in accumulation of savings. With a healthy work force, employers can reduce the number of work days lost and so are able to achieve their production targets well in time. Healthy staff also induces the employers to invest more in staff training, and exploit the benefits of specialization. In developing economies as a big proportion of the population lives below the poverty line or is

vulnerable and any contingency of life in terms of diseases or any medical expenditure pushes them below the poverty line. On the other hand, the social benefits of health expenditure are immense in these economies, therefore, it is suggested that at least the primary health care should be provided by the public sector and there should be adequate measures of social security to save these poor masses.

Here, it is important to discuss the efficiency-wage model. The primary implications of efficiency wage models can be illuminated in a simple model in which a worker's physical health and therefore productivity is assumed to depend positively on the real wage paid. This formulation was advanced by Leibenstein (1957) to highlight the linkages among wages, nutrition, and health in less-developed countries. Firms, in this context, get healthier, more productive workers if they pay higher wages. This implies that there may be some people who are so poor and so unhealthy that they are too costly to be employed. The studies show that people in better health (measured by BMI, for example), are more likely to undertake strenuous tasks

4.2 Health and Education: Health and education are closely related with each other, especially in a growing economy. Greater achievements of the health standards improve the returns to education as well. Longer and healthy life raises the returns to education during the life time of an individual. Higher life expectancy induces the families to invest in higher education. With mutual reinforcement of expenditure on health and education, the societies are able to achieve higher levels of growth through healthy and educated personnel. On the other hand, poor health and nutrition reduce the gains of schooling in three areas: enrollment, ability to learn, and participation by girls. Children who enjoy better health and nutrition during early childhood are more ready for school and more likely to enroll. Health and nutrition problems affect a child's ability to learn. Nutritional deficiencies in early childhood can lead to lasting health problems. Anemia reduces cognitive function, iodine deficiency causes irreversible mental retardation, and vitamin A deficiency is the primary cause of blindness among children. All these factors reduce the education standards in a country. Since in the poor families, the girl children are always discriminated in their diets, they are more likely to suffer from these problems and hence, the first one to drop out of school. In the developing countries, poor health is considered to be one of the major reasons behind the school drop-outs.

4.3 Health and Eradication of Poverty: The goal of reducing poverty provides a different but equally powerful case for health investments. The adverse effects of ill health are greatest for poor people, mainly because they are ill more often, but partly because their income depends exclusively on physical labor and they have no savings to cushion the blow. They may therefore find it impossible to recover from an illness with their human and financial capital intact. This further impairs their capacities to earn their livelihood on sustainable basis. The poor are exposed to greater risks from unhealthy and dangerous conditions with poor sanitation facilities, both at home and at work. Malnourishment and the chronic diseases in the family mean that they are more likely to fall ill and slower to recover, especially as they have little access to health care. Whenever an earning member

of the family falls ill, other members of the household may at first cope by working harder themselves and by reducing consumption, perhaps even of food (WDR, 1993). Both adjustments can harm the health of the whole family. If free health care is not available, the costs of treatment may push a household deeper into debt. In this perspective, it is widely suggested that for mitigating the financial distress of the poor, the importance of health should be explicitly recognized in formulating policies. Investments to reduce health risks among the poor and provision of insurance against catastrophic health care costs should be important elements in a strategy for reducing poverty. Spending on health is a productive investment as it can raise incomes, particularly among the poor, and it reduces the toll of human suffering from ill health. Good health is no doubt a fundamental goal of development as well as a means of accelerating it. By targeting health as part of development efforts the countries are also able to reduce the economic inequalities. In this context, it is important to cite WHO's Commission on Macroeconomics and Health [2001] states that improving the health and longevity of the poor is an end in itself, a fundamental goal of economic development. But it is also a means to achieving the other development goals relating to poverty reduction. The linkages of health to poverty reduction and to long-term economic growth are powerful, much stronger than is generally understood. The burden of disease in some low income regions, especially sub-Saharan Africa, stands as a stark barrier to economic growth and therefore must be addressed frontally and centrally in any comprehensive development strategy.

4.4 Demographic Achievements: All the developing economies wish to attain certain demographic standards for better quality of life. These standards are to reduce the death rate, birth rate, equitable sex ratio, higher life expectancy, lower infant and maternal mortality etc. along with low growth rate of population. But the targets related to each of these indicators are hard to achieve without improving the health status. From the growth experience of several economies, we can simply find that the birth rate is also a function of death rate. A high death rate among the children always forces the families to have more of the children but if the life expectancy is high, child mortality is low, or say if there are enough chances to survive, only then the families can be convinced to adopt the family planning measures. So, raising the health status is the first condition to control the population growth. Moreover, in an equitable society the male-female ratio should be equitable but this is possible only if the mortality rate among the girl children as well as adult females is controlled which needs higher health and nutritional standards for females.

5. Health Policy: Keeping in perspective the benefits of health for poor segments of the population and poor countries, WHO suggests that the developing countries should begin to map out a path to universal access for essential health services. It is also suggested that the civil society and the government should join hands to identify the health needs of the masses and to provide the required services in a more effective manner. It suggested the establishment of a National Commission on Macroeconomic Health (NCMH) with the following tasks (WHO, 2001):

(1) Identify the priority areas for health interventions and the financing strategies to address those priorities;

(2) Designate a set of essential interventions to be made universally available to the entire population on the basis of public financing;

(3) Initiate a multi-year program of health-system strengthening, focused on service delivery at the local level and including training, construction, and bolstering of infrastructure, and management development to enable the health sector to achieve universal coverage of essential interventions;

(4) Establish quantified targets for reductions in the burden of disease based on sound epidemiological modelling;

(5) Identify key health synergies with other sectors (e.g., education etc); and

(6) Ensure consistency of the strategy with the overall macroeconomic framework.

Realizing that the developing economies face the shortage of funds and there will be a need of external help to carry on any nation-wide programme, WHO has suggested that the donor countries should fulfill their commitments of health financing of around 0.1 per cent of their combined GNP but it is also emphasized that the actual disbursements should depend upon the capacity and performance of the developing economies in designing sound, credible and monitorable strategies for scaling up essential health interventions. The global strategy will require bolstering the operations of existing international institutions such as WHO and its sister concerns. The WHO has a crucial role to play in several areas. It advises member governments on appropriate health strategies and it helps them to strengthen their on-the-ground capacity.

In addition, WHO will be critical in establishing epidemiological baselines in each country and at the world level, which will be needed as critical inputs to global disease control efforts. The WHO and the World Bank together have a shared responsibility in the analysis and dissemination of best practices in health systems reform as well as in giving support for policy reforms to address existing resource imbalances in the health sector. Finally, the IMF will be important in assisting donors and recipient countries to take account of the scaling up process in the macroeconomic policy framework of low-income countries, particularly with respect to the absorption of additional international funding. Fighting disease is to be considered the truest test of the common capacity of international institutions to forge a true global community. It has been estimated that people need just $34 per person to cover essential health services and hence there is no excuse in today's world for millions of people to suffer and die each year for lack of resources. Thus, it is well established that the public health and epidemiological programs help to improve the current 'state of affairs' of LDC, in creating complementarities on other forms of human capital, as education or sustainable fertility rates for families. Indeed, it is today well documented that increases in life expectancy affects parents decisions to invest in their children's education, by lowering the expected losses from infant mortality. In addition, more educated and healthy population increases productivity and hence the per capita income.

6. Summary: Health is an important form of human capital. It can enhance workers' productivity by increasing their physical capacities, such as strength and endurance, as well as their mental capacities, such as cognitive functioning and reasoning ability. We expect to see a positive relationship between health and productivity for both unskilled and skilled workers. In recent years, substantial progress has been made in documenting the existence of a causal impact of health on wages and productivity in low-income settings using both experimental and non-experimental methods. Economists have identified several channels through which health affects the level of output in a country. Obviously, healthier people are better workers. They can work harder and longer, and also think more clearly. Health also has a number of indirect channels through which the aggregate output of the economy is affected. Improvements in health raise the incentive to acquire schooling as investments in schooling can be paid off over a longer working life. Healthier students also have lower absenteeism and higher cognitive functioning, and thus receive a better education for a given level of schooling. Improvements in mortality may also lead people to save for retirement, thus raising the levels of investment and physical capital per worker. Physical capital per worker may also rise because the increase in labour input from healthier workers will increase capital's marginal product. The effect of better health on population growth can also be observed through higher child survival rates and lower infant and child mortality and Net Rate of Reproduction along with an increase in life expectancy. In sum, it can be stated that the investments in health not only leads to higher growth but also better living standards and better quality of life.

7. References

Bloom, David E. and David Canning (2005). Health and Economic Growth: Reconciling the Micro and Macro Evidence. Working Paper. Harvard School of Public Health. February.

Fogel, Robert (1994). Economic Growth, Population Theory and Physiology: The Bearing of Long- Term Processes on the Making of Economic Policy, *American Economic Review*, June, 84(3), pp. 369-95.

Savedoff, W. D., and T. P. Schultz. (2000). *Wealth from Health*. Washington, D.C.: Inter-American Development Bank.

Waaler, Hans (1984). Height, Weight and Mortality: The Noiwegian Experience, Acta Medica Scandinavia, Supplement No. 679.

World Health Organisation (2001). *Macroeconomics and Health: Investing in Health for Economic Development.* Report of the Commission on Macroeconomics and Health. World Health Organisation, Geneva.

World Bank. (1993). World Development Report 1993: Investing in Health. Washington, D.C.

8. GLOSSARY

Life Expectancy: an indication of how long a person can expect to live, it is the number of years of life remaining to a person at a particular age if death rates do not change. This measure only provides an indication of length of life not quality of life ie. It is regardless of any illness or disability.

Anaemia: a condition in which there is a deficiency of red cells or of haemoglobin in the blood, resulting in pallor and weariness.

Obesity: the state of being grossly fat or overweight.

BMI: BMI or Body Mass Index is a measure of whether someone is over- or underweight, calculated by dividing their weight in kilograms by the square of their height in meters.

Malnutrition: Malnutrition is the condition that develops when the body does not get the right amount of the vitamins, minerals and other nutrients needed to maintain healthy tissues and organ functions.

Epidemiology: The branch of medicine which deals with the incidence, distribution, and possible control of diseases and other factors relating to health.

9. Suggested questions

Q 1. Discuss briefly the role of health in economic development.
Q 2. Give various methods to measure the health status in an economy.
Q 3. Why health matters?
Q 4. Give policy suggestions for improvement of health status in developing countries.

SUSTAINABLE DEVELOPMENT

2.3.1 Introduction

2.3.2 Objectives

2.3.3 Defining Sustainability and Sustainable Development

2.3.4 Evolution of the concept of Sustainable Development

2.3.5 Dimensions of Sustainable Development

2.3.6 Objectives of Sustainable Development

2.3.7 Sustainable development and environmental accounting

2.3.8 International engagement towards sustainable development

     2.3.8.1. The Rio Earth Summit

2.3.8.2 Earth Summit 2002

2.3.8.3 Earth Summit 2012

2.3.8.4 The UN Sustainable Development Summit 2015

2.3.9 Conclusion

2.3.10 References

2.3.11 Suggested readings

2.3.12 Questions for practice

2.3.1 Introduction

Development is much more than simply economic growth. Development is a multi-dimensional processthat goes far beyond the objective of increased average income, to include reduction in social and political inequalities, eradication of poverty, improved health-care and education access and changes in the entire social and political system so as to bring about an over-all improvement in the standard of living of the people. Development is a holistic concept and one of the main elements of development also

includes ecological viability of the process. Concept of 'Sustainable Development'takes such a rounded approach towards development

The concept of 'Sustainable Development' developed around 1960s, when people became more and more aware of the detrimental effects of industrialisation. It provided as a framework where the future actions and directions of policies could be decided appropriately at international, national as well as individual level. Sustainability involves a long-term planning process which ensures utilization of resources in such a way so that it leads to a much larger user-base of these resources and more people are able to enjoy these for a much longer period.

Earlier theories of development did not focus on the connections between an individual and his environment. The concept of Sustainable Development is a way forward in this regard. It has emerged out of the fear of depleting natural resources. This depletion has resulted from voracious misuse of earth's valuable and limited resources by those who have control over production systems. We now understand, that in the process of development, communities may accidentally be exhausting the resources on which they depend for their existence. Faster economic growth may mean more damage to the natural environment. The rising pressure on earth's resources may have serious consequences on the future growth potential of the countries. Reduced productivity of resources due to environmental degradation and increased health-related expenses, imposes high costs on countries which undermines the pace of economic development. Farm productivity has also reduced due to environmental degradation caused by population pressures and the lower-income groups are the ones that suffer the most.Thus, the definition of economic development must also include attaining environmentally sustainable growth so as to ensure productivity of resources and improving living conditions, especially among the poor.

2.3.2 Objectives

After studying this chapter, the students should be able to:

- Explain the meaning and origin of the concept of sustainable development
- Describe the fundamental principles of sustainable development
- Understand the interaction between the Economic, Ecological and Social aspects of development
- Understand the importance of global action with regards to sustainable development

2.3.3 Defining Sustainability and Sustainable Development

Nature provides humans with abundant physical resources that are necessary for the sustenance of economies. Nature has supported and maintained life since the beginning of time and should continue to do so. But the nature's resources are limited, with many of the essential physical resources taking billions of years to get replenish. With an increase in the developmental activities, this natural wealth is fast depleting. Development is an essential

feature of any economy, but in the process, we are depriving the future generations of the benefits of nature's resources as the rate at which humans are consuming them is much higher than the rate at which nature can replace these resources. Thus, the challenge before mankind is to determine a developmental framework that can be sustained over time keeping in mind the limits of the nature's carrying capacity.

2.3.3.1 Sustainability

In simplest terms, 'Sustainability' refers to a process or a state that can be maintained indefinitely. It includes a prudent use of natural resources in ways that it does not overexploit the productive capacities of earth.

The *Caring for the Earth: A Strategy for Sustainable Living* (1991) by International Union of Conservation of Nature and Natural Resources (IUCN) says that 'sustainable use refers to using renewable resources at a rate within their capacity for renewal'.[1]

2.3.3.2 Development

Development refers to process of holistic growth of human and natural environment. It involves a whole gamut of changes in the economic and social sphere ensuring availability of choices to all individuals and freeing them from any form of dependence. It comprises social and economic improvements, creation of opportunities, increased prosperity, removal of inequalities and an overall improvement in the quality of life.

2.3.3.3. Sustainable Development

The need to develop the concept of **Sustainable Development** arose from the untenable development strategies that were being pursued. Sustainable Development combines two words, 'Sustainability' and 'Development' to imply a growth pattern that also focusses on maintaining the ability of natural systems to continually provide natural resources upon which the economy and society depend. The most widely used definition of Sustainable Development has been provided by the Brundtland Commission in its report *Our Common Future*(1987); it defines Sustainable Development as '***development, which meets the needs of the present without compromising the ability of future generations to meet their own needs'.***[2]The report went on to argue that the concept "contains within it two key concepts: the concept of 'needs', in particular the essential needs of the world's poor, to which priority should be given; and the idea of limitations imposed by the state of technology and social organisation on the environment's abilityto meet present and future needs.''[3]

Thus, sustainable development lays emphasis on the relationship between humans and their environment. It encourages a development path that focuses on the conservation and preservation of natural resources and the environment.

2.3.4   Evolution of the concept of Sustainable Development

The earlier conceptions of sustainability can be found in the theories of Malthus (1766-1834) and William Stanley Jevons (1835-82). Malthus was concerned about the shortage of resources in the face of rising population, and Jevons postulated a theory concerning the energy (coal) shortages. However, it was not until the 1960s that the concerns about environmental degradation rose significantly and reached the public domain. This was due to the fact that the adverse impacts of industrialisation, especially the health hazards, were now becoming more apparent and there was now a demand for a shift from the conventional, growth-oriented economic development to a more holistic approach.

In *The Limits to Growth* (Meadows *et al.* 1972) published in 1972, the authors concluded that given the current trends of world population, industrialisation, pollution, food production and resource use, the carrying capacity of the planet would be exceeded within the next 100 years and there would be a collapse of the global system by the mid to latter part of the 21st century.[4]Similarly, Herman Daly developed the 'steady state economics', recognising the absolute limits to economic growth.[5]

However, the term 'sustainable development' became popular with the *World Conservation Strategy* (IUCN 1980) presented by the International Union for the Conservation of Nature and Natural Resources in 1980. But its focus was rather limited to the ecological sustainability and didn't link sustainability with social and economic issues.

In 1987, the World Commission on Environment and Development (WCED) published its report titled *Our Common Future*. The WCED was chaired by Gro Harlem Brundtland, and *Our Common Future* is therefore most commonly referred to as the Brundtland Report. It was for the first time that the inter-connections between the economic, ecological and the social dimensions of development were clearly laid out and defined. By linking these three aspects, the report conceived a new model of development of sustainable development by putting together 'development', a traditional social and economic goal and 'sustainability', an ecological goal.

One of the key proposals made in the Brundtland Report was for the UN to hold an Earth Summit., which was held in Rio de Janeiro in 1992 and has been followed by twosuch summits in 2002 and 2012. These will be discussed later in the chapter.

In addition to this, various specialised UN agencies and programmes like Food and Agriculture Organisation, United Nations Development Programme, United Nation Environment Programme have been working to promote sustainable development at a global level.

2.3.5 Dimensions of Sustainable Development

The term 'sustainability' originally belonged to the realm of ecology. But relating the concept of sustainability with the notion of 'development', shifted the focus of analysis from that of the ecology to that of the society. The main focus of sustainable development is on the direction of societal changes and the corresponding environmental considerations. Development therefore can be perceived as an integration of three essential and inter-related aspects: **Economic, Environmental** and **Social** dimensions.

Economic Sustainability:

This concerns the allocations and distribution of scarce resources. Sustainability of an economic system would be defined as its capacity to achieve a constant growth of its economic indicators, especially generation of employment and income in order to sustain its population. It involves production through the most judicious resource-mix that ensures generation of income not only for the present, but also the future generations.

Social Sustainability:

The Social aspect relates to human morals and values, relationships and institutions. It includes extending welfare services to all individuals and ensure their equitable distribution. A sustainable society is the one living within the restrictions established by environmental limits, but connected to the ideas of social equity and justice.

Environmental Sustainability:

The Ecological aspect involves the contribution of both the economic and the social and their effect on the environment and its resources. It involves conserving and enhancing our resource base. Environmental Sustainability can be defined in the terms of our capacity to preserve the resource-supply function of the environment; ensuring the protection and renewal of the natural resources.
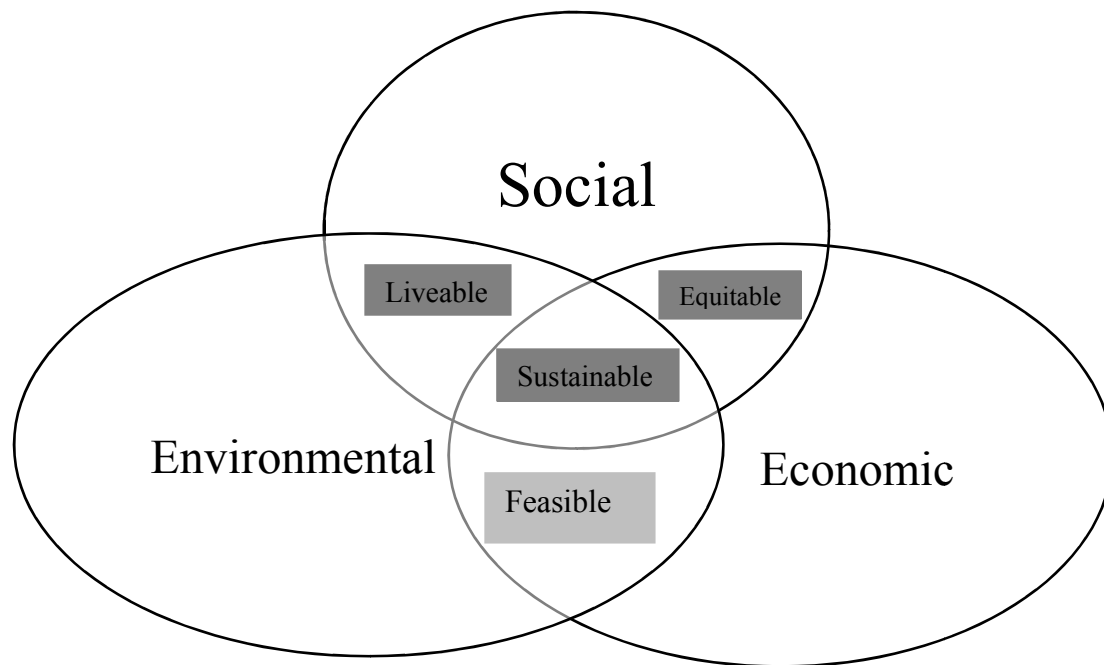
Fig. 1. Interaction between three dimensions of Sustainable Development are depicted through three circles:

A sustainable community balances its social, economic and environmental components while improving the quality of life for both existing and future generations. It therefore becomes vital to assure such an economic development that ensures the balance between three E's: **Ecology, Economy** and **Equity.** To achieve Sustainable Development, the respective governments must ensure a complete interaction between Environment, Society and Economy.

