68577

United Nations
Economic Commission
for Africa

Workshop and Training Course on Public Intervention for Poverty Reduction and the PRSP in Ethiopia

Organized by Economic and Social Policy Division (ESPD) of ECA in collaboration with the Ministry of Economic Development and Cooperation (MEDaC)

Issues of Targeting for Poverty Reduction

by
Alem Abreha

31 July - 2 August 2001

The views expressed represent those of the author and not necessarily those of the ECA or the UN.

TARGETING INTERVENTIONS AGAINST FOOD INSECURITY AND POVERTY

INTRODUCTION

One of the major problems in designing tax-transfer schemes in developing countries, especially in the rural areas, is inability to assess incomes of households as a basis for targeting taxes and transfers. Some studies have suggested that targeting problems in tax-transfer schemes (income redistribution schemes) in rural areas of developing countries can be addressed reasonably well using land-contingent means testing (Ravallion 1989).

The attempt to find indicators is, a worthwhile exercise as it has been shown (in India, for instance) that even crude measurement can help guide policy where reliable data on poverty incidence, intensity, and distribution are not available (Lipton and van der Gaag 1993). This note attempts to identify the major factors that need to be included in designing alternative means-testing schemes using examples of land-contingent targeting. It reviews the empirical and theoretical literature regarding the relationships between the variables and construction of robust indicators for targeting taxes and transfers.

THEORY AND METHODOLOGY

QUANTITATIVE ISSUES IN POVERTY ANALYSIS–FUNDAMENTAL QUESTIONS

Measurement: Motivation

In defining the agenda for quantitative measurement of food insecurity/poverty, an important consideration is assisting policymakers in assessing alternative policies. Lipton and Ravallion (1995) identify three tasks for poverty analysis:

- "to define and describe 'poverty',
- to understand its causes, and
- to inform policy".

The criteria used to assess the options in interventions against food insecurity/ poverty include:

1. overall congruence with the macroeconomic and social objectives;

2. the likelihood that the program would reach the poor;
3. ease of implementation;
4. cost effectiveness;
5. sustainability; and,
6. administrative feasibility.

The choice of a measure of poverty/food insecurity is important for poverty policy since it implicitly sets the target. For instance, the head-count ratio does not show how poor the poor are and gives no indication of the effect of intra-poor transfers. Similarly, the poverty gap ignores intra-poor distribution. Furthermore, "Correct measurement, used in allocating resources among groups of the poor, can save lives. ... [This] has required research that goes beyond incidence– beyond counting the poor" (Lipton and van der Gaag 1993: 2). Attributing poverty (as between incidence, intensity, and intra-poor distribution) tells us a lot about appropriate steps towards curing it.

How to use information from secondary sources or household-level income and expenditure surveys to rank regions depends on (i) the precise objectives of the government, and (ii) how expenditure devoted to each group translates into individual incomes. For instance, if the objective is to have maximum impact on national poverty gap, the relevant regional ranking would be by incidence of poverty in each region (not by the regional poverty gap). This is justified by "poverty alleviation efficiency" (Besley and Kanbur 1993).

Basic Questions: Measurement

The most fundamental questions in food security/poverty analysis include

1. What variable should we use to measure welfare? A large number of variables have been identified in the literature, and the task here is to select the best under particular circumstances.
2. What levels of this variable would indicate food insecurity/poverty? This is a question of defining, in terms of the chosen indicator of welfare, a food insecurity/poverty line which defines the boundary between food insecure/poor individuals (or any other socio-economic grouping such as households, communities) and others.
3. How do we aggregate a measure of food insecurity/poverty to obtain a summary indicator of welfare in a group?

Basic Questions: Interventions

In the design of interventions, the most important quantitative issues/questions that data should speak about are:

1. given patterns of characteristics among groups, what levels of benefits are appropriate?
2. how should boundaries between groups be defined?
3. how many groups should there be?

MOTIVATION: WHY TARGETING

TARGETING: DEFINITION

The problem of targeting was characterised by J.S. Mill (Besley and Kanbur 1993) as "one of giving the greatest amount of needful help with the smallest amount of undue reliance on it. ... directing them [resources] as much as possible toward those who need them most." Targeting can be direct, as when a set of criteria must be met by the household/group to qualify for assistance; or self-targeted as when an in-build mechanism prevents non-qualifying households from claiming benefits. These notes deal with direct targeting. Direct targeting can be done:

1. by identifying structural characteristics of a target group, such as those who are disabled, women-headed households.
2. using a welfare criteria such as incomes, assets, or the absence of assets.

REASON 1. EFFECTS OF IMPERFECT INFORMATION I

An important source of uncertainty that calls for targeting using indicators is uncertainty about the socio-economic environment in which policymaking takes place. Governments may be uncertain about key behavioural parameters and levels of income. These problems are pervasive in developing countries, and "[W]hen the administrative structure is weak and information poor, it may be worthwhile to implement policies that economise on information and play to the strengths of the existing administration. History may matter in important ways " (Besley 1993: 375).

