CHAPTER-4

INCOME INEQUALITIES: DECOMPOSITION
BY INCOME DETERMINANTS

In the preceding chapter it was noted that wide-spread inequalities existed in the possession of land, ownership of traditional and modern productive assets, live-stock and incomes amongst the cultivating households of our sampled population. One can seize this opportunity to dwell into the sociology, politics, political-economy or economics of this distribution profile. While the first three aspects have engaged the attention of many researchers, it is the fourth aspect viz. economics of inequality, that needs a more rigorous treatment for devising policies in consonance with India's social objectives. Such a line of attack would call for the measurement of the contribution of the various components to total income inequality. Accordingly, in this chapter, an attempt has been made to decompose inequality through simulation via multiple regression models.

4.1 Income Inequality and Endowment Inequalities

Our interest in studying the structure of inequality amongst the cultivating households led us to the measurement of inequality in incomes, land and other productive assets in Chapter 3. Mere existence of high levels of inequality would not reflect the empirical correlates of disparity,
much less the identification of important explanatory or contributing factors of inequality. To gain some insight into the processes at work, an indirect decomposition procedure was adopted. All the households were arranged in an ascending order of net household incomes and decile distribution was obtained for the accompanying profile of land (area owned and area operated), work force, consumption pressure, live-stock, draught power, liquidity, and various types of other productive assets. The results are shown in Table 4.1.

This kaleidoscopic view of rural economy makes us believe that it is the uneven distribution of land, productive material assets, demographic traits, quantity and quality of work force that may account for prevailing inequalities of income amongst the agricultural economy under consideration. The top ten per cent of the households account for over 38 per cent net household incomes, 34 per cent crop-output, 25 per cent milk-output and 35 per cent of total farm output as compared to 0.50 per cent net household income, 4 per cent crop-output, 7 per cent milk-output and 3 per cent farm-output in the case of the bottom ten per cent of the households.

Given an almost equal proportion of consumers and biological units, in the two polar classes, the level of inequalities in per capita disposable income is really appalling.

On the endowment side, our top income bracket has 30 per cent land, 33 per cent modern productive assets, 22 per
cent milch cattle, 30 per cent liquidity (financial resources), 38 per cent tractors and 17 per cent of the total engines/motors. Whereas the bottom rung of the ladder has just 7 per cent milch cattle, 6 per cent liquidity, 0.54 per cent tractors, 8 per cent engines/motors and just 4 per cent modern productive assets. Here lie the causes and the consequences of heterogeneous levels of prosperity of the peasantry. It seems that a skewed distribution of land, working force and other productive assets gets manifested in an uneven distribution of incomes. No doubt, the divisible content of new farm technology, such as HYV seeds and fertilizers, had no ex-ante scale bias but unequal access to funds brought in this bias in an ex-post sense.

In our area of study we find that (Table 4.1) over 24 per cent of the total outstanding debt is concentrated in the top decile as against 16 per cent for the bottom decile. There is indebtedness in poverty as well as in plenty but the underlying causes are not identical. These loans mean bridging the consumption gap for the poor and augmentation of productive capacity for the well to do. Given such an institutional support, a sound irrigation base and efficient draught power, the wealthy farmers stand in an advantageous position in the setting of modern agriculture. Even the productivity advantage, hitherto enjoyed by the small farms under traditional agriculture, via higher intensity of cultivation seems to benefit the better irrigated and
mechanized large farms. So the causes of variability in household incomes are to be found in the existing inequality of factor endowments itself. And as it is, the concentration of productive wealth appears to be self-perpetuating and self-reinforcing but for the offsetting impact of meta-economic processes. The forces at work may, perhaps, be fairly suggestive but not all of these can possibly be taken up for empirical investigations. However, to the extent it is possible, one would like to quantify these forces and measure the relative contribution of the various factors to total inequality. Such an approach would help in understanding the sources of current disparity and, thereby, evolve measures aimed at inequality reduction. Hence, decomposition of inequality by income determinants or income sources holds much promise here.

4.2 Approaches to Decomposing Inequality

The interest in the problem of economic inequality in the past three decades has led to two approaches to isolate determinants of inequality. Traditionally, celebrities like Kuznets (1955), Chenery (1960, 1968), Adelman and Morris (1971), Chiswick (1971) and Ahluwalia (1976) have engaged themselves in measuring the degree of inequality in a cross-section of nations, and relating the same to other characteristics of these countries thereby arriving at various associates/determinants of inequality.
Another approach that has found favour with authors like Theil (1967), Fishlow (1972), Fei and Ranis (1974), Fields (1977, 1979), Fei, Ranis and Kuo (1978), and Fields and Schultz (1980) is to measure inequality within a country and decompose the same by income source, sector or place or economic activity, regional or spatial divides or employing demographic and personal attributes/characteristics of earning units. Such an approach allows us to obtain functional decomposition as well as micro-economic decomposition by income-determinants which is relevant for tackling the problem of economic inequality amongst households. For once the contribution of different factors towards inequality stands estimated, the choice of policy instruments becomes easy and almost automatic. Therefore, in the present study, we shall adopt the latter scheme for decomposing overall income inequality, found to exist amongst our target population, into the various determinants thereof.