2.3.6 Objectives of Sustainable Development

Sustainable development could be referred to as "equitable and balanced" development, which means that in order for development to be sustained, or to continue indeterminately, it must strike a balance between the interests of various groups, both within the same generation and among different generations. And this must be done simultaneously in three major interrelated areas–economic, social, and environmental. The idea of Sustainable Development can be seen as a model for social change that has ecological sustainability as an objective, in addition to the traditional developmental objectives. Some of the objectives under the three main dimensions can be listed as:

Economic objectives- 1. Growth 2. Efficiency 3. Stability4. Employment

Social objectives- 1.Full employment 2. Security 3. Equity 4. Education 5. Health

Environmental objectives- 1. Healthy Environment 2. Rational use of renewable resources   3. Conservation of non-renewable resources.

To ensure sustainability, these inter connected objectives have to be kept in mind. Ignoring these may slow down or even reverse the development process.

The broad objectives of sustainable development are:

1. To maintain the standard of living of the largest number of people with equity and justice.
2. To conserve and protect earth's natural resources from misuse and wasteful consumption.
3. To innovate new technology, which works in harmony with laws of nature.There needs to be a consideration of sharing risks and benefits from development policies undertaken by different nations.
4. To involve local and indigenous communities for a more grassroots oriented development policies
5. To plan international institutions, which recognise the requirements of poor nations and support them to achieve their growth targets without destroying their natural wealth and environment

Self-Check Exercise

1. What do you understand by Sustainable Development?
2. What are the three pillars or dimensions of Sustainable Development?

2.3.7Sustainable Development and Environmental Accounting

 For some economists, a development path is sustainable "if and only if the stock of overall capital assets remains constant or rises over time.''[6]But in order to pursue short-term economic goals, we are indiscriminately destroying the natural endowments which are an assent not only to the current generation but also for all the future generations as well. Therefore, it becomes important for the policy makers to integrate some sort of **environmental accounting** into the policy decisions. For instance, the loss of valuable resources should be factored in to the national income accounting and the estimates of economic growth.

David Pearce and Jeremy Warford have explained the concept of environmental accounting.The overall capital assets of a country include not only manufactured capital

(machines, factories) but also human capital (knowledge, skills, experience) and environmental capital (forests, minerals, soil quality). Now, the correct measure of national income, corresponding to the idea of sustainability mentioned above, would mean 'the amount that can be consumed without running the stock of capital down'.[7] Thus, the measure of *sustainable national income* or *sustainable NNP** would be:

$$NNP^* = GNP - D_m - D_n$$

Where, NNP* is sustainable national income, D_m is the depreciation of manufactured capital, D_n is the depreciation of environmental capital (the monetary value of the environmental decay over the accounting year)

An even more accurate measure would be-

$$NNP^* = GNP - D_m - D_n - R - A$$

Where, D_m is depreciation of man-made capital, D_n is the depreciation of natural capital, R is restorative capital (expenditure required to restore environmental capital) and A is aversive capital (expenditure required to avert/avoid destruction of environmental capital)

2.3.8 International engagement towards Sustainable Development

As discussed earlier, sustainable development received global attention with the publication of the Brundtland Report in 1987. An important feature of the report was that it proposed the United Nations to hold an Earth Summit. This led the UN General Assembly to call for the UN Conference of Environment and Development.

2.3.8.1 The Rio Earth Summit

In 1992, the first conference of its kind, the United Nations Conference on Environment and Development (UNCED), commonly referred to as the Rio Conference or Earth Summit, was held at Rio de Janeiro, Brazil. It was attended by 172 nations and some 2400 representatives of various NGOs. The main focus of the Summit was on two issues: the connection between environment and development; and the practical issues surrounding the advancement of sustainable development. Some of the important agreements reached during the Summit were:

1. The Rio Declaration on Environment and Development
2. Agenda 21
3. The UN Framework Convention on Climate Change (UNFCCC)
4. The UN Framework Convention on Biological Diversity (CBD)
5. The Forest Principles

The Rio Declaration on Environment and Development presented with 27 principles of sustainable development. It defines the rights of the people to be involved in the development of their economies, and the responsibilities of human beings to safeguard the common environment. The Rio Declaration states that long term economic progress is only assured if it is linked with the protection of the environment. It emphasised development, national sovereignty over natural resources and cooperation between states.

The Agenda 21 is the blueprint of sustainability for the 21st century. It consist of an analysis of the causes of the unsustainable forms of development and outlines a set of ideas on how to promote sustainable development. Agenda 21 provides a detailed action plan for sustainable development and establishes targets for actions that combine economic development and environmental protection. It emphasises on community-based approaches and grassroots level participation.

Two legally binding conventions were signed, viz. the **UNCCC**and the **CBD**, dealing with climate change and protection of biological diversity respectively. The framework dealt with the limits on the use of fossil fuels.

The statement on Forest Principles emphasised on national sovereignty over forests.

The Rio Summit also led to the establishment of new institutions, specifically the **Commission on Sustainable Development**. The primary role of the Commission is to monitor the progress on the agreements reached at the Rio Summit. It was created in December 1992 to provide an effective follow-up of UNCED. The CSD is a functional commission of the UN Economic and Social Council (ECOSOC). The Commission ensures UN's co-ordination of development and environment activities. It holds annual sessions to which governments are required to report, which has helped build a picture of global progress towards sustainable development.

2.3.8.2Earth Summit 2002

The **World Summit on Sustainable Development**, WSSD or **the Earth Summit 2002** took place in Johannesburg, South Africa, from 26 August to 4 September 2002.It was held 10 years after the Rio Summit, and is also referred to as Rio +10. The summit was convened with two main goals: 10 year review of the 1992 Summit and to reconcile the global commitment towards sustainable development. However, as implementation-focussed Summit, it did not produce particularly dramatic outcomes. The Summit resulted in the Johannesburg Declaration on Sustainable Development. It focussed on the need to promote sustainable development through multi-level policy actions and encouraging broad participation. But it lacked the intellectual sophistication and authority that the Rio Declaration commanded. Nevertheless, some important targets were set, such as: to halve the proportion of people without access to basic sanitation by 2015, to use and produce chemicals by 2020 in ways that do not lead to significant adverse effects on human health

and environment, to achieve by 2010 a significant reduction in the current rate of loss of biological diversity.

2.3.8.3 Earth Summit 2012

The **United Nations Conference on Sustainable Development** (**UNCSD**), also known as **Rio 2012**, **Rio+20, Earth Summit 2012,** was held at Rio de Janeiro from 13 to 22 June, 2012. It was held as a 20 year follow-up to the Rio Summit held in 1992 and the 10 year anniversary of World Summit on Sustainable Development, held in Johannesburg in 2002.

The conference had three objectives:

1. Securing renewed political commitment for sustainable development
2. Assessing the progress and implementation gaps in meeting previous commitments.
3. Addressing new and emerging challenges.

The Rio +20 renewed its commitment to Sustainable Development based on decisions taken in the UNCED in 1992; the WSSD, 2002; and the Johannesburg Plan of Implementation. The Rio +20 conference acknowledged the need to mainstream sustainable development at all levels. For this, an outcome document entitled 'The Future We Want' was released. One of the key outcomes of the Conference was the agreement by member states to launch a process to develop a set of Sustainable Development Goal, which will build upon the Millennium Development Goals and converge with post 2015 development agenda.

2.3.8.4 The UN Sustainable Development Summit 2015
The UN Sustainable Development Summit was held from 25 to 27 September, 2015 at New York where more than 150 world leaders adopted new sustainable post-2015 development agenda. The new development agenda is envisaged in the form of **Sustainable Development Goals** as defined in '*Transforming Our World - the 2030 Agenda for Sustainable Development*'

The 17 new SDGs aim to end poverty, hunger, take action on climate change and environment, and improve access to health care and education, and more. The Goals are:

1. No poverty: End poverty in all its forms everywhere
2. Zero hunger: End hunger, achieve food security and improved nutrition and promote sustainable agriculture
3. Good health and well-being: Ensure healthy lives and promote well-being for all at all ages
4. Quality education: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
5. Gender equality: Achieve gender equality and empower all women and girls[

6. Clean water and sanitation: Ensure availability and sustainable management of water and sanitation for all
7. Affordable and clean energy: Ensure access to affordable, reliable, sustainable and modern energy for all
8. Decent work and economic growth: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all
9. Industry, innovation and infrastructure: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
10. Reduced inequalities: Reduce inequality within and among countries
11. Sustainable cities and communities: Make cities and human settlements inclusive, safe, resilient and sustainable

12. Responsible consumption and production: Ensure sustainable consumption and production patterns
13. Climate action: Take urgent action to combat climate change and its impacts
14. Life below water: Conserve and sustainably use the oceans, seas and marine resources for sustainable development
15. Life on land: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
16. Peace, justice and strong institutions: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
17. Partnerships for the goals:Strengthen the means of implementation and revitalize the global partnership for sustainable development

Each goal has specific targets to be achieved over the next 15 years.

The United Nations Development Programme (UNDP) will support governments around the world in tackling the new agenda and taking it forward over the next 15 years.

Self-Check Exercise

1. Explain the concept of Environmental Accounting.
2. Discuss the importance of Rio Summit in promoting Sustainable Development at the global level.

2.3.9 Conclusion

Sustainable development developed as a concept in the early sixties when the callous industrialisation started showing signs of natural degradation and problems of pollution and ill health. Promoting sustainable development is about creating a new development paradigm, outlined within the environmental boundaries of the planet. The concept became

popular after the Brundtland Commission Report of 1987. It advocated the basic principle of inter-generational as well as intra-generational equity and justice. It brought together the traditional developmental objectives and the objective of ecological sustainability to create a new development paradigm of 'Sustainable Development.' The global efforts since then have been remarkable. The UN and its specialised agencies have been the front-runners in this regard, especially through the Earth Summits and the adoption Sustainable Development Goals

Although, the developed and less-developed countries often try to define the term to suit their own requirements, there is a basic value inherent in the concept of sustainable development which everyone agrees to, that is the respecting the human needs only relative to the ecological capability to support large number of people for the longest period of time.

2.3.10 References

1. *Caring for Earth. A Strategy for Sustainable Living* (1991), IUCN/UNEP/WWF, Gland Switzerland, p. 10

2. Brundtland Commission Report (1987), World Commission on Environment and Development, *Our Common Future*, Oxford University Press, Oxford.

3. ibid.)

4. Meadows, D.H., Meadows, D.L., Rander, J. and Behrens, W.W. (1972)*, The Limits to Growth*, Universe Books, New York

5. Daly, H.E. (1977), *Steady-state Economics*, Freeman, San Francisco, CA

6. Pearce, D. W. and Warford, J.J. (1993*) World without End: Economics, Environment, and Sustainable Development- A Summary*, The World Bank, Washington D.C., p.2.

7. ibid., pp. 2-3

2.3.11 Suggested Readings

1. Susan Baker   : Sustainable Development

2. Todaro and Smith   : Economic Development (9ᵗʰ Edition)

3. https://sustainabledevelopment.un.org/

2.3.12 Questions for practice

1. What do you understand by sustainability and sustainable development? What are its various dimensions? Explain their inter-relationship.

2. How important is global action with respect to promoting sustainable development? What have been the various international efforts in this regard?

3. Write short notes on:

    i) Environmental accounting

    ii) Sustainable Development Goals

    iii) Earth Summit 1992

    iv) Objectives of sustainable development

Conservation and Use of Exhaustible Resources

Structure
2.4.1 Introduction
2.4.2 Sustainable Development
2.4.3 Causes of Environmental Degradation
2.4.4 Neo-Classical Theory and Environmental Issues
2.4.5 Use and Conservation of Exhaustible Resources
       2.4.5.1 Conservation of Minerals
       2.4.5.2 Conservation of Water
2.4.6 Policy Measures for Environmental Protection
2.4.7 Conclusion
2.4.8 Short Answer Type Questions
2.4.9 Long Answer Type Questions
2.4.10 Suggested Reading Material

2.4.1 Introduction

Life appeared some five billion years ago on earth and man appeared much later in the process. Since then, man has lived on the earth for hundreds of thousands of years. He has never threatened so perilously its environment as he has been doing during the last two or three centuries of modern era. It is matter of great concern that for the sake of development, man has always disrupted the ecological balance by exploiting natural resources and the precious flora and fauna. The industrial revolution in the nineteenth century witnessed the large scale use of fossil fuels for industrial activities. Thereafter, continuous upward trend in industrial growth, urbanization and unexpected rise in population led to sharp rise in green house gases. In 1980s, the developed countries pushed the drive of neo-liberalism into the developing countries through the Structural Adjustment Programme (SAP) of International Monetary Fund and World Bank. The SAP forced the less developed countries to adopt the policy of liberalization, privatization and globalization. The free game plan of market forces that neo liberal globalization advocates, has overexploited the environment in its drive to maximize profits. The Environmental Kuznets Curve (EKC) is often used to describe the relationship between economic growth and environmental quality. It refers to the hypothesis of an inverted U-shaped relationship between economic output per capita and some measures of environmental quality. As per capita incomes rose, pollution and other forms of environmental degradation would first

rise and then fall in an inverted-U pattern. According to the theory, as incomes rise, societies will have both the means and the willingness to pay for environmental protection. This theory implies that level of environmental degradation is high in developing countries due to low levels of per capita income. Rising pressures on environmental resources in developing countries have severe consequences for self-sufficiency, income distribution, and future growth potential. Environmental degradation also detracts from the pace of economic development by imposing high costs on developing countries through health related expenses and the reduced productivity of resources. It is therefore, extremely important to prevent further degradation of natural resources and use them in a wise and judicious manner to ensure their sustainable utilization. Natural resource conservation involves wise use of natural resources so that they are not wasted, depleted or degraded and are available to both present and future generations.

2.4.2 Sustainable Development

The process of overexploitation of natural resources, have not only changed the natural environment but in some cases, destroyed it. The modern industries, factories, cities, towns, roads, railways, dams etc. have replaced the natural habitats of plants and animals. Thus, the natural resources are depleting gradually and a day will come when most of these will not be available for our future generation. So it is high time to think about maintaining a balance between environment and development so that both present and future generations can derive proper benefits out of these resources. This can only be achieved by the process of sustainable development. The term, sustainable development, was popularized in *Our Common Future*, a report published by the World Commission on Environment and Development in 1987. Also known as the Brundtland report, *Our Common Future* included the "classic" definition of sustainable development: "development which meets the needs of the present without compromising the ability of future generations to meet their own needs." It is generally accepted that sustainable development calls for a convergence between the three pillars of economic development, social equity, and environmental protection. David Pearce and his colleagues, in their report on sustainable development to Margaret Thatcher, then British Prime Minister, said that the principles of sustainable development meant recognizing that 'resources and environments serve economic functions and have positive economic value.' As a component of the economic system, the environment is seen to provide raw materials for production and to be a receptacle for its wastes. Figure 1 shows traditional model of economic system in which inputs of production i.e. labour and other inputs are provided by households to firms for carrying out production. The products which are being produced by firms are consumed by the households. Environment is not included in this model.

Figure 1: Traditional model of an economic system
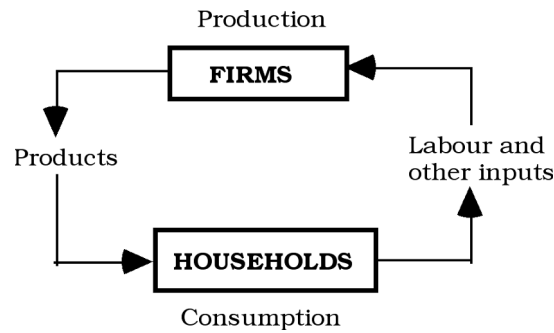
The above model has now been modified to include environmental resources. Figure 2 shows modified model of economic system in which inputs are provided by both households and natural environment. The products produced by firms are consumed by households. The waste generated by both firms and households are absorbed by the natural environment.

Figure 2: Modified model of an economic system
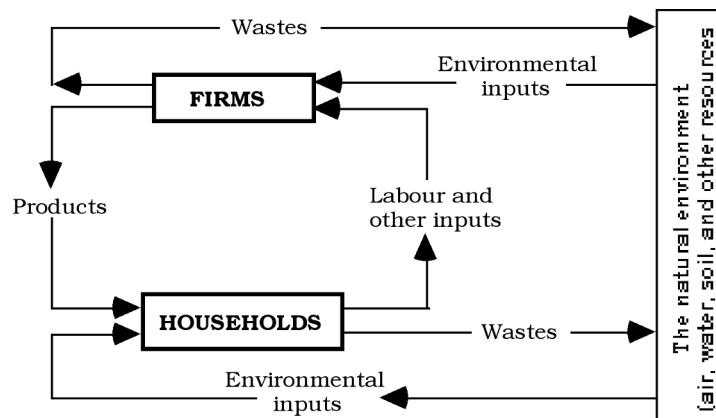
Clearly, the natural environment is an important component of the economic system, and without the natural environment the economic system would not be able to function. Hence, we need to treat the natural environment in the same way as we treat labour and capital; that is, as an asset and a resource.

Standard measures of National Income such as gross national product (GNP) and gross domestic product (GDP) fail to capture important environmental and social factors. This can result in misleading measurement of national well being, potentially ignoring important environmental problems. It is therefore important that development

policymakers incorporate some form of environmental accounting into their decisions. We can identify sustainable development path as one in which the stock of overall capital assets remains constant or rises over time. These capital stocks include manufactured capital, human capital and natural capital. The total capital stock of an economy can thus be presented as:

$$K = K_m + K_h + K_n + K_n^*$$

where, K_m is manufactured capital, K_h is human capital, K_n is non critical natural capital and K_n^* is critical natural capital. The correct measure of sustainable net national income (NNI*) is the amount that can be consumed without diminishing the capital stock. Symbolically,

$$NNI^* = GNI - D_m - D_n$$

where NNI* is sustainable national income, D_m is depreciation of manufactured capital, and D_n is depreciation of natural capital—the monetary value of environmental decay over the course of a year. NNI* includes costs of activities to reverse or avert environmental decay.

An even better measure, though more difficult to calculate with present data collection methods, would be

$$NNI^{**} = GNI - D_m - D_n - R - A$$

where D_m and D_n are as before, R is expenditure required to restore natural capital (forests, fisheries, etc.), and A is expenditure required to avert destruction of natural capital (air pollution, water and soil quality, etc.).

12.3 Causes of Environmental Degradation

Growing economic activity (production and consumption) requires larger inputs of energy and material, and generates larger quantities of waste by-products. Increased extraction of natural resources, accumulation of waste and concentration of pollutants will therefore overwhelm the carrying capacity of the biosphere and result in the degradation of environmental quality and a decline in human welfare, despite rising incomes. Furthermore, it is argued that degradation of the resource base will eventually put economic activity itself at risk. The various causes which lead to environmental degradation are as follows:

1. **Pollution:** Pollution, in whatever form, whether it is air, water, land or noise is harmful for the environment. Air pollution pollutes the air that we breathe which causes health issues. Water pollution degrades the quality of water that we use for drinking purposes. Land pollution results in degradation of earth's surface as a result of human activities. Noise pollution can cause irreparable damage to our ears when exposed to continuous large sounds like honking of vehicles on a busy road or machines producing large noise in a factory or a mill.

2. **Overpopulation:** Rapid population growth puts strain on natural resources which results in degradation of our environment. Mortality rate has gone down due to better medical facilities which have resulted in increased lifespan. More population means more demand for food, clothes and shelter. You need more space to grow food and provide homes to millions of people. This results in deforestation which is another factor of environmental degradation.

3. **Landfills:** Landfills pollute the environment and destroy the beauty of the city. Landfills come within the city due the large amount of waste that gets generated by households, industries, factories and hospitals. Landfills pose a great risk to the health of the environment and the people who live there. Landfills produce foul smell when burned and cause huge environmental degradation.

4. **Deforestation:** Deforestation is the cutting down of trees to make way for more homes and industries. Rapid growth in population and urban sprawl are two of the major causes of deforestation. Apart from that, use of forest land for agriculture, animal grazing, harvest for fuel wood and logging is some of the other causes of deforestation. Deforestation contributes to global warming as decreased forest size puts carbon back into the environment.

5. **Natural Causes:** Things like avalanches, quakes, tidal waves, storms, and wildfires can totally crush nearby animal and plant groups to the point where they can no longer survive in those areas. This can either come to fruition through physical demolition as the result of a specific disaster, or by the long term degradation of assets by the presentation of an obtrusive foreign species to the environment. The latter frequently happens after tidal waves, when reptiles and bugs are washed ashore.

6. **Poverty:** Poverty is a major cause of environmental problems and amelioration of poverty is a necessary and central condition of any effective program to deal with environmental concerns. Living in poverty is associated with poor sanitary conditions, unabated sewerage system, lack of clean water resources, and increased exposure to environmental risks. Urban poverty is a challenge in all developing countries, where not only the number of urban poor is increasing, but also the divisions among social groups within cities. There is also evidence that rural households use environmental resources quite extensively. The poor are traditionally taken as the agents for causing environmental degradation. It is generally accepted that environmental degradation, rapid population growth and stagnant production are closely linked with the fast spread of acute poverty in many countries of Asia. Both poverty and environmental degradation have been increasing in many developing countries. Ethnic minorities, migrants and refugees are the sources of environmental degradation. Their needs often receive far less attention, and they cannot always be reached through the usual hygienic and welfare channels. This problem is increasing in many developing countries.

