NORTH-HOLLAND

Journal of Policy Modeling
24 (2002) 237–255

Journal of Policy Modeling

Geographical targeting of poverty alleviation programs: methodology and applications in rural India

David Bigman [a,*], P.V. Srinivasan [b]

[a] *International Service for National Agricultural Research (ISNAR), P.O. Box 93375, 2509 AJ, The Hague, The Netherlands*
[b] *Indira Gandhi Institute for Development Research (IGIDR), Mumbai, India*

Received 1 February 2001; received in revised form 1 October 2001; accepted 1 January 2002

Abstract

The paper presents a methodology for mapping poverty within national borders at the level of relatively small geographical areas and illustrates this methodology for India. Poverty alleviation programs in India are presently targeted only at the level of the state. All states includes, however, many non-poor households, whereas many poor households who live in states that have not been selected for the targeted programs are left out. The paper applies the methodology for mapping poverty at the level of districts in India, thus increasing the target areas from the 24 states (with over 40 million people in each) to 466 districts (with only around 2 million people in each). © 2002 Society for Policy Modeling. Published by Elsevier Science Inc. All rights reserved.

Keywords: Geographical targeting; Poverty alleviation

1. Introduction

A significant constraint on most poverty alleviation programs is the lack of accurate and up-to-date information on the level of income and other characteristics of individual households that are necessary to identify the poor. Untargeted programs that cover the entire population are generally beyond the budgetary means of

* Corresponding author. Tel.: +31-70-349-6235; fax: +31-70-381-9677.
E-mail address: d.bigman@cgiar.org (D. Bigman).

0161-8938/02/$ – see front matter © 2002 Society for Policy Modeling.
PII: S0161-8938(02)00108-4

most developing countries, and the benefits of many of these programs go disproportionately to the non-poor (Besley & Kanbur, 1993). More accurate targeting requires, however, detailed information in order to identify eligible recipients. Narrow targeting at the level of individual households, for example, requires very detailed data on all households and a complex and expensive means of testing in order to identify the eligible households (Ravallion & Chao, 1989). In most developing countries, these detailed data are not available and the costs of these programs are blown up by widespread leakage to non-poor households (Kanbur, 1987). In addition, these programs often suffer from ineffective implementation and high administrative costs, and their overall costs are further augmented due to the incentive they give to households to change their characteristics (e.g. the number of children) or provide wrong information in order to qualify for the program.

Alternatively, targeting government programs on the poor can be based on certain visible and clearly identifiable characteristics that approximate the household's income or standard of living. Typical examples are the households' housing conditions, access to drinking water, availability of electricity, etc (Ravallion and Wodon, 1997). Other targeted programs focus on the elderly, on young children, or on lactating mothers. Many non-poor households share these characteristics, however, and they also benefit from these programs, whereas poor households that do not have these characteristics remain uncovered. Moreover, once the program determines these characteristics as criteria for eligibility, all households are given an incentive to change their relevant characteristics or misinform about them in order to be eligible for these benefits.

Geographical targeting that determines the place of residence as the main eligibility criterion can be an effective way of reaching the poor in countries where there are substantial disparities in the living conditions between geographical areas. The administration of geographically targeted programs is relatively less complex and their implementation can be facilitated by the existing local administrations. In India, the central government disburses funds across states in large measure according to the income disparities among states, and the implementation of rural development projects in the country's poorer regions has become the center of India's poverty-oriented agricultural development strategy (Datt & Ravallion, 1993).

Targeting at the level of large administrative areas can entail, however, considerable leakage to the non-poor who live in these areas. Although programs targeted on large geographical areas can still provide better coverage than the general programs, the savings achieved by these programs may not be very large (Baker & Grosh, 1994; Ravallion, 1993, 1998). Baker and Grosh concluded that targeting on selected regions can be an effective mechanism for transferring benefits to the poor but the smaller the target areas then the greater will be the reduction in poverty that can be achieved with the same overall budget. In their analysis, the greatest reduction in poverty is achieved when the target areas are municipalities or villages (Jalan & Ravallion, 1998).