REASON 2. EFFECTS OF IMPERFECT INFORMATION II

Another source of uncertainty that justifies targeting using indicators is imperfect information that arises because of incompatible incentives between agents in an

exchange in a situation where one agent has private information the other hasn't. In this case, households have information that the "government" would to use to design and implement a tax-transfer scheme. These are "problems of imperfect information that are *strategic*"[1] In this context, the need to target the poor using indicators/correlates of poverty is justified by the following (Ravallion 1993)

1. It is difficult to identify the poor from a large set of households, especially when survey managers and respondents know that the answers determine eligibility for benefits.
2. The task of targeting at the household level requires more time, training, and resource than project managers have.
3. Project managers, like potential beneficiaries, have incentives to mis-report or short-circuit.

REASON 3. WEAKNESSES OF MARKET-WIDE POLICIES

Market-wide poverty alleviation policies that lower consumer prices also lower producer prices and lead to efficiency losses by diverting resources towards production of lower-valued goods. It is also argued that provision of free/subsidised goods and services on a broad scale creates demand for them in excess of what it would have been otherwise. That is, distortion of allocation of resources occurs when demand is not established on the basis of willingness-to-pay for the full cost of certain commodities or services[2]. "In making the choice between an income transfer program to a target group and a trade intervention that distorts the price of food to producers, it is necessary to compare the cost of administering the first with the efficiency losses of the second in relation to how each increases food security" (p. 30). The benefits of direct targeting anti-poverty are that targeting: prevent generation of wrong incentives: targeting minimises the distortionary impact of interventions on incentives, such as price of labour and output; market structures and product flows, and; consumption habits of the population.

REASON 4. EFFECTS OF IMPERFECT INFORMATION I

The government and donors have too much to be done, but only limited resources to execute their objectives. It is thus essential to provide assistance to those most in need.

---

[1] These issues are the motivations for the analysis of optimal income taxation pioneered by Mirrlees and the staple of economics of information.

Direct targeting has the benefit of holding budgetary costs down (helps reduce fiscal cost of interventions) and concentrate allocation of resources to those in need. Well targeted interventions avoid under-coverage, i.e., they provide support to all or most individuals/ groups that are affected by food insecurity so that the main objective of food aid intervention, that of saving lives and minimising and preventing the loss of human capital of the poor is achieved. By minimising leakage, targeting increases the efficiency of resource use by maximising the impact of a given budget on poverty and food insecurity, or minimising the cost of achieving a given target level of food security.

REASON 5. EFFECTS OF IMPERFECT INFORMATION I

Subsistence agriculture is not liable to market forces to any significant scale. Thus , the supply response of peasants to "improved market incentives is limited due to non-price factors such as deep-seated structural problems of the agricultural sector such as dualistic production structure, atomised risk averse producers that lack resources to invest in innovation, malnutrition, seasonal labour shortages". It follows that "given the structural poverty of food-insecure households there is no simple, market-based solution waiting to be tapped".

TARGETING USING INDICATORS: THE PROBLEM

With reference to regional targeting, the problem of targeting is stated by Ravallion (1993: 455) as:

> We do not know the standard of living of any one individual, but each person in a population can be assigned precisely to one of a number of mutually exclusive groups ... with fewer groups than people. Thus the information signal each person supplies in the present setting is place of residence, though one can also allow more complex multivariate signals. ... This distribution [of the signal] is empirically derived from a sample survey. The policymaker's problem is then how to regionally target his policy instruments so as to minimise aggregate expected poverty, as measured by a suitable index.

TARGETING: CRITICISM I

The issue of targeting, answering the question "Who are the Poor?" is central to the World Bank's approach and policy recommendations for poverty alleviation (see World Development Report 1990). This position, coming from the World Bank, is criticised since "policy directions to reduce poverty in the context of adjustment programs cannot

---

[2] These effect might be mitigated by a prudent choice of the type of assistance, for instance, through distribution of money or tradable vouchers rather than food.

be taken seriously for the simple reason that their general drift runs counter to the SAP policy package" (Ali 1994: 60). Ali (1994: 61) further argues that the recommendations for targeting are made

> based on its [the World Bank's] 40 percent estimate [of head count ratio] ... when the value of the head count index goes to above 70 percent it seems reasonable to argue that designing poverty reduction policies on the basis of identifying and targeting the poor is blatantly irrelevant. ... In such cases, it would seem reasonable to argue that perhaps a cost-effective way to go about designing the poverty reduction policies is to identify and target the rich who form only 30 percent of the population. If targeting is such a cherished term then what is needed is could be termed "reverse targeting" or some similar name!!
>
> The objective of reverse targeting would be to generate the required real resources to finance the various items of additional government expenditure to reduce poverty. ... Besley and Kanbur (1990) argue the case for an optimum policy design strategy that lies between the ideal solution of perfect targeting and the extreme of complete universal programmes on the basis that the ideal solution of perfect targeting is not available to policy makers in developing countries. *Our position is that the optimum policy design lies closer to the complete universal programmes* {italics added).

Though this is an important point, it is more a critique on the style of targeting than on targeting *per se*, and is silent on how the negative consequences of universalist schemes is to be addressed.

TARGETING: CRITICISM II

In some cases, well designed market-wide interventions may perform better. Market-wide interventions are (i) quickly and easily implemented, (ii) can be designed to make their burden on public finances small (the costs being transferred to other sectors), and (iii) are popular (and politically more sustainable) because they benefit a large segment of the population.

However, these interventions have not proved to be unambiguously better than direct targeting; in fact, calls for increased targeting have been intensifying due to the failure of market-wide interventions. Cost-effectiveness of market-wide interventions against food insecurity/ poverty depends on

1. extent of leakage to unintended beneficiaries,
2. income losses of the poor who are adversely affected by price changes;
3. inefficiency losses in production that arise due to price distortions.