7. **Urbanization:** Another major of environmental degradation is urbanization. The rate of urbanization and its attendant impacts differ in regions across the globe. Asia contains almost half the world's mega cities and continues to urbanize rapidly. The origins of many global environmental problems related to air and water pollution are located in cities. It is argued that means of transports like cars, buses etc are more intensively used in urban areas as compared to rural parts of the economy. Moreover food and other consumer goods have to be transported into cities, which again should lead to higher pollution that affect human health. The other sources of water pollution come from various other different situations which tend to occur in urban areas. For example, soil particles from construction and demolition sites, and also oil and toxic chemicals from car maintenance and runoff from road surfaces are also causes of water pollution. In addition, land in urban cities is covered with buildings, asphalt and concrete which bring a large amount of quick runoff, contributing to water pollution. Such an urban lifestyle and an urban design are also factors contributing to water pollution. However, much of the sewage in urban areas goes untreated and is dumped into rivers and lakes. As a result, surface water and ground water have been increasingly polluted.

8. **Globalization:** In general, globalization can affect the environment through trade liberalization by several channels such as trade affects the overall scale of the economy, the techniques of production and composition of industries. The most direct effect of trade liberalization on the environment would be through the composition of industries and hence much of the focus of the literature has been on dissecting the composition effects of trade. Trade liberalization leads to specialization. The countries that specialize in less pollution-intensive goods will have cleaner environments and countries that specialize in more pollution-intensive goods will have dirtier environment. This discussion suggests that developing countries may be able to achieve high levels of economic growth and high levels of environmental performance long before they reach the income levels of the industrialized countries. This is not to say that there are no trade off between growth and the environment. Even with good environmental policies and clean technologies, continued increases in output may tend to increase the total volumes of various kinds of pollutants in many cases. Every society has to decide for itself on the relative value it places on economic output and the environment.

2.4.4 Neo Classical Theory and Environmental Issues

Neoclassical theory has been applied to environmental issues to determine what conditions are necessary for the efficient allocation of resources and how market failures lead to inefficiencies and to suggest ways in which these distortions can be corrected. Figure 3 demonstrates how the market determines the optimal consumption of a natural resource. Finding the optimal market outcome involves maximizing the total net benefits

to society from a resource, which is the difference between the total benefits derived from a resource and the total costs to producers of providing it. This is equal to the shaded area in Figure 3. Total net benefit is maximized when the marginal cost of producing or extracting one more unit of the resource is equal to its marginal benefit to the consumer. This occurs at Q^*, where the demand and supply curves intersect. In a perfectly competitive market, the "invisible hand" will ensure that Q^* is the quantity produced. The marginal cost curve in Figure 3 is upward-sloping because extraction costs increase as a resource becomes more scarce. The resulting producer surplus is area aPb, and the consumer surplus is area DPb. Together they yield a maximum net benefit equal to Dab.

If a scarce resource is publicly owned and is freely available to all, as is the case with a common property resource, any potential profits or scarcity rents will be competed. Common property resource is a resource that is collectively or publicly owned and allocated under a system of unrestricted access, or as self-regulated by users. Neoclassical theory suggests that in the absence of scarcity rents, inefficiencies will arise. Using a somewhat different framework, we will investigate the misallocation of resources under a common property system. Figure 4 describes the relationship between the value per unit of labor on a given piece of land and the number of laborers cultivating it. Suppose for the moment that this piece of land is privately held. Conventional wisdom tells us that the landowner will hire additional labor to work the land until the marginal product of the last worker is equal to the market wage, W, at point L^*. The workload is shared equally among the employees, each of whom produces the average product. However, assuming decreasing returns to labor, each new worker hired reduces the average product of all workers. The marginal product of each additional worker is thus equal to his average product minus the decrease in the average product across all other workers. If an additional employee is hired beyond L^*, his cost to the producer, W, will be greater than his marginal product, and the difference will represent a net loss to the landowner. A profit maximizer will thus hire L^* workers, with a total output equal to average product AP^* multiplied by the number of workers, L^*. Scarcity rents collected by the landowner will equal AP^*CDW.

Society's total net benefit from land will be lower under a system of common property unless workers can coordinate their resource use decisions in a cooperative manner. Generally, if land is commonly owned, each worker is able to appropriate the entire product of his work, which is equal to the average product of all workers. Worker income will continue to exceed the wage until enough workers are attracted so that the average product falls to the level of the wage, at which point the labor force equals LC. Though total output may either rise or fall (depending on whether MPL is positive or negative—it is negative as drawn in Figure 4), the marginal product of the additional workers is below the wage. Because we are assuming that all workers could be employed elsewhere with productivity equal to or greater than W, it follows that social welfare must fall when marginal product falls below W. This situation is sometimes referred to as the "tragedy of the commons." No scarcity rent is collected at LC. The implication of the common property resource model is that where possible, privatization of resources will

lead to an increase in aggregate welfare and an efficient allocation of resources.

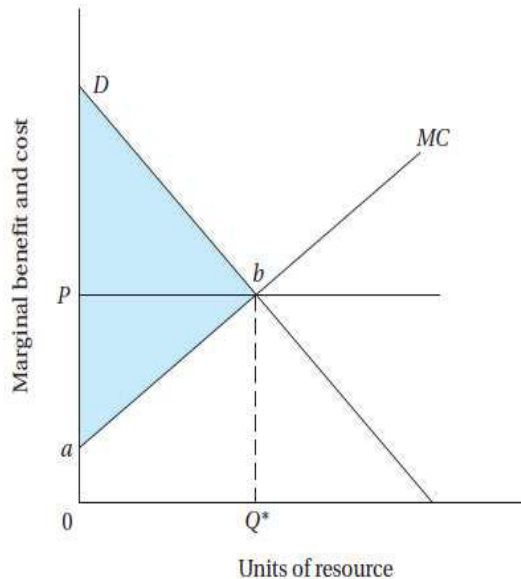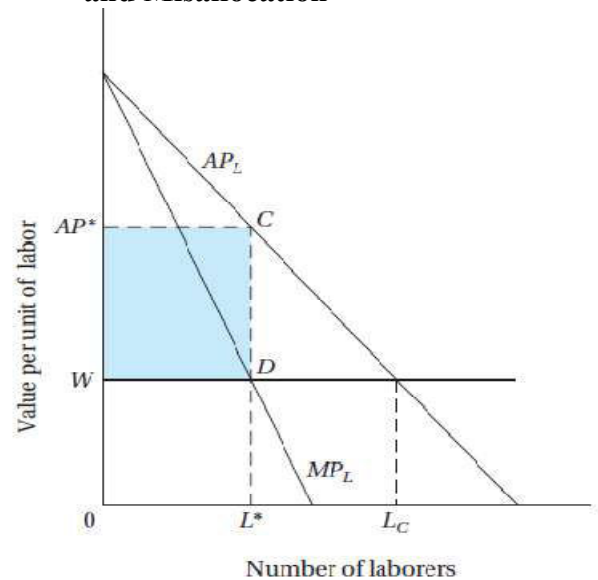| **Figure 3: Static Efficiency in Resource Allocation** | **Figure 4: Common Property Resources and Misallocation** |

2.4.5 Use and Conservation of Exhaustible Resources

Exhaustible resources are found in finite quantities which cannot be renewed if exhausted during one's lifetime. Their renewal or formation may require millions of years, not within the human life scale i.e. they are replaced slowly than they are used. Once these resources are consumed, they are practically gone forever and these include minerals and mineral oils. With increase in population and industrial growth, water is being degraded day by day. Under this present resources set up, conservation and economical exploitation of exhaustible resources and fuller utilization along with maintenance of the quality of renewable resources are considered very much essential for accelerating the pace of economic development of the country. In this section we will discuss the conservation and use of minerals and water resources.

12.5.1 Conservation of Minerals: Minerals play a significant role in industrialization and economic development of a country; hence they must be used carefully and judiciously so that they are conserved for future. Minerals are broadly divided into two groups- metallic and non-metallic minerals. Metallic minerals are further subdivided into ferrous and non-ferrous. Ferrous metallic minerals include iron, manganese and chromite. Non- ferrous metallic minerals include gold, silver, copper, tin, zinc and aluminum (Bauxite ore). Non metallic minerals include lime

stone, dolomite, mica, gypsum, phosphate and kyanite. The volume index for all mineral resources has increased from 100 in 1951 to 270 in 1986. The same index has increased from 100 in 1980 to 269 in 1996-97 and then from 100 in 2004-05 to 124.7 in 2013-14. The idea of sustainable development in the context of nonrenewable resources, in particular mineral resources, may seem a contradiction if a one-dimensional view is taken. Mineral resource development is unsustainable only if we ignore the complex interaction of economic growth, social development, and the environment. It is not always self-evident that our present modern technological society requires an ongoing supply of minerals. Mineral production, although having environmental impacts, is and will continue to be an essential part of ensuring the economic well-being of our society. To satisfy the present global mineral needs without compromising the mineral resource needs of future generations, it is imperative that we approach mineral resource development within a holistic framework comprising all components of the complex interaction between humans and the ecosystem on which they depend. By using nonrenewable resources for capital formation that will be reinvested in social, economic, and environmental activities, the concept of sustainability and mineral resource development would no longer seem to be a contradiction. A mineral becomes "economically depleted" when it costs more to find, extract, transport and process the remaining deposit than its worth. The various ways to reduce minerals depletion are as follows:

- To check and reduce depletion of minerals, we must recycle or reuse existing supplies, wasteless as well as useless minerals.
- When a resource becomes scarce, its price rises. This can encourage exploration of new deposits, stimulate development of better mining technology and make it profitable to mine lower-grade ores.
- We must encourage research in the field of minerals so that we can search for substitutes and promote resource conservation.
- Substitution of more abundant materials like plastic and glass, for scarce minerals is an important way to check depletion. The amount of lead and steel used in telecommunication has decreased with the use of plastic. Glass fibers have replaced copper wiring in telephone cables. Synthetic substitute of mica has reduced its export as well as production.
- One way to improve mining technology is to use microorganisms to extract metals from its ores known as "biomining" or 'ecological engineering', which may be an economical and environmentally preferable way to mine metals. Presently 30 percent of all copper produced worldwide, comes from such biomining. Biomining is economically feasible especially with low- grade ores.
- The science of nanotechnology has immense potential of using atoms in producing or manufacturing everything from medicines to solar cells to automobile bodies.

Thus the job of many metals can be taken over by new materials produced by nanotechnology.

- To reduce the misuse of natural gas and petroleum, subsidies given on their price should be eliminated or reduced which will help to reduce the use these minerals.
- The ownership of minerals should not be given in the hands of private players because private players use resources very extensively to get current profits. It should be in the hands of the government only then we can save minerals for our future generations.

12.5.2 Conservation of Water: Water is a finite but renewable natural resource and like other natural resources, it is an integral part of the environment. It is essential for survival of all living beings on this planet and also for socio-economic development of households, communities and nations all over the world. Water is required for several purposes such as drinking, bathing, washing, agricultural production, industrial production, generation of hydro-power, abatement of pollution, navigation, recreation and maintenance of eco-systems. Water resource of India is under great biotic and abiotic pressure. Most of the rivers, lakes, tanks and ponds are polluted and the groundwater aquifers are being over-exploited in most of the arid and semi-arid regions and are on the verge of complete exhaustion/depletion. Besides, in many areas the groundwater aquifers have been polluted. Droughts and floods also have been a bane of India's economy since time immemorial. All these factors together have adversely affected the quality of the environment. The various water related problems in India are as follows:

- The gap between availability and requirement of water is growing day by day.
- There exists lack of well-defined property rights in water.
- Pricing of water is irrational.
- Water resources are depleting and degrading day by day.
- Allocation of water from public reservoirs between irrigation and urban use is highly inequitable and is in favour of the latter.
- There are several negative externalities associated with the use of water like canal water irrigation etc.
- Water logging and salinity are the two serious problems in many of the canal command areas.
- Floods and droughts are the serious problems which have several adverse economic and environmental impacts.

Conservation and management of water are essential for the survival of mankind,
plants and animals. This can be achieved by adopting the following methods:

a. Growing vegetation in the catchment areas, which will hold water in the soil and allow it to percolate into deeper layers and contribute to formation of ground water.

b. Constructing dams and reservoirs to regulate supply of water to the fields, as well as to enable generating hydroelectricity.

c. Sewage should be treated and only the clear water should be released into the rivers.

d. Industrial wastes (effluents) should be treated to prevent chemical and thermal pollution of fresh water.

e. Rainwater harvesting should be done by storing rainwater and recharging groundwater.

f. Water supply planning should be done which involves the assessment of ground water, surface water supplies and the design of means to capture, treat, and distribute the water to users. Water supply planning is beneficial which helps communities take a proactive approach to assess existing and projected water demand, and then to identify potential resources by asking how much water is available, how much water will be needed, and what alternatives are available to meet that need. It helps to address the cumulative effects of a region's water withdrawals, which can be especially beneficial for areas that are rapidly growing.

g. Developing and implementing water pricing strategies that help to achieve economic efficiency and demand management could be one of the most beneficial options to balance water supply and demand. . When consumers do not pay the actual costs for water, or when they pay a decreasing or flat rate when consuming more water, water tends to be wasted. Effective water pricing can enhance water conservation by inducing consumer efficiency and conservation.

h. Fundamental to pricing water effectively is the implementation of metering. Universal metering is a necessary tool for enhancing water system management and water conservation. Communities that begin metering after charging a flat rate for water, regardless of the amount used, have seen water use reduced by approximately a third. Metering can help to accurately determine customers' water usage, and thus ensures that their water bill charges correspond appropriately. This can not only help to price water efficiently, but can also inform customers of their exact amount of water usage. Often providing this information encourages greater water conservation from consumers.

i. Water utilities can play a vital role in water conservation. A number of measures are available that utilities can use to design a water conservation program:

- **Water Accounting and Loss Control:** Utilities must account for water, and thus should implement at least a basic water accounting system. Accounting for water provides a basis for developing strategies to control any water losses. A water accounting system can help water utilities track water throughout the system. This can be particularly helpful if there are large volumes of water that are not accounted for, such as water that is unmetered or water that is metered but not billed. Water accounting can also help to identify any leaks in the system, and these leaks should be repaired. Even though repairing larger leaks can be costly, it can produce substantial savings in the long run. Utilities can also implement comprehensive leak detection and repair strategy. Not only should leaks be repaired, but a proactive loss-prevention program can also be beneficial. This can include "pipe inspection, cleaning, lining, and other maintenance efforts to improve the distribution system and prevent leaks and ruptures from occurring.

- **Water Use Audits:** Water-use audits can provide water systems, as well as their customers, with information about how water is used and how water usage might be decreased through specific conservation measures. Large-volume users, both commercial and industrial, are prime candidates for water audits. Water audits for these large volume users should identify the categories of water use and areas in which overall efficiency of water use efficiency could be improved by utilizing alternative technologies or practices. Large-landscape audits, including audits of irrigation practices, can also be helpful. End-use audits can be molded to the usage practices within particular user groups. Residential water audits, for instance, could focus on plumbing fixtures, lawn and garden water practices, and the behavior of the customer. Such audits could be used to make repairs and retrofits. The end result of all water audits should be a written report to the particular customer that contains specific suggestions for improved water conservation.

- **Recycling Municipal Effluent:** The recycling of municipal effluent is another option that could be used to satisfy new water demands. Treated wastewater can be used as an alternative water source. Most household water, such as water from toilets or showers, is normally filtered and then discharged by wastewater treatment facilities. However, effluent now has economic value. Industry can often use effluent as a water supply, and farmers can use it for crops. Additionally, communities can put the effluent to use for watering golf courses, parks, cemeteries, and road medians

or rights-of-way. Water utilities should work with customers to identify potential opportunities for reuse or recycling.

- **Pressure Management:** Reducing excessive pressures in a utility's distribution system can save a substantial amount of water. Reduced water pressure can lead to decreased leakage, decreased flow through open faucets, and less stress on pipes and joints which may result in leaks. A reduction in water pressure can also decrease system deterioration, which in turn lowers the number of repairs needed and extends the life of existing facilities. Lower pressures can also reduce the wear on end-use fixtures and appliances. Utilities can offer technical assistance to customers to help them address any pressure problems, and can install pressure-reducing valves for customers to lower their water pressure. Such assistance can prove particularly beneficial for customers that use large amounts of water.

2.4.6 Policy Measures for Environmental Protection

A range of policy options is available for governments in developing countries to protect the environment. These are as follows:

- Proper resource pricing should be done by eliminating the subsidies which encourage wasteful and unsustainable use of resources.
- Community involvement should be encouraged in the environmental protection programs to make them most effective.
- Programs should be launched for improving economic alternatives for the poor. Environmental devastation in rural areas may be avoidable in many cases through on-farm investments in irrigation and sustainable farming techniques, the use of alternative fuels, and the creation of barriers to erosion.
- Economic status of women should be improved. Community-based environmental programs must work closely with women because their own day-to-day activities may largely determine patterns of resource use and their ability to meet the needs of their families is dependent on the sustainable management of water and fuel supplies.
- A range of policy options can be adopted by the government of developing countries for the purpose of limiting industrial pollution which are taxation of emissions, tradable emissions permits, quotas, and standards. Additional incentives to adopt clean technologies may be provided through tax credits and subsidies specifically tied to the purchase or development of pollution abatement technologies.

- Developing nations can implement and continuously improve early warning systems to anticipate environmental emergencies; promote reforestation; restore natural ecosystem barriers such as mangroves; improve micro insurance programs; and construct storm shelters, flood barriers, and protected roads and bridges.
- The empowerment of the poor and their organizations can play an important role in protecting the natural resources on which many of the poor depend for their livelihoods.

Industrial countries can help developing nations in their efforts to improve the environment of development in following three areas:

- Developed countries should eliminate trade barriers against developing country exports in order to break the cycle of poverty and environmental destruction in developing countries.
- Heavy debt servicing drastically reduces funds available to developing-country governments for domestic social programs, including those designed to alleviate poverty and reduce environmental degradation. Debt relief should be provided by developed countries to developing countries which will help them to achieve sustainable development.
- Substantial new development assistance should be provided by developed countries to developing countries which is necessary to achieve sustainable development.

2.4.7 Conclusion

Both economic growth as well as environment protection at the national level should go hand in hand. This can lead to economic development in the real sense. Environment action is not idealism; it should be the blue print of new economic policy to help unlock the vast income creation potential inherent in environmental policy. The preservation, conservation and protection of physical or "green" environment pose a crucial challenge to economy. Many experts believe that environmentalism will become one of the critical issues that economy will need to integrate into its activities for successful operations in the twenty-first century.

2.4.8 Short Answer Type Questions

(a) What do you mean by Sustainable Development?
(b) Mention two ways with which we can reduce mineral depletion.
(c) Explain Sustainable Net National Income.

12.9 Long Answer Type Questions

a) What do you mean by exhaustible resources? Explain various methods with which we can conserve water.
b) What do you mean by environmental degradation? Explain various causes of it.

12.10 Suggested Reading Material

(1) Gaurav Datt and Ashwani Mahajan   : Indian Economy, 71th edition.
(2) M.P. Todaro and S.C. Smith   : Economic Development, 12th edition.
(3) Kartar Singh and Anil Shishodia   : Environmental Economics, Sage Publications, 2009.
(4) Ulaganathan Sankar   : Environmental Economics, Oxford University Press, New Delhi.

Lesson 2.5   Author Dr. Manisha

Land market in Agriculture

2.5.1 Introduction

A proper functioning of the land market is very important for the overall development of the economy. If land is held unequally and many individuals fail to obtain access to it, they are likely to leave agriculture in search of a less precarious source of living.Quite apart from political acceptability, there is the narrower question of economic efficiency in agriculture. Input markets such as the land market exist to bring the ratios of various inputs into line for efficient production**.**

2.5.2 Objectives of the lesson

In this lesson we will study ownership and tenancy, land rental contracts, contractual forms, contracts and incentives, risk, tenancy, and sharecropping.

2.5.3 Ownership and tenancy

Although there is substantial inequality in Asia, land inequalities in Latin America are higher by an order of magnitude. But it is also true that average landholdings are smaller

in Asia and the rural population density is very much higher, which perhaps explains, to some extent, why there are limits to inequality. After all, there is some lower bound to the smallest farm size that can be profitably used in cultivation. Latin American levels of inequality in Asia would surely drive the smallest plots to sizes that are just not feasible to cultivate. In this sense, a high population density places limits on inequality.

The African countries are somewhat, of an outlier in this respect. Much of the land is held under forms of group or communal tenure, and individual claims on such plots are weak. Thus we see that a small proportion of land is under owner cultivation simply because property rights are not well defined.

Also note that several countries provide for ownership or use rights to tenants who have worked the land for some pre-specified number of years. This legal stipulation often lowers the amount of tenancy, and in the case of Asia there may be a substantial amount of informal tenancy that goes unrecorded in the data. Several countries in Latin America also uphold as a basic principle that the land belongs to those who farm it and have legalized this principle by regarding tenancy as a basis for granting use rights or ownership. This is true, for instance, in countries such as Mexico or Brazil. Such legislation has not always had a potent effect in turning land over to the tiller. Often, the reaction in Latin America has been in the direction of tenant eviction, followed by large-scale mechanized farming.

Whereas tenancy exists all over the world, there are variations in the form of the tenancy arrangement. Latin American tenancy is largely of the fixed-rent variety: the tenant pays a fixed sum of money to the landlord in return for the right to cultivate the land. In contrast, Asian tenancy is characterized by a high incidence of sharecropping, in which the tenant yields to the landlord an agreed-upon share of the crop. Asian fractions of tenanted, land under share tenancy range from around 30% (Thailand), through 50% (India) or 60% (Indonesia), all the way up to 90% in Bangladesh. In contrast, the corresponding percentages in Latin America are much lower.