In rural areas in particular, the smaller the geographical area on which the program is targeted, the more homogeneous its population is likely to be. At the village level, income differences between households are due, in large measure, to differences in the size of their landholdings. All households in the village are affected, however, by the same agro-climatic and spatial conditions and even the employment opportunities of the landless households depend on these conditions. At the district level, differences between households are due also to differences among villages in their distance to the market, their road conditions, their access to drinking water, the availability of health services, electricity, the telephone connection, etc (Glewwe & Kanaan, 1989).

The larger the geographical area, the larger the impact of spatial and agro-climatic factors such as the distance from the main urban centers and ports, the soil conditions, the access to irrigation facilities, rainfall, etc. and the larger, therefore, the income differences *between* villages and households *within* this area. Consequently, in a large geographical area, the overall income differences are due primarily to income differences between villages whereas income differences between individuals within the same village have a much smaller weight. The smaller the target areas, therefore, the more homogeneous the population in terms of its standard of living; accordingly, targeting on small areas such as districts or even villages in which the majority of the population is poor can reduce significantly the leakage to the non-poor. Targeting on small areas requires, however, rather detailed information on the socio-economic characteristics of the households that reside in these areas. However, in most developing countries this detailed information is not available (Bigman & Deichmann, 2000).

In India, information on the socio-economic characteristics of households in different geographical areas is available from three main sources. The most comprehensive source is the Population Census, which has been conducted every decade since 1871 and provides information on certain demographic and socio-economic characteristics of all households in the country. The census provides, only very limited information on the socio-economic conditions of the households and, in particular, it does not provide any information on the living standards or the income of the households. Moreover, in most countries the data of the census for individual households are not available to researchers due to their political sensitivity and they are regarded as highly confidential. In India, the small size of this data set for individual households is also a severe constraint (Murthi et al., 1999).

The second most important source is the National Sample Survey (NSS) of households' income and expenditures. This survey has been conducted every 5 years since 1950–1951 and it provides detailed information not only on households' consumption expenditures, but also on many other socio-economic characteristics. The NSS covers all states and regions, but only a small sample of villages and a small sample of households within each village. The sample size in this survey is, therefore, too small to allow an estimation of the incidence of poverty at the level of the village or even the district. Mapping the spatial distribution of poverty on the basis of this survey is, therefore, possible only at the level of the country's main regions and states. Another source of information is the Agricultural Census that provides detailed data on households' landholdings, agricultural production, farming systems, etc. along with climatic and soil conditions data (Dreze & Srinivasan, 1996).

The third source is the Cost of Cultivation Survey that provides detailed information on production costs and yields. These data are available only for a sample of districts and a relatively small sample of households and they do not provide information on the standard of living of the households. A direct application of the data from all these sources for estimating the prevalence of poverty in small geographical areas of the district or the village is, therefore, impossible.

The objective of this paper is to present a methodology for narrow geographical targeting and illustrate its application in India. The method is based on the use of data from several sources after bringing these data together at the level of the target area—the district or the village—by using their geographical coordinates. The overall goal of this analysis is to evaluate the effect and cost-effectiveness of poverty alleviation programs that are targeted on small geographical areas. In this paper, however, the analysis is incomplete because the data set that was available to us was not complete and the paper's main objective is to illustrate the use of the methodology (Hentschel et al., 2000; Bigman et al., 2000; Bigman & Fofack, 2000).

2. Data sources and methodology

The NSS is the main source of data for estimating the poverty incidence in India. This survey collects data on households' consumption expenditures at periodic time intervals and the sample size is large enough to provide estimates of the incidence of poverty in geographic areas called "NSS regions." The borders of these regions are determined within the states, mostly on the basis of their agro-climatic characteristics, and are considerably larger than the districts. In the recent NSS survey, the 24 Indian states were divided into 77 NSS regions and 466 districts. Hence, the size of the population is, on average, 40 million people in a *state*, 15 million in a *region* and 2 million in a *district*.