Hence, cost-effectiveness of trade interventions on specific commodities/food items increases if the commodity is consumed more intensively by the poor and if the commodity is less responsive to relative price changes. This is what is expected under commodity-based self-targeting schemes. Yet, the undeniable fact is that targeted interventions are more demanding in terms of management, identification of target

groups is more difficult, and can give rise to negative effects (disincentives) to the supply of labour. On the other hand, food security can be achieved with minimum interfering with markets and other objectives.

TARGETING USING INDICATORS: SELECTING INDICATORS

Having established the major reasons for targeting, we need to determine the type of data that should be collected. What is needed now is "a refinement of the methodology for selecting and weighting of indicators of household and regional distress" (Webb and von Braun 1994: 132). Moreover, there are ambiguities as well as differences in the way a given indicator is interpreted in empirical data collection. The literature on poverty and food security documents the arguments regarding the alternative methods of defining even the most direct predictors such as wealth and income. Hence the need to identify the most workable definitions and representations of such issues in a location specific manner.

Data on household socio-economic characteristics holds a central position in the search for indicators. Firstly, "entitlements" approach and later works that build on it focus on the access of households to food through own production, income, gathering of wild foods, community support (claims), assets and migration. This led to new ways of thinking which "focused on the household as the unit of analysis for food security ... [and] placed intra-household power and resource allocation issues in the front line of analysis" (Maxwell 1994: 2). Thus, a greater recognition is given to the role a number of socio-economic variables play in household's access to food.

Secondly, the paradigm looked at victims of stress not as passive agents, they exhibit a sequence of responses to economic stress. Thus, a greater recognition of the importance of coping mechanisms — behavioural responses to economic stress. Nevertheless, the focus on household data and vulnerability in inter- and intra-household levels, does not in any way reduce the importance of vulnerability at regional and local levels.

Advances in analytical and data processing capacity that contribute to the paradigm shift away from the "trickle down approach" have not materialised in most African countries to be of use in decentralised decision making. Hence, there is a need to identify a few indicators that are robust, easy to collect, analyse, and verify. The specifics of the methods used are detailed below.

CATEGORIES OF INDICATORS

Indicators of food insecurity can generally be classified as process and outcome indicators. The first group of indicators includes indicators of availability/production of commodities and indicators of household access to them. Indicators of supply of food, for instance, include factors in the physical environment that affect agricultural production and institutions and features of the market for food items.

This includes meteorological data, such as the volume, variability, and timeliness of rainfall, and temperature; agro-climatological data such as characteristics of the soil; information on natural resources and common property resources. Common property resources are given a special place because they affect the availability of resources for housing, fuel, food supplements, wild foods used during periods of crisis, and off-farm income generating alternatives.

Furthermore, market data such as prices, the pattern of flow of food items into and out of the market, and the volume of transactions in a market are valuable indicators of impending shortage of food. Related to this are the institutions and physical infrastructure supporting the market. Efficiency of the market is found to be an important limiting factor on household responses to food insecurity in Ethiopia (Webb, von Braun, and Yohannes, IFPRI).

Household access to food means availability of food for consumption from own production, and other means such as off-farm-income, and community support. The access of households is indicated by the type and volume of assets it owns, the amount of off-farm income it has been able to earn, and its access to common property resources and community support mechanisms. In addition to its access to food, the coping mechanisms it adopts and the change in these coping mechanisms are valuable indicators of the level of insecurity the household faces.

Outcome indicators include all direct and indirect indicators of household consumption of food items. The direct indicators include household income and consumption obtained from the household budget and consumption surveys, subjective judgement of households regarding food security, and assessments of food frequency. Indirect indicators of household consumption of food include storage estimates, nutritional assessments and calculation of various indices that indicate availability of food in a household such as subsistence potential ratios.

Another class of indicators that need to be collected includes intermediate variables that affect the conversion of food consumed into energy and better nutritional status. This

includes health status of individuals and the availability of health care, the level of education of the care-taker in the household (usually the mother), and environmental factors that affect the degree of risk individuals in a household face.

Indicators might also be categorised temporally in relation to the actual event they represent. Early indicators (also called leading indicators) include variables and conditions and changes in conditions and variables that appear before the actual decline in access to food and food consumption. Stress indicators (also called concurrent indicators) are those that occur simultaneously with the decline in access. Late outcome indicators (also called trailing indicators) are those that appear after access has declined and show the extent of suffering by households.

DIRECT VS INDIRECT INDICATORS

Direct indicators "[of food consumption] include those indicators which are closest to actual food consumption rather than to a marketing channel information or medical status" (Frankenberger 199?: 96). They include (i) household budget and consumption surveys; (ii) household perception of food security; and, (iii) food frequency assessments.

The major weaknesses of household consumption and budget surveys include (i) underestimation of food expenditure because value of home-produced and gathered foods is underestimated or unrecorded; (ii) data are not available on a timely basis; and (iii) inaccessible areas are underrepresented.

Indirect indictors are used when direct indicators are either unavailable or too costly. For instance indicators of outcome are usually proxies for adequate food consumption.. Selection of proxies should take into account the fact that the proxies might be measuring more than food (e.g. nutritional status).