Typically, richer tenants engage in fixed-rent tenancy, because the landlord is relieved of all risk: the rent is the same whether the crop does well or not. Thus in this sense, fixed-rent tenancy requires that the tenant be willing and able to bear the risks of agricultural production. This is generally so if the tenant has substantial wealth of his own. This is (admittedly indirect) evidence for the assertion that Latin American tenancies are held by large farmers, and perhaps even evidence for the conjecture that many tenancies flow from relatively poor farmers to relatively rich farmers.

In a country with large landholdings, agriculture may take on a highly mechanized and

capitalistic form, using wage labour where labour is required. In such a regime, it may be better for smaller landowners to give up their land to large owners in exchange for a rent.

Contrast this with Asia, in which the bulk of tenancy is in the form of sharecropping. Sharecropping is an arrangement that has particular value when the tenant is small and averse to risk: if a given fraction of output is paid as rent, then the tenant is, to some extent, insulated against output fluctuations, because he can share some of these fluctuations with his landlord. This suggests that Asian tenancy probably reflects, on the whole, land leases from relatively large landowners to relatively small landowners. However, be careful not to treat this as a general rule, even in Asia.

2.5.4 Land rental contracts
2.5.4.1 Contractual forms
When a landowner wishes to rent out his land to a potential tenant. Several contractual forms are available. The simplest form of tenancy contract is what is called a fixed-rent contract, one in which the landlord charges a sum of money (per year or per season) for the rental of the land and, in turn, allows the tenant to carry out production. This sort of contract is found wherever land rentals are observed, but by no means is it the only form of contract that we observe, or even (depending on the region of observation) the dominant form. A second type of contract is commonly referred to as sharecropping. Sharecropping comes in many flavors, but all of them involve the sharing of the tenant's output in some preassigned proportion between the landlord and the tenant. The proportions vary from country to country and across regions within a country, although a 50-50 division is commonly observed. Variations on the sharecropping contract include different proportions of division of the output depending on whether input costs are also shared between the landlord and the tenant, and tied credit arrangements.

There is a simple but useful way to write down a class of rental contracts that contains fixed rent and sharecropping contracts as special cases. If Y denotes agricultural output on the rented land, then write the total rent as
(12.1) $\qquad\qquad\qquad R = aY + F.$
If $a = 0$ and $F > 0$, this is a fixed-rent contract with rent F. If $F = 0$ and a lies between 0 and 1, then this is a sharecropping contract, where the share to the landlord is a and the share to the tenant is $1 — a$. Finally, if $a = 0$ and $F < 0$, this can be interpreted as a "pure wage contract," where the wage is simply $w = —F$: the tenant is not a tenant at all, but a labourer on the landlord's land.

2.5.4.2 Contracts and incentives
There is a long tradition in economics that argues that sharecropping is essentially an inferior system to that of fixed-rent tenancy. A fixed-rent contract has the property that the

tenant pays a fixed sum to the landlord no matter how much output is produced. Another way of saying the same thing is that the tenant retains 100% of any extra output that is produced. In contrast, sharecropping effectively leaves the tenant with some fraction of any additional output—a percentage such as 50 or 60%, depending on the exact form of the contract. Thus, if the effort of the tenant cannot be monitored and controlled by the landlord, the tenant has an incentive to undersupply his effort, because, under the sharecropping contract, part of the output produced by him gets siphoned off to the land-lord. It would be better, instead, to extract this rent up front by charging a fixed payment and then leave the tenant alone.

To summarize then, the use of contracts other than a fixed-rent contract leads to a distortion of the tenant's input supply away from the efficient level: In particular, it appears that (i) sharecropping leads to undersupply of the tenant's inputs' and (ii) a rational landowner trying to maximize the earnings from land lease will always prefer a suitable fixed-rent contract to any share contract.

It is this last point that gives rise to the famous Marshallian puzzle of sharecropping. If a fixed-rent system is demonstrably superior to a sharecrop-ping arrangement, not- only from a social efficiency angle, but also from the point of view of the landlord's individual rationality, then why does share-cropping enjoy such enduring popularity in real world practice? Surely, there is more to the story than currently has been outlined.

At one level, the persistence of sharecropping seems merely a puzzle to be explained. Nothing of great importance appears to hang on it, only that contracts are in reality more diverse than theory predicts. However, this view is wrong for several reasons.

First, if we do observe sharecropping where theory tells us there should be none, then there is something wrong with the theory. At any rate, the theory needs to be augmented by a fuller description of reality. This enrichment may assist us in understanding other situations where the theory of incentives is also important. Second, and at a more practical level, if share-cropping exists despite the production losses that it appears to generate, it suggests that there are other compensating factors that necessitate such an arrangement. If these factors can be corrected by appropriate policy, the re-suiting inefficiencies will decline with the resulting decline in sharecropping. Third, these contractual relationships may have implications for other kinds of landlord and tenant behavior, such as the provision of credit to the tenant, the tendency to evict tenants, and the incentives to make long-run improvements on the land.

2.5.5 Risk, tenancy, and sharecropping
An individual is risk-averse if he prefers a certain (i.e., known or deterministic) sum of

money A to a lottery with the same expected, value A. Thus the very fact that a given amount is the expected value of some uncertain lottery is intrinsically displeasing to a risk-averse person. This is not to say that risk-averse persons cannot be compensated for taking risk. They can, but the greater the risk aversion, the greater will have to be the compensation (over and above the expected value of the lottery).
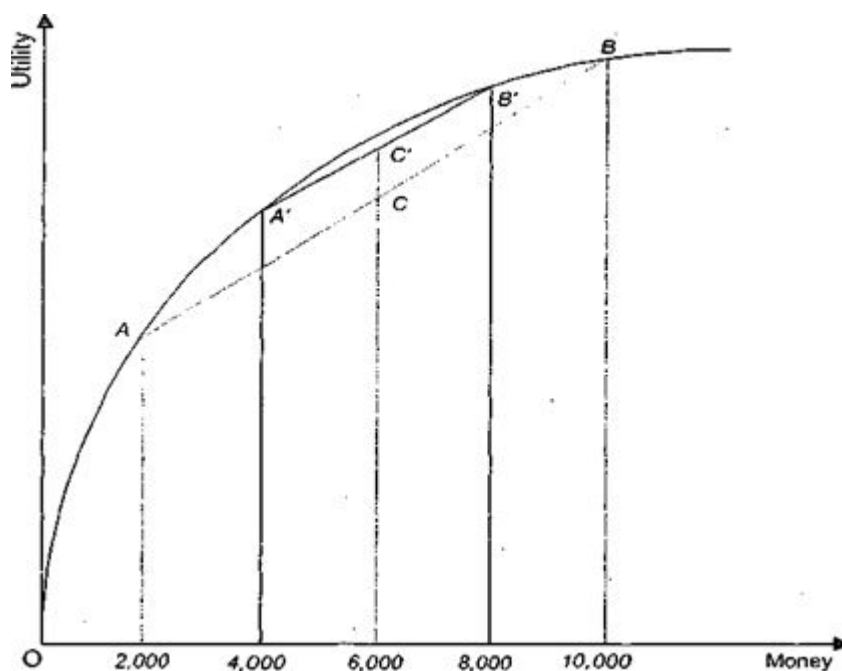

Figure 2.5.1.Comparison of two risky projects.

We will apply this idea directly to land contracts. To do so, we will abstract entirely from theproduction function approach and assume that somefixed amount of labour and other inputs is being applied to the tenanted land.

Even if all the inputs that a farmer can reasonably control are properly applied, the size of the harvest is still heavily dependent on Nature and will vary. For simplicity, suppose that only two levels of output are possible. We denote the values by G (for "good") and B (for "bad"). Suppose the probability that the good output will be produced is given by p.

In what follows, we consider a situation where land is leased out by a relatively wealthy landlord to a relatively poor tenant (much, if not most, of land lease is of this form). Thus it is reasonable to assume that as far as the return from this one particular plot of land is concerned, the landlord is less risk-averse than his tenant. If the landlord

is wealthy, this income is just one of many sources of income, so he does not attach the same importance to it. In addition, the other sources of income may be uncorrelated with the present source, which permits diversification for the landlord. We in fact simplify by assuming that the landlord is risk-neutral, whereas the tenant is risk-averse.

First consider a fixed-rent contract in which the tenant is required to pay a rent of R to the landlord, irrespective of the fortunes of the plot. In this case, the tenant receives G — R' if things go well and B — R otherwise. The landlord receives a sure payment of R.

Now imagine replacing this contract with a sharecropping contract, where the share is purposely chosen to provide exactly the same expected return to the landlord as before. (The reason for doing this will soon become clear.) Thus, if s is the share of the crop accruing to the landlord, then the expected return to the landlord is

$$psG + (1 - p)sB.$$

If this is equated to the landlord's return from fixed-rent tenancy, R, we see that the share of the landlord under the new contract must be given by the equation

$$pG + (1-p)B$$

Now observe, of course, that the expected return to the tenant is also the same under the two contractual modes. However, the tenant is not risk-neutral, so we still need to figure out which of the two contracts he prefers. Indeed, we may think of these two situations as analogous to the two projects of Nazim :—and draw upon the analogy. To see this, first compare tenant returns in the good state under the two contracts. With fixed rent, it is G — jR, whereas with sharecropping, it is (1 — s)G.

$$(1-s)G-(G-K)=R-sG=R_{-pG+}{}^{\wedge}{}_{\_p)B}<0,$$

by the fact that G > B. We may therefore conclude that the sharecropping contract loiuers the return to the tenant in the good state. Because the share has been chosen so that expected monetary values are the same to either party under the two contracts, we may conclude that the sharecropping contract increases the return to the tenant under the bad state (relative to fixed rent). Now the analogy with Nazim's example should become clear: share-cropping and fixed-rent tenancy are like two projects with the same expected value, but the "spread" of returns to the tenant is narrower under sharecropping.

It follows that if the tenant is risk-averse, he should prefer the share-cropping contract over the fixed-rent contract. The landlord can play on this preference by cutting the tenant's share a bit more, but not too much, so that the tenant still prefers the sharecropping contract. Now the landlord, who is risk-neutral, enjoys a larger expected

payoff, so he should switch from fixed rent to sharecropping.

Thus sharecropping emerges as a way to share, not just the output of productive activity, but the risk that is associated with it as well. A tenant who pays fixed rent is forced to bear the entire uncertainty of production. He may be more than happy to yield some of this uncertainty to his landlord. Sharecropping accomplishes this by essentially varying the rent payable with the size of the harvest.

2.5.6 Forms of Tenancy: Other considerations
So far we have studied the two major ingredients that affect the structure of land contracts: incentives and risk. There are several related considerations that are important as well. We discuss some of them in this section.

The double-incentive problem
Is leased land farmed only by the tenant, his or her family, and the labourershired by them? It depends. If land is leased out by a small landowner to alarge tenant or by an absentee landlord who is only interested in maintaininga-secure source of rental income, the landlord usually will not be involved with the leased land in an ongoing way. Typically such leases are carried -out on the basis of fixed rent, because minimal activity on the landlord's part (such as verification of tenant output) is required. Indeed, the landlord would not care whether the land is even cultivated or not, as long as the rent is paid.

On the other hand, there are situations in which the landlord is deeply involved with the crop grown on the land, the methods used for cultivation, the inputs used, and the proper maintenance and care of the leased plot. The landlord may be in a position to make suggestions, to provide managerial care, and supply inputs of production. Some of these inputs may be non-contractible, just as the tenant's labour is noncontractible because it cannot be observed or verified by the landlord.

Return to the Marshallian inefficiency of sharecropping. Ignore risk or assume that everybody is risk-neutral to display this incentive problem in its starkest form. Recall that the problem arises from the observation that if the tenant does not get to keep the entire marginal output resulting from production, he will have an incentive to undersupply effort. However, what if both landlord and tenant are required to apply effort, as in the discussion of the previous paragraph? Think of this as the double incentive problem.

If the tenant gets to keep the entire marginal output from the land, the landlord keeps none of it: (This is fixed-rent tenancy.) Of course, the tenant will then work very hard,

but the landlord will have no incentive to put in effort on the leased land. Now suppose the landlord gets to keep the entire marginal output from the land, but the tenant keeps none of it. (This is the case of wage labour where the landlord is really an employer and the tenant is really an employee.) In this case, the landlord will have all the incentive to put in effort and the tenant-labourer will have none.

Cost sharing of inputs

Sharecropping may be the preferred contract when input costs are shared between landlord and tenant. To see this, reconsider the Marshallian argument. Because tenant labour is non contractible, its marginal cost is borne entirely by the tenant. However, sharecropping means that some fraction (say half) of the marginal output is taken by the landlord. Thus the tenant, instead of equating the marginal product of this labour to its marginal cost, effectively equates half the marginal product of his labour to its marginal cost. This means that he stops applying labour at a point when marginal product still exceeds marginal cost, so the resulting outcome is inefficient.

Limited liability and sharecropping

If a tenant is poor and his output is uncertain, then quite, apart from considerations of risk aversion, there may be states of the world in which the tenant will not be able to pay a fixed rent. This constraint, stemming from the tenant's small wealth and the small output that he might produce, is known as limited liability.

Landlords who charge fixed rent will therefore know that such rent cannot always be paid. If the tenant is poor and the harvest fails, the rent will have to be forgiven or essentially advanced as a loan. However, there is no guarantee that the loan will be repaid in the future, so part or all of the rent may truly have to be forgiven.

The problem with this arrangement is that it creates an incentive for the tenant to overinvest in risky methods of production. This is because if production fails, rent is forgiven, whereas if it succeeds, the tenant gets to retain all the excess (under fixed-rent tenancy). One way to counterbalance this tendency is for the landlord to lower the rent in bad states and raise it in good states. This gives a tenant a stake in thebad outcome as well and reduces his tendency to overinvest in risky forms of farming.

Screening

A sharecropping contract offered along with other kinds of contracts may be a suitable screening device to obtain high-quality tenants. Suppose that a landlord is uncertain about the true ability and productivity of a prospective tenant, although the tenant himself knows precisely what his abilities are. In such a situation, it is sometimes possible to separate the two kinds of tenants by offering a menu of contracts.

The idea behind the argument is that high-ability tenants will prefer contracts in which they can retain a larger share of their (high) marginal product, whereas low-ability tenants would like to divide their (low) marginal product between the landlord and themselves. Now landlords would like to ferret out the high-ability tenants in a world where tenant abilities are largely unknown. By doing so, the landlord can use their higher productivity to extract more rent. Note well that all the implicit extra surplus cannot be extracted as increased rent, because in that case no high-ability tenant will reveal himself through an appropriate choice of contract. This is where a cunningly chosen menu of contracts can make a difference.

Specifically, suppose that the landlord can ask the tenant to choose between two contracts: one in which a share of the output is offered and another in which fixed rent must be paid. It is possible, under some circumstances, to choose this menu so that the following conditions are met.
(1) A high-ability tenant will prefer the fixed-rent contract to the share-cropping contract, even though the implicit rent in the fixed-rent contract is higher. The reason is that he gets to keep his high marginal product entirely. In the sharecropping contract, he must give some of this away.
(2) A low-ability tenant will prefer the sharecropping contract to the fixed-rent contract. The fixed rent is too high relative to the extra marginal product that he would get to retain.

If these two conditions are met, then the two types of tenants will "separate" by choosing different contracts. Under this view, sharecropping is a sifting device to leave fixed-rent contracts in the hands of the most productive tenants. This screening theory also offers an explanation for the coexistence of sharecropping with other contractual forms.

2.5.7 Land contracts, eviction, and use rights
So far, we have considered the problems of incentives and risk sharing in a static context. We have neglected the possibility that a tenancy contract may or may not be renewed. When nonrenewal or eviction is a possibility, the landlord acquires an additional instrument through which effort incentives are provided.

It is easy enough to see how such an option might work. Suppose that a tenant is offered a sharecropping contract for one or more of the reasons that we have already discussed. To some extent, the problem of Marshallian inefficiency might be overcome by threatening to evict the tenant in the case of poor performance.

The possibility of eviction has several consequences. For one thing, it introduces a new

form of risk for the tenant, and he will have to be compensated for this risk; otherwise he will not accept such a contract. For another, the going value of the contract must exceed the tenant's next-best opportunity; otherwise the threat of eviction has no bite. The landlord will have to examine closely whether he is willing to pay the additional compensation required to acquire the instrument of eviction.

In addition, there are other potential sources of loss for the landlord. Chief among these are activities that increase the long-run earning potential of the land, which the tenant will now be less willing to carry out. But there are gains as well. For instance, the threat of eviction can be used to ensure better performance in production, as well as to enforce loan repayment

There are two particular cases where contracts with eviction may be widely employed.
(1) Limited liability. Consider the limited liability contracts studied in the previous section. We have already shown that because of the need to provide appropriate incentives, such limited liability contracts often provide the tenant with a return that exceeds his opportunity cost. In such situations, the threat of removing the contract tomorrow if output is not at a satisfactory level today carries bite. Faced with such a threat, the tenant will indeed raise effort even further. Recall that poor tenants are more likely to be in a situation where the limited liability constraint is binding. It follows that the threat of eviction can be more profitably used against poor tenants.
(2) (Non veriftable) Information regarding tenant effort. Suppose that the landlord can obtain some information regarding tenant efforts, but cannot write such observations into a contract because they are not verifiable in a court of law.

On the other hand, such non-verifiable information can be taken into account in the decision to renew a contract. If the tenant is no worse, intrinsically than other tenants, then the landlord will have no perverse incentive to take undue advantage of this (informal) understanding. At the same time, the tenant will be aware that it is certainly credible for the landlord to not renew the contract if he, the tenant, shirks. Thus in this case, the decision to no trenew a contract can be based on information in ways that an ordinary contract cannot mimic. Such information is potentially valuable for contractual efficiency.

The effect of potential eviction on the welfare of the tenant needs to be examined with caution. The first sentiment regarding eviction is the usual gut reaction: if tenants can be evicted, they must be worse off. Like many gut reactions, this is probably correct, but there is a paradoxical counterargument that needs to be addressed first.

The argument is this: eviction can only provide additional incentives if the tenant is

strictly better off working with the current landlord than being relieved of his current contract. This is possible if the contract package is more attractive than one in which the tenant is indifferent between working for the current landlord and seeking out alternative opportunities. Put another way, if the landlord does decide to use a contract that employs eviction as a background threat, a new tenant has to be made better off relative to his outside opportunities. If eviction is banned and there is an excess supply of potential tenants, a new tenant will not be given any more than his next best alternative. Thus it is not surprising to observe that, despite the vicissitudes and uncertainties of tenancy, tenancy is still preferable to landless labour.

2.5.8 Land ownership
The path to specific property rights has seldom been smooth. Rights to land have historically been subject to challenge, largely backed by force. It was only natural that a class of overlords or rulers would emerge, who would exact tribute or rent from cultivators in return for patronage and protection. Such payments were subsequently enshrined in tradition, in social custom, in religious norms, and last (but not least) in the legal dictates of the stale. As populations swelled the world over, these norms, regulations, and traditions were supplanted and reinforced by the power of the market. Land rents and prices rose as land became scarce relative to labour. With a cheap supply of labour, the laws that supported slavery gradually were dispensed with.

2.5.9 Land size and productivity: Concepts
Consider two notions. One is total factor productivity: do small farms have a production function that lies "beyond" that of large farms? This is a hard concept to get at for two reasons. First, small and large farms do not typically use the same inputs. Thus, we need to compare the inputs in some way, presumably by multiplying by market prices, to get an overall aggregate. Second, there is the problem of valuing nonmonetized inputs such as family labour. In the following text we will have no theoretical reason to believe that small farms are more efficient in this narrow technological sense. We are after a broader notion of productivity anyway.

The second notion of productivity is "productivity in the sense of market efficiency." This is a vaguer notion, and we can roughly translate it by asking, Do small farms produce an output per acre that is closer to the "efficient market" output than large farms? However, what do we mean by market efficiency? In a world where several markets are inefficient or nonfunctional, it is unclear whether the standard rules of efficiency in one market promote economic efficiency overall.[17] Very tentatively, we may say that production efficiency is achieved when the values of the marginal product of all inputs equal their true marginal costs. This is the viewpoint we will use to make the subsequent theoretical arguments.

This second, subtle notion is often tested very bluntly: simply ask if output per acre (perhaps correcting for land quality) on small plots exceeds those of large plots. This presumes that whenever the preceding efficiency conditions are violated, they are violated in the direction of under application of inputs. The arguments we consider in the upcoming text suggest that this may be the case, but it helps to keep the conceptual distinctions of this subsection in mind throughout.

Consider, first, the technological angle. Obviously, there are minimum sizes below which land cannot be usefully cultivated, at least for certain crops. Moreover, large plots are suitable for mechanization and cultivation with capital-intensive methods, in a way that small plots are not. .Thus, from the pure technological point of view, it is reasonable to suppose that land either exhibits constant unit productivity (with all other inputs being expanded in proportion) beyond a certain minimum scale or displays increasing productivity once large-scale techniques of cultivation can be brought to bear.

Pooling land
Small farms must always win over large farms because of the incentive argument and because small, landowners, if they so wish, can pool their land to take advantage of technological returns to scale.
This argument is problematic, however. The pooling of land recreates the incentive problem to some extent, although how much of it can be avoided depends on the precise source of returns to scale. For instance, suppose that the source of increasing returns is large-scale marketing. Then it is likely that the incentive problems can be dealt with through reform. The land can still be farmed separately and the fixed costs of setting up a marketing group can be pooled. It is not at all surprising that cooperatives whose main incentive to form was marketing have been relatively successful (the dairy cooperatives in India are a particularly visible example).

