

## Some conceptual and statistical issues on measurement of poverty

T. Krishna Kumar<sup>a,\*</sup>, Anil P. Gore<sup>b</sup>, V. Sitaramam<sup>c</sup>

<sup>a</sup> *Economic Analysis Unit, Indian Statistical Institute, Bangalore-560059, India*

<sup>b</sup> *Department of Statistics, University of Poona, Pune-411007, India*

<sup>c</sup> *Biotechnology Unit, Zoology Department, University of Poona, Pune-411007, India*

Received 2 February 1993; revised 30 October 1993

---

### Abstract

One of the major topics that attracted the attention of econometricians in recent years is measurement of poverty. This paper reviews critically the conceptual and statistical issues that have been examined by the econometricians. The paper provides a comprehensive review of major recent approaches and results on measurement of poverty. It devotes one section to outline a new approach to the measurement of poverty that is based on the actual consumption behaviour of the people instead of on arbitrary choice of either a poverty line or a deprivation function. It devotes two sections for suggesting fruitful areas of research, one addressed to economists on synthesizing poverty measurement with applied welfare economics, and another to statisticians on problems of statistical inference associated with functional estimation. The paper also highlights the importance of reliability theory and risk assessment in translating consumption deprivation into a poverty measure. A new index of poverty that depends on risk of consumption deprivation is also proposed.

*AMS Subject Classification:* Primary 90A19; Secondary 62F10

*Key words:* Poverty; Inequality; Income distribution; Consumption deprivation; Functional estimation; Hazard rate; Reliability theory

---

“..... important conceptual issues in the measurement of poverty remained undiscussed for too long. What was needed was a greater degree of vertical integration between the statistical measurement of poverty on the one hand and welfare economics on the other”.

A.B. Atkinson (1987)

## 1. Introduction

It was nearly ten years ago that one of us overheard a conversation between two professional economists, both concurring with the view that 'poverty' is an over-worked topic. The rate at which economic journals are currently publishing articles in this field testify that, contrary to one's expectation, it is an area that has great potential for further research. We feel that there are at least three specific aspects of poverty which have not been adequately explored in this allegedly 'over-worked' topic. These three aspects are: (i) overcoming the arbitrariness and subjective elements in the choice of the poverty index, (ii) synthesizing the poverty measurement with applied welfare economics and public policy, and (iii) statistical issues relating to measurement of poverty index based on sample survey data on income and expenditure distributions. This review of econometric methodology pertaining to poverty measurement is therefore to be regarded as a review by late comers to the field with emphasis on the three specific issues just cited that have a great potential for future research. Several important contributions to the poverty literature such as empirical analysis of poverty indices and their variation are excluded in this review in order to focus on issues relating to methodology and in order to maintain a moderate length for the paper.

This paper is organized as follows. The next section introduces the basic conceptual issue — 'what is a Poverty Index?'. Section 3 traces methodological developments in poverty measurement upto the development of an axiomatic approach by Sen (1973). Section 4 deals in some detail with some recently developed measures of poverty. Section 5 offers a critique of the existing methods of measuring poverty and suggests an alternative approach. In Section 6 measurement of poverty is viewed from a policy perspective and it is integrated with applied welfare economics and public policy. Finally, Section 7 presents a few statistical issues such as welfare comparisons of income distributions, efficient estimation of poverty index, incorporation of variability of consumption deprivation in devising new measures of poverty.

## 2. What is a poverty index?

The question of what a poverty index is, cannot be answered unless we state clearly what is meant by poverty of an individual. Poverty connotes the notion of a poor state of economic well-being or a state of economic ill-being. It connotes a state of economic deprivation. Deprivation can be based on comparing an individual's economic state with either an absolute norm, in which case it is called an absolute deprivation, or a normative or relative norm, in which case it is called a relative deprivation. An individual's economic state can have several dimensions. Hence the notion of economic deprivation implies that the individual is comparing, or introducing a partial ordering on, various economic states he is confronted with. Thus, the notion of economic deprivation of an individual is closely related to the partial

ordering of economic states by an individual. The economic state can be redefined in terms of three states, better than the norm, worse than the norm, or equivalent to the norm. From these states one can say whether a person is deprived or not. It is thus clear that the norm used for determining deprivation has to be specific to each individual or each specific group of individuals. For example, this norm should be very low for a priest who vows to live in poverty and thus chooses poverty voluntarily.

When there are several individuals in a community, it is necessary to arrange them into different groups such that persons within a group can be expected to have the same set of preferences or partial ordering and persons in different groups have different partial orderings. Within each homogenous group one can determine the number and/or proportion of persons who are deprived. There will be as many such deprivation indices, viz., the proportion of persons who are deprived, as there are groups. The question then arises as to how one can combine the group deprivation indices into a single community deprivation index. Such a conversion of group deprivation indices to a single deprivation index requires a partial ordering of economic well-being of different groups — viz., a social welfare function.

It was assumed above, just for convenience, that each individual's economic state is represented in two categories — deprived, not deprived. But one can also introduce a partial ordering of different economic states of an individual relative to an absolute or a relative norm. Then under certain regularity conditions on the individual's preferences an individual's deprivation can be represented by a deprivation function as demonstrated by Debreu (1959). Using an analogous reasoning, if the social partial ordering of group deprivations satisfy certain regularity conditions there exists a community deprivation function.

A group poverty index can be defined as the mean level of deprivation for that group. The community poverty index can be defined as a 'subgroup consistent' aggregate of the various constituent group poverty indices. The meaning of the term subgroup consistency and the problem of aggregation and decomposition of poverty indices are dealt with in detail in Section 4.

### 3. Early developments in measurement of poverty

As indicated in the previous section measurement of poverty depends on a norm with respect to which the economic state of an individual is compared. The standard approach which goes at least as far back as Rowntree (1901) is to define a poverty line in terms of a minimum level of income needed to purchase the basic necessities of life and use the income distribution to see what percentage of the people have an income less than such a poverty line. This measure is called Head Count ratio ( $H$ ). Some other contributors who used this approach are Bowley and Burnett-Butt (1915), Townsend (1954), Weisbrod (1965) and Atkinson (1970b). Another measure which also depends on the poverty line and income distribution is the poverty-gap used by Batchelder

(1971). It is the aggregate income short-fall from the poverty line of all those persons whose incomes are below the poverty line. A slightly modified and normalised version — viz., per person percentage gap of all the poor, called the income-gap ratio,  $I$ , is in popular use.