The government poverty alleviation programs are targeted on selected states and they are implemented by the state's administration. At this level, however, the population is not very homogeneous and targeting poverty alleviation programs on *states*, therefore, involves considerable leakage to the non-poor. Targeting the programs on *districts* and implementing them by the district administration may, therefore, be much more cost-effective and can reach a larger number of poor people. However, to determine the target districts it is necessary to have robust estimates of the incidence of poverty in all districts in the country and the sample size in the NSS survey at the district level is too small to provide statistically significant estimates. The objective of the methodology presented in this paper is to bridge this gap and provide estimates for the incidence of poverty in all districts.

This methodology is based on an econometric estimation of the poverty indices at the district level by using data from a wide variety of sources that include the census, the NSS, the Agricultural Survey, the Cost of Cultivation Survey and various geographical surveys. This methodology is based on a five-step procedure:

(I) Econometric estimation of the impact of district-specific characteristics based on the probability that the households residing in a given district are poor. This estimation is based on the entire NSS sample of households and the characteristics of districts from all other sources.
(II) Predictions of the incidence of poverty in *all* the districts of the country based on the characteristics of these districts which are available in the other data sources, and the relationships estimated in step I.
(III) First validation of the predictions: this validation is based on a comparison of the ranking of *states* established by the predicted values of the incidence of poverty in the states with the ranking established by the levels of poverty in the states computed directly from the NSS survey data. High rank correlation for the states indicates that the corresponding rank correlation for the districts is also likely to be high.
(IV) Ranking districts from the poorest to the least poor according to the predicted values of the incidence of poverty in each district and grouping the districts into broad poverty groups, each containing an equal share in the general population. The group of the *poorest* districts includes the districts that can be the target of poverty alleviation programs; the group of the *least poor* districts includes the districts that could be the target of cost-recovery programs.
(V) Second validation of the predictions: this validation is based on a comparison of the predicted levels of the poverty incidence in *groups* with the actual levels of poverty in *these groups* computed directly from the NSS survey data.

The first step of deriving estimates for the impact of the district-specific characteristics on households' poverty status combines household level data (consumption expenditures as well as other socio-economic data) from the NSS with data on districts from the other sources. In India, in addition to the data on households' expenditures, the NSS also provides information on other characteristics of the households, including the following:

- Household location—state, district, rural/urban
- Household size
- Occupation of the head of the household
- The size of the household's land holding (hectares)
- The household's sources of income (land cultivation, wage income, fishing, non-agricultural enterprise)
- Participation in public works—(y/n)

- Recipient of government assistance (IRDP)—(y/n)
- Education level of the household's members.

The data in the Population Census are available to researchers only at the level of the *district*—although, in principle, these data could also be available for smaller geographical areas and even for individual villages. The census data in India include information on the following:

- Population size in the district (male, female, rural, urban)
- Fertility rate
- Literacy rate (male, female, rural, urban)
- Infant mortality rate
- Female participation rate in the labor force.

Additional sources of data on other characteristics of the districts include the 'Statistical Abstracts' and the 'Season and Crop' reports. Not all these data are available for all states, though. The data included in the Statistical Abstract and the data compiled by the Center for Monitoring the Indian Economy in its publication 'Profiles of Districts' are listed in Appendix A.

In step I, the impact of district-specific characteristics on the probability that the household's per capita consumption expenditures (PCE) fall below the poverty line is estimated. There are two sets of explanatory variables: the household's individual characteristics (obtained from NSS) and the characteristics of the 'area' in which the household resides (obtained from all other sources). Data on certain 'area' characteristics may not be available at the level of the district, primarily agro-climatic data that are typically available only at the level of the *region*. This estimation is made by means of a probit or logit regression analysis where the dependent variable is binary and it indicates whether or not the household's PCE is below the poverty line. The dependent variable assumes value 1 if the household's PCE is below the poverty line and value 0 otherwise. The probability that the level of per capita consumption of an individual household with the characteristics specified by the explanatory variables falls below the poverty line is measured by:

$$\begin{aligned}\mathrm{prob}(Y=1) &= \mathrm{prob}\{(\boldsymbol{X}^{\mathrm{H}})'\beta^{\mathrm{H}} + (\boldsymbol{X}^{\mathrm{A}})'\beta^{\mathrm{A}} + \varepsilon \le Z\} \\ &= F\{Z - [(\boldsymbol{X}^{\mathrm{H}})'\beta^{\mathrm{H}} + (\boldsymbol{X}^{\mathrm{A}})'\beta^{\mathrm{A}})]\} \qquad (1)\end{aligned}$$