On the other hand, suppose that the source of increasing returns lies at the production level (say, through the use of tractors). Then the incentive problem returns with full force: as long as production is joint, individual owners will have a tendency to slack off. Indeed, the production activity must be joint if returns to scale are intrinsically embodied in the cultivation process.

A simple story that explains the incentive problem with joint farming has the same flavor as several disparate models that we have studied before. Joint production creates an externality. Additional effort by any one farmer leads to additional output, but the added output is shared among the whole team. That part of the marginal output that accrues to other members of the cooperative is an externality as far as our farmer is concerned, and

unless he internalizes this through a complete sense of altruism, he will undersupply effort. To be sure, everybody else will do the same as well and productivity will be reduced below the efficient level. This is a free-rider problem characteristic of situations like the Prisoners' Dilemma.

Technology is at best neutral toward land size and may even favor large land sizes because of the possibility of extensive mechanization. On the other hand, with imperfect insurance markets, incentives favor the cultivation of Jand by owners and family labour: hired labourers and tenants typically must be given contracts that create an efficiency loss. Imperfect labour markets with unemployment heighten this proclivity by reducing the opportunity cost for family labour relative to that of hired labour. Finally, the technological gains of large farms cannot be achieved easily by small farms simply by pooling land, because that creates a free-rider problem. Hence there genuinely are two opposing sets of forces, and which effect dominates is ultimately an empirical question.

2.5.10 Land sales
When credit markets are imperfect, land value consists of two components. The first is the discounted sum of income streams that will emanate from working the land. The second component comes from imperfect credit markets: land can be used as collateral, and this ability has value (measured by the profitability of the additional loans that can be obtained as a result of mortgaging the land). Note that this second value would be zero if competitive credit markets prevailed and one could obtain all the loans one wanted at some going rate of interest. Under normal conditions, then, a seller will therefore want to sell the land for a price that is no less than the sum of these two values. Now consider what a buyer is willing to pay. If he can buy the land outright from his own funds, then he reaps both these gains as well, but if he must obtain a loan to buy the land and if he must mortgage that very piece of land for the loan, then he cannot reap the collateral value until the loan is paid off. Thus the collateral advantages of a land purchase are pushed back to a distant future when the loan has been repaid, so the buyer's present valuation of the land must be less than that of the seller. Thus no sale will occur;

But the preceding argument has a loophole. It does not take into account the productivity gain that we discussed in the previous section. If small farmers are truly more efficient than large farmers because of the incentive problem, then the discounted sum of income streams (per acre) will be higher for small farmers than for large farmers. It remains to be seen whether the difference in this income stream outweighs the collateral value. If it does, then the absence of a land market cannot be explained with this argument.

2.5.11 Land reform
Putting together all that we have discussed so far, it seems that (i) productivity is higher

on smaller plots than on larger plots, (ii) that these productivity gains cannot be realized by tenancy, because the tenancy contract itself erodes the productivity gain, and (iii) that land sales markets cannot adequately substitute for land tenancy markets.

To realize these gains, we are then left with the option of land transfers from rich to poor, either without adequate compensation or with full compensation paid by the government or by foreign donors, but not entirely by the beneficiaries. This last phrase is important. If full compensation is paid by the beneficiaries, then this is no different from a land sale, which, as we've already seen, would not have occurred spontaneously.

Although such reforms undoubtedly go in the right direction and can have significant effects on productivity, they fall short of reaping the full gains in an environment of imperfect credit markets. Without ownership rights, land cannot be pledged as collateral for productivity-enhancing investments. Seized land also may be redistributed as collectives, as in Mexico or Peru, but we have already remarked on the possible incentive problems associated with collective farming. Finally, land ceilings may be used to curtail ownership of large plots. Even though such ceilings are in force in many developing countries, they are easily sidestepped by holding land in several parcels under the names of friends and relatives.

2.5.12 Summary

In this chapter, we studied land markets. We began with a study of land rental contracts. A landowner leases out his land to a tenant for cultivation and charges rent. The contract may be in the form of a fixed-rent tenancy, in which a constant sum is paid by the tenant to the landlord regardless of the output that he produces, or a sharecropping tenancy, in which a share of the output is relinquished by the tenant to the landlord as rent. To be sure, these are extreme forms, and we observed several variations, such as the provision of credit as part of the tenancy contract or the cost sharing of production inputs between the landlord and the tenant.

Fixed rent allows the tenant to retain the full .marginal product of his efforts and, therefore, does not distort the tenant's choice of inputs. In contrast, sharecropping lowers the marginal product of effort, or at least that part of it that accrues to the tenant. Thus sharecropping should be associated with lower land productivity.There are other roles for sharecropping as well. We viewed eviction as another instrument that the landlord might use to provide incentives and we discussed situations in which eviction clauses may be implicit or explicit in tenancy arrangements. Of particular interest in this context are situations in which the tenant is granted use rights to the land (while he does not have ownership as in a full land reform, the landlord cannot evict him). We discussed how this form of partial land reform can affect land productivity.

2.5.13 Short answer type questions
Write short notes on
1. Rent
2. Sharecropping
3. Land reforms
4. Land contracts

2.5.14 Long answer type questions
1. Explain land contracts and risk involved in these contracts.
2. Which one is better fixed rent or sharecropping?

2.5.15 Recommended books
1.   Development Economics by Debraj Ray
2.   Development Economics by R. Barro
3.   Development Economics by M. P. Todaro

Lesson 2.6   **Author: Dr. Manisha**

Labour market in agriculture

2.6.1 Introduction

The land markets may not be able to adequately bridge the gap between ownership and operations. A flourishing labour market arises, when small owners or landless individuals supply labour to large landowners who need more than just family labour to adequately carry out cultivation.

We also noted that in some situations, land and labour markets may be complements instead of substitutes. This is especially so with agricultural activities that are mechanized and must take place on a large scale. In this situation both land and labour are employed by those individuals with access to capital.At the same time, labour markets do not work perfectly either. Many of the problems raised in the preceding chapter find corresponding expression in labour markets. Thus, even in situations in which land and labour markets are substitutes and not complements, land markets will have some role to play.

2.6.2 Objectives of the lesson

The task of this chapter, then, is to understand the market(s) for labour, the various forms that such markets assume, and the implications for the development process. Our goal in this chapter, then, is to study how the casual and long-term labour markets function and to bring out some key differences between the two kinds of markets.

2.6.3 Labour categories

Even in a geographically contiguous region, there is no single, homogeneouslabour market. We can very broadly distinguish between two types of hiredlabour: (1) labourers that are hired on a casual basis, perhaps on some dailyarrangement or for some prespecified short duration (such as the harvestingperiod) and (2) labourers that are under some (implicit or explicit) long-termcontract with their employer. As we will see, the distinction between thesetwo categories is important, because their markets work quite differently. Casual labour is normally hired to carry out tasks that are easily amenableto observation. Harvesting and weeding fall into this category. The tasks of long-term labour are somewhat more mixed. On large plots of land, a long-term employee may serve in a supervisory capacity, along with the owners of the farm. They might be responsible for tasks that require special care and are relatively difficult to monitor, such as the application of fertilizer and pesticides or the application of water. In addition, long-term employees might work at "standard" tasks along with their casual counterparts, participating in the harvesting process, for instance.

Common sense tells us that this division of tasks is to be expected. In a long-term relationship, an employee can be held accountable for errors or deliberate mismanagement that are only known after the passage of sometime (such as the wrong application Of fertilizer or pesticide). With casual employees, even those resident in the same village, this may be far more difficult. It is interesting to think about just what the source of the difficulty is. It isn't that the culprit cannot be identified just because he was a casual employee (although with large numbers of employees this might happen as well). It is that the scope for "punishing" the casual employee is much narrower. With a long-term employee, future employment may be denied or the terms of employment may be modified.

This sort of argument leads to a puzzle. Standard supply and demand models of the labour market tell us that the labour market will "clear" at a wage that mirrors accurately the opportunity cost of the worker's time: if denied employment, the worker can find employment elsewhere at the same wage, or even if the worker is unemployed, the utility of the additional leisure just compensates for the loss in wages (see following text for more details on this). In that case, the employer has no additional power over a long-term employee, because the denial of employment has no adverse consequences. This suggests that the standard model may be inappropriate for thinking about long-term relationships. Long-run contracts must involve payments that exceed the alternative expected returns from defaulting on the contract.

2.6 .4 A familiar model

It will be useful to begin with the standard supply-demand story, and then amend this for

the particular features of rural markets as we go along. Figure 2.6.1 summarizes the usual paradigm: According to this account, the demand for labour depends, among other things, on the "going" wage (captured by w) that is paid to hired employees. It stands to reason that if the going wage falls, the demand for labour should be further stimulated (or at least not lowered), so that the resulting demand curve is downward sloping. On the other hand, the supply curve of labour is derived by a calculation of the costs and benefits of working. A higher going wage serves as better compensation for the use of labour, so this should elicit a greater supply of labour from each worker, as well as encourage a larger number of workers to enter the labour market. For both these reasons, it makes sense to assume that the supply curve of labour is upward sloping. As you know, the intersection of the supply and demand curves then gives us the equilibrium wage.
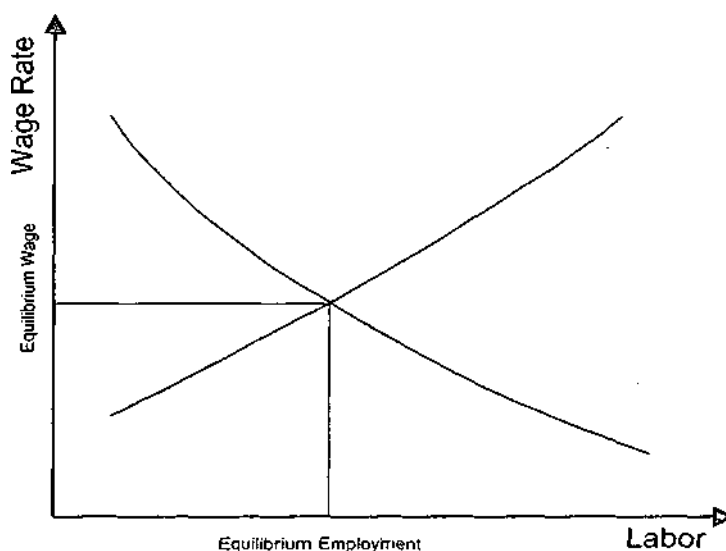

Figure 2.6.1. Supply of and demand for labour.

There are several disturbing features in this seemingly harmless description of labour markets. Observe, first of all, that this model does not (and in its present form, cannot) make a distinction between casual and long-term labour. It is as if different periods of time are neatly separated: what happens in tomorrow's labour market has no bearing in what occurs today and vice versa. Later, we will see why this is a bad assumption to make in some contexts.

Second, the model fails to make a useful distinction between labourpozoer and labourers. We already know from Chapter 8 that this distinction can matter. For instance, some individuals may be excluded from the labour market because their work capacity does not permit them to participate at an adequate level.

Third, each labourer in the equilibrium of this model will be perfectly indifferent between working for his current employer and entering the labour market to search for another employer. When tasks are difficult to supervise and shirking can only be punished by termination of employment, the state of affairs described in Figure 13.1 may not persist. Put another way, the standard story assumes that all work is perfectly monitor able.

Fourth, and this follows on the preceding point, an equilibrium of the standard model in which some workers don't find jobs must leave every worker indifferent between working and not working at all. There is no such thing as involuntary unemployment in the model. Failure to incorporate this phenomenon leaves out a central aspect of reality.

Fifth, rural labour markets are characterized by substantial uncertainty and/or seasonality in agricultural production. At one level, this can easily be captured by the standard story. For instance, suppose that rainfall levels are uncertain and that this will affect the size of the harvest. In that case, total demand for harvesting labour will be affected as well. Collapsing this train of events, we may think of the labour demand curve as itself being uncertain; it fluctuates between the highs and lows described by the dotted lines in Figure 2.6.2. Of course, the corresponding equilibrium wages fluctuate as well. In Figure 2.6.2, the wage fluctuations lie between a band of w_H and iu_L.

Now, although this account certainly captures some of aspects of uncertainty, it fails to tell the whole story. Most important, it does not illustrate the possible ways in which employees and employers cope with this uncertainty "before the fact," by writing contracts or making informal agreements that insure one party or the other from the effects of rainfall variation (in this example).

In the same way, workers may wish to smooth out seasonal fluctuations in their wage income, and employers who are willing to provide such income smoothing may be preferred by employees. The standard model is too simplistic to take these features into account.
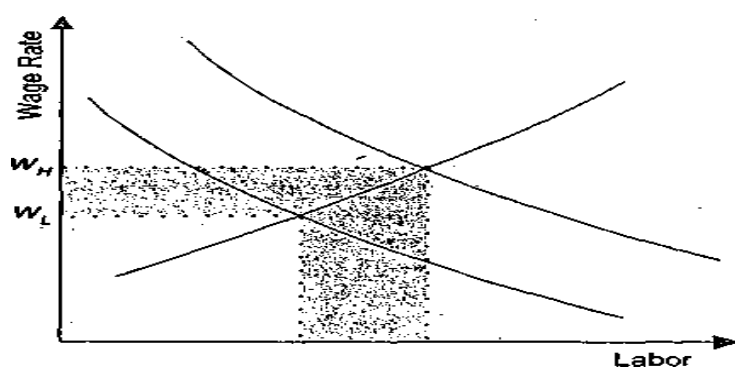

Figure2.6.2. Labour market equilibrium under uncertainty.

Our goal in this chapter is to explore these variations on the standard model, in the context of rural labour markets in developing countries. We begin by taking up a theme initiated in Chapter 8: the connection between nutrition and labour markets. Initially we do not concentrate on any particular type of labour market. Subsequently, we examine the case of casual labour-markets more carefully.

2.6.5 Poverty, nutrition, and labour markets

The basic model

The capacity curve

In its simplest form, it introduces a relationship between nutrition and work capacity, which we call the capacity curve. Figure 2.6.3 displays a typical capacity curve.

Here is a quick summary of the discussion in Chapter 8. Because the horizontal axis of Figure 2.6.3 is labeled "income," our implicit assumption is that all income is converted to nutrition. Nothing of substance is lost by amending this to a more realistic situation where, say, 70% of income is spent on nutrition. The vertical axis is labeled "work capacity": think of it as a measure of the total number of tasks an individual can perform during a given period, say, the number of bushels of wheat that he can harvest during a day. The capacity curve is found by linking different nutrition (or income) points to the corresponding levels of work capacity that are generated by the individual.
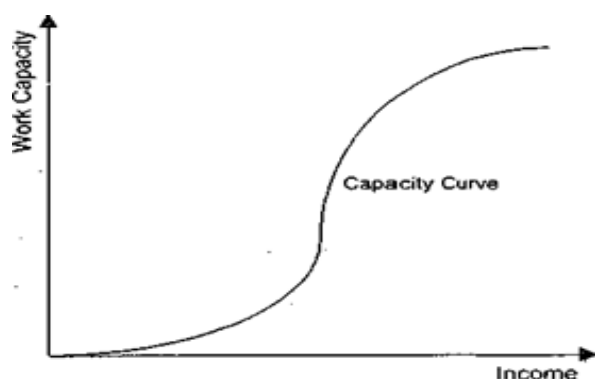
Figure 2.6.3The capacity curve

To understand the shape of the capacity curve, recall that most nutrition initially goes into maintaining the body's resting metabolism. In this stretch very little extra energy is left over for work, so work capacity in this region is close to zero and does not increase too quickly as nutrition levels change. Once resting metabolism is taken care of, however, there is a marked increase in work capacity with nutrition. Finally, there is a phase of diminishing returns, as natural bodily limits restrict the conversion of increasing nutrition into ever-increasing work capacity.

Piece rates

Thus income determines work capacity, but work capacity determines income as well. Let's try to capture this in the simplest possible way. Imagine that incomes are generated by working in a labour market where piece rates are paid. This is payment on the basis of tasks completed, such as 10 rupees per harvested bushel or 100 pesos per acre weeded.

Piece rates have an easy pictorial representation along the same axes used for Figure 2.6.3. If income is paid per unit of task—say 10 rupees per bushel harvested—then we see that there is a relationship between the number of tasks that are performed (bushels harvested) and total income. Figure 2.6.4 draws the obvious graph that portrays this relationship. A piece rate, then, appears as a relationship between the number of tasks performed and the total income -of a person.

Now we superimpose Figure 2.6.3 on Figure 2.6.4. The result is shown in Figure 2.6.5. Four piece rates are shown. I have called them v_u, v_2, v_3, and $z>_4$. Thus v, means something like "you get paid $v_x$pah for every bushel you harvest." If you have done the exercise above, you will immediately see that Vi is larger than v_2, which in turn exceeds v_3, which is larger than v_4.

Labour supply

Now suppose that a labourer (call him Mihir) tries to obtain the highest possible level of income that he can possibly earn, given the constraints imposed by his capacity curve.

Suppose, first, that the going piece rate'is w-,.Mihir will clearly choose the point A, which yields the largest possible feasible income for him. As the piece rate drops to v_z, this maximum income falls. On the graph, Mihir now slides down to the point B, which involves less total work and lower income.

At v_3, something interesting happens. This is a piece rate that has the property that it is just tangent to the capacity curve along its hump. At this
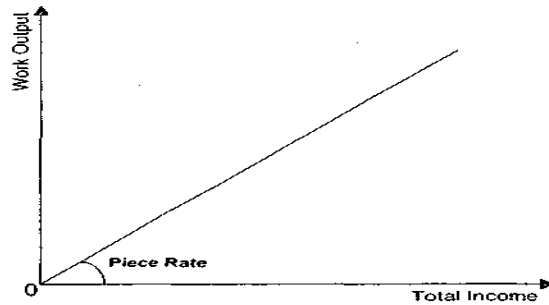

Figure 2.6.4.A piece rnte.

piece rate Mihir can just about choose the point C. If the piece rate drops a little more, then the amount of work that Mihir can supply drops dramatically, jumping, as it were, from a point like C to a point like D (which is the intersection of the lowest piece rate with the capacity curve). This jump occurs precisely because the capacity curve is shaped the way it is shaped,with low levels of nutrition permitting only very low levels of work, and
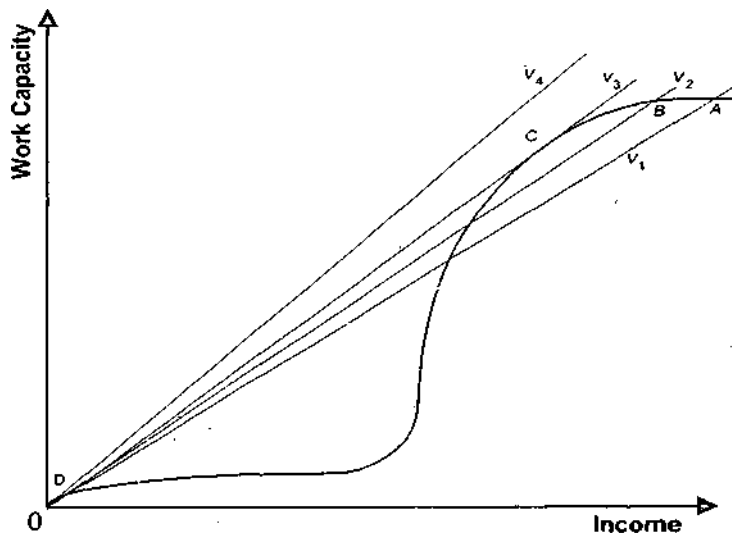

Figure 2.6.5.Piece rates and ivork effort.

.moderate to high levels creating a rapid increase in work capacity.

We can use all this information to generate a supply curve of labour, which tells us the

different levels of labour power supplied at different piece rates. All we have to do is put our information on an appropriate graph, after multiplying Mihir'slabour supply (at each piece rate) by the number of labourers like him in the economy. Figure 2.6.6 shows the transition from individual labour supply to aggregate labour supply.

The left-hand panel of Figure 2.6.6 shows Mihir'slabour supply, by simply transplanting the information gleaned from Figure 2.6.5. The gap in labour supply at the piece rate v_3 captures our previous discussion that after a certain threshold wage, Mihir'slabour supply must jump discontinuously. The right-hand panel effectively multiplies this individual supply curve by the number of labourers. Two things happen. First, the horizontal axis gets blown up by the number of labourers like Mihir, so you're supposed to think of these numbers as larger numbers than Mihir'slabour supply, for each piece rate. Second, I've "filled up" the gap with dots; the meaning of these dots will become clear presently.