Suppose there are  $n$  individuals with incomes  $y_1, y_2, \dots, y_n$  arranged in an increasing order and let  $z$  be the poverty line. Let  $q$  be the maximum index  $i$  such that  $y_i \leq z$ . Then,  $q$  denotes the number of the poor. The head-count ratio  $H$  is given by

$$H = q/n. \quad (3.1)$$

Let  $g_i$  be the income gap defined as  $g_i = z - y_i$ . Then the per person percentage gap of all the poor,  $I$  is given by

$$I = \frac{1}{qz} \sum_{i \leq q} g_i = \frac{1}{q} \sum_{i \leq q} \left( \frac{g_i}{z} \right) = \frac{1}{q} \sum_{i \leq q} \frac{(z - y_i)}{z}. \quad (3.2)$$

If one defines the mean income of the poor as  $m$  then

$$I = \frac{z - m}{z}. \quad (3.2a)$$

The numerator of (3.2a) represents the minimum cost of poverty alleviation under the assumption that the policy maker knows who the poor are and what their incomes are. The denominator represents the maximum cost of poverty alleviation if the policy maker only knows who the poor are and does not know what their incomes are and he needs to give away a per capita income of  $z$  to all the poor. It is clear that the head count ratio gives only a count of the poor and it gives no account of how poor the poor are relative to the poverty line. The income-gap ratio tells how poor the poor are, as a group, in relation to the poverty line. But neither of these indices give any idea of how the income is distributed among the poor. It is desirable that the poverty index reflect the inequality of income distribution among the poor. For this one can use an inequality index introduced by Lorenz (1905) and Gini (1912). This index is given by:

$$G = \frac{1}{2q^2m} \sum_{i=1}^q \sum_{j=1}^q |y_i - y_j|. \quad (3.3)$$

One may conclude that the best thing to do for measuring poverty is to present the triplet  $H, I$ , and  $G$ , where  $H$  tells us how many persons are poor,  $I$  tells us by how much the mean income of the poor falls short of the poverty line, and  $G$  tells us the inequality of the income (and income gap) among the poor. The first item measures the extent of poverty, the second measures severity of poverty and the third measures the distribution or incidence of poverty. The weights a policy maker attaches to these three components must be commensurate with the policy maker's concern for these three policy dimensions of poverty — extent, severity and distribution or incidence.

There was, and there has been, a considerable debate on how to determine a poverty line. This debate is quite significant because the poverty indices depend on this choice of the poverty line. Here there are three different points that one may note.

First, the poverty line need not have a scientific basis. It can be chosen administratively using certain objective criteria.<sup>1</sup> Second, one can possibly consider a ranking of people according to the degree of deprivation in such a way that the ranking is invariant to the choice of the poverty line. This suggestion was made by Atkinson (1987). Third, there may not be a unique deterministic minimum requirement on the basis of which a poverty line can be determined. This last point is closely related to a criticism made by Sukhatme (1978) of a method which uses an average requirement in choosing the poverty line.

Dandekar and Rath (1971) used a nutritional norm and then used the minimum total expenditure needed to meet that norm as the poverty line. For this they used the *average* daily energy requirements of an individual of a given age, sex, body weight, and physical activity. Sukhatme (1978) proposed that one must instead use a *minimum* daily requirement rather than an *average* requirement. He had also argued that the energy requirements of an individual would follow a first order autoregressive stochastic process with a constant variance. Survival risk depends not on nutritional deficiency on a given day. Instead, it depends on a history of nutritional deficiency. A part of this history is contained in the individual's body weight and state of health, and the rest is contained in the cumulative nutritional deficiency over the recent past. This is an important point to note as it leads to certain statistical issues that will be dealt with later in Section 7. In particular, if the actual energy intake and the required energy intake are both random the nutritional deficiency is determined by the probability that actual energy intake falls below the requirement. The survival risk due to nutritional deficiency is then a function of the probability that the actual energy intake falls short of the requirements over a period.

#### 4. Later developments in measurement of poverty

Sen (1974) used an axiomatic approach for developing a normative poverty index along the lines of his earlier axiomatic approach for arriving at normative measures of income inequality (Sen, 1973). He extended this work by developing a new ordinal approach to the measurement of poverty (Sen, 1976). It was mentioned in the previous section that a poverty index can be derived from three indices  $H$ ,  $I$  and  $G$ . Sen introduced three axioms that a poverty index must satisfy and from them he deduced that the only poverty index that satisfies those three axioms is of the form:

$$P_S = H\{I + (1 - I)G\} \quad (4.1)$$

There are three other major contributors besides Sen who made significant contributions to the development of poverty indices. These are Atkinson (1970a, b, 1987),

<sup>1</sup> Usually there is a hierarchy of needs that are to be satisfied in a given order. Which items of this hierarchical need structure should enter into the minimum needs that are to be met may vary from society to society. Similarly the desired levels of consumption also may vary from society to society. Hence, the choice of poverty line may be different for different societies.

Kakwani (Kakwani and Poddar, 1976; Kakwani, 1980a, b), and Foster (Foster, Greer and Thorbecke, 1984; Foster and Shorrocks, 1991; Foster, 1984). Some of these developments are described below in a modified form and in some detail as they lead us to some useful and researchable issues that will be discussed in Sections 5–7.

Let us assume that deprivation and poverty are measured in terms of income. Let  $z$  be the poverty line and let  $y$  be the income of a typical individual. Let  $f(y)$  denote the density while  $F(y)$  is the cumulative distribution function associated with the income distribution. Following Sen (1976) and Kakwani (1980a) the poverty index may be expected to satisfy the following three axioms:

*Monotonicity axiom:* Given other things, an increase in income of a person below the poverty line must decrease the poverty.