4.2(a) Determinant Decomposition—Alternative Procedures

In the literature on decomposition, different decomposition procedures have been used for different decomposition problems. Even, with respect to decomposition of inequality by income-determining factors, several procedures are available. Faced with multiplicity of income determinants across households, analysis of variance, the comparative $R^2$, and simulation approach have been
recommended because of their intuitive appeal, the availability of tests of statistical significance and handy computational procedures. Fields (1979) puts his weight with analysis of variance by virtue of its ability to give results with the following desirable properties:

1) Decomposes overall inequality into within-factor and between-factor components;

ii) Measures the gross contribution of each explanatory factor to total inequality;

iii) Tests the statistical significance of these main effects;

iv) Measures the marginal contribution of each explanatory factor;

v) Tests the statistical significance of the marginal effects;

vi) Measures the effects of interactions between pairs (and higher order combinations if needed) of explanatory factors;

vii) Tests the statistical significance of the interaction effects;

viii) Estimates the magnitude of the effect on income of each category of each explanatory variable.
The list is, no doubt, impressive and any one would be inclined to adopt this procedure. However, this approach assumes that variance of logarithms is the only inequality measure of interest. But, as argued in the preceding chapter, we are convinced that no single measure of inequality is an all purpose 'best'. Since the analysis of variance procedure ties us down to a single inequality measure, as such, it loses much of its appeal.

4.2(b) Decomposition of Inequality—The Comparative $R^2$, Regression Procedure

The comparative $R^2$ regression procedure advocated by Wise (1975), Fields and Schultz (1980) has much in common with the above approach. Herein more emphasis is placed on high and low contributions of a given variable in terms of the coefficient of determination for income determination models. If our specified relation determining income is:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + U \quad \ldots (4.2.1)$$

where $Y$ is income, $X_i$'s are independent variables and $U$ is the error term, with $R^2$ as the coefficient of determination, then high and low contributions for the $i$th independent variable are defined as follows:

High Contribution ($X_i$) = $R_{i1}^2$ \quad \ldots (4.2.2)

Low Contribution ($X_i$) = $R^2 - R_{i2}^2$ \quad \ldots (4.2.3)
where \( R^2_{11} \) and \( R^2_{12} \) represent the coefficients of determination obtained from the regression equations:

\[
Y = \alpha + \beta_1 X_1 + U \tag{4.2.4}
\]

and

\[
Y = \alpha + \beta_1 X_1 + \ldots + \beta_{i-1} X_{i-1} + \beta_{i+1} X_{i+1} + \ldots + \beta_n X_n + U \tag{4.2.5}
\]

respectively, i.e., the relations obtained by only including \( X_1 \) and then by only excluding \( X_1 \) in the regression equation. The use of logarithms of the dependent variable \( Y \) (income) in the regression equations and finding these high and low contributions to \( R^2 \), gives us limits on individual independent variables as contributants to income inequality measured by the variance of logarithms.

This approach is informative only if the inequality measure of interest is variance log income and not any other measure. Besides this, the other limitations of this procedure are self-evident. For instance, the sum of high contributions of the \( n \) independent variables need not exceed the sum of the low contributions (Hendry and Marshall, 1983); the two sums would be equal if and only if all the independent variables are orthogonal; existence of even slight collinearity amongst independent variables would influence the pattern of these contributions; and with as simple a measure as variance of logarithms, the procedure...
requires 2n+1 regressions to get at low and high contributions.

4.2(c) Decomposition of Inequality—The Simulation Approach

In contrast, the Simulation approach (Behrman, Knight and Sabot, 1983) provides information on decomposition irrespective of the measure of inequality employed. Therefore, the procedure seems to be more useful and a superior alternative to the comparative R² regression approach. More so, because the simulation procedure also answers the same type of questions like:

1) "What would be the impact on inequality if the dispersion due to each determining characteristic were in turn eliminated?"

   OR

2) "What would be the impact on inequality if the dispersion due to all determining characteristics except one were in turn eliminated?"

and is applicable for all measures of inequality that might interest the researcher.