where Y is the household's poverty status, $\boldsymbol{X}^{\mathrm{H}}$ the vector of explanatory variables that describe the household's characteristics, $\boldsymbol{X}^{\mathrm{A}}$ the vector of explanatory variables that describe the characteristics of the 'area'—the district (or the region) in which the household resides, Z the poverty line and F is the cumulative distribution function which is standard normal in the case of probit and logistic in the case of logit regression.

The regression analysis is conducted over the entire data set of the NSS after incorporating the vector $\boldsymbol{X}^{\mathrm{H}}$ of individual households' characteristics from the NSS and the vector $\boldsymbol{X}^{\mathrm{A}}$ of the districts' characteristics from the Population Census and from the other sources. For a given poverty line Z and a given set of observations on $\boldsymbol{X}^{\mathrm{H}}$ and $\boldsymbol{X}^{\mathrm{A}}$, the estimates of β can be obtained by maximizing the corresponding likelihood function.

Step II is the prediction of the incidence of poverty in *all* districts. These predictions are made on the basis of the 'area' characteristics only and the relationships between these characteristics and the probability that households residing in these areas are poor—based on the estimates obtained in the first step. In other words, in this step the probability that households in a given district are poor is predicted on the basis of the district characteristics, and this prediction is given by:

$$\mathrm{prob}(Y = 1) = \mathrm{prob}\{(\boldsymbol{X}^{\mathrm{A}})'\beta^{\mathrm{A}} + \varepsilon \leq Z\} = F\{Z - (\boldsymbol{X}^{\mathrm{A}})'\beta^{\mathrm{A}}\} \tag{2}$$

where coefficients β^{A} are the sub-group of coefficients of $\boldsymbol{X}^{\mathrm{A}}$ that have been estimated in the first step. The probability that households residing in a given district are poor is given by the estimate of the incidence of poverty in that district—established in (2). The prediction error in these estimates depends on how detailed and how accurate the information available is for these districts in all the other sources. When this information is not very detailed, the prediction error can be quite large. In these cases, the predictions may not be used directly for *individual* districts and instead the analysis in the subsequent steps is conducted for *groups* of districts. In addition, the predictions must meet the tests established by the two validation steps III and V in order to assess their reliability.

To determine target groups of districts for the application of suitable policies, the districts are ranked in step IV according to the values of their predicted probabilities from the *poorest* district, in which the (predicted) incidence of poverty is the highest, to the *least poor* district, in which the (predicted) incidence of poverty is the lowest. The districts are then divided into several target groups that have equal shares in the *general* population. The districts in the first group, in which the predicted probabilities are the highest, are the highest priority targets for poverty alleviation programs; the districts in the last group, in which the predicted probabilities are the lowest, could be the target of cost-recovery programs. For simplicity, we divide the districts into four groups and refer to them as "poorest," "highly poor," "poor," and "least poor." The number of households in the NSS in each of these groups is sufficient to provide statistically significant estimates of the actual incidence of poverty in each group on the basis of the NSS data.

In the first test to validate the reliability of the predictions, we evaluate the reliability of the district ranking by comparing the ranking of *states* established by these predictions with the ranking established by the NSS data. The higher the coefficient of rank correlation for the states the higher the likelihood that the ranking of districts established by these predictions is also highly correlated with the ranking established by the NSS data. Moreover, by dividing the districts into four groups and targeting policies on the two extreme groups, a clear and significant difference is thereby established between the "poorest" districts and the "least-poor" ones. This division can, therefore, significantly reduce the error of including non-poor districts—according to the ranking determined by the NSS data—in the group that would be the target of poverty alleviation programs and including poor districts in the group that would be the target of cost-recovery programs.