*Transfer axiom:* Given other things, a pure transfer of income from a person below the poverty line to any other person with a higher income must increase the poverty.

*Transfer sensitivity axiom:* Given other things, if a transfer  $t > 0$  of income takes place from a poor household with per capita income  $y$  to another poor household with per capita income  $y + d$  ( $d > 0$ ) then the magnitude of increase in poverty decreases as  $y$  increases.

From the earlier discussion in Section 2 it is clear that an individual's deprivation  $d$  can be expressed as a function of either absolute or relative deviation of his income from the poverty level:

$$d = g(z - y), \quad (4.2a)$$

$$d = g\{(z - y)/z\}. \quad (4.2b)$$

In general terms the deprivation can be expressed as a function of  $y$  and  $z$ :

$$d = d(y, z). \quad (4.2c)$$

Atkinson (1987) has shown that most of the poverty indices proposed in the literature can be represented by a poverty index defined as follows:

$$P_A = \int_0^z d(y, z)f(y) dy. \quad (4.3)$$

From Fig. 2 in Atkinson's paper, and from Table I of the same paper it is clear that in all those cases the deprivation function is a non-increasing function of  $y$  and convex to the origin, or strictly convex to the origin. It can be shown that Atkinson's class of poverty indices satisfy the three axioms listed above if the deprivation function is a decreasing function of  $y$  and strictly convex (Kumar, 1993).

It is often desirable, either due to analytical necessity or convenience and for policy purposes, to divide the entire population into different groups of persons and define

poverty index for each group. This gives rise to the basic conceptual problem of consistent aggregation or decomposition. In order to tackle this issue Foster and Shorrocks (1991) introduced another axiom termed 'subgroup consistency' which only means that the concept of monotonicity applies to groups of persons, viz., if the poverty index of all groups except one group, group  $i$ , remains the same and if the poverty of group  $i$  increases then the aggregate poverty index must increase.

Suppose that we partition the people into a finite number of mutually exclusive and collectively exhaustive groups  $G_1, G_2, \dots, G_k$ . These may refer to different regions, different income groups, different ethnic groups, etc. Assume that  $w_i$  is the proportion of households who belong to group  $i$  and  $f_i(y)$  and  $F_i(y)$  refer to the conditional density and conditional cumulative distribution function of income distribution, respectively, associated with group  $i$ . Let  $d_i(y, z)$  denote the deprivation function associated with group  $i$ . We can then define a sub-group poverty index  $P_i$  as follows:

$$P_i = \int_0^z d_i(y, z_i) f_i(y) dy \tag{4.4}$$

The interesting question to pose is: how can one combine the sub-group poverty indices into an aggregate poverty index so that aggregate poverty index satisfies the three axioms listed earlier and also the sub-group consistency axiom of Foster and Shorrocks (1991). Let the aggregate poverty index  $P$  be given by

$$P = \int_0^z d(y, z) f(y) dy \tag{4.5}$$

where  $z$  and  $d$  are yet unknown and  $f(y)$  is given by

$$f(y) = \sum_{i=1}^k w_i f_i(y) \tag{4.6}$$

If one notes that  $d(y, z)$  is consumption deprivation then what is needed is a condition for consistent aggregation or consumption over individuals that will give rise to the following consistency relation:<sup>2</sup>

$$d(y, z) = \sum_{i=1}^k d_i(y, z_i) \tag{4.7}$$

where  $z = \max\{z_i\}$

If the above condition is satisfied then employing (4.6) and (4.7) one can write:

$$P = \int_0^{z - \max z_i} d(y, z) f(y) dy + \sum_{i=1}^k w_i \int_0^{z_i} d_i(y, z_i) f_i(y) dy \tag{4.8}$$

---

<sup>2</sup> Here it is being assumed that  $d(y, z_i) = 0$  for  $y > z_i$ . For conditions on consistent aggregation of consumption expenditure the reader may refer to Muelbauer (1975).

i.e.

$$P = \sum_{i=1}^k w_i P_i. \quad (4.9)$$

A poverty index that satisfies (4.9) is called (an additively) decomposable poverty index.

There is no compelling reason why one should have this *additive* separability or *additive* decomposability of the poverty index. One can define poverty index for each homogeneous group. Such individual group poverty indices can be combined into an overall poverty index in more than one way satisfying the axiom of subgroup consistency. Two points that must be noted here are that  $d_i(y, z_i)$  is the mean deprivation of group  $i$  and that each group may have its own poverty line. These two points are related to the issues raised by Sukhatme and these are quite important for some statistical issues raised in Section 7. The analysis given above is a synthesis of ideas and results contained in Atkinson (1987), Foster and Shorrocks (1991) and Kumar (1993).

## 5. Poverty measure based on consumption deprivation

The literature on measurement of poverty reviewed in the previous section suffers from a few major shortcomings. A large segment of the economics profession seems to operate with a few gestalt switches, such as poor and non-poor, deprived and not-deprived, poverty line, etc. The prevalence of these terms in the economics literature, and the profession accepting them unquestionably, seems to suggest the fallacy of reification — i.e., since something is merely named, it exists. Some of the gestalt switches must be abandoned because they do not lead to measurable concepts. Others must be rejected as they are inconsistent with empirical observations. The notion of poverty which connotes a fine distinction between the poor and non-poor is one such concept. Poverty line is another concept. Both of them can be dispensed with.