The procedure calls for setting an income determination equation like 4.2.1; estimating the parameters directly, and then using those estimates to simulate the contribution of any given independent variable by hypothesizing zero variance for it. For example, if Gini coefficient is selected as a
measure of inequality, then $G(\hat{Y})$ can be calculated at the first instance, employing the estimated equation. Now holding $x_i = \bar{x}_i$ (the mean value of $x_i$) for the sample as a whole), predicted income $\hat{Y}_i$ can be used to obtain $G(\hat{Y}_i)$ at the next stage. The contribution of each $x_i$ to inequality is then estimated as $G(\hat{Y}) - G(\hat{Y}_i)$ and the relative contribution of the same as a percentage of $\sum_i [G(\hat{Y}) - G(\hat{Y}_i)]$. The relative contribution of each $x_i$ might change with the adoption of a different measure of inequality but this is no weakness of the procedure itself. The difference stands attributed to the measure because of its thrust.

This approach has its limitations but as the authors put it "The advantages of the simulation procedure appear to outweigh the disadvantages", we, therefore, intend using the same procedure for decomposing income inequality in our study. Setting up of a suitable income determination regression equation is the first step in this direction, to which we now turn.

4.3 Determinants of Household Incomes—Multiple Regression Models

As observed by various authors, our farming households defy any clear categorisation to fit into the classical moulds of a 'firm', and 'organisation' or a 'household economy'. Equally diverse are the motivations for taking decisions on the farm: objective output or profit maximisation
survival and growth, or a subjective labour—consumption pressure balance. It could be none or all of these put together. In the same vein the receipts (actual or potential) of a family unit have elements of wages, profits, rent and interest, of course, without a clear demarcation. Consequently, the observed differences in this composite flow can be attributed to chance, ability, human capital and a whole host of other factors.

For want of a clear theory encompassing these multitudes, we shall investigate how the variability in net household incomes (for a year) is explained by differences in resource endowment, the levels and manner of utilization thereof, demographic and societal traits of the units, and finally, the technology and decision mix as reflected in the variability of farm productivity, employing simple regression equations.

4.3(a) Description of the Variables Used

In the ensuing regression exercises, we have employed several combinations of different explanatory variables associated, a priori, with the determination of the net household income (our dependent variable). The list includes both conventional and non-conventional variables. A brief description of these variables as well as their anticipated relationship with the regressand will be in order.
i) **Net Household Income:**

It is defined as the value of all farm output (i.e., crop output, plus dairy output and minus intermediate inputs from the farm), plus income from outside sale of labour, renting-out land and hiring-out of machinery, minus all actually paid-out costs, and minus depreciation.

This definition of income corresponds with current flow of net operating surpluses available with the households at the existing scale of operations, with all the capital stock intact. Since our aim is to gauge economic inequalities across cultivating households, the definition would best serve our purpose. After all, the same very surpluses are available for utilization, may it be consumption or accumulation. Of course, the allocation would depend upon socially acceptable standards of living, demographic characteristics of the family and economic objectives of the production unit.

The net household income would be retained as the dependent variable.

ii) **Size of the Holding and Related Variables:**

It is an accepted proposition of economic theory that the size of an economic unit and its economic fate are closely related. However, there is
no universal agreement either on the concept of size or on the nature of causation linking the two, particularly in the case of agricultural units. In this context, researchers in India have employed several measures of size, like:

1. Size of the land holding—owned or operated (acres/hectares);
2. Total cropped area (acres/hectares);
3. Total gross farm output (value units);
4. Total Investment in fixed capital (value);
5. Total man-equivalent days utilised;
6. Number of permanent farm workers (Family Farm Workers, plus Permanent Hired Workers) engaged full-time on the farm;
7. Total bullock labour days utilised on the farm.

Amongst these, net operated area has been the most popular one in studies involving size. However, in the present study we rely on various sets of size related variables rather than picking up just one variable. Such a procedure is necessary for getting at a meaningful decomposition of income inequality. Since land size directly or indirectly influences several other attributes of the farm; leaving out
other variables would result in an overplay of land size itself as a determinant of income, and hence income inequality.

Accordingly, we have decided to include in our list (1) Land owned (acres); or Land operated (acres); (2) Number of family farm workers (man equivalent units); (3) Number of permanent full-time farm workers; and (4) Number of milch cattle (Thousand rupee equivalents).

We expect each of these variables to influence the net family income positively.

iii) Family Characteristics:

Barring some exceptions, the general practice among economists has been to treat the 'farm' and the 'firm' on the same footing and hence concentrate mainly on the input-output relations alone, to the exclusion of organic unity between 'farm' and 'family'. There is every justification of including certain elements of the internal structure of the family itself which is tied with the farm. Equally important is the position of the family unit, vis-a-vis, other units in the village because, "the traditional village economy, society and polity are almost always internally unequal, exploitative and far from idyllic...." (Lipton, 1980, pp.22).
Both, the 'internal structure' and the 'relative-position' of the family have a bearing on the setting-up and achievability of its goals. Consequently, the list of regressors should include suitable proxies to capture the impact of variability of the twin phenomenon. Type of the family, headship of the household, sex-ratio of the unit, age-composition and work force constitution, dependency ratio, age of the unit and social standing of the family immediately strike the mind. Of these characters, family type and sex of the head of the household showed no variability in our sample and had to be ignored. The same cannot be said about the education of the head (Ed.-Head) of the household. So far as the 'age of the cultivating unit', variable is concerned, there is no doubt that peasant families pass through a life-cycle like any biological organism, and all the units under consideration may be passing through various phases. It is also accepted that life-cycle of the unit can affect its activities in a number of ways starting with the necessity to satisfy the demands of its consumers to coping with other social and economic pressures. But in an empirical study the problems get compounded if the analytically neat scheme is devoid of neat observable divisions. In the context of our sampled population,
where each unit of an extended family offers a pool of embryo nuclear families at varied stages of their respective life-cycles, defining the age or stage in the life-cycle of the overall unit is plagued with ambiguity. Therefore, we have to contend with consumption pressure and labour force as observable variables rather than hanging on to the life-cycle.¹⁹

Coming to the village hierarchy, there exists a strong feeling amongst some keen observers²⁰ of peasantry that, "Cultivator's initial resource position determines his 'bargaining' capacity in the credit, land and labour markets. His past and current involvements in the markets, in turn, influence his production 'decision' and through these, his future resource position. Furthermore, the economic status of the cultivator and his position in the nexus of property relations also define for him the 'Objectives' he sets himself to achieve." (Krishna Bhardwaj, 1974).

This thought is echoed in the works of Bardhan (1980), Bhaduri (1973), Prasad (1974), Hudra (1982) and Vyas (1979) but, to our knowledge, there has been no attempt to formulate this phenomenon into a quantifiable variable.²¹ Given the importance of this phenomenon, an attempt would be made to assess the strength of these forces by inclusion of a suitable
proxy variable, representing 'position' in the village hierarchy, in the regression equations.

Thus, we shall include Ed.-Head, Dependency Ratio and Position Hierarchy Proxy as variables to capture the influence of family characteristics towards income variability. A brief description of these variables follows:

iii.(a) Ed.-Head

It is taken as the number of years of formal education received by the head of the household. While the inclusion of education as an explanatory variable would find ample support in the literature (see for instance, Griliches (1964), Dandekar (1967), Schultz (1975) and Chaudhri (1979) among others), the particular variant chosen might require further justification. More so, because a comprehensive study by Chaudhri (1979) spanning over a decade for this region, comes up with seven alternative ways to define education producing interesting results regarding levels of education and the diffusion of Green Revolution technology in the two neighbouring states of Punjab and Haryana.

But now that the green revolution technology has almost become the 'New Traditional Technology', we
feel that even elementary education becomes relevant and useful under these conditions. The author himself is aware of such a stage, for he writes "this does not suggest that elementary education is not important and relevant for regions with technological conditions requiring use of inputs with which farming communities are already familiar." So the choice of this definition of the variable stands justified. The variable is expected to contribute positively towards net household income.

iii. (b) Dependency Ratio

It is taken as the ratio of the number of standard consumer units in the household to that of the number of family workers in standard units.

While the adult equivalent scales developed by the Indian Council of Medical Research, New Delhi, were employed to obtain standard consumer units in the household, workers were standardised following the procedure by Economic and Statistics Organization (Punjab). The resulting ratio indicates the consumption pressure per family worker. A high consumption pressure might goad the workers to put in a greater labour effort and earn more for the unit, thereby exerting a positive influence on the
net family income.\textsuperscript{23} But this positive impact may be more than off-set by the negative account of extra consumption burden. The latter effect would be more pronounced if the excess consumption pressure has already eroded the productive base of the family enterprise. In any case, the net impact could show up definitely, though differently in several complex situations.

iii.(c) Proxy for Position in the Village Hierarchy

It is defined as the ratio of average land owned per male worker in the household to the average cultivable land per agricultural worker in the village. In our view, such a ratio can be safely taken as the nearest quantifiable approximation of the power enjoyed by the relatively bigger farmers in the village. The actual diffusion of power for the distribution of favours might follow somewhat different route but in our profession it is accepted that relative resource position of the family, in a village economy, also determines its social position\textsuperscript{24} and hence the control over common and uncommitted resources.

By virtue of their having easy access to information and inputs, it is expected that a high
soore on this proxy signifies relative advantage sine-quâ-non. Hence, it is expected to be positively associated with our dependent variable.

iv) Off-farm Employment Mix

The modernisation of the Punjab agriculture has created many ancillary non-farm activities besides creating more employment opportunities for male agricultural labourers. These developments have provided opportunity for supplementary employment and income to hard pressed labour surplus households. The spread of technical and non-technical formal education in rural areas, and its aftermath, might be serving as "........a huge sieve, through which the ablest young people pass to the cities, there to help the urban elite" (Lipton, 1980, p.259). But it has also provided the necessary motivation and skill amongst the ruralites to grab these opportunities, near home, without being cut-off from the natal units and leading to the 'rural skill drain'.