CRITERIA FOR SELECTING INDICATORS

The choice of indicator depends on the underlying reality of concern (functional outcome) (Habicht et al 1982). If for instance, one measures mortality as a functional outcome of malnutrition the sensitivity and specificity of the indicators will depend, among other things, on the time interval between the measurement of malnutrition and the measurement of mortality. Some indicators are better in predicting short term survival, while others predict better long term survival. A nutritional indictor might be quite sensitive if mortality, but not if school performance, is considered as a functional

outcome. If different outcomes are under consideration this may also request the determination of different cut-off points of a certain nutritional indicator in order to maintain its discriminatory function (Kostermans 1994: 14).

Costs of collection and non-collection

An indicator might be usefully evaluated in terms of the costs of collection and the costs of non-collection. The costs of non-collection are essentially the benefits derived from the collection of the indicator, for instance, the savings in the budget required to bring a given reduction in food insecurity that may accrue due to the use of an indicator. On the other hand, "costs of collection include the fixed and variable costs associated with the design, collection, analysis, and sustainability of the data collection effort as well as the costs of acting on the data" (Haddad, Kennedy, and Sullivan 1994: 330).

One approach to the conceptualisation of the notion of 'costs of non-collection' of an indicator is as any resource saving (say, calorie saving) achieved using this indicator to target transfers compared to untargeted transfer, given some target reduction in food insecurity. Haddad et al (1994) cite Ravallion 1989, Glewwe and Kanaan 1989, Haddad and Kanbur 1991 as studies that adopt this conceptualisation. This conceptualisation may not be appropriate to Ethiopia since (i) the volume of resources available for transfers related to food insecurity are exogenous, being determined by donors; and (ii) resources available through aid are largely 'incremental', leading to a lack of benefit to minimising resources devoted for food security transfers.

Hence, we will conceptualise the costs of non-collection as the rise in the share of those who are not food insecure when transfers are not targeted from that which would occur if the indicator was used to target benefits. Costs of non-collection depend fundamentally on the society's nutritional objectives and goals, and its sensitivity to under-nutrition (Haddad, Kennedy, and Sullivan 1994: 332).

Besley and Kanbur (1993) approach the problem of setting the criteria for selection of indicators in terms of a trade-off between the cost components of antipoverty programs: administrative costs (A), cost of benefits that accrue to the poor (P), and cost of benefits that accrue to the non-poor (N). They argue that there is decreasing returns in A, and that can $P/(P+N)$ be raised at the cost of higher A. They show a method of optimising $P/A+P+N$ so as to maximise P given total cost (A+P+N). Their model shows that the

more imperfect the indicator of poverty, the severer the trade-off between A and $P/(P+N)$ will be.

'Optimal ignorance'

Chambers' position is that the identification and collection of alternative indicators should embody the principles of 'optimal ignorance' –not trying out to find out more than is needed– and 'appropriate imprecision' –not measuring more accurately than is necessary for practical purposes (Haddad, Kennedy, and Sullivan 1994: 330). These concepts would imply restrictions on, for instance, the level of measurement of the data and the number of questions asked for every observation. For instance, if ranking household land in a five-scale categorical scale is sufficient to determine the eligibility of households, then exact measurement of household land should not be undertaken.

Positive Predictive Value

A common criteria for assessing indicators their sensitivity and specificity (Kostermans 1994: 11). Sensitivity of an indicator refers to its capacity to detect correctly those who are truly food insecure; and its specificity is its capacity to correctly detect those who are not food insecure. Lack of sensitivity leads to under-coverage of deserving households. On the other hand, lack of specificity leads to leakage of benefits to non-deserving households. Hence, the sensitivity-specificity issue can be approached as a problem of maximising the benefits that accrue to a target population while simultaneously minimising the proportion of benefits that accrue to non-target groups. Sensitivity and specificity, and hence under-coverage and leakage, are inversely related[3]. Under this framework, indicators can be ranked in terms of the values of:

$$\text{Sensitivity} = TP / (TP + FN);$$

$$\text{Specificity} = TN / (FP + TN);$$

$$\text{Positive Predictive Value} = TP / (TP + FP);$$

$$\text{Real Prevalence of Malnutrition} = (TP + FN) / \text{All};$$

$$\text{Measured Prevalence} = (TP + FP) / \text{All}.$$

[3] These are similar to type I (or α) and type II (or β) error. A small type I error (which involves rejecting the null hypothesis when it, in fact, is true) means a high specificity and a small type II error (which involves failing to reject the null hypothesis when it, in fact, is false) means a high sensitivity.

Where TP is true positive, correct diagnosis of the food insecure; TN is true negative, correct diagnosis of the non-food insecure; FP is false positive, wrong diagnosis of the non-food insecure; and FN wrong diagnosis of the food insecure. The best diagnostic test is the one with the highest proportion of correct diagnosis, that is, the least FP and FN results.

Administrative capacity

When the administrative structure is weak and information poor, it may be worthwhile to implement policies that economise on information and play to the strengths of the existing administration. History may matter in important ways. ... One must be wary of ignoring administrative considerations in policy design. A policy that is unsophisticated yet well administered may be better than a more sophisticated policy administered incompetently (Besley 1993: 375).