There is no qualitative difference between a person just below the poverty line and the one just above the poverty line. If there is such a difference the expenditure pattern between the two, on either side of the line should differ discernibly. But the observed consumer expenditure do not support such a phenomenon. There is no biological basis for the poverty line either in that it represents a critical food requirement or a debilitating deprivation of a sharp nature. These concepts of poverty line and the poor and non-poor must therefore be rejected. On the otherhand, it is quite meaningful to group people in such a way that within each group the consumer behaviour is quite similar and between groups it is dissimilar. For each such group one can observe the pattern of consumption expenditure. A detailed empirical and theoretical investigation made by the authors, which will be reported shortly in another paper, reveals that consumption expenditure equilibrium for



any commodity can be regarded as a steady state equilibrium in response to an income stimulus. This steady state equilibrium is being pulled up by a desire to consume as close to a saturation level as possible while it is pulled down by a shortfall in income. As this phenomenon bears a close resemblance to fluid dynamic equilibrium and as hyperbolic curves represent fairly well such equilibrium responses we postulated that consumer expenditure equilibria can be represented by hyperbolic relations.<sup>3</sup>

In particular, if one considers the consumption expenditure on foodgrains it represents a hyperbolic relation of the form

$$C = \frac{Vy}{K + y}, \quad (5.1)$$

where  $C$  is the consumption expenditure on foodgrains and  $y$  is the total expenditure (a proxy for income that is not observed). This is the equation for a rectangular hyperbola with  $V$  and  $-K$  as its asymptotes.  $V$  represents the saturation level of foodgrains consumption expenditure. From this observed behaviour one can define, quite objectively, consumption deprivation for foodgrains as the shortfall of the actual consumption expenditure from the saturation level  $V$ .

For a group of  $n$  consumers one can then define a normalized per capita consumption deprivation for foodgrains as:

$$D = \frac{1}{nV} \sum_{i=1}^n \left[ V - \left( \frac{Vy_i}{K + y_i} \right) \right] = \frac{1}{n} \sum_{i=1}^n \left( \frac{K}{K + y_i} \right). \quad (5.2)$$

It can be easily verified that this foodgrain consumption deprivation index lies between 0 and 1. It assumes the lowest value 0 for infinitely large values of income and the highest value 1 for zero incomes.

If one were to measure the degree of poverty among a homogeneous group this foodgrain consumption deprivation index can be employed as the poverty index for that group. It may be emphasized that (5.2) used as a poverty index has the following desirable properties which the other indices reviewed in Section 4 do not have: (i) It does not depend on an arbitrarily chosen poverty line, (ii) There is no abrupt distinction between the poor and the non-poor, and (iii) This index depends on the observed and measurable behaviour of the people. This index also satisfies the three axioms that a poverty index must satisfy, as deprivation function is a decreasing function of  $y$  and it is convex (Kumar, 1993). Since this index is

<sup>3</sup>One of the authors (V.S.) used such hyperbolic curves in his research on biochemical phenomena. He postulated in early eighties a similar model for consumer expenditures while examining the National Sample Survey data and National Nutrition Monitoring Bureau data. In actual practice the sample data may give rise to non-hyperbolic curves for some categories of expenditures. This only means that the observed incomes and behaviours are at lower ranges, too far from saturation levels.

related to consumption and it is derived from the familiar Engel curves of economic theory of consumer behaviour it is possible to treat poverty as the opposite of welfare and integrate studies on poverty, poverty alleviation programs and policies with welfare economics.

## 6. Poverty measurement and applied welfare economics

By its very nature measurement of poverty was motivated by welfare economics because it is through such measurement that one can identify the need for welfare measures and also devise suitable methods for needs targeting (Kanbur, 1987; Keen, 1992). However, different methodological approaches were followed in poverty measurement, needs targeting, and other applied welfare economics. In particular the approaches of poverty measurement invariably depended on income distribution and a deprivation function. Recent literature on needs targeting (Kanbur, 1987; Thorbecke, 1989; Keen, 1992) that used such poverty indices were hence focused on changing the incidence of poverty through direct income transfers. The arbitrary choice of deprivation function based on certain axiomatic considerations has placed poverty-based welfare economics into an altogether different compartment from other microeconomic theory based welfare economics. In the latter type of analysis one would normally take social welfare as a function of levels of individual consumption, and this level of individual consumption as a function of income and prices confronting the individual.

A synthesis between these two seemingly different approaches is possible by noting, as mentioned in the previous section, that consumption deprivation can be taken as the deprivation function in the poverty index. Bhanoji Rao (1981) had already drawn the attention to the relation between consumption deprivation and poverty. Since for necessities, such as foodgrains, the Engel curves are increasing functions of income and strictly concave, it follows that the deprivation function so defined is a decreasing function of income and strictly convex.<sup>4</sup> Thus, the poverty index based on consumption deprivation of necessities satisfies the three axioms suggested by Sen (1976), and Kakwani (1980a). The parameters of Engel curves are functions of own and cross prices. If one writes the consumption deprivation functions in full, using aggregate consumer demand functions for necessities, they will be functions of prices and income. One can then devise policies through which deprivation can be reduced employing price and income policies. Such policies affect both the level of consumption deprivation and the resulting income distribution, these two being the major components of the poverty index.

<sup>4</sup> In the consumption basket of persons who are near the officially defined poverty line of India such essential items or necessities constituted in 1983-1984 more than 75% of the total consumption expenditure. See Technical note on the Seventh Five Year Plan, p. 10.

Such a synthetic approach to applied welfare economics can be illustrated by considering the following enlarged definition of the poverty index

$$P = \sum_{i=1}^k w_i \int_0^{z_i} d_i(y, z_i, p_i) f_i(y, p_i) dy, \tag{6.1}$$

where  $p_i$  is a vector of prices encountered by persons in group  $i$ .

Let  $f_{i0}(y, p_{i0})$  denote the distribution of income in group  $i$  prior to introduction of public policy measures, some of which are directly meant to be poverty alleviation programs while others are policies that do have an impact on the poverty level. Let  $f_{ia}(y, p_{ia})$  denote the distribution of income of group  $i$  after the introduction of public policy measures. The policy measures are in general aimed at asset redistribution policies, changing factor incomes, pure income transfer policies, and price policies such as rationing and public distribution, etc.  $p_{i0}$  and  $p_{ia}$  represent the price vectors without and with new policy interventions. It is possible that there are direct and indirect effects of programs. The poverty levels can be favourably affected by the poverty alleviation schemes but they can be adversely affected by other public policies. Hence one must consider the over-all welfare impacts of all public policies put together. Such an evaluation is eminently carried out by describing the functioning of the economy as a temporary equilibrium model with price controls and rationing. One can consider an extended input output model with different types of households as different endogenous sectors that supply factor inputs to the other sectors and receive factor incomes. Similarly, in this set-up one must include government as a separate sector that provides government services to various sectors, including different household sectors, and it receives as inputs taxes and services of factors.<sup>5</sup>