Our economic units, due to their acquisitive character, compulsions of consumption pressure or lack of adequate work for their work-force on the farms, allow (nay, at times compel) male family
hands to take up employment on other farms, nearby factories or even in tertiary activities, by way of adding 'second string to their bows'. But, this off-farm sale of labour becomes, invariably, an important constituent of household incomes. Inclusion of off-farm employment as an explanatory variable, therefore, is in order. This off-farm employment is divided into two categories—rural and urban—in the wake of existing geographic wage differentials. Thus the number of household members holding rural and urban jobs are treated as separate variables for isolating the contribution of off-farm employment. Undoubtedly, both these variables are expected to show up positively in terms of their estimated coefficients.

v) Farm Productivity

Having taken into account the size, structure and conduct (in terms of off-farm participation of labour) of the cultivating households for studying income variability, it is logical to consider performance related variables as determinants of variations in income and hence income inequality. Though the literature offers no unique indicator of gauging performance
efficiency of farming households, yet it is customary to consider some or the other measure of productivity as a relevant quantifiable variable for this purpose.

Therefore, some differences of opinion notwithstanding, we also employ farm productivity as a summary variable for performance. Here again, several alternative variables like: gross farm output per acre, net farm income per acre, crop yield index, total family or farm business income per unit of land or labour, and output-input ratio for the farm business as a whole, are available. Out of these, value of farm output (crop output plus dairy output) per unit of net area operated serves our purpose the most. Since it takes into account all the activities on the farm, the impact of qualitative differences in factor endowments, cropping intensity differentials attributable to labour input or use of tractors and other implements, the use of land-augmenting inputs like high yielding seeds and fertilizers facilitated by assured irrigation through pumping sets or tube-wells, and, above all, it is unrelated to size of the holding. In short, this variable accounts for input quality, their mix and intensity, as determinants of output
variability. Higher productivity, ceteris-paribus, would result in higher income for the economic units.

The selection of these explanatory variables is mainly guided by theoretical backing, ease of computation compatible with data availability and the statistical pre-requisites for the application of classical regression analysis.

4.5(b) Alternative Models of Income Determination: Results of Multiple Regression Exercise

In Chapter 1 and also in the preceding paragraphs we have outlined the basic principles deemed to be responsible for motivating or regulating the operations on farming units. The paucity of agreement on the guiding principles applicable to the entire farming community was indeed baffling. But Lyudogovskii's classical definition that all agricultural undertakings are described by "the kind and manner of combining quantitatively and qualitatively land, labour and capital" provides the necessary core. Accordingly, we can think of several variants of the income variability model, without disturbing the basic spirit of Lyudogovskii's definition, viz:

A. The Basic Endowment Model (BEM);
B. The Extended Endowment Model (EEM); and
C. The Mixed Model (MM).
The above nomenclature has been adopted to express the essential characteristics of the dependent variables. For instance, BEM—I includes area owned (acres) and number of family farm workers (adult equivalent units) as the only explanatory variables because it is the starting point of a farming unit. Thus we can call it the basic endowment of a cultivating household.

To account for the quality of decision making, Ed.-Head is also included in the original regression model, BEM—I, to obtain BEM—II. This way, moving in a step-wise manner, alternative specifications for these models were made. Inclusion of theoretically sound independent variables for elimination of specification bias and improvement of the explanatory power of the models has been the main guiding principle for trying alternative forms. Due care has been taken in the selection of variables to minimize the incidence of multicollinearity a-la Heady and Dillon (1961, pp.134-136).

In all, eleven regression equations have been estimated incorporating different combinations of the above described explanatory variables. The results of our exercise are presented in Table 4.2. We notice that there is an improvement in the coefficient of multiple determination, $R^2$, at almost every step. Furthermore, all the regression equations are statistically significant as
shown by high F-ratios.

**Remark**: Incidentally, all the F-ratios (say, $F(v_1, v_2)$) in the last column of table 4.2 have very large degree of freedom (say, $v_2$) in their denominators. We could as well apply a limiting test-statistic $\chi^2$ in place of F-statistics.

Though, these F-ratios only permit us to reject the null hypothesis that $\beta_1 = \beta_2 = \ldots = \beta_k = 0$, we also find that the individual regression coefficients are also significant in most cases (as shown by significant t or z values, the latter statistic to be used when the sample size is large) and have the expected signs.