Grosh and Baker (1995) selected indicator variables based on two criteria: close correlation with welfare and *verifiability*. Some fine-tuning of the basic system, such as calibrating for the poorest half of the population, improves results considerably. Besley and Kanbur (1993) using an optimising model, show that indicators of poverty should be:

1. closely correlated with poverty,
2. capable of being readily assessed, and
3. hard to simulate or mis-report.

The indicators to be selected should be relevant, usable, and reliable, sensitive to change, replicable, timely, and cost-effective. It has been suggested in many studies that (See for example Grosh and Baker 1995 Haddad, Kennedy and Sullivan 1994) a small number of indicators that satisfy these criteria, and selected at localised levels may serve well in targeting the food insecure as well as more complex indicators.

> Simple indicators such as number of unique foods consumed, household dependency ratio, household rooms per capita, incidence of illness, vaccination status, age at weaning of the preschooler, and household drinking water and sanitation facilities ... were able ... to identify households ... at risk of food and nutrition insecurity (Haddad, Kennedy, and Sullivan 1994).

Moreover, using a wide range of indicators would serve as a type of sensitivity analysis and selecting indicators that lead to a convergence of evidence is essential.

STATISTICAL METHODS AND ISSUES

There are alternative measures of association among indicators and between indicators and food security: overlap technique, two-way tables, correlation coefficients, classification analysis, factor analysis, cluster analysis, etc.

Multivariate or univariate?

The ability of indicators to locate food insecure households could be improved by interacting indicators. Note that interacting indicators can have the following limitations (Haddad, Kennedy, and Sullivan 1994: 337-338):

1. Because the number of indicator combinations is endless, the specific combinations used will almost certainly be determined locally; our analysis is purely suggestive.
2. The best combination of indicators may not necessarily include indicators that performed well in first-round overlaps [using single variables].
3. As we combine indicators, we run into sample size problems.

In attempting to make the indicator more specific (to reduce type II error), we run the risk of making them too specific to be of help in classifying large segments of the population (we may well increase type I error)."

Ravallion and Chao (1989) developed a transfer scheme using two or three variables. The scheme allowed equal transfers to all individuals within a group, but different transfers for groups with different characteristics. Results of three studies that applied the approach have shown that single indicators at macro level are of limited use, especially when constraints are put on the transfers to restrict the optimal solutions to something that might be politically feasible.

Variable weighting in indexes

Giving items in an index equal weight (as for instance in UNDP's human poverty index) may create problems. For example, in a country with a high literacy and low income, it may be inappropriate to assign the same weight to each variable because a variation in income has a higher impact on poverty than the same variation in education (Vella and Vichi 1997).

Contingency tables

Two-way tables that investigate the characteristics of households classified as food insecure are suggestive, but they began with the premise that these households have been identified (Haddad, Kennedy, and Sullivan 1994: 333). Using percentiles of income/expenditure may not be meaningful if incomes/expenditures are low, have a low variation and cannot sustain even basic needs (Vella and Vichi 1997).

Overlap technique. The overlap technique has been used by Haddad et al (1994) in an assessment of indicators of food security. The overlap technique involves asking, "what percent of households/pre-schoolers who have a certain indicator value are food or nutrition insecure?" The overlap technique has some major weaknesses: "an indicator group could have a high percentage of households who are food insecure yet cover only a small percentage of those at risk. Similarly, some indictors will be so general that they will contain all the food insecure household but the food insecure households will only represent a small percentage of the household with that indicator value" (Haddad, Kennedy, and Sullivan 1994: 333).

Proxy means tests

"Proxy means test" is used to describe a situation where information on household or individual characteristics correlated with welfare levels is used in a formal algorithm to proxy household income, welfare, or need. Means tests is the best way to determining eligibility in theory. In practice, it is rather difficult because of problems of under-reporting and less-than-perfect correlation between income and welfare (Grosh and Baker 1995).

Grosh (1994) found that among all targeting mechanisms, proxy means tests produce the best incidence outcomes. Proxy means tests use household or individual characteristics to proxy a means test, thus avoiding the problems involved in relying on reported income.

The proxy systems all have significant errors of under-coverage, but they cut down on leakage so much that the impact on poverty is better with imperfect targeting than with none. Some fine-tuning of the basic system, such as calibrating for the poorest half of the population, improves results considerably (Grosh and Baker 1995).

Discriminant analysis

If your observations are in known groups and you want to predict group membership based on a set of independent variables you can use the Discriminant procedure.

Discriminant is used to classify cases into one of several known groups on the basis of various characteristics. You must have information available for a sample which have been pre-assigned to a category. To use the Discriminant procedure the dependent variable must have a limited number of distinct categories. Independent variables that are nominal must be recoded to dummy or contrast variables.

Factor and cluster analysis

The goal of cluster/factor analysis is to identify relatively homogeneous groups of cases based on selected characteristics. Cluster analysis can be used to classify the sample into socio-economic groups. The first step in cluster analysis algorithm is selection of a measure to evaluate degrees of dissimilarity between objects (e.g. households). The squared Euclidean distance is often used to measure the distance between points that describe characteristics of the object.