The welfare economic problems can then be posed as follows: Minimize

$$P_a = \sum_{i=1}^k w_{ia} \int_0^{z_i} d_i(y, z_i, p_{ia}) f_{ia}(y, p_{ia}) dy, \tag{6.2}$$

subject to the conditions

$$P_a \leq P_0 = \sum_{i=1}^k w_{i0} \int_0^{z_i} d_i(y, z_i, p_{i0}) f_{i0}(y, p_{i0}) dy, \tag{6.3}$$

$$x_d = Ax + Bx + e, \tag{6.4}$$

$$y_{ia} = r'_{ia} p_{ia} + t_{yi} + t_{ai}, \tag{6.5a}$$

$$y_{i0} = r'_{i0} p_{i0}, \tag{6.5b}$$

$$p_{ia} = p_{i0} + D_i(x_d - x) + p_{qi} \tag{6.6}$$

<sup>5</sup>It is no doubt a very difficult task to allocate between various sectors the public goods that the government produces. One approach is to treat public goods as a factor of production that enters into all sectors at the same levels.

where  $x_d$  is a  $n \times 1$  vector of demand for outputs of the various sectors while  $x$  is the supply;  $A$  and  $B$  are  $n \times n$  matrices of current and capital coefficients,  $e$  is  $n \times 1$  vector of net imports,  $D_d$  is a  $n \times n$  diagonal matrix of coefficients that reflect the speed of market adjustments between demand and supply, the diagonal elements of  $D_d$  can be zero for those sectors whose price is controlled.  $r_{i0}$  denotes the initial endowment of factors of the  $i$ th person while  $r_{i\sigma}$  is the final endowment of factors. Some of the additions to factor endowments is through markets for factor services while other additions are due to asset distribution policies. Thus in Eq. (6.5a)  $t_{pi}$  is a pure transfer income while  $t_{ai}$  is a transfer of an asset (its income equivalent). In Eq. (6.6)  $p_{gi}$  is that component of change in price which is set through government's price control. In the above optimization problem the government policies regarding choice of elements of the price vector  $p_{gi}$  and its activity levels (certain components of  $x$  vector denote the activities of the government) are instruments of policy,  $e$  is exogenous and all other variables are to be determined endogenously.

Thorbecke (1989) presented this kind of model to assess the impact of programs targeted for poverty alleviation. His approach also included both the direct and indirect impacts employing the Social Accounting Matrix (SAM). But his approach involved estimating the income-equivalence of such direct and indirect impacts and including the post-policy incomes in a Foster-Greer-Thorbecke index of poverty (Foster et al., 1984). What is suggested here is a more direct approach in which the policy impacts, both direct and indirect, are traced through a computable general equilibrium model. The resulting consumption levels are used to derive consumption deprivation and the consumption deprivations are used to measure poverty.

## 7. Statistical issues on measurement of poverty

There are mainly three distinct ways in which statistical issues arise in measurement of poverty. First, since the poverty index depends on the density of income some problems confronting poverty measurement are similar to the statistical problems associated with probability distributions. Second, as we do not have, except in rare occasions, a census of incomes poverty measurement is quite often made employing information on income distribution provided by sample surveys. This gives rise to estimation of population poverty index using a sample. Third, following the critical contribution of Sukhatme one can interpret deprivation, whether it is a nutritional deprivation or a consumption deprivation, as not being deterministic but as being stochastic. Then even if we have a census of incomes the poverty index must be based on a stochastic deprivation.

Some of the statistical issues first arose with the measurement of inequality. Since there is a close relation between income inequality and poverty it is useful to cover the methodological issues relating to income inequality also. One early application of properties of density functions was by Levine and Singer (1970). They consider the closed-form expression for the income inequality in terms of the income density

function. Using an exponential distribution for incomes they reach the conclusion that under such an income distribution a proportional tax does not change income inequality but if a lumpsum tax is imposed after a proportional tax then the effect of that on income inequality depends on the proportional tax also. The usual definition of Lorenz curve is in terms of two equations both having a common parameter (decile or fractile) as given in Kendall and Stewart (1969). Gastwirth (1971) presents the equation of Lorenz curve as a single function of the percentile employing the inverse function of the cumulative distribution function. Such an approach provides an elegant analytic scheme for studying the effects of income transfers, through taxation, on income inequality. With the PC revolution and the availability of software for numerical integration this technique has great potential applications for studies on impacts of taxation on income inequality and welfare.

Another study that has a significant potential for future research on measurement of poverty is by Singh and Maddala (1976). They noted that the two distributions that are quite often used in income distribution studies, viz., the Pareto and the log normal, were not quite suited for the graduation of incomes. These authors capitalized on the well-known result that the larger the income the more is the probability of having more income. This phenomenon is just the opposite of the phenomenon of probability of survival with aging. The authors argued that upto a point, i.e., at lower levels of income, the failure rate or hazard rate defined by  $f(y)/\{1 - F(y)\}$  is increasing and then it is decreasing. And they also noted that this property is shared by the log normal distribution of incomes. This approach of Singh and Maddala to relate probability distributions in economics to those in reliability theory of Barlow and Proschan (1965) has a great potential for further research on poverty measurement. This is because the deprivation function used in the poverty indices must be related to failure rate or hazard rate that measures the risk of survival. We shall return to this topic when we discuss the stochastic nature of deprivation.

There were quite a few studies that addressed to the question of dominance of one income distribution over another in terms of the dominating distribution giving rise to an unambiguously more equal income distribution than the other (Atkinson, 1970a; Dasgupta et al., 1973; Rotschild and Stiglitz, 1973; Kanbur and Stromberg, 1988). Bhatty (1974) and Moothathu (1991) examined a similar issue of deriving conditions for one Lorenz curve being entirely above another Lorenz curve. Monotonicity axiom plays a central role in measurement of poverty. This axiom imposes certain conditions on the cumulative distribution function as shown by Atkinson (1987) and Kumar (1993). Spencer and Fisher (1992) point out that this condition is same as the condition of stochastic ordering or majorization. Hence, there is a need to further explore in detail the applicability of the concepts of stochastic dominance and majorization developed by Marshall and Olkin (1979) to the problems of measuring and comparing poverty.