Out of all these estimated equations the following need special mention:

$$\hat{Y} = -4582.49 + 1585.45 \times_1 + 1969.25 \times_3 \quad \text{(29.16)*} \quad \text{(3.57)*} \quad R^2 = 0.72 \quad \ldots(4.2.1)$$

$$\hat{Y} = -9057.29 + 1182.95 \times_2 + 1445.25 \times_3 + 347.49 \times_5 + 786.28 \times_6 \quad \text{(2.04)*} \quad \text{(8.14)*} \quad R^2 = 0.79 \quad \ldots(4.2.2)$$

$$\hat{Y} = -21887.98 + 1270.19 \times_2 + 1362.38 \times_4 + 112.22 \times_5 + 563.50 \times_6 + 2318.47 \times_8 \quad \text{(2.12)*} \quad \text{(6.35)*} \quad (1.79) \quad R^2 = 0.84 \quad \ldots(4.2.3)$$

and

$$\hat{Y} = -26808.19 + 1059.62 \times_2 + 1805.97 \times_4 \quad \text{(14.31)*} \quad \text{(3.60)*} \quad \ldots(4.2.4)$$

$$+ 325.45 \times_5 + 293.40 \times_6 + 4056.47 \times_9 \quad \text{(2.12)*} \quad \text{(3.08)*} \quad (1.07)$$

$$+ 31.56 \times_{10} + 3531.92 \times_{11} + 3542.60 \times_{12} \quad \text{(5.44)*} \quad \text{(2.60)*} \quad (3.01)*$$

$$+ 5.08 \times_{13} \quad \text{(10.53)*} \quad R^2 = 0.87$$
They, respectively, contain two, four, six and nine explanatory variables and the one with the largest number of independent variables has the highest $R^2$ as well.

While, it would be too simplistic to accept equation 4.2.1, it might sound tautological to argue in favour of the relationship with the highest $R^2$. But since equation 4.2.4, which has the highest explanatory power, does not contain redundant or illogical independent variables, it can be considered to be the best fit. More so, it has fairly robust variables (Ref. Table 4.2) and does not inflate the regression estimates, as is the case with equation 4.2.1, by virtue of omitting relevant but not totally uncorrelated explanatory variables. Consequently, we shall retain equation 4.2.4 (second version of the variant-C) for carrying out the decomposition exercise.

This version of the income determination model contains four sets of explanatory variables:

Size Related Variables— Area Operated (acres), full-time farm workers (adult male equivalents), number of milch cattle (standardised units);

Household Characteristic Variables — Education of the head, dependency rate, hierarchy proxy;
Off-farm Employment Mix Variables — Number of household members employed in rural labour market, number of household members employed in the urban labour market; and

Farm Productivity Variable — Value of Farm output (Rupees) per acre.

One can immediately notice the strains of the Russian Organization and Production school, the Neo-classical school and the Marxian school in this model. This had to be so, because the regression equation, as formulated and estimated here, is not simply a technological relation but, as a necessary supplement, contains important behavioural, demographic and socio-economic elements as well.

The resulting estimated equation, as shown in

Table 4.2, comes out to be:

\[ Y = -26808.19 + 1059.62^* \text{(Area Operated)} \]
\[ + 1805.97^* \text{(Number of Farm Workers)} \]
\[ + 325.45^* \text{(Ed.-Head)} + 293.40^* \text{(Number of Milch Cattle)} \]
\[ + 4056.47 \text{ (Dependency Ratio)} \]
\[ + 31.56^* \text{ (Relative Position Proxy)} \]
\[ + 3531.92^* \text{ (Off-farm Rural Employment)} \]
\[ + 3542.60^* \text{ (Off-farm Urban Employment)} \]
\[ + 5.08^* \text{ (Farm Output Per Acre)} \]
with 87 per cent of the variation, in net household income, explained by it. All the estimated coefficients show the expected sign and, barring that for dependency rate, are statistically significant at 5 per cent level.

A perusal of the results, in Table 4.2, makes it obvious that inclusion of theoretically well conceived relevant explanatory variables has enabled us to obtain accurate idea about the contribution of a broad spectrum of factors, besides increasing the explanatory power of the model.

4.4 Decomposition Results—A Close Look

Using the procedure explained in Section 4.2, the above-mentioned regression equation (4.2.4) was deployed to find the estimated contribution of each of the nine variables towards income inequality as measured by the Gini coefficient.