In step V, we conduct a second test to validate the reliability of the predictions. In this test we compare the incidence of poverty and the "error of inclusion" for the four groups that are calculated from the NSS survey with the poverty incidence and the error for these groups that are calculated in step IV on the basis of the predicted values for the individual districts. Since the sample of households in each group of districts is sufficiently large, we can calculate the incidence of poverty in the group from the NSS survey. In step V we can, therefore, compare the NSS estimates of the poverty incidence in the groups with the estimates for the groups obtained from the econometric estimations. If the difference is not very large, we can conclude that, despite the possibly high prediction error of the econometric estimations for *individual* districts, the estimations for groups of districts are sufficiently reliable.

3. The econometric analysis

As noted in the Section 1, the objective of the empirical analysis in this paper is primarily illustrative. In large measure, this is because the data we had was incomplete: the census data that we had were only for 15 of the 24 states and 340 of the 466 districts and we had only limited data on the characteristics of districts from the other surveys. For the econometric analysis, we use the household survey data of 1987–1988 and the data of the other sources are also from this year. The data of the Population Census are, however, from the census of 1991.

Table 1 gives the estimated coefficients of the probit regression that estimates the probability that the household is poor as function of various household level characteristics and various characteristics of the area in which the household resides. These estimates indicate that variables such as household's size and ownership of land, milk and draught animals have a significant impact on the probability that the household is poor. Access to the public distribution system (PDS) also has a significant impact on this probability. The positive and significant value of the coefficient for the dummy variable for sex of the household's head (1 = female head) indicates that female-headed households face a higher risk of being poor.

The characteristics of the districts that were included in this analysis, primarily the population density and female literacy, also have a significant impact on this probability. The R^2 value in the probit regression with the area characteristics only, in regression Eq. (1), is less than 0.3, indicating that the explanatory power of the area characteristics is quite low. In part, this is due to the specific and rather limited group of 'area' characteristics that were available to us at the time we conducted this analysis. In large measure this was due, however, to the fact that the 'area' characteristics can only explain the districts' average incomes

Table 1
Coefficients and *t* values of the probit regressions[a]

Explanatory variables	Coefficient (*t* value)	
Household characteristics		
Sex of household head		0.141 (6.74)
Household size		0.135 (35.88)
Land owned		−0.103 (−6.45)
Dummy for ownership of milch animals		−0.375 (−20.33)
Dummy for ownership of draught animals		−0.055 (−2.69)
Dummy for PDS beneficiary		−0.158 (−8.94)
Area characteristics		
Population density	0.0006 (16.86)	0.0004 (10.48)
Female literacy	−0.0063 (−8.85)	−0.007 (−9.47)
Medical infrastructure	−0.0023 (−9.00)	−0.0029 (−8.49)
Under 5 child mortality	0.002 (5.88)	0.0023 (6.44)
Urbanization	−0.006 (−6.79)	−0.004 (−4.03)
Irrigation	−0.0024 (−4.91)	−0.0032 (−6.16)
Fertilizer consumption	−0.0018 (−7.22)	−0.0015 (−5.87)
Constant term	−0.196 (−3.46)	−0.462 (−7.40)

[a] The dependent variable takes the value 1 if the household's monthly PCE fall below the poverty line, and the value 0 otherwise.

and income differences between districts, but not income differences between individuals within districts. Since the size of the population in a district is still quite large, income differences between individuals in a district are the main contributors to the overall income inequality. For this reason, it is important to narrow down the target area in this analysis as much as possible. In the present context, this makes it necessary to conduct the subsequent analysis for groups of districts.

A comparison of the *t* values in regression Eqs. (1) and (2) shows that in some of these variables, primarily the 'population density,' 'infant mortality,' and 'urbanization,' the magnitude of the *t* values change quite considerably, although they do not change their sign. These changes indicate that the 'area' characteristics are not 'orthogonal' to individual characteristics and one cannot learn, therefore, from the R^2 value of the regression Eq. (1) on the *marginal* impact of the area characteristics in regression Eq. (2). The significant *t* values of the area characteristics in the regression equations indicate, however, that they contribute significantly to explain the probability that the household is poor. This probability increases with the rise in population density in the district and with a rise in infant mortality. The higher the share of the urban population in a district and the higher the rate of female literacy, the lower the probability that households that reside in this area are poor. The larger is the size of the area under irrigation in a district and the larger is the consumption of fertilizers, the lower the incidence of poverty in that district.