The problem of utmost significance in devising sustainable poverty alleviation programs is to treat income over individuals and time as a multivariate stochastic

process and to derive conditions under which one stochastic process dominates another in terms of reducing the poverty index over a period. Kanbur and Stromberg (1988) examined this problem for a particular case in which a conditional transitional density was used. Employing the statistical theory of stochastic dominance and majorization of Marshall and Olkin (1979) some more research can be undertaken on this topic which has great policy relevance.

Most of the applied and empirical work on poverty measurement that one comes across is of the 'economic statistics' variety such as construction and comparison of economic indices. Several authors had computed poverty indices over time and space and made comparisons not realising or ignoring the fact that such indices were constructed on the basis of sample income distributions. They are therefore only estimates of the underlying population poverty indices.

Iyengar (1960, 1964) derived a consistent estimator of the Lorenz ratio under the assumption that the underlying income distribution was log normal and that this sample distribution was available in grouped data form. Similarly Maiti and Pal (1988), and Kakwani and Poddar (1976) developed efficient methods of estimating the population Gini coefficient using grouped frequency distributions. However, it is very rare to find instances where a comparison of estimates of Gini coefficient over time or space is done against the standard errors of such estimates. Similar statement can be made regarding the comparison of estimates of poverty indices over space and time. There is a need to judge whether the observed differences in estimates of poverty over time and space or groups of persons are significant in relation to their standard errors which reflect the inherent variation attributable to mere random variation. One important statistical problem that needs immediate research attention is the problem of estimation of population poverty index employing grouped sample frequency distribution of incomes with a deterministic deprivation function. In this case the population poverty index can be written as:

$$p = \int_0^{\infty} d(y, z) f(y) dy \quad (7.1)$$

Its estimate, based on grouped sample frequency distribution, can be written as

$$\hat{p} = \int_0^{\infty} d(y, z) f_n(y) dy, \quad (7.2)$$

where  $f_n(y)$  is an estimated sample density function. The sample density function can be estimated either parametrically, assuming a specific functional form for the density with unknown parameters or non-parametrically employing any one of the non-parametric Kernel estimators. If one interprets  $d(y, z)$  as the conditional mean of deprivation given an income of  $y$  then the poverty index  $P$  of (7.1) has the interpretation of unconditional mean deprivation. One can then treat (7.1) as being quite similar to a non-parametric regression and employ methods suggested by Nadaraya (1964) and Watson (1964). More importantly, one needs to know the properties of estimators  $\hat{P}$  of (7.2) under various types of density estimators  $f_n(y)$ .

If a parametric approach is followed assuming a specific functional form for  $f(y)$  in (7.1) then one can derive an estimator of  $P$  based on appropriate sufficient statistics of the parameters of  $f(y)$  thereby assuring oneself of consistent and efficient estimator. If analytic methods become intractable for drawing inferences on the sampling distribution of the poverty estimator given by (7.2) one can employ bootstraps technique of Efron (1982) to generate replicated samples and generate sampling distributions of the poverty estimators. If non-parametric approach is adopted for estimating  $f_n(y)$  then the problems of inference under that set-up are more challenging. One may refer to Prakasa Rao (1983) for inference in situations of non-parametric functional estimation.

There are some attempts in the literature for 'correcting' or improving the estimate of Gini coefficient which was calculated from grouped sample frequency: Gastwirth and Glauberman (1976) proposed an interpolation formula employing Hermite interpolation. This approach is of the nature of curve-fitting and smoothing rather than efficiently estimating the Gini coefficient. Suryanarayana (1991) compared the Gini coefficients estimated by the traditional trapezoidal method and by the Kakwani and Poddar procedure and concluded that the estimates provided by the trapezoidal method underestimate the Gini index and that the estimate was sensitive to the method of estimation. It is possible to apply Gastwirth's method to the NSS data to improve the estimates of the Gini coefficient. But it is more interesting to assume a parametric form to the income density function and to obtain an efficient estimator of the Gini coefficient based on sufficient statistics of the parameters of the density.

Suryanarayana and Geetha (1992) estimate Foster, Greer, and Thorbecke's poverty index using grouped sample income distribution and specifying a two parameter log normal distribution. The authors, while realizing this is a sample estimate of the population poverty index do not examine the properties of this estimator and its sampling distribution. Suryanarayana and Geetha express the Foster, Greer, Thorbecke index as a function of the moments of the log normal distribution. This approach opens up the doors for a series of interesting statistical problems. Some of these may be outlined here.

First, any deprivation function  $d(y, z)$  can be approximated by a sufficiently high degree polynomial in  $y$ , for a given  $z$ . Hence any general poverty index can thus be expressed as a function of the moments of the income distribution. Second, even if a given deprivation function cannot be easily approximated by polynomial of a low degree, since the poverty index is only ordinal, one can ask if there is any monotonic increasing function of  $d(y, z)$  that can be approximated by a low degree polynomial. Third, since the statistical literature provides efficient methods for estimating the moments, one can take such estimates of the moments and derive efficient estimators of the poverty index. For instance one can take the maximum likelihood estimators for the moments. Then poverty index can be estimated as a function of such consistent estimators. Such an estimator of the poverty index which is a function of the maximum likelihood estimators itself will be consistent and efficient under certain very general conditions.