The results are shown in Table 4.3. These findings support the popular belief that widespread inequalities of incomes have their genesis in the unequal distribution of land and other productive assets. We find that over 40 per cent of the said inequality gets attributed to land area with number of farm workers and milch cattle accounting for another 25 per cent and 6 per cent, respectively. Thus, the size related variables taken together explain the bulk of income inequality—a startling seventy one per cent.
<table>
<thead>
<tr>
<th>Description of Control Variable</th>
<th>( G(\hat{Y}_1) )</th>
<th>( G(\hat{Y}) - G(\hat{Y}_1) )</th>
<th>( \frac{G(\hat{Y})}{\sum [G(\hat{Y}) - G(\hat{Y}_1)]} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Operated (Acres)</td>
<td>0.36</td>
<td>0.24</td>
<td>40.68</td>
</tr>
<tr>
<td>Number of Farm Workers</td>
<td>0.45</td>
<td>0.15</td>
<td>25.42</td>
</tr>
<tr>
<td>Number of Milk Cattle (Standard Units)</td>
<td>0.57</td>
<td>0.03</td>
<td>05.08</td>
</tr>
<tr>
<td>Education of the Head</td>
<td>0.59</td>
<td>0.01</td>
<td>01.69</td>
</tr>
<tr>
<td>Dependency Rate</td>
<td>0.60</td>
<td>0.00</td>
<td>00.00</td>
</tr>
<tr>
<td>Relative Position of the Household</td>
<td>0.50</td>
<td>0.10</td>
<td>16.95</td>
</tr>
<tr>
<td>Off-farm Rural Employment (Number)</td>
<td>0.60</td>
<td>0.00</td>
<td>00.00</td>
</tr>
<tr>
<td>Off-farm Urban Employment (Number)</td>
<td>0.59</td>
<td>0.01</td>
<td>01.69</td>
</tr>
<tr>
<td>Productivity (Farm Output Per Acre)</td>
<td>0.55</td>
<td>0.05</td>
<td>08.48</td>
</tr>
</tbody>
</table>

\[
G(\hat{Y}) = \sum_{i=1}^{9} [G(\hat{Y}) - G(\hat{Y}_1)]
\]

Total \( -0.60 \quad -0.59 \quad 99.99 \)

Source: Computed.
Of the remaining variables, the position hierarchy proxy—embodying households' land-man ratio relative to the same ratio obtaining for the village economy—is another major contributor. It accounts for some 17 per cent of income inequality. Thanks, mainly, to the wide diffusion of new technology, the productivity differentials get only 3 per cent of relative contribution towards inequality, associated with them. Education of the head of the household, dependency rate, and off-farm sale of labour prove to be only minor irritants.

In short, size of the holding and other size related variables along with relative resource position in the village prove to be the major factors responsible for the highly skewed income distribution observed in the area. These results compare favourably with a comprehensive study by Bliss and Stern (1982). The authors observed (pp.209-10) that, "......output was roughly proportional to land cultivated", and, "All the variables, land owned, number of adult males, value of animals owned, and number of pumping sets owned, were significant explanations of income." Their message is clear—that causes of income inequality are to be found in the distribution of resources itself. Hence, any commitment for the removal of inequalities of incomes boils down to a pledge for radical redistribution of productive resources. However, such a
blanket approval has to be weighed against the other basic tenets of democracy and distributive justice besides their effect on overall production and also the feasibility of such a move in the changed culture of agriculture as a way of life. In any case, these 'value-laden' questions need a separate discussion and thorough probe. We will not venture into these inconvenient questions at this stage but instead now focus our attention on the twin problem of absolute poverty because a high or low relative poverty, by itself, is no insurance against absolute poverty existing amongst the agriculturists of India's granary—the Punjab. This would be taken up in the Chapter 5 devoted to the measurement of poverty.
Growth and social justice has been a policy goal in each successive five year plan since India's independence because of the country's commitment towards a 'socialistic pattern of society'. Over the years the plan emphasis has shifted from 'rapid growth' to 'growth with redistribution' and finally, to 'faster economic growth with removal of poverty'. For a review of these policies see Bhagwati, Jagdish and Desai, Padma, *India Planning for Industrialisation and Trade Policies*, OECD, Oxford University Press, 1970; Chaudhuri, Pranit, The Indian Economy Poverty and Development, Vikas Publishing House, New Delhi, 1982 pp. 275-311; Raj, K.N., 'Growth Models and Indian Planning' in Wadhwa, C.D. (ed.) *Some Problems of India's Economic Policy*, Tata McGraw-Hill Publishing Company, New Delhi, 1977.


3 'It is unnecessary to recapitulate the causes of debt, but in examining them we saw that debt was allied to prosperity and poverty alike, and that, while its existence was due to poverty, its volume was due to prosperity', wrote Darling (1925, p. 246) concluding his celebrated work, *The Punjab Peasant in Prosperity and Debt*, Geoffrey Cumberlege, OUP, Bombay, Fourth ed. 1947. While the same might be valid even now, commercialisation has added a new dimension to the causes of debt at the top echelon of peasantry, see, Bagai and Soni (1985, op. cit.) for the utilization pattern of credit by various categories of farmers.