4. Predicting the incidence of poverty for all districts

Step II is an application of the estimates of the coefficients (β^A) that were obtained in the econometric analysis and the area characteristics ($\boldsymbol{X}^A$) that are available for all districts from the Population Census and from the other sources in order to predict the incidence of poverty for all the districts in the country on the basis of the functional relations in Eq. (2).

Step III is first validation of the estimates by comparing the ranking of *states* according to the incidence of poverty that is established by these estimates with the ranking established by the NSS data. The focus of the test is on the *ranking* of the states, since this is the relevant criterion for determining also the ranking of the districts which in turn determines their eligibility for targeted programs. The rank correlation of the poverty levels in the states obtained for the rankings was .757. Although our analysis includes only 15 states, this coefficient is sufficiently high to ensure that the error of inclusion in policies targeted on the two extreme groups is small.

In step IV, the districts are divided into four groups ranging from the "poorest" to the "least poor"—each group having a (nearly) equal share in the general rural population (in the 15 states for which we have data).

Table 2 summarizes the main characteristics of the four groups on the basis of these estimates. The Table also includes the NSS estimates head-count ratios in the groups. The objective is to conduct the second validation test in step V that compares the NSS estimates of the incidence of poverty in the groups to the predicted levels of poverty in the groups based on econometric estimates for districts. It should be noted, however, that the relevant comparison to test the reliability of econometric estimates is not the levels but the extent to which the incidence of poverty in the group of the "poorest" districts is higher than that in the group of the "least poor" districts—according to the two estimates. The reason is that the higher the reliability, the smaller the differences between the ranking of

Table 2
Estimates of the incidence of poverty in the four groups[a]

	Poorest	Highly poor	Poor	Least poor	Total[b]
Total population (% of the total population)	143.8 (24.7)	146.2 (25.1)	146.1 (25.1)	146.5 (25.1)	582.6 (100)
Number of poor (% of the poor population)	69.7 (33.5)	57.8 (27.7)	47.3 (22.7)	33.6 (16.1)	208.4 (100)
The econometric estimates of the head-count ratio	48.5	39.0	32.4	22.9	35.7
NSS estimates of the head-count ratio	51.8	43.1	35.9	24.3	38.7

[a] Normalized estimates.

[b] The total rural population in the 15 states at the time of the survey in 1991.

districts determined by the NSS and by the econometric estimates. In our analysis, the average head-count ratio in the group of the "poorest" districts is 112% higher than in the group of the "least poor" districts, according to the NSS estimates, and 113% higher, according to the econometric estimates. The average head-count ratio in the group of the "highly poor" and "poor" districts is 77.4% and 47.7% respectively higher than in the group of the "least poor" districts, according to the NSS estimates, and 70.3% and 41.5% respectively higher, according to the econometric estimates.

In India, the level of poverty in individual states is one of the main criteria that guide the transfer of resources from the central to the state governments. At present, reliable estimates of poverty are available only at the level of the states and resource transfers from the state administration to the local bodies are currently based on informal methods which can be politically motivated. Estimates of the incidence of poverty in smaller geographical areas can, therefore, aid and de-politicize the transfer of resources from the state government to the local administration.

In general, poverty alleviation programs that are targeted on states are likely to include districts from all the four categories and thus may reach the poor less than programs that are targeted on the "poorest" districts only. In India, however, the differences in the incidence of poverty between the "poorest" group and the two median groups is not very large and the gains from a program targeted on the "poorest" districts rather than on the two poorest states of Bihar and Orissa will, therefore, be quite limited. (We will discuss this option later.) It should also be noted that even in the "least poor" districts in India there are a considerable number of the poor and cost-recovery programs that are targeted on the "least poor" districts may, therefore, be prohibited because they include so many of the poor.