Equilibrium
To-complete our description of this labour market, we introduce a demand curve for labour. This is a perfectly standard demand curve. In general, it is downward sloping to capture the fact that if labour is cheaper, employers will demand a larger quantity of it. Figure 2.6.7 shows the demand and supply curves on the same diagram. Two cases are of interest.
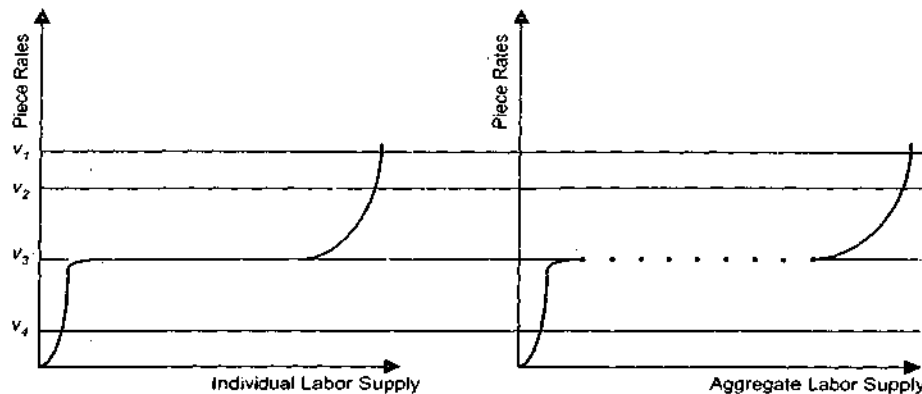In the first case, represented by the left-hand panel of Figure 2.6.7, the demand curve for

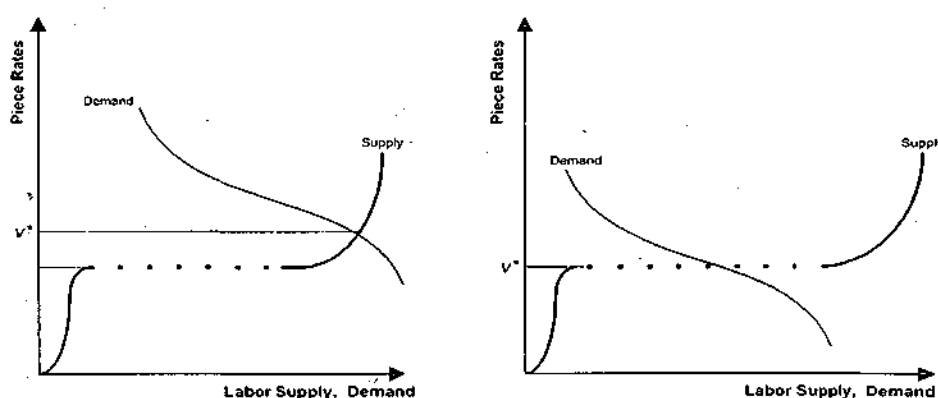Figure 2.6.6.Individual and aggregate labour supply.

labour cuts

Figure 2.6.7."Equilibrium" in the labour market.

the supply curve at a point that is beyond the gap in the supply curve. This case is perfectly norma): the market determines an equilibrium piece rate v* and everybody gets to supply a "high" level of work effort, that is, a level of work effort that is somewhere on the hump of the capacity curve for each labourer. The market clears in a standard fashion. This case obtains if demand is large relative to supply.

The interesting case is represented by the right-hand panel. Here supply is large relative to demand/so that the demand curve passes through the dotted gap in the aggregate supply curve. Now we have a problem with determining the equilibrium piece rate. If the rate is any larger than v*, we have excess supply, which brings the piece rate down. On the other hand, for piece rates below this critical level, there is excess demand, so that wages are bid up.

However, note that a piece rate of exactly v* can be thought of as an equilibrium, provided that we admit the idea of unemployment. Just as in Figure 14.6, we can "fill in" the gap in the aggregate supply curve by having some people work and restricting labour market access to others. This unemployment is involuntary in the sense that unemployed people are strictly worse off than their counterparts who are lucky to find employment. However, the piece rate cannot be bid down because no one can "credibly" supply the same amount of labour at any lower piece rate.

We see, then, that the vicious cycle is complete in this little model. Lack of labour market opportunities makes for low wages, but it is not only that wages determine work capacity: a low capacity to work feeds back on the situation by lowering access to labour markets!

Nonlabour assets and the labour market
If you have absorbed this basic model, we can take it a step further. What we are going to do in this section is introduce the realistic possibility that people may have other sources of income. This means that, strictly speaking, it is not correct to equate total income with

wages paid by the labour market. For instance, in rural settings, some individuals may have tiny landholdings that are leased out for cultivation. To the extent that such assets augment income possibilities, such individuals are more easily able to participate in the labour market.

This is expressed diagrammatically in Figure 14.8, which compares two individuals. The left-hand panel of this figure depicts another worker, Timir, who has access to a source of non labour income, of size R let us say (think of this as rent from his own landholding). Now work capacity depends on rent plus wages. We can easily draw this on the same kind of graph, but if the horizontal axis only involves wage income, this is done by "shifting" the capacity curve, as it were, horizontally to the left by the amount R. This is exactly what we have done in the left-hand of Figure 2.6.8.

The right-hand panel superimposes this diagram on the corresponding picture for Mihir, who has no sources of non labour income. Of course, this curve is what we have been studying all along. Note that although Mihir may be biologically just the same as Timir, his capacity curve lies to the right and below that of Timer, who enjoys some land rents. Two piece rates are drawn, v-$\pm$ and v_2- Note first that under v_uMihir is only able to supply a small amount of labour, for the reasons discussed in the previous section; he is effectively excluded from the labour market. Not so Timir, who can supply labour at v_x. Even if piece rates are so high that both can supply labour (as in the case of v_2), note that Timir is still earning a larger income than Mihir. It is. important to note that the larger size of Timir's income is not just because of his nonlabour assets: he earns
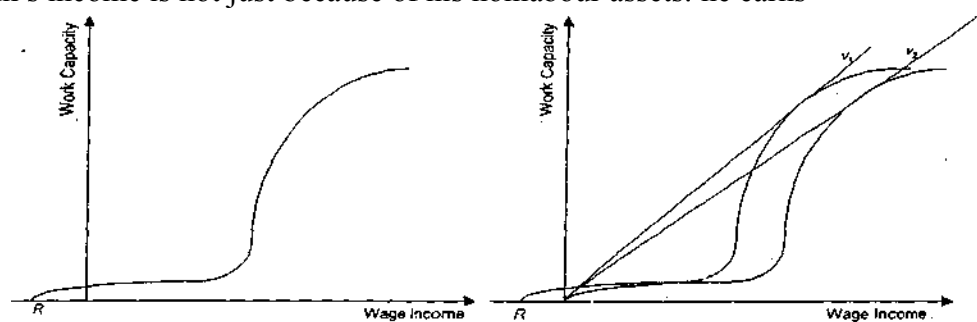

Figure 2.6.8.Hoivnonlabour assets affect labour income.

higher wage income. Thus inequalities in the asset market magnify further into labour market inequalities, at least among the poor.People without assets are doubly cursed. Not only do they not enjoy nonlabour income, they are at a disadvantage in the labour market relative to those who do possess assets.

2.6.6 Asset inequality, nonlabour income, and the labour market

Let us probe further the effects of disparate asset holdings. The easiest way to think about this is to imagine that there is only one commodity produced in the economy, which is food. Food is produced using land and labour power. Now suppose that there are many individuals in this simplified economy and that each person has access to the same capacity curve that we described in the preceding sections. The only other asset is land. It is natural to suppose that different individuals own different quantities of land; in particular, some individuals may be completely landless. This is a rarefied world, but it is enough to provide us with some important insights that survive generalization to more realistic situations.

Observe, to begin with, that the demand curve for labour (from the previous sections) is just the sum of the demands for labour that all households have to cultivate their land. To be sure, for some households (with plenty of land), their demand will exceed what they can supply on their own: these will be the employers of labour. Households with little or no land will be net suppliers of labour. Finally, those households that have approximately the land needed to absorb their own labour will be the family farms, who do not participate actively in the labour market. For the purpose of this exercise, we can think of everyone as participating in one giant labour market, where self-employment is treated as part of the overall market.

Now let us draw on an idea from the previous section. For each person with or without landholdings, let us keep track of the minimum -piece rate at which he will be able to supply labour to the labour market. The left-hand panel of Figure 2.6.9, which essentially captures ideas from the previous section, shows the obvious: people with greater amounts of nonlabour income (rental income from land, in this case) are able to supply their labour at a lower threshold piece rate, simply because their rental income takes care of some of their nutritional needs. The right-hand of the figure plots this minimum against people arranged in increasing order of land income. People up to the index i* are landless, so for them the minimum piece rate is unchanging. Thereafter, the minimum falls as land income increases.
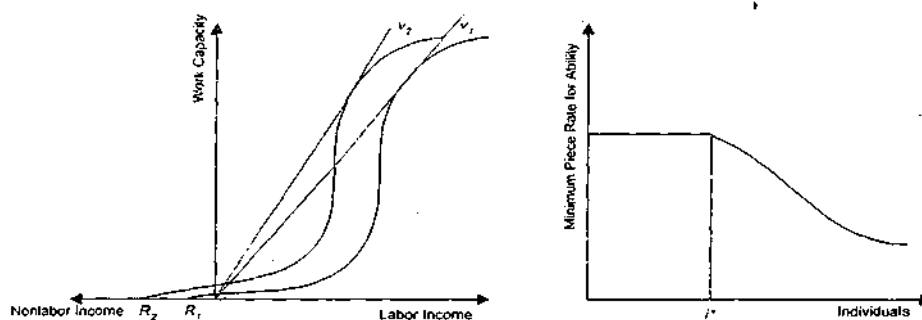
Figure 2.6.9.The minimum piece rate that determines ability to work.

This minimum piece rate represents the least amount for which an individual will be able to work on the labour market. However, there is an additional consideration. Presumably, the minimum wage at which a person will be willing to work rises with the amount of nonlabour income. This is because a person who has other sources of income, will value leisure more highly and will be willing to sacrifice it only for high enough compensation.

Thus two opposing forces are at work here, but we can say something reasonable about the way they interact. At very low levels of nonlabour income, people will be willing to work for anything, so that the consideration that really binds is the minimum piece rate at which they can work. As non-labour income increases, this "ability-based" minimum rate falls, and at some point the willingness to work becomes the binding constraint: ability is no longer an issue.

We may therefore combine the two minimum piece rates, shown as the heavy line in Figure 2.6.10. The resulting U-shaped curve represents the minimum piece rate at which individuals are willing and able to work. Indeed, we can figure out which parts of the curve correspond to which regime-Given that individuals are arrayed in order of increasing nonlabour income, the falling portion of the curve corresponds to the zone in which ability is the operative constraint. The rising part of the curve represents the zone in which willingness is the operative constraint.

Using Figure 2.6.10, the supply curve of labour can be derived in a very simple yet general way. Figure 2.6.11 shows how this works. For each piece rate in the market, the supply of labour is given by the amounts worked by all those whose minimum piece rates liebeloio the going market wage. These are the individuals who are willing and able to work at the going piece rate. By varying the piece rate, we trace out a supply curve. The demand curve
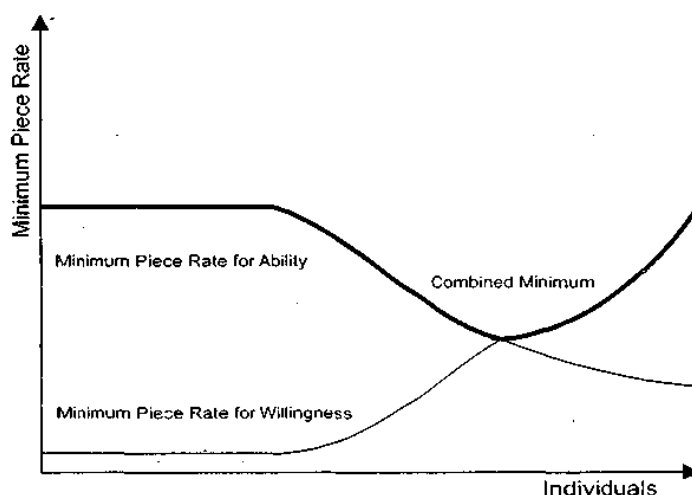
Figure 2.6.10. Ability and willingness: the combined effect.

is drawn just as before. The intersection of the two curves represents market equilibrium. The left-hand panel of Figure 2.6.11 displays one such piece rate and the segment of people who do supply labour. at that piece rate, shown by the line AB. (The exact quantity of labour supplied will depend on the capacity curve, the size of the nonlabour incomes, as well as the going piece rate itself, in a way that we discussed in detail in the earlier sections.) People to the "left" of A and to the "right" of B are unemployed, but take a closer look and you will see that they are unemployed for very different reasons. People to the right of B are able to work, but they do not wish to. We can call them the voluntarily unemployed or, in keeping with the rural flavor of this model, the landed gentry. Their nonlabour incomes, derived from land rent, are too high for them to be attracted by the going piece rate. Contrast this with
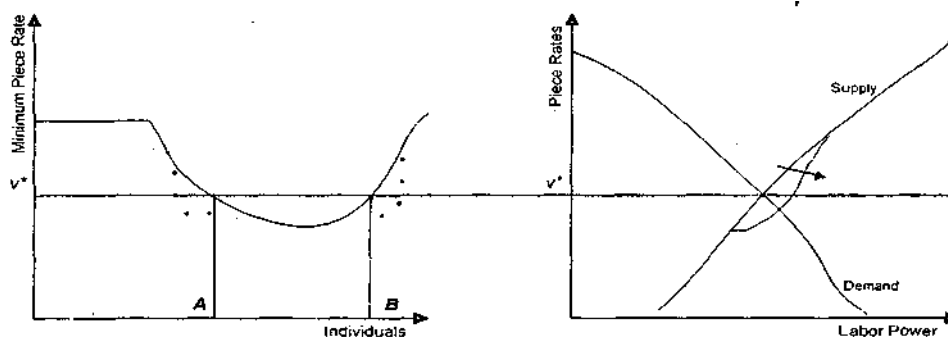

Figure 2.6.11.Market equilibrium.

the individuals to the left of A, who are unemployed not because they are unwilling to work, but because they are unable to work at the going piece rate (thieir resulting incomes are not high enough to reproduce the needed work capacity). We can call them involuntarily unemployed. Of course, you should not take this type of unemployment

completely literally. Perhaps such individuals can work in the short run. The point is that the going wage rate does not allow them to carry out sustained work over time without seriously impairing their health and physical strength. For more on these matters, see the next section.

We can use this model to analyze the effects of changes in the distribution of wealth holdings. In the case considered here, this is tantamount to looking at land reforms, in which land is taken away from those who have a lot of it and given to those who have none or little of it. Certainly, such land reforms hurt those who lose land and benefit those who gain it; that is only to be expected. However, we can say something about what happens to total output as a result. To see this, refer to Figure 2.6.11 once again. Suppose that land holdings are transferred from the landed gentry just to the right of B to the involuntarily unemployed just to the left of A.

There are two immediate effects of this transfer. First, the beneficiaries of the reform become "more able" to work at the going market rates of remuneration. That is, their minimum piece rates come down, because their nonlabour income has increased. Second, the losers of land become more -willing to work, because their nonlabour income has decreased, so their minimum piece rates decline as well! Thus land reform has the effect of bringing downthe minimum piece rate for all who are directly affected by the reform. This is shown in Figure 2.6.11 by the dotted bulges that appear to the left of A and to the right of B.

So a judicious land reform has the power to increase overall output in the economy. Such reforms have three effects. First, the unemployed become more attractive to employers as their nonwage income rises. Second, those among the poor who are employed are more productive to the extent that they, too, receive land. Finally, by taking away land from the landed gentry, their reservation wages are lowered, and if this effect is strong enough, this could induce them to forsake their state of voluntary unemployment and enter the labour market. For all these reasons, the number of employed labour units in the economy rises and pushes the economy to a higher output equilibrium.

2.6.7. Nutrition, time, and casual labour markets
Writing in 1943, Paul Rosenstein-Rodan observed that one of the fundamental features of labour markets that lack contractual structure is its neglect of possibly beneficial externalities. The example that he stressed was on-the-job training. Firms that impart on-the-job training to their workers not only contribute to their own profits, but raise the level of skills and proficiency throughout the economy.

The problem is that more often than not, firms fail to capture the entire benefit of their

training activities. After all, workers might change jobs. This leads to an externality. On-the-job training requires substantial upfront investments in the worker, with no guarantee that the worker will be aroundto give back any of the fruits of such investments to the firm. The point is that if the firm has to incur such expenses, it would like to reap the gains.

Although Rosenstein-Rodan focused on training, his perceptive comments apply just as strongly to the nutritional status of workers that supply labour in a casual market, with no regulations or safeguards. Well-nourished workers are of great long-term advantage to their employers, provided that there is some way to guarantee that such workers remain in the employer's keep. In the absence of such guarantees, the collapse of nutritional status in a poor rural labour market can be comprehensive.

Nutrition is only a parable for all sorts of long-run investments that a firm could make in a worker. We have already discussed on-the-job training. Other investments that have a beneficial impact include firm-provided health insurance, as well as financing for technical training and higher education.

This is the curse of a casual labour market, and the curse is especially harsh in poor countries. To make sure that firms recoup their investments, there must be restrictions on labour movements, and these restrictions have their own costs (slavery being the extreme example). What one does about it is unclear.

2.6.8 Permanent labour markets
2.6.8.1 Types of permanent labour

We will use the terms "permanent labourer," "tied labourer," and "attached labourer" interchangeably to identify any labourer who commits his labour to an employer for an extended period. The period is in contrast to that for casual labourers, who are frequently hired by the day and sometimes to complete an operation lasting for a few days at best.

Think of two broad categories of attached labourers. There are those who perform special tasks that require some judgment and precision—tasks that might be difficult to monitor. Plowing, regulating the flow of irrigation water from pump sets, driving and maintaining tractors, supervision and recruitment of casual labour, and operating threshers are examples of agricultural tasks that pose monitoring problems, simply because they have an effect on the final output (or on the upkeep of machinery) that may be inseparable from a host of other effects, such as bad weather or the failure of some other complementary activity. In contrast, an activity such as weeding, harvesting, or basket weaving lends itself to natural observation, and often such tasks can be paid for on a piece-rate basis.

There are essentially three ways in which an employer can carry out nonmonitorable

tasks. First, he might entrust them to family members, who have a spontaneous interest in the welfare of the farm. This is a good idea forsmall farms, but if the scale of operations is large, outsiders will have to be hired. Second, the employer might hire casual labour to carry out these tasks, but in that case direct supervision of labour becomes a necessity. Even with direct supervision, however, it is not possible to keep track of every passing moment of the labourer's activities, so a judgment of success or failure must rely on the final output. However, as I have just said, the final output is an imprecise indicator. In addition, it is often a late indicator. Many slack season tasks such as plowing can be judged on this yardstick only after the harvest is finally realized. By that time casual labour hired during the slack season will already have been paid. This brings us to the third option, which is to hire a subset of the labour force of the farm on a "permanent" or "attached" basis, under the implicit or explicit understanding that their employment is long term, but can be revoked if the indicators of their performance are consistently low.

In the second category of attached labour, there are no special tasks performed. Attached labour might be used to perform the tasks of casual labourers. However, what is the purpose of engaging such labourers in long-term contracts? Put another way, do long-term contracts provide some extra value to either employer or employee (or both) that short-term contracts cannot? We will see in the following text that there are situations where they indeed provide this value.

Permanent labour: Nonmonitored tasks
In this section, we study the first kind of permanent labour. We set up the simplest model that helps us capture the basic issues involved.

Production conditions
Consider a farm that carries out production with the help of different techniques that it can choose. Each technique comes with different proportions of tasks that must be monitored and tasks that can be left to casual workers. For instance, a technique of production that relies heavily on the correct and timely application of irrigation and fertilizer (as is the case with certain types of high-yielding varieties of seeds) is more likely to be intensivein tasks that are hard to monitor, such as the proper regulation of irrigation water. Think of producing the same level of output with different combinations of monitored and nonmonitored tasks. Presumably the amount of physical capital that is heeded changes as well from technique to technique, but for the moment we ignore this complication.
Imagine that to carry out each unit of each task requires one unit (say a man-hour) of labour. The nonmonitored tasks will be carried out by permanent labourers, and just how they are paid will be discussed in detail later. The monitored tasks will be carried out with the use of casual labour. Thus the isoquant that describes various combinations of nonmonitored tasks and monitored tasks (applied to the given plot of land) to produce a

given quantity of output may as well be depicted by various combinations of permanent and casual labour that are required to achieve the same level of output. Denote these quantities by L_p and L_c, respectively. Figure 2.6.12 shows us the relevant isoquants.

Now imagine that the employer is faced with different wage rates for permanent and casual labour. (Presently we will-be concerned with the determination of these wages, but ignore this for now.) Let's call these wages
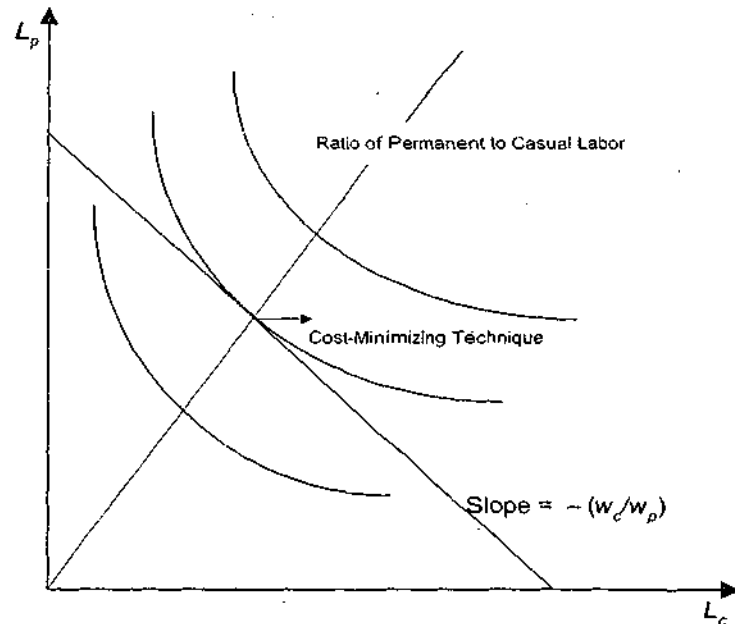

Figure 2.6.12. Combinations of permanent and casual labour needed for some fixed output level.

zv_p and zv_c respectively. Efficiency in production requires that a technique be chosen that minimizes the total cost of inputs, which is $w_p L_p + w_c L_c$, for any given choice of output. This is perfectly standard: we see that the cost line (which has a slope of $-w_c/w_p$ in Figure 2.6.12) must be tangent to the corresponding isoquant.