Many practical problems of poverty alleviation can be converted into policies aimed at modifying the income distribution so as to reduce poverty. Since the poverty index can be expressed in terms of the moments of the income distribution the policy problem reduces to that of choosing public policies that affect the moments of the income distributions in suitable ways. One can generalize the concept of regression which is an expression for conditional expectation or conditional mean. Similarly one can define higher order regressions representing conditional variance and conditional third moment, etc. These moments of income distributions can be expressed as functions of certain factors such as public policies aimed at poverty alleviation and other market forces. We thus have a hierarchy of regression:<sup>6</sup>

$$E(y|X) = h(X; \beta), \quad (7.3a)$$

$$E[\{y - E(y|x)\}^2|X] = k(X; \gamma), \quad (7.3b)$$

⋮

From the coefficients of these hierarchy of regressions one can devise suitable policies such as altering the values of certain components of vector  $X$  which are under the policy makers' control. It is possible that some policies increase not only the mean of the income but also variance while some policies change only the mean and do not change the variance. This type of knowledge will help in designing optimum policies for poverty alleviation.<sup>7</sup>

The poverty indices defined in Section 4 assumed a deterministic deprivation function. Very little attention is paid in the literature regarding operational procedures for deriving the deprivation function. It was suggested earlier that consumption deprivation can be taken as the deprivation function. The question then arises as to whether consumption deprivation function can be regarded as deterministic. Consumption deprivation in fact should be regarded as stochastic since the amount of consumption expenditure for a given amount of income is not deterministically known. The usual interpretation of the consumption deprivation function is that it is the mean value of consumption deprivation at a given level of income. The real problem with poverty, is not the mean level of consumption deprivation, but it is the variability of the consumption deprivation. The lower income persons are more susceptible for deprivation as the spread of actual consumption is so wide due to high

<sup>6</sup> It may be noted that the same way  $y_i$  is a sample information on the mean  $\{y_i - (y_i + y_j)/2\}^2$  is the sample information on the variance. Hence the above expression can be calculated for every pair of observations and those calculated values can be used as values assumed by a dependent variable for the regression of the variance. The mean values of the corresponding pairs of values for  $x$  can be taken as the values of the independent variables.

<sup>7</sup> Suryanarayana (1986) derived analytically the partial derivatives of the Head Count Ratio with respect to the location and scale parameters of a log normal income distribution (Suryanarayana, 1986, pp. 234-237). The type of hierarchical regressions given above together with such partial derivatives provide the necessary information to devise policies to reduce the income inequality.



variance that it can go below the consumption requirements more frequently. Even if the variability is same at different income levels the probability that a person's consumption falls below the minimum required consumption is more for a lower income person than for a higher income person. This is because the mean deprivation is a decreasing function of incomes at all levels of income. It is this variability in the consumption deprivation, and that too the possibility of differential variability at different income levels, that causes a major problem for the poor.

The existing literature on poverty indices does not take into account this aspect of variance in deprivation level. A more useful measure of poverty can be defined as an  $(E, V)$  measure that is frequently used in portfolio choice:

$$P = \int_0^{\infty} E[d(y, z)] f(y) dy + \sigma \int_0^{\infty} V[d(y, z)] f(y) dy \quad (\text{with } \sigma > 0). \quad (7.3)$$

The problem raised by Sukhatme is a special case of this problem. Sukhatme treated  $z$  to be random (variability of intra-individual energy requirements). In general both actual consumption and the requirements may show variability giving rise to variability in  $d(y, z)$ .

In fact the above suggestion to replace  $E\{d(y, z)\}$  by  $E\{d(y, z)\} + \sigma V\{d(y, z)\}$  is a standard approach of measuring risk. One can generalise this procedure and introduce alternative measures of risk in place of  $d(y, z)$ . Another promising approach seems to be the approach of reliability theory. Based on the distribution of  $d(y, z)$  one can determine a hazard rate or failure rate function (Barlow and Proschan, 1965) and use that function instead of  $E\{d(y, z)\}$ . From a policy perspective the sustainability of a poverty alleviation program depends on reducing the vulnerability of the poor people towards consumption deprivation. Hence sustainable poverty alleviation as a policy objective calls for devising a measure of poverty which captures such vulnerability. Eq. (7.3) provides such an alternative measure of poverty.

The problems of statistical inference cited above with a deterministic or known deprivation function need to be modified to incorporate the stochastic deprivation function. The population poverty index may now be defined as:

$$P = \int_0^{\infty} E\{d(y, z)\} f(y) dy \quad (7.4)$$

In this case the sample estimate may be written as

$$\hat{P} = \int_0^{\infty} \hat{d}(y, z) f_n(y) dy. \quad (7.5)$$

One needs to examine the sampling properties of this estimator. In the absence of any analytic approach to derive the sampling distribution of this estimator one can compute this estimator of the poverty index for given sample estimates  $\hat{d}(y, z)$  and  $f_n(y)$  employing numerical integration, for which appropriate computer software is available. The bootstrap technique developed by Efron (1982) can be employed to generate sampling distribution of the poverty estimator given by (7.5). Similarly one

should consider suitable modification of (7.5) if the population poverty index is given by (7.3).

Another important statistical problem is that of classification. If one knows that different groups of persons must have different poverty lines, the determination of which is quite difficult, and uses for convenience a uniform poverty line for all, what will be the extent of the errors in classification of the poor in different groups?

This review is only selective and not exhaustive. The main objective, which is hopefully satisfactorily achieved, is to give an over-all survey of recent developments in econometric methodology of poverty measurement and to suggest a few directions for future research.