Sarma (1981, op. cit.) hints at some of these forces influencing the distribution of land holdings for (on page 25) he writes "These trends are largely the result of demographic pressures, implementation of land reform measures that involve fixation of ceilings on holdings (or anticipation thereof)...." and Michelsen, Norman A., in his 'Landholding, Agricultural Modernisation, and Local Institutions in India', Economic Development and Cultural Change, April, 1984 asserts, "........the analysis presented here suggests that the sources of growing rural inequality in Punjab are quite independent of the newly introduced technology". But we feel that fusion and fission of farm families are equally important determinants of the size distribution of land holdings. Refer to Heady (1964, op. cit.), Chayanov (1966, op. cit.) and Shalin (1972, op. cit.) for rare insight into the internal structure of various types of farms.


8 For a synoptic treatment of the various decomposition procedures and a comprehensive account of the ANOVA procedure, see, Fields (1979, op.cit.).

9 Wise, D.A., in his paper 'Academic Achievement and Job Performance' AER, June 1975 and Fields and Schultz (1980, op.cit.), have used the comparative $R^2$ regression procedure for decomposing inequality.


11 We have already alluded to this aspect in the first Chapter, see note 25 and 47-49, Chapter 1.


13 Infact, the choice of the life-span itself will have a bearing on the definition. If one were to imagine a farm-family surviving for just one agricultural year, then, one does not need to bother about depreciation or outstanding payments i.e., it is the gross entity that is relevant. Prof. Mellor in a personal communication to Robert Stevens (1963) wrote that in the choice of a measure of farm income, the object should be to limit the field as no measure can take care of
all possible purposes. So also we derive justification for this measure from the objectives of the present study.


16 In the recent years, almost all studies on farm management, size and performance have employed net operated area as a measure of farm size (A scanning of literature, particularly IJAE and EPW (Review of Agriculture) for the past ten years gives this impression.

17 See, the reference cited in note 14 for this organic unity between the 'farm' and the 'family'.

18 For the 'life-cycle' hypothesis and related characteristics of farming units, refer to note Chapter-1. Further, see Nelson, J., The Mobility of Farm Families, Manchester University Press, Manchester,
1968, for the relationship between the life-cycle of farm operators and the rate of amalgamation and fragmentation of agricultural holdings.

We feel that 'age of the head', a customary choice for studies involving life-cycle phenomenon, is too poor a representative of the 'age of the household'. This is so because a typical cultivating household has a multitude of families, each at a separate stage of its life-cycle, and therefore, defies simplistic aggregative analysis on the pattern of nuclear families.

This type of feeling, 'that the purely technical divisibility of productivity-enhancing innovations is denied in practice by the social economic, and political disabilities of depressed classes.....', is widespread in the literature. For a good collection of illuminating papers see, Desai, Meghnad, Rudolph, S.H., and Rudra, Ashok (eds.) Agrarian Power and Agricultural Productivity in South Asia, Oxford University Press, Delhi, 1984.


This forms the central theme of Chayanov's work on the organisational structure of peasant households, see Chayanov (1966, op.cit.) Ch.3.

This argument forms an integral part of the Marxian thesis on economic differentiation of peasantry. Bhardwaj's (1974, op.cit.) comment that "There are no motivational forces which are definable a-priori for any operator independently of the entire gamut of production and market relations in which he is involved" can be taken as a typical argument. Also, see Vyas (1979, op.cit.); Joshi (1982, op.cit.); Kurien (1985, op.cit.) and Desai (1984, op.cit.). Lipton (1980, op.cit.) p.248 talks about the big landlord-employer-moneylender's control.

For rural skill drain, see, Lipton (1980, op. cit.) pp.259-269.


For a comprehensive review of all these measures, see, Kahlon, A. S. and Karam Singh, Economics of Farm Management in India: Theory and Practice. Allied Publishers, New Delhi, 1980, Ch.3.

The underlying logic is simple: the state of art on the farm and the quality of resources must show up in the size of output. Further, since both, the crop as well as dairy, activities have been included in the farm output, no direct relationship between size and productivity was observed.


C.f., Chayanov (1966, op. cit.) p.90.
Frequently, $R^2$ is taken to be a measure of goodness of fit and the regression equation yielding the highest $R^2$, corrected for degrees of freedom, is accepted as the best fit. It is only when the dependent variable is not the same or the functional forms are different, that one needs to adopt a different procedure for comparing alternative equations. See, Rao, P. and Roger Le Roy Miller, *Applied Econometrics*, Prentice-Hall, New Delhi, 1972, pp. 16-20.

In the choice of these variables, we have depended upon all the three schools mentioned here. For details one can see Heady and Dillon (1961, op.cit.); Heady (1964, op.cit.); Chayanov (1966, op.cit.); Bhardwaj (1974) and Desai et.al. (eds.) (1984, op.cit.).

This view, well supported by empirical findings, is held by a large majority of Economists, irrespective of their ideological background. See, for instance, Vyas (1979, op.cit.); Shah (1981, op.cit.); Bhalla and Chadha (1981, op.cit.); Joshi (1982, op.cit.); Bliss and Stern (1982, op.cit.) and Tendulkar (1983, op.cit.).