Note that as the ratio of permanent to casual wages changes, employees choose different techniques of production and, therefore, different ratios of permanent to casual labour on the farm. In particular, note that as the ratio of permanent to casual wages falls, the cost line becomes steeper. The tangency point for cost minimization then shifts from A to A', which leads to a higher ratio of permanent to casual labour in the workforce. We therefore record a simple but important observation that will be useful in what follows: the ratio of permanent to casual labour depends inversely on the ratio of the permanent wage rate to the casual wage rate. Note that this in itself does not tell us much, because the casual wage rate is an endogenous variable. In what follows, we uncover this endogeneity and look deeper.

Determining the permanent wage
Consider a typical employer who faces a given wage for casual labour in the market. With competitive markets, this employer can do nothing to change the casual wage. The question is, Whatis the minimum wage rate that he must pay a permanent labourer, so that Such a labourer will have the incentive to adequately carry out nonmonitorable tasks?

The permanent wage will serve as a premium to the worker which he loses in the event that he supplies inadequate effort. The determination of this wage will therefore critically depend -upon the importance that the worker attaches to future gains and losses. We introduce a simple and tractable story that incorporates that worker's mental time horizon: the extent to which the future worries him when he makes current decisions.

Specifically, we suppose that at each date, the worker thinks N dates into the future, and factors in the consequences of his current decisions on gains and losses in the coming N periods. The current decision is whether to put in effort or to shirk. The future consequences come from nonrenewal of the permanent contract.

2.6.8.2 Permanent labour: Casual tasks Seasonality and income
We take a leisurely approach to motivating the second type of permanent labour. Imagine agricultural life in a developing country. The first fact of cultivation is the existence of seasons. Roughly speaking, we can divide an agricultural production cycle into a slack season and a peak season. In the slack, agricultural activity is at a relative low, whereas the peak season contains the bulk of the physical activity: harvesting. Thus the demand for labour in the slack season is relatively low and spikes sharply in the peak season. This

fluctuation is naturally mirrored in the behavior of spot market (casual) wages: they follow a zigzag pattern, exhibiting lows in the slack and highs in the peak season.

We can add to these observations by noting that, in addition to the natural fluctuations imposed by the presence of the seasons, there is another source of fluctuation: uncertainty. From the vantage point of the slack season, it is difficult to predict what the peak season wage will be, except to observe that it is likely to be higher than the slack wage. The uncertainty arises from the fact that the peak season demand for labour depends on the abundance of the harvest, and the bounty of the harvest is affected by the weather as well as a host of other factors that may be out of the control (or predictive range) of the farmer.

It is possible to temper these observations somewhat by noting that in all parts of the 'world, multiple cropping is on the rise. This is especially so in regions where the availability of irrigation or new varieties of seeds serves to reduce dependence on the season. To the extent that multiple cropping occurs, the spikiness of the agricultural cycle is smoothed out and seasonality plays less of a role, but the role is still prominent enough to deserve our full attention.

Fluctuation aversion

Risk aversion is closely related to what we might call fluctuation aversion, a state of affairs in which an individual reacts to the prospect of a fluctuating but perhaps perfectly deterministic stream of income. In Figure 2.6.13, we depict a two-income stream: a wage of Rs 100 per month in the slack and Rs 200 per month in the peak. Assume for simplicity that the two seasons are of equal duration. The utility that the worker receives from Rs 100 is given by point A and the utility he gets from Rs 200 is shown by point B.

What is the average utility received by the worker? By our simplifying assumption that both seasons are of equal lengths, it is, of course, the average height of points A and B. This is easy enough to depict on the diagram. Simply draw the chord joining A and B and
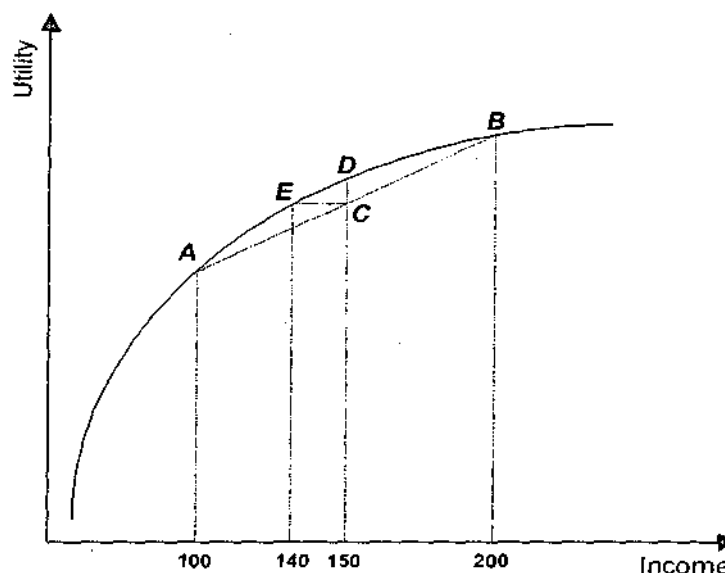

Figure 2.6.13.Fliictiintion aversion.
then look at the height of the
midway point along this chord, marked C. This is the average utility generated by the fluctuating stream of incomes (100, 200).

Now ask a slightly different question: What is the utility received from an income stream that does not fluctuate and provides a constant flow of Rs 150 per month? This is clearly the height of point D in Figure 2.6.13. Notice that if the utility function displays diminishing marginal utility, trie height of point ,D must exceed that of C. In other words, the utility of the average income under a fluctuating stream must exceed the average utility received from that stream. It sounds tricky, but it is actually very simple, and if you reread the section on risk aversion you should have no difficulty following this. Fluctuation-aversion and risk-aversion are closely related.

Permanent contracts that smooth income
Suppose that an employer of labour is risk-neutral and does not mind fluctuating payments or receipts as long as they have the same average value. Suppose, furthermore, that he desires to hire labour in both slack and peak seasons. Then he can make the following speech to a prospective -worker: "Imagine I pay you Rs 150 per month throughout the year, whether it is slack or peak. This is the average amount you are

receiving anyway. Because you are fluctuation-averse, you will prefer this to a varying income stream with the same average value. Indeed, if I shade the amount a bit, say to Rs 140 per month, you should still prefer this to the fluctuating stream (compare points C and E in Figure 2.6.13).

The argument is very simple. If employers are risk-neutral and workers are'risk-averse, then similar attitudes apply to fluctuations as well, and there should be scope for contractual smoothing of incomes over the slack and peak seasons. This is a form of (interseasonal) permanent labour, the second kind of permanent labour that we referred to at the beginning of this section.

The reasoning in favor of such contracts is persuasive, perhaps a bit too persuasive. Why don't we observe the entire labour force "tied" in this fashion? Casual labour should be the exception, not the rule. Indeed, as the box on labour tying in India tells us, there was a time when this was indeed the case. However, labour tying of this sort has been most definitely on the decline. The box briefly documents this trend for India, and the same is true for other countries as well as for developed regions such as Europe at earlier points in time. Today, tied labour, although present, is certainly not dominant.

Why are permanent contracts not dominant?
There are probably several reasons that we can advance (some fallacious) for the observation that labour tying of this sort is not universal.

First, our argument appears to apply only to employers who are intending to hire a particular labourer both in the slack and in the peak. If labour demand is dramatically different over the two seasons, this may not be true of all employers and all labourers.

Second, seasonality itself may have declined over time, which obviates the need for the kind of permanent contract that we are discussing. The reduction in seasonality might stem from technological or infrastructural improvements (such as irrigation) that permit multiple cropping and thereby smooth out agricultural activity over the year.

Fourth, consider the worker's access to credit, specifically consumption credit. Fluctuating incomes pose little or no problem if they can be smoothed by taking recourse to the credit market. Workers might be able to borrow against a high peak income in order to consume more in the slack. If this can be done at a reasonably low interest rate, then consumption is smoothed even though income is not.
This argument, is sensible provided we do not go the whole way and imagine that credit markets are perfect.

2.6.9. Summary
In this chapter, we studied rural labour. Like land markets, labour markets are devices of adjustment between the ownership pattern of factors and the way they are put to operational use.

Casual labour is normally hired to carry out tasks that are easily amenable to observation. The tasks of long-term labour are somewhat more mixed. On large plots of land, a long-term employee may serve in a supervisory capacity and might be responsible for tasks that require special care and are relatively difficult to monitor. In addition to these, long-term employees might work at "standard" tasks along with their casual counterparts, for instance, participating in the harvesting process.

We moved on to these considerations. An entire section was devoted to the interrelationship between poverty, nutrition, and labour markets. There are two main building blocks: first, an individual's work capacity affects his income on the labour market and, second (and the object of our particular focus), individual income in turn affects work capacity. The first effect is captured by supposing that the market pays piece rates for observable tasks, so that capacity multiplied by the piece rate yields total labour income. The second effect is summarized in the capacity curve, which links nutrition to work capacity. We argued that a minimum threshold level of nutrition is required before a person can carry out productive work. This feature generates an individual labour supply curve that exhibits a discontinuity: at low piece rateshe is unable to profitably function in the labour market, but there is a critical piece rate at which he can supply labour.The nutrition-based model of the labour market also brings out the beneficial impact of land reform. Not only does land reform lower inequalities, it also has a supply-side effect on total output, which rises.

2.6.10 Short answer type questions
Write short notes on
5. Labour demand
6. Labour supply
7. Permanent labour
8. Wage determination

2.6.11 Long answer type questions
1. Explain labour categories.
2. Explain poverty, nutrition and labour market.

2.6.12 Recommended Books
1.   Development Economics by Debraj Ray
2.   Development Economics by R. Barro
3.   Development Economics by M. P. Todaro

Credit market in agriculture

2.7.1 Introduction
2.7.2 Objectives of the lesson
2.7.3 Sources of demand for credit
2.7.4 Rural credit markets
      2.7.4.1 Institutional lenders
      2.7.4.2 Informal lenders, information, and collateral
2.7.5 Some characteristics of rural credit markets
2.7.6 Default and fixed-capital loans
2.7.7 Informational asymmetries and credit rationing
2.7.8 Default and enforcement
2.7.9 Interlinked transactions
2.7.10Interlinkages and information
2.7.11 Alternative credit policies
2.7.12 Microfinance
2.7.13 Summary
2.7.14 Short answer type questions
2.7.15 Long answer type questions
2.7.16 Books for reference

2.7.1 Introduction
Two features of this market make it problematic. First, it is often very difficult to monitor exactly what is being done with a loan. A loan may be taken for an ostensibly productive reason, but may be used for other needs (such as consumption) that cannot be easily transformed into monetary repayment. Alternatively, a loan may be put into a risky productive activity that may fail to pay off. This creates the problem of "inability to repay" or involuntary default, at which point there is little that a lender can do to get his money back.
The second problem is voluntary or strategic default, a situation in which the borrower can repay the loan, in principle, but simply does not find it in his interest to do so. Such a state of affairs is especially pertinent in contexts where the legal system of loan enforcement is weak. Two examples of weak enforcement come to mind. One is in the sphere of international debt. An effective international court of law does not exist, and disgruntled lenders must take recourse to punitive measures that are often limited, such as the threat to

advance no further loans (which often lacks credibility) or the threat to cease trading relationships (which are subject toobjections from other economic actors in the lending country that gain from such trade). The second example, which we focus on here, comes from developing countries. Internal courts of law are often weak or absent, and many lenders must rely on the same sorts of punitive mechanisms as in the case of international debt, such as the threat to advance no future loans. The less effective these threats, the more they constrain the operation of credit markets in the first place.

2.7.2 Objectives of the lesson
In this lesson we will discuss sources of credit, rural credit markets, characteristics of rural credit markets and grameen banks.

2.7.3 Sources of demand for credit
The demand for credit or capital can be divided into three parts. First, there is the capital required for new startups or a substantial expansion of existing production lines. The credit market that services these needs is called the market for fixed capital: capital that is poured into the purchase and organization of fixed inputs such as factories, production processes, machines, or warehouses. In contrast, there is the credit required for ongoing production activity, which occurs because of a substantial lag between the outlays required for normal production and sales receipts. Thus, merchants who buy handicrafts from poor producers advance or "put out" sums of money that are used to purchase various materials. When the product is finally produced, these credit advances are deducted from the price that the merchant pays for his wares. This market is called the market for working capital. Finally, there is consumption credit, which typically is demanded by poor individuals who are strapped for cash, either because of a sudden downturn in their production, or a sudden fall in the price of what they sell, or perhaps because of an increase in their consumption needs caused by illness, death, or festivities such as a wedding.

It is this last source of demand that also underlies the demand for insurance. The possibility of a sudden crop failure or illness can give rise to groups of individuals who band together in some form of reciprocal relationship. Certainly, such reciprocity can only be sustained over time by the recurrent possibility that each participant can some day be in the same unfortunate position.

Although fixed capital credit is of great importance in determining the overall growth of the economy, working capital and consumption credit are fundamental to our understanding of how an economy supports its poor and disadvantaged. In no sector is this more the case than in agriculture. The seasonality of agricultural production and the low incomes of those who live and work in the rural sector heightens the importance of

working capital in production.

At the beginning of the crop cycle, the peasant faces a considerable need for working capital: money to purchase seeds, fertilizers, pesticides, and so on. These expenditures are bunched up front, and the farmer is often without sufficient funds to finance it. Hence, there is a need to borrow, with theloan repaid after the crop is harvested and sold. The repetitive taking and repayment of loans is an intrinsic feature of life, and the ease with which such loans can be taken fundamentally affects the economic productivity and well-being of millions of individuals.

When we add to seasonality the uncertainty surrounding productive activity, consumption credit also takes on great importance. An individual's harvest might fail, which causes immense temporary hardship that can only be alleviated through loans. Farm wages are typically lower in the lean season relative to the harvesting season, when the demand for labour is high. Moreover, there is often a high rate of unemployment in the slack season. Peasants, and landless labourers in particular, who rely on wages as a means of livelihood, find considerable fluctuations in their earnings from month to month. Credit is required for such people to smooth consumption over time to cover their needs in periods of low income by borrowing against higher expected earnings during times when the going is good.

2.7.4 Rural credit markets
2.7.4.1 Institutional lenders

First, there are the formal or institutional lenders: government banks, commercial banks, credit bureaus, and so on. Often special banks are set up, as in Thailand, the Philippines, and India, and in many other countries, to cater especially to the needs of rural production. The main problem with formal lenders is that they often do not have personal knowledge regarding the characteristics and activities of their clientele. Often, these agencies cannot precisely monitor just how the loans are used. The problem is not just production versus consumption: for example, the fear that a loan taken ostensibly for some productive purpose may be squandered to meet the expenses of a wedding. There are other, more subtle reasons for a systematic divergence between what lenders want done with the money and what borrowers want.

Many governments in the nascent postwar developing countries recognized the importance of their agricultural sectors, so rural policy was an integral component of government policy in general. Rural credit programs, in turn, formed a fundamental component of rural policy.

Traditionally, poor households in the countryside could obtain loans only from local

village moneylenders, who charged exorbitant rates of interest. Borrowing was restricted only to cases of desperate need. The purpose of government intervention in rural credit markets was presumably twofold: the egalitarian motive of making cheap credit available to poor rural families and the motive of raising agricultural productivity and efficiency through financing the adoption of new inputs and technology on a wider scale.

2.7.4.2 Informal lenders, information, and collateral

In the previous section, we discussed the problem of collateral and that some economic agents offer collateral in forms that are unacceptable to formal lenders. However, the "right" sort of informal moneylender may be willing to accept collateral in these forms. A large landowner who has land adjacent to that of a poor farmer may be interested in the tiny plot as collateral (indeed, perhaps more interested in the plot than in getting the loan back). An employer of rural labour will accept labour as collateral, in case the labourer-borrower fails to repay.

It is no surprise, therefore, to find that formal banks cannot effectively reach out to poor borrowers, whereas informal moneylenders—the landlord, the shopkeeper, the trader—do a much better job.

There is another reason for the dominance of informal moneylending. Quite apart from the ability to accept collateral in exotic forms, the informal moneylender often has much better information regarding the activities and characteristics of his clientele. A trader who advances loans for working capital often has first claim on the farmer-borrower's output; he arrives with his truck at the field on the day of the harvest. A "landlord has a better chance of knowing what his tenant is doing with a loan than any commercial bank can hope to have. Thus, even in countries where government efforts to extend rural credit are strong, the informal credit sector flourishes. Take the case of India. Right after the country's political independence, in the early 1950s, a majority of rural households borrowed from the village moneylender. The All India Rural Credit Survey, published by the Reserve Bank of India, reveals that in 1951 only 7.2% of all borrowing was from government sources, banks, and cooperatives. By 1981, this number had jumped to 61.2% (Bell [1993]), thanks mainly to the Indian government's substantial drive to extend rural credit through official channels. Individual moneylendersby.nO means vanished; 24.3% of all debt was still owed to them. Similarly in Thailand, the share of the informal sector in total rural credit declined from a steep 90% in 1975 to a still influential 50% in the late 1980s (see box). On the other hand, there are countries in which most of the rural population relies almost entirely on informal-sector moneylenders. Nigeria is a case in point. A study from the late 1980s (see Udry [1994]) reveals that only 7.5% of all loans (in value) came from banks, companies, or projects. Our next box_studies informal moneylenders in more detail for the Philippines.

2.7.5 Some characteristics of rural credit markets

If rural credit markets were perfectly competitive and smoothly functioning, there would be no need for this chapter. As in the case of any commodity, there would be a demand curve for credit and a corresponding supply curve of credit, and the intersection of the curves would determine the volume of credit and its equilibrium "price" as well, which is simply the interest rate. There would be little else to write about. Unfortunately, rural credit markets are pretty far removed from perfect competition. You should already have an inkling of this from the previous section.

Informational constraints

The fundamental feature that creates imperfections in credit markets is informational constraints. We have already made this argument (and there is much more to come). To summarize the discussion so far, recall that informational gaps occur at two basic levels. First, there is lack of information regarding the use to which a loan will be put. Second, there is lack of information regarding the repayment decision. This deficiency includes limited knowledge of the innate characteristics of theborrower that may be relevant in such a decision, as well as limited knowledge of the defaulter's subsequent needs and activities, which place limits on his incentive to default. All the important features of credit markets can be understood as responses to one or the other of these informational problems.

Segmentation

A characteristic of the rural credit market is its tendency toward segmentation. Many credit relationships are personalized and take time to build up. Typically, a rural moneylender serves a fixed clientele, whose members he lends to on a repeated basis; he is extremely reluctant to lend, outside this circle.

Interlinkage

A third feature, which may be considered an extension of the second, is the existence of what we might describe as interlinked credit transactions. Indeed, this is a good time to banish from your mind the image of a crafty moneylender, whose sole purpose is to lend money at-exorbitant rates of interest to hapless borrowers. A majority of village moneylenders do not pursue usury as their sole occupation. Most of them are also wealthy landlords, shopkeepers, or traders dealing in the marketing of crops. Given a segmented market, it probably won't come as a surprise to learn that landlords tend to give credit mostly to their tenants or farm workers, whereas traders favor lending to clients from whom they also purchase grain (see the box on the Philippines). Thus segmentation often takes place along occupational lines, and the complementarity of some production relationship (tenant and landlord or farmer and trader) facilitates the credit relationship. This interlocking of markets—people who conduct their business in different markets (land, labour, credit, etc.) 'with the same trading partners, and indeed make the terms of

transaction in one market depend on the terms and conditions in the other—is a far cry from the impersonal and independent functioning of markets that characterizes most textbook economic theory.

Interest rate variation

Segmentation has a natural corollary: informal interest rates on loans exhibit great variation, and the rates vary by geographical location, the source of funds, and the characteristics of the borrower. Sometimes the rate of interest is extraordinarily high. Aleem's survey of the Chambar region of Pakistan showed that the average annual interest rate was as high as 78.7% and involved substantial dispersion. The rate in specific cases varied from a low of 18% (which is nevertheless higher than the 12% charged by formal sector banks) to an astonishing high of 200% per annum, that in most parts of Thailand, the informal-sector interest rate varied between 5 and 7% per month, which is dramatically higher than the 12% per annum charged by formal-sector banks.

Rationing

Informal credit markets are characterized by widespread rationing; that is, upper limits on how much a borrower receives from a lender. At first sight this appears natural: why would any moneylender advance infinite quantities of money? However, note that by rationing, we mean that at the going rate of interest, the borrower would like to borrow more but cannot. In this sense credit rationing is a puzzle: if the borrower would like to borrow strictly more than what he gets, there is some surplus here that the moneylender can grab by simply raising the rate of interest a wee bit more. This process should continue until the price (interest rate) is such that the borrower isborrowing just what he wants at that rate of interest. So why does rationing in this sense persist?

Exclusivity

Finally, many informal credit transactions are characterized by exclusive dealings. Moneylenders typically dislike situations in which their borrowers are borrowing from more than a single source. They insist that the borrower deal with them exclusively; that is, approach no other lender for supplementary loans.

2.7.6 Default and fixed-capital loans

The analysis in the earlier section is deficient in one serious respect. It assumed that the default probability is independent of the amount to be repaid.Larger amounts to be repaid may lead to a greater risk of default. This statement suggests that certain loans will not be given at all under any circumstances, irrespective of the interest rate premium, because the premium itself affects the chances of repayment. Likewise, large loans themselves raise the chances of default and will, therefore, not be made. What is "large" depends, of course, on the particular circumstances of the society, such as per

capita wealth and the availability of alternative opportunities such as migration to another area (in the case of default).