## References

- Atkinson, A.B. (1970a). On the measurement of inequality. *J. Economic Theory* 2, 244-263.
- Atkinson, A.B. (1970b). *Poverty in Britain and the Reform of Social Security*. Cambridge University Press, Cambridge.
- Atkinson A.B. (1987). On the measurement of poverty. *Econometrica* 55, 749-764.
- Barlow, R.E. and F. Proschan (1965). *Mathematical Theory of Reliability*. Wiley, New York.
- Batchelder, A.B. (1971). *The Economics of Poverty*. Wiley, New York.
- Bhanoji Rao, V.V. (1981). Measurement of deprivation and poverty based on the proportion spent on food: an exploratory exercise. *World Development* 9(4), 337-353.
- Bhatty, I.Z. (1974). Inequality and poverty in rural India. *Sankhya* 36, 291-336.
- Bowley, A.L. and A.R. Burnett-Hurst (1915). *Livelihood and Poverty*. Bell, London.
- Dandekar, V.M. and N. Rath (1971). *Poverty in India*. Indian School of Political Economy, Poona.
- Dasgupta, P., A.K. Sen and D. Starrett (1972). Notes on measurement of inequality. *J. Economic Theory* 6, 180-187.
- Debreu, G. (1959). *Theory of Value*. Wiley, New York.
- Efron, B. (1982). *The Jackknife, the bootstrap, and Other Resampling Plans*. Society for Industrial and Applied Mathematics, Philadelphia.
- Foster, J.E. (1984). On Economic Poverty: A Survey of Aggregate Measures. In R.L. Basman and G.F. Rhodes, Eds., *Advances in Econometrics*, Vol. 3, JAI Press, Connecticut.
- Foster, J.E., J. Greer and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica* 52, 761-766.
- Foster, J.E. and A.F. Shorrocks (1991). Subgroup consistent poverty indices. *Econometrica* 59, 667-709.
- Gastwirth, J.L. (1971). A General definition of a Lorenz curve. *Econometrica* 39, 1037-1039.
- Gastwirth, J.L. and M. Glauberger (1976). The interpolation of the Lorenz curve and Gini index from grouped data. *Econometrica* 44 (3), 479-483.
- Gini, C. (1912). *Varabilita e mutabilita*. Bologna.
- Government of India, Perspective Planning Division, Planning Commission (1986): A Technical Note on The Seventh Plan of India (1985-1990).
- Iyengar, N.S. (1960). On the standard error of Lorenz ratio. *Sankhya* 22, 371-386.
- Iyengar, N.S. (1964). A consistent method of estimating the Engel curve from grouped survey data. *Econometrica* 32, 591-618.
- Kakwani, N.C. (1980a). On a class of poverty measures. *Econometrica* 43, 437-446.
- Kakwani, N.C. (1980b). Income Inequality and Poverty: *Methods of Estimation and Policy Applications*. Oxford University Press, London.
- Kakwani, N.C. and N. Poddar (1976). Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations. *Econometrica* 44 (1), 137-148.
- Kanbur, S.M.R. (1987). Transfers, targeting and poverty. *Economic Policy* 4, 112-136, 141-147.

- Kanbur, S.M.R and J. Strömberg (1988). Income transitions and income distribution dominance. *J. Economic Theory* **45**, 408-416.
- Koen, M. (1992). Needs and targeting. *Economic J.* **102**, 67-79.
- Kendall, M.G. and A. Stuart (1969). *The Advanced Theory of Statistics*, Vol. 1, Hafner, New York.
- Kumar, T.K. (1993). Measures of poverty suited to design and evaluate poverty alleviation programmes. *Economic Political Weekly* **28** (52), 2893-2895.
- Levine, D.B. and N.M. Singer (1970). The mathematical relation between the income density function and the measurement of income inequality. *Econometrica* **38** (2), 324-330.
- Lorentz, M.O. (1905). Methods of measuring Concentration of wealth. *J. Amer. Statist. Assoc.* **9**.
- Marji, P. and M. Pal (1988). A unified approach to estimation of Lorentz ratio from a finite population. *Sankhya B* **50** (2), 215-223.
- Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, San Francisco.
- Muothalhu, T.S.K. (1991). On a sufficient condition for two non-intersecting Lorentz curves. *Sankhya Series B* **53**, 268-274.
- Muelbauer, J. (1975). Aggregation, income distribution and consumer demand. *Res. Economic Studies* **42**, 525-543.
- Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl.* **9**, 141-142.
- Prakasa Rao, B.L.S. (1983). *Nonparametric Functional Estimation*. Academic Press, New York.
- Rotschild, M. and J.E. Stiglitz (1973). Some further results on inequality. *J. Economic Theory* **6**, 188-204.
- Rowntree, B.S. (1901). *Poverty: A Study of Town Life*. Macmillan, London.
- Sen, A.K. (1973). *On Economic Inequality*. W W Norton, New York.
- Sen, A.K. (1974). Poverty, inequality and unemployment: some conceptual issues in measurement. *Sankhya Series C*, Parts 2 & 4, **36**, 67-82.
- Sen, A.K. (1976). Poverty: an ordinal approach to measurement. *Econometrica* **44** (2), 219-231.
- Sen, A.K. (1979). Issues in the measurement of poverty. *Scandinavian J. Economics* **81**, 285-307.
- Sen, A.K. (1981). *Poverty and Famines*. Oxford University Press, Oxford.
- Singh, S.K. and G.S. Maddala (1976). A function for size distribution of incomes. *Econometrica* **44**, 963-970.
- Spencer, R.D. and S. Fisher (1992). On comparing distributions of poverty gaps. *Sankhya, Series B, Part 1*, 114-126.
- Sukhatme, P.V. (1978). Assessment of adequacy of diets at different income levels. *Economic Political Weekly* **13**, 1373-1384.
- Suryanarayana, M.H. (1986). *The Problems of Distribution in India's Development*. Economic Analysis Unit, Indian Statistical Institute, Bangalore, April (a thesis submitted for the award of Ph.D. degree).
- Suryanarayana, M.H. (1991). A note on Lorentz ratio based on NSS consumption distributions. *J. Quantitative Economics* **7** (1), 179-183.
- Suryanarayana, M.H. and S. Geetha (1992). Estimation of  $P$  - a poverty measure for Grouped data using two parametric log normal distribution. Indira Gandhi Institute of Development Research, Bombay, August. Discussion Paper, No. 78.
- Thorbecke, E. (1989). Planning techniques for social justice. In Irma Adelman and Sylvia Lane, Eds., *The Balance between Industry and Agriculture in Economic Development*, Vol. 4: Social Effects, Macmillan, New York.
- Townsend, P. (1954). Measuring poverty. *British J. Sociology* **5**, 130-137.
- Watson, G.S. (1964). Smooth regression analysis. *Sankhya, Ser. A*, **26**, 359-372.
- Weisbrod, B.A. (1965). *The Economics of Poverty*. Prentice-Hall, Englewood Cliffs, NJ.