Factor and cluster analysis techniques do not work well with categorical values or continuous variables that are not normally distributed (Haddad, Kennedy, and Sullivan 1994: 333). Measurement errors and outliers might seriously affect correlation coefficients between indicators. Even if a correction is made for this using correlation of ranks, high correlation coefficients could arise due to association at the upper end of the indicators, say calorie adequacy distribution, while for monitoring purposes policy makers are interested in households at the lower end of the distribution (Haddad, Kennedy, and Sullivan 1994: 333).[4]

Non-linear principal components analysis

Characteristics of the data such as high correlation between variables and need for non-equal weights calls for the use of non linear component analysis (Vella and Vichi 1997). In one study, NLPCA was used to (i) remove variables that were correlated with each other; (ii) remove variables that were not correlated with the most important factors defined by NLPCA; (iii) transform original variables into a few composite indices (factors) which were proxies of the original variables; and (iv) to give category quantifications of the original variables (Vella and Vichi 1997). This method enables one to use variables with different levels of measurement.

[4] This applies for the significance of the chi-square statistic of contingency tables. This problem is overcome by partitioning the chi-square statistic and testing the significance of residuals.

OLS Regression vs. direct poverty minimization

Glewwe and Kanaan (1989) used regression to predict welfare levels based on several combinations of variables that are fairly easy to measure. Glewwe (1990) used a method that minimises poverty directly to find a targeting mechanism. The method is theoretically more appropriate, but difficult to compute.

The results are not different from those based on regression analysis. Ravallion and Chao (1989) algorithm is better than OLS since it directly minimises poverty. There is also a problem of endogeneity of the right-hand-side variable. But OLS is sufficient since we are concerned with identifying the poor rather than explaining why they are poor.

Correlated variables: consequences and solutions

If variables determining the index are correlated with each other, the same information might be duplicated. For instance, people with high education have also high income, thus the two variables are correlated and their use in building an index produces a duplication of information summarised by the index (Vella and Vichi). Methods such as NLPCA turn out to be useful under such circumstances.

Measurement errors: consequences

Uncertainty in poverty measurement that arises due to factors such as (a) errors in living standards data and in indicator variables; (b) differences in needs between households with similar values for indicator variables; and (c) arbitrariness & uncertainty about the cut-off point of poverty and best indicator of poverty.

Measurement errors will influence the sensitivity and by consequence the positive predictive value of an indicator: if for instance the sensitivity and specificity of an indicator are reduced because of errors while the true rate of prevalence is unchanged, then the Positive Predictive Value will be lower. The PPV appears to be very sensitive to changes in sensitivity and specificity of the indicator (Kostermans 1994: 12).

A random error will not influence the estimates of the mean or median, but will increase the variance. Random error will always–except in one case– cause the prevalence rate to be overestimated because the tails of the distribution become fatter due to the increased variance. The exception is the very rare case that the means of the Z-scores of a population is below the cut-off point the prevalence estimate will decrease because of the increased variance caused by the random error. The random error is important because malnutrition is a phenomenon of one tail of the distribution: the lower one. a regression

analysis misses this point because it utilises the mean of the distribution of variables which is not affected by random error. The random error in the dependent nutritional variable would affect the proportion of unexplained variation, i.e., lower the R^2, and decrease the power to detect statistically significant associations with explanatory variables (Kostermans 1994: 19).

A systematic error, or bias, will change the mean of the normal distribution, say of height or weight. The prevalence rate would be under or over-estimated depending on the direction of the bias of systematic error.

For instance, a positive bias in measuring weight will lead to an underestimation of the prevalence of malnutrition. The impact on the estimate of prevalence depends on, first, the size of bias: the larger the bias the bigger the influence. Secondly, the impact will depend on the value chosen as cut-off point for the definition of malnutrition. The further the cut-off point is form the median, the lower the influence on the prevalence of malnutrition will be in absolute numbers. Or in other words, the impact of the bias will be smaller in percent points if the real prevalence of malnutrition becomes lower in the population. Third, the influence will be larger at a lower age, because younger children weigh less, hence the bias forms a larger proportion of the child's weight, assuming that the bias has a constant value. The maximum influence of the bias occurs when the average of the medians of the reference population and the measured population is chosen as cut-off point (Kostermans 1994: 20).

How much can "method" substitute for data?

Some of major data problems in economic analysis are (i) flawed surveys, (ii) too few surveys, (iii) surveys that are incomparable across time, and (iv) lack of complementary data. Under such conditions, a counterfactual experiment becomes useful as a way of testing how well data problems can be circumvented using "method".

The objective in such experiments is to try the method in a setting where it is not needed because the required data are already available and see how well it performs (see Ravallion 1996 for examples). Moreover, counterfactual experiments could be used in assessing the robustness of poverty comparisons to changes in assumptions, for instance, "drawing on and developing results from the theory of stochastic dominance" (Lipton and Ravallion 1995: 2582).

LAND CONTINGENT TARGETING

Gittinger et al (1990 cited in Maxwell and Smith 199?: 21-22) identify the main causes of household food insecurity in terms of "variations in the amount of food provided by the work and wealth of the household. The level of food consumption can vary because of shocks in work, in production, or in assets. The shock can change the quantity available or change in the price". In conditions of subsistence farming, the effect of any of these shocks is mediated/conditioned by the land rights (access to land) of households.

LAND AS AN INDICATOR IN A TAX-TRANSFER SCHEME

Traditionally, taxes contingent on land ownership have been advocated on the grounds that (i) land taxes are non-distortionary, and(ii) since land is generally held by richer individuals, such taxes are desirable on equity grounds as well. These conclusions have been challenged in terms of their applicability under imperfect risk markets, in particular in developing countries.