We can extend this line of reasoning, not just to the size of the loan, but also to the kind of use to which the loan will be put. If the loan can be usedby the borrower to permanently put himself in a situation in which he never has to borrow again, then such loans may not be forthcoming. Suppose, for instance, that a rural labourer wishes to borrow money so that he can migrate to the city and set up a small business there. Indeed, given his contacts and entrepreneurial spirit, this may be efficient from a societal viewpoint, but it would be surprising indeed if any rural moneylender advanced him such a loan. In the absence of a legal enforcement mechanism, often the only instrument that a moneylender has is the threat of not advancing loans when needed in the future. But if future loans will never be needed, then the threat has no value.

It is therefore reasonable that in the presence of strategic default, the overwhelming provision of informal loans will be for working capital or consumption purposes, rather than for fixed investments that may permanently reduce the borrower's future need for credit.

2.7.6.1 Default and collateral

The fear of default also creates a tendency to ask for collateral, whenever this is possible. Collateral may take many forms. Certain property rights may be transferred while the loan is outstanding: land may be mortgaged to the lender, and use rights to the output of that land may be in the hands of the lender as long as the loan is outstanding. Labour may be mortgaged as well, and later used, if necessary,, to pay off the loan. More exotic forms of collateral are not uncommon. For instance, Kurup's [1976] study of Kerala, India, showed that ration cards (which are used to purchase subsidized food from the public distribution system) are often handed over to the lender for the duration of the loan. This is a weird special case of what is generally thought of as the mortgage of usufructuary rights. For instance, a borrower who has coconut trees on his land might mortgage the rights to the coconut output as long as his loan remains outstanding. This idea applies just as well to land that produces other crops.

Fundamentally, collateral is of two types: one in which both lender and borrower value the collateral highly, and another in which the borrower values the collateral highly, but the lender does not (the third variant is obviously not observed). From the point of view of strategic default, it is irrelevant whether the first or the second form of collateral is employed. The pawnbroker who accepts your favorite wristwatch (given to you by your grandmother) as collateral for a loan may not be particularly concerned with selling the watch if you do not repay, and may not be able to sell it for a high price either. But he

knows that you attach sentimental value to that watch and hence will pay back a loan even if the rate of interest is high.

Collateral that is valuable to both parties has the additional advantage that it covers a lender against involuntary default as well. For these typesof collateral, credit may simply be a veil for acquiring collateral, as the following simple model reveals. As a by-product of this model, we obtain an alternative view of usurious interest rates.

Suppose that a small farmer is in need of a loan of size L, perhaps to tide over a family emergency. He approaches the local large landowner, who is known to lend money for such purposes. The landowner asks him to pledge land as collateral for the loan. Our farmer has a plot of land adjacent to that of the large landowner and this is -what he pledges.

Now let us keep track of some of the relevant variables by introducing some notation. Let i be the interest rate charged on the loan L and let V_s (S for "small") be the (monetary) value that the small farmer places on his land. Likewise, let V_B (B for "big") be the value that the big landowner attaches to the same plot of land. Because the plot is adjacent, V_s is not a negligible value: one can certainly imagine cases in which V_B exceeds V_s. Thus we have an example here of the first kind of collateral, one which is of value to both parties.

Next, let us place a monetary value on the loss to the farmer from default, over and above the loss of his collateral. Such losses may include the fear of not receiving future loans or even the threat of physical retribution. Summarize the money value of this loss by F.
When the time comes to return the loan, we can conceive of two possibilities:
(1) The borrower may be in a state of involuntary default: he simply does not possess the wherewithal to repay the money. In that case, he certainly loses the land, which passes into the hands of the large landowner.
(2) The borrower may contemplate willful default and take his chances with the landless labour market or even with migration to the city. The total loss to the borrower in this case is $V_s + F$, 'whereas the gain is that he gets to keep the principal plus interest that he owed. Thus the borrower will prefer to return the loan if
(2.7.2) $\qquad\qquad L(1 -j- i) < V_S + F.$
Consider, now, the lender's preferences. Does he prefer to see his loan returned or does he prefer to keep the collateral? He will prefer his money back if
(2.7.3) $\qquad\qquad L(1 + i) > V_B.$
Combining (2.7.2) and (2.7.3), we may conclude that loan repayment is in the interest of both parties only if
(2.7.4) $\qquad\qquad V_B < V_s + F,$
" The observations to follow are based on Bhaduri [1977].

which is a way of saying that the lender's valuation must not exceed the borrower's valuation by too much. In the special case where F = 0, so that collateral is the only way to force loan repayment, (2.7.4) says that the lender's valuation of the collateral must be less than that of the borrower.

Turn this around. Suppose that (2.7.4) does not hold, so that $V_B > V_s + F$. In this case, it follows that whenever the borrower prefers to repay the loan, the lender actually wants him not to do so! The lender would actually like the credit transaction to be an excuse to acquire the collateral (cheap). Thus collateral that is of high value to both lender and borrower may (paradoxically) result in credit transactions with excessive rates of default.

2.7.7 Informational asymmetries and credit rationing

Not all borrowers bear the same amount of risk. There are high-risk borrowers and there are low-risk borrowers. A diligent farmer with a significant amount of land to cultivate may be considered low risk by a lender, because the chance of crop failure and bankruptcy is low. A landless labourer in poor health is high risk. A farmer who owns a pump set or has access to assurd irrigation carries lower risk than one who doesn't. Alternatively, the crop that is being grown may be more or less prone to the vagaries of the weather. Lending risk may vary significantly from borrower to borrower.

Risk may be correlated with characteristics of the borrower that are observable to the lender (such as landholdings or access to irrigation). However, it may substantially.depend on other qualities that are not observable (perhaps farming skills or mental acumen in the face of a crisis, thriftiness, the quality of his land, and so on). When the factors that make for risk are observable, the lender can select his clients or charge appropriately higher rates for the high-risk clients. However, to the extent that clients bear different risks that cannot be discerned by the lender, an additional dimension is added to credit market transactions—the interest rate now affects the mix of clients that are attracted (and hence, the average probability of default). Tlnis new dimension might give rise to a situation in which at prevailing rates, some people -who want to obtain loans are unable to do so; however, lenders are unwilling; to capitalize on the excess demand and raise interest rates for fear that they will end up attracting too many high-risk customers.

The risky borrower is willing to pay a higher rate of interest than the safe borrower, and this interest rate is independent of his probability of success, p. The reason is that bankruptcy yields zero, and in such a situation he defaults on the loan anyway, so his expected profits depends only on the success state. In this sense the risky borrower acts as if he does not care about failure.Of course, the lender cares.

2.7.8 Default and enforcement

Typically, the borrower will have access to more than one moneylender. He may therefore be tempted to default on the loan from the current lender and switch to another source when the current lender refuses to deal with him any further. Indeed, precisely this fear forced the lender to offer the borrower some premium or surplus on the loan over and above his opportunities elsewhere. Nevertheless, the existence of alternative sources of credit strengthens the incentive to default.

If a borrower defaults in his transactions with one lender, this may destroy his reputation in the market and mark him as a bad risk. As a result, other lenders may be reluctant to lend to him in the future. Clearly, this requires that information about the borrower's default action be spread throughout the lending community, so it is no surprise that a lender will eagerly make a default public, and he will certainly want to make this willingness known in advance to the borrower.

As societies develop, mobility increases and traditional ties fall apart. Over time, informal information networks are replaced by the anonymous devices that we see in present-day industrialized societies. However, the replacement may be a long time coming. Hence, there is a large intermediate range of cases where the flow of information slows to a trickle. This is the transitional stage in which many developing countries find themselves. Indeed, it is perfectly reasonable to postulate that information flow follows a U pattern: both traditional and economically advanced societies have a lot of it, whereas societies in transition do not.

As countries begin to develop and industrialize, the traditional rural structure of closely knit, isolated village communities begins to disintegrate. A great deal of mobility is created: people move from village to village, from village to town, and from town to city in response to the growing demands of commerce and trade. Access to markets in distant regions develops and people increasingly enter into transactions with strangers instead of neighbors. At the same time, the introduction of modern inputs into agriculture (e.g., fertilizers, pesticides, pump sets for irrigation, etc.) creates a surge in the need for credit and working capital.

In this environment of relative anonymity, the problem of loan recovery becomes particularly acute. Traditional community pressure can no longer be relied upon nor are there well-developed channels of information flow as in developed countries (e.g., computer networks tracking the credit histories of each individual customer, which banks and credit agencies invariably check before advancing a loan or credit line to a customer). Consequently, a farmer or worker may default on a loan from a moneylender in one town or village and approach another lender in another town for future loans, with very little

risk that his past crime will be known in the new place.

It seems that limited information and the associated hazards of lending have prompted moneylenders to build up tight circles of trusted clients, and they are unwilling to lend outside the circle. It is this sharp segmentation of the market that induces most borrowers to comply with contractual terms: a defaulting borrower, who is removed from the good books of his current lender, will find it extremely difficult to find a new loan source. Thus, apparent competition between lenders and free access to multiple sources is actually restricted due to informational limitationsand this restriction, in turn, helps to solve the moral hazard problems that such informational limitations give rise to.

2.7.9 Interlinked transactions

A common feature of many loan transactions in developing countries is that credit is linked with dealings in some other market, such as the market for labour, land, or crop output. For instance, it is commonly observed that landlords are often the principal source of credit for their tenants, using their labour or even their rights to tenancy as some form of collateral. On the other hand, traders are the principal source of funds to owner-cultivators, especially those who lack access to the formal sector. Traders usually combine such credit dealings with purchase of their borrower's crop.

Interlinked contracts are formed in a variety of other ways. While loans are outstanding, the lender may have use-rights to the land or other assets of the borrower, as already discussed. To the extent that the lender can directly benefit from the sorts of assets owned by the, borrower, this makes credit transactions easier to enforce. However, for the direct benefit to be present, it is often necessary that the borrower and lender be engaged in similar or complementary occupations.

2.7.10   Interlink ages and information

It is possible that an interlinked bargain is struck because in this way, the lender can dispense with some of the costs of keeping track of the activities of the borrower. A rice trader who makes funds available to a farmer may demand repayment in terms of the output because such repayment is easier to enforce under the normal routines of the trader-lender. At harvest time, the trader might arrive at the fields of his suppliers to pick up the crop for transportation. If this is something the trader has to do anyway, a useful by-product is that he gets to place first claim on the crop. Such claims are extremely powerful, because other debts are effectively pushed to a secondary position. Likewise, a labourer or tenant farmer who works on the estate of a large landowner under normal circumstances presents a relatively economical credit prospect. In the case of a default on the monetary terms of the loan, the loan can be worked off (and the implicit wages deducted as payment). These are all ways to reduce the chances of involuntary default

without having to incur the extra costs of monitoring or tracking.

Interlink ages and enforcement
Interlinked relationships are sometimes useful to prevent strategic default as well. To see this, recall two stories with very similar features. First, think about the model of strategic default that we considered in an earlier section of this chapter. We noticed that to prevent default, the moneylender cannot drive down the borrower to his participation constraint. To avoid default, a certain surplus over the next best option had to be provided. The borrower presumably trades off the loss in this surplus at future dates with the onetime gain to be had from default.

2.7.11 Alternative credit policies
There has been a growing realization that the needs of rural credit cannot be adequately served with the use of large financial institutions such as commercial banks. As already discussed, the micro information that is required for these operations precludes efficient market coverage on the part of these large organizations.

Two kinds of policies can arise in response to this observation. One is to recognize explicitly that informal lenders are much better placed to grant and recover loans from small borrowers than formal institutions are. The idea, then, is not to try to replace this form of lending, but to encourage it by expanding formal credit to economic agents who are likely to use these funds in informal markets. The second approach is to actually design credit organizations at the micro level that will take advantage of local information in innovative ways. We discuss these two avenues.

Vertical formal—informal links
Expansion of formal credit to informal lenders
The lending of formal funds to informal markets is not a new phenomenon by any means. Large landowners or traders are in a much better position to put up collateral: from the point of view of the banks, they are good credit risks. They then use the funds to cash in on their informational advantage in informal markets. We have already seen evidence of this in our Philippine and Pakistan examples. The expansion of formal sector credit to these agents generates competition among them, and this hopefully improves the borrowing terms faced by individuals who fall outside the ambit of the formal credit system. Similarly, loans might be made to cooperative groups who are better placed to lend the funds because of social or religious ties among the members.

According to this view, informal lenders are not bank competitors. Rather, they complement the lending activities of the banking sector. According to the official,
The informal lenders operate in a different segment of the market. The bank's market

lies largely in corporate and commercial accounts which are generally fully collaterized. In fact, some of the informal lenders are, in effect, conduits of bank funds. This is true with our big grain miller and trader clients who provide advances to the paddy farmers in the regular course of their business.

Now, an expansion of formal sector credit has two opposing effects: on the one hand, it tends to expand competition because each trader now has an incentive to undercut his rival. On the other hand, it tends to reinforce collusive practices by increasing the severity of credible punishments when deviations occur. The net effect depends on which of these two forces dominates.

There are actually good reasons to argue that the possibility of strategic cooperation may indeed improve with the credit expansion. For instance, a larger invasion of rival territory is easier to detect. Thus invasion gains will not go up as quickly as potential punishments, and this will promote strategic cooperation. Essentially, the ratio of invasion profits to punishment losses declines.

Other variations of the model yield similar results. For instance, consider a situation where trader 1 specializes in processing a particular crop A, whereas trader 2 specializes in crop B. Imagine a (proposed) collusive outcome whereby trader 1 (respectively 2) lends money to small borrowers specializing in the production of crop A (respectively crop B). Now trader 1, say, offers a loan to a borrower specializing in crop B on the understanding that the farmer will divert some of his land to the production of crop A, which is then sold to the trader. Suppose that the production function for the farmer exhibits diminishing returns in the production of each crop (although overall it may be more suited to the production of crop B). Then the outcome will be a large gain for trader 1 if the amount of the invasion is small, but with diminishing marginal gains as the size of the intervention becomes larger.

Although the deviation gains are small, the loss that can be imposed on the deviating trader following a reprisal can still be substantially large. As his clients shift the pattern of crop production, this leads to a fairly small but positive marginal gain for the rival, but it still imposes large marginal losses on the incumbent trader. In this scenario, the expansion of credit lines to traders also leads to a decrease in the ratio of invasion profits to reprisal losses, leading to an increase in the chances of collusion.

The foregoing examples emphasize the importance of understanding the nature of interaction between lenders. The impact of any (credit) policy initiative on the welfare of borrowers crucially depends on the resulting effects on competition (or collusion).

Differential information.

Lenders may have differential information regarding borrowers. For instance, lender A may be able to easily distinguish between good and bad credit risks, whereas lender B cannot. Then lender A essentially contaminates the pool for lender B, because he siphons off the good borrowers. Of course, if A faces a credit ceiling, he cannot siphon off all the good borrowers. Thus lender B faces a mix of borrowers. The mix might still be profitable enough for lender B to operate. Now an expansion of formal-sector credit will allow lender A to siphon off even more good borrowers. This might contaminate the available pool to lender B to the extent that he no longer finds it profitable to operate. The increase in A's activities may be more than offset by the shutdown in B's lending.

2.7.12 Microfinance
The Grameen Bank

It is also possible for institutional lending to closely mimic and exploit some of the features of informal lending. For instance, it may be possible to design an innovative rural credit scheme in a rice-growing area in which a formal-sector institution acts both as a lender and a miller. The combined activity will permit the formal institution to accept rice output as repaymentfor loans. By mimicking the activities of a trader-moneylender, the formal institution may actually be able to carry out lending activities that reach small borrowers and are profitable at the same time.

Even if such mixed institutions are difficult to put into practice, there are ways in which the information base of a community can be put to use by a cleverly constructed lending institution. An example of this is the Grameen Bank of Bangladesh.

In Bangladesh, as in many other developing countries, subsidized rural credit offered by government banks laid the basis for its rural credit policy. It didn't work. Repayment rates were low, and loans (or the forgiveness of past loans) were often used for political reasons, undercutting the effects of the policy.

The Grameen Bank, started in the mid-1980s by Mohammed Yunus, lends to very poor households, and lends to groups of borrowers rather than individuals. A typical group consists of five borrowers, and lending to individuals within the group occurs in sequence.[34]The average loan size is around $100. No collateral is required and the nominal rate of interest is around 20% (roughly 12% real). Over 90% of the borrowers are women).

Group lending and the use of information

The central feature of the Grameen Bank's lending policy is that in the event of a default, no group member is allowed to borrow again. This means that a group has to be formed very carefully by the individuals to weed out bad borrowers who could jeopardize the

creditworthiness of the group as a whole. Observe how information is being used by the Bank even though the Bank itself has no a priori access to it. The point is that the borrowers themselves have the incentive to use this information to form groups, and this induces a form of self-selection that no individual-based banking scheme can mimic.[35]Consider some of the specific implications.

Peer monitoring. The argument so far assumes that borrowers are intrinsically safe or risky. This is a valid assumption in many situations: group members may not be able to control what a fellow member does with her money, although they might still be able to predict what she is going to do (for instance/because they know her personally). In other situations, group members might be able to monitor and influence the choice of individual projects. In that case, the group will want these projects to be safe, relative to the safety levels they would choose as individual borrowers.

Viability and performance
Viability. The Grameen Bank lends to the poor. There are limits on the interest rate that it can charge without seriously affecting repayment capacity, and loan sizes are very small. This means that each loan dollar comes with a high administrative overhead. There are costs of providing, tracking, and ensuring repayment of the loan, and these are fixed costs per borroiuer, so that small loan sizes raise these costs. In addition, as we've already seen, the Grameen Bank provides several social services as well, and despite the extraordinary commitment of its staff, these are not free.

Grameen has therefore functioned under a significant amount of subsidy, both from foreign donors as well as from the Central Bank of Bangladesh. In the period 1991-92, subsidies amounted to about twenty-two cents for every dollar lent. According to calculations carried out by Morduch [1997] for the period 1987-94, the Grameen Bank charged annual interest rates that ranged between 12 and 16.6%, but it would have had to charge borrowers between 18 and 22% to cover operating costs without any direct grants. However, there are subsidies on capital costs as well, in the form of lower-than-market interest rates on loans to Grameen, and these are more important than the direct grants. To cover these subsidies, Morduch estimated that interest rates would have to climb to 32-45%.[38]

Performance. What effect has Grameen had on the lives of people? This amounts to the empirical question, in the absence of 'the program, how much worse off would the borrowers have been? This turns out to be a very hard question to answer, precisely because Grameen has been adept at identifying good borrowers! In other words, it is not enough to look at the incomes of a borrowing household relative to the incomes of nonborrowers, even after controlling for the observable variables, such as age and

education, that we (the statistical observers) can measure. The reason is that we are unable to observe characteristics such as drive, creativity, inventiveness, or entrepreneurship— traits that the borrowers themselves use when forming groups. Thus it is possible that only the more "able" borrowers have been able to access Grameen, but the incomes of such borrowers would have been relatively high even if no Grameen had existed! It follows that a regression destined to capture the effects of credit on participation will overestimate the positive impact of the program.

This sort of selection bias may also run the other way. Not only does Grameen look for good borrowers, it looks for poor ones: in particular, such borrowers may be poor for unobservable reasons that are not captured by the characteristics that we can measure. Now a cross-sectional regression will underestimate the effect of the Grameen Bank, because the incomes of those treated under the program will, on average, be lower (at least before the program). This happens because "program placement" isn't random, which is a general problem with assessment of any program that targets the underprivileged, such as setting up health services in areas with substantial undernutrition and illness, or crop extension services that are set up in areas of low know-how, or family planning programs in areas of high fertility.

One way to get around this problem is to find an identifying variable or instrument: one that affects participation in the program but does not itself affect the outcomes of participation. These are hard to find, in general. Variables that correlate with participation usually affect incomes as well, which resurrects the specter of selection bias.[39]

2.7.13 Summary
We studied the main features of informal credit markets. Typically credit is taken to finance ongoing production {working capital) or to finance shortfalls in consumption {consumption credit). Credit may be sought for a new venture, such as the overhead expenditures needed to switch to a new crop {fixed capital).Two fundamental problems characterize credit markets: the problem of "inability to repay" {involuntary default) and the problem of "unwillingness to repay" {strategic default). In turn, both these problems are outcomes of informational imperfections. For instance, if the use of a loan cannot be monitored, it may be frittered away in consumption or in an overly risky project, leading to involuntary default. The fear that this might happen could shut down a potentially profitable credit transaction. Likewise, the absence of a legal system to honor credit transactions and a possible absence of information flow regarding defaulting borrowers might exacerbate the problem of strategic default.

We distinguished between institutional or formal lenders—government or commercial banks, credit bureaus, and so on—and informal lenders—village moneylenders, landlords,

traders, and shopkeepers. Two considerations typically put formal lenders at a disadvantage relative to their informal counterparts: (i) they lack information regarding the characteristics, history, or current activities of their clients, and (ii) they cannot accept collateral in some nonmonetary forms, such as labour or output.

2.7.14 Short answer type questions
Write short notes on
1. Interest
2. Purpose of loan
3. Sources of loan
4. Grameen bank

2.7.15 Long answer type questions
1. Explain rural credit markets.
2. Explain sources of demand for credit.

2.7.16 Books for reference
1.   Development Economics by Debraj Ray
2.   Development Economics by R. Barro
3.   Development Economics by M. P. Todaro