But if there are imperfect risk markets, then land-contingent tax-transfer schemes have the disadvantage that the tax liabilities they impose do not vary with farmers' output; they are not "state-contingent". There may be high costs associated with making the rural sector bear this risk–in terms of not only reduced welfare but, in some cases, reduced output as well. Taxes that are related to output represent a sharing and pooling of risks (Hoff and Stiglitz 1993: 326). It is argued that state-contingent tax-transfer schemes are better than land-contingent tax-transfer: it has been observed that sharecropping contracts are often preferred to rental contracts because rental contracts are not state-contingent

> If workers and landlords decide to use state-contingent contracts (sharecropping) in the agreements they arrive voluntarily, shouldn't that be an indication that in the "contract" between the government and its citizens, a state-contingent contract might be desirable? (Hoff and Stiglitz 1993: 326).

The main issue that to be addressed is the extent to which land ownership serves as a proxy to incomes of farming households. Size of farmland owned is highly correlated to household income, as land is the most important factor of production. Land is perhaps the scarcest productive input, especially in the central and northern cereal producing regions–the ox-plough culture–and the *enset* producing regions. The size of farmland affects the efficiency with which other inputs are utilised and the possibility of adopting

modern or/and improved technologies and inputs[5]. *Ownership* of sufficient farmland is especially important where markets for farmland are non-existent or highly imperfect.

Size of land owned by households, however, is not the only variable that determines household income. The factors that modify the relationship between household income and size of land owned include characteristics of farmland (such as quality of land), access to complementary productive inputs, the nature of the market for non-land productive inputs (such as labour and oxen), and demographic structure of households. These complications raise the following questions: Is size of farmland a robust indicator of income of farm households? That is, does size of farmland serve as a good proxy for income after accounting for these variables? If not, how should we calibrate size of farmland to develop a better instrument? The relation of each of these variables is discussed below.

In post-1974 Ethiopia, size of land owned by households changed frequently due to redistribution by peasant associations (the lowest level of public administration) as well as by households themselves, for instance, due to formation of new households as a result of marriage of offspring. Land redistribution undertaken since the land reform, in spite of local differences in the specifics of application, generally had the three major characteristics.

First, household size was an important factor determining size of land given to a household, but the formula for giving allowance to family size varied from place to place. Second, differences in quality of land were accounted for by compensating households with lower quality land by larger land size, and/or giving households plots from every type of land such that more than one plot were given to each to ensure that each household gets some from land of each quality. Third, 'ability to farm' and availability of complementary inputs was generally ignored: in some cases, however, larger land was given to those that are better endowed.

As a result, households with relatively large land may or may not have higher income depending on the relationship of size of land with quality of land, availability of complementary inputs, and family size. Under these conditions, implementing tax-transfer schemes based on means testing that exclusively use size of farm land may act contrary to proclaimed objectives of programs.

---

[5] Alemayehu Seyoum (citing Yeraswork (?)) claims that the minimum size of farmland required to attain efficiency in production is 2(?) hectares.

CHARACTERISTICS OF FARMLAND

The effect of land size on household incomes may arise due to economies of scale. However, evidence from many countries shows that economies of scale are either very small or non-existent in traditional agriculture where lumpy inputs such as machinery are not used. A similar result was found by Assefa et al (1996) in Basona Worana in a cereal monoculture region of Ethiopia. Yet, "the large farmer may have better access to production credit (particularly significant as purchased inputs become more important in modern agriculture), to information and marketing networks, and to the capacity to diffuse and insure against risk" (Bardhan 1995: 66).

Characteristics of farmland that may significantly affect the income-farmland size relationship include fertility of land, slope of land, fragmentation of land and distribution of plots among different micro-climates, such as distribution between micro-climates where double-cropping is possible and where it is not. These factors determining the efficiency with which productive resources are employed, intensity of use of farmland, the product-mix that is feasibility, and vulnerability to disasters that arise due to variations in climate and other elements of the environment. Characteristics of farmland thus determine the level and variability of income in conjunction with size of farmland (see Table 1).

Number of different plots may be more sensitive indicator than total acreage since households with fragmented landholding can take advantage of different micro-climates more than households with larger but often less diverse landholding. It has been found that Access to seasonally flooded lowlands is an important buffering mechanism in drought-prone areas.

Slope of agricultural land affects income of farming households since (i) plots with higher slopes have very thin top soils and lower fertility as they have been exposed to continuous erosion due to minimal/no soil conservation; (ii) plots with higher slopes have lower water retention capacity; (iii) plots with in higher slopes are more difficult to plough using oxen. Where soil conservation measures have been used, some land had been forgone for construction of bunds, etc.

Table 1. Relationship between characteristics of farm-land and farm income

Characteristics	Variable Affected	Direction of Effect
Fertility	Productivity	+
	Temporal variability of output	-
	Number of crops in the product mix	+

	Possibility of double cropping	+
Slope	Productivity	-
Fragmentation	Productivity	-
	Temporal variability of output	-
	Number of crops in the product mix	+
	Possibility of double cropping	+

Studies on land distribution state that in most areas quality of land was taken into account during distribution, and several plots were allotted to a household to ensure that each gets land of varying quality (Dessalegn 1984). Except in the few cases where this was not the case, thus, we would not expect adjustments for land quality to be significant improvements over land-contingent tax-transfer schemes that use land size only. When the comparison is over sites, however, land quality might be an important consideration since sites with larger land per-household may have the additional benefit of having more fertile land. In this case, the income differential between sites might well be larger than what is implied by differences in farmland.

NON-LAND INPUTS

Availability of complementary inputs, especially of labour and draught oxen, affects the level and variability of farm incomes. Markets for these services are highly imperfect and seasonal. In the cereal producing areas in the central and northern highlands, there is a severe shortage of oxen with more than a third of farm households having no oxen and about two of five farm households having only one ox. Seasonal shortage of labour, especially adult male labour, is prevalent in many areas, its prevalence being highest in female-headed and old households.

Lack of complementary farm inputs, oxen and adult male labour, forces households to lease-out their land to share tenants. Hence, households that lack complementary inputs may have less income than what corresponds to the level of land ownership. On the converse, households with extra oxen and adult-male labour may take land owned by other households as share tenants or in the form of rent and their income would be higher than ownership of farmland would indicate. This suggests that size of land cultivated might be a better proxy of household income than size of land owned[6].

The positive relationship that is likely to arise naturally has in some cases been reinforced by land distribution practices in which ownership of non-land inputs led to larger

allotment of land. Dessalegn (1984: 48), for instance, found that in some weredas "[H]ouseholds with the requisite oxen, tools, and 'farm hands' were given enough land for their maintenance; those without those basic items of cultivation were given smaller plots". In such cases, the size of land owned would underestimate household income since land size is positively correlated with wealth in other productive inputs.

FAMILY SIZE

Larger size of land may reflect higher need of households than greater income-generating capacity *per se*. That is, households with larger size of land may have a larger family size–a larger number of mouths to feed–such that the benefits that apparently accrue to households with larger land size do not carry to individual household members. Hence the conclusion "[H]ousehold-level data are not of much use as measures of poverty or well being because of differences in size between households".

Adjustment of household size and composition to recurrent food insecurity is a common strategy. Disaggregated analysis in terms of adult equivalents (consumer units) and per capita give better (and similar) results, though the former is claimed to give "somewhat better" results (Lipton 1983c cited in Lipton and van der Gaag 1993).

Moreover, it has been claimed that systematic changes in household size and demographics in response to food insecurity/poverty[7]. The major trend observed include:

(a) During prolonged crisis the tendency is towards smaller consumption units.

(b) Larger/extended households are more likely than smaller/nuclear households to be associated with greater diversification of assets, income sources and crop cultivation, and less vulnerable to illness/ death of breadwinners.

(c) Poorest households tend to have large young families.

---

[6] However, using size of land cultivated as an indicator of income is not necessarily better than using size of land owned because it is harder to verify information on rented-in. There is no official record like land ownership and households have incentive to withhold information.

[7] Analysis of these issues should involve consideration of the fact that household size/composition changes with household biological life cycle.

(d) Households with female heads are often, but not always, disadvantaged.

REFERENCE

Binswanger, Hans P., Klaus Deininger and Gershon Feder. 1995. "*Power, distortions, revolt ad reform in agricultural land relations*". In Jere Behrman and T.N. Srinivasan. 1995. **Handbook of Development Economics**, Volume III.

Lipton, Michael, and Martin Ravallion. 1995. "*Poverty and policy*". In Jere Behrman and T.N. Srinivasan. 1995. **Handbook of Development Economics**, Volume III.

Dessalegn Rahmato 1984. **Agrarian Reform in Ethiopia.** Scandinavian Institute of African Studies, Uppsala.

Stefan Dercon and Pramila Krishnan 1996. "*A consumption-based measure of poverty for rural Ethiopia in 1989 and 1994*" in Bereket Kebede and Mekonnen Taddesse 1996. **The Ethiopian Economy: Poverty and Poverty Alleviation,** Proceedings of the Fifth Annual Conference on the Ethiopian Economy.

Philippa Bevan and Bereket Kebede 1996. "*Measuring wealth and poverty in rural Ethiopia: a data-based discussion*" in Bereket Kebede and Mekonnen Taddesse 1996. **The Ethiopian Economy: Poverty and Poverty Alleviation,** Proceedings of the Fifth Annual Conference on the Ethiopian Economy.

Bardhan, Pranab 1995. "*Research on poverty and development twenty years after redistribution with growth*" in Michael Bruno and Boris Pleskovic (1995) (eds.) **Annual World Bank Conference on Development Economics.**

Kostermans, Kees 1994. "*Assessing the Quality of Anthropometric Data: Background and Illustrated Guidelines for Survey Managers*" Living Standards Measurement Study Working Paper No. 101, The World Bank, Washington, D.C.

Vella, Venanzio and Maurizio Vichi 1997. Identification of Standards of Living and Poverty in South Africa. Unpublished.

Besley, Timothy and Ravi Kanbur 1993. "*The principles of targeting*" in Michael and Lipton and Jacques van der Gaag (1993) **Including the Poor,** Proceedings of a Symposium Organised by the World Bank and the International Food Policy Research Institute, The World Bank.

Kakwani, N. 1993. "*Measuring poverty: definitions and significance tests with application to Côte d'Ivoire*" in Michael and Lipton and Jacques van der Gaag (1993) **Including the Poor,** Proceedings of a Symposium Organised by the World Bank and the International Food Policy Research Institute, The World Bank.