Table 3 illustrates the potential cost-savings with district targeting of poverty alleviation programs. The first three columns provide the relevant statistics of the four poorest states in India. In Bihar, for example, there are 35 districts and 68.7 million people. The third column indicates the percentage of the population in Bihar who are not poor; these are people who will be included in a poverty alleviation program that is targeted on the entire state. The subsequent five columns provide the relevant statistics for the poorest districts in these states. In Bihar, only 25 of the 35 districts are in the "poorest" category and they include nearly 70% of the state's population and 75% of the state's poor. Half the population in the poor districts is not poor, however. Hence, poverty alleviation programs targeted on the poor districts will reduce the "error of inclusion" from 53 to 50% but the reduction is quite small. The reason is that the other districts of Bihar also include a large number and a high percentage of poor people. In this case, the gains from district targeting versus state targeting are relatively small.

The gains from district targeting are somewhat larger in the state of West Bengal where district targeting includes half of the districts and half of the state's population. The error of inclusion will be reduced from 57 to 52% and the resources that were used under state targeting to cover the less poor districts in

Table 3
Targeting poverty alleviation programs on poorest states vs. poorest districts

State	Poorest states			Poorest districts				
	Number of districts	Population (millions)	Error of inclusion ($1-H$)	Number of districts	Population (millions)	% of the state population in target districts	% of state poor population in target districts	Error of inclusion ($1-H_{\mathrm{D}}$)
Bihar	35	68.7	0.53	25	48	69.9	74.8	0.50
Orissa	12	26.6	0.56	9	20.4	76.7	79.6	0.54
Madhya Pradesh	44	49.7	0.57	24	25.1	50.4	55.4	0.53
West Bengal	15	49.8	0.57	7	25.1	50.5	56.8	0.52

West Bengal can be used, under district targeting, for poorer districts in India. Although the reduction in the error of inclusion is rather small, district targeting will make it possible to exclude a considerable number of the non-poor from the program and thus reduce its costs. If a subsidy program targeted on the four poorest states could be replaced by a program targeted on the poorest districts only, around 7 million non-poor households could be excluded and the subsidies given to these households under state targeting could be saved. We will later discuss this option in more detail.

The rather small reduction in the error of exclusion is due to the fact that district targeting is still too wide. This is the case in which additional criteria for targeting that identify selected sub-groups of the rural population can be highly desirable. In our subsequent analysis, we plan to examine the possibility of narrowing the targeting to selected socio-economic sub-groups within these regions. In particular, we plan to explore the possibilities and the potential gains from narrowing the targeting to three socio-economic sub-groups: landless rural households, small landowners and larger landowners. The land owned by the households is relatively easy to measure and monitor and ‘cheating’ under this program may not be very easy. Better targeting can also be achieved by narrowing the target geographical areas. In principle, targeting can be made at the sub-district or even the village level if appropriate data are available.

In Table 4, the option of targeting a cost-recovery program is evaluated. In this case, the goal is to reduce as much as possible the number of poor people who will be included in the program. The Table shows, however, that even in the most affluent state of Punjab, where all districts are in the “least poor” category, one-fifth of the state’s population are poor. Geographical targeting of cost-recovery programs is, therefore, very difficult and additional socio-economic criteria will be required to reach the group of non-poor consumers and reduce the error of inclusion. District targeting offers a reduction in the error of inclusion, but the implications are that replacing a cost-recovery program targeted on the four most affluent states by a program targeted on the “least poor” districts will only reduce the number of the poor that are included in this program by around 2.5 million people.

5. Effectiveness of narrow geographical targeting

To provide a more complete evaluation of the effectiveness of narrow geographical targeting at the district level relative to targeting at the state level, we compared a subsidy program targeted on several of the poorest states with a program targeted on the “poorest” districts when the two programs have the same overall budget. The two programs thus cover the same total number of people and the criterion of effectiveness consists of the increase in the number of the poor that can be covered with district targeting. The district targeting in this comparison is referred to as ‘equivalent district targeting.’

Table 4
Targeting cost-recovery programs on least poor states vs. least poor districts

State	Least poor states			Least poor districts				
	Number of districts	Population (millions)	Error of inclusion (H)	Number of districts	Population (millions)	% of state population in target districts	% of state non-poor population in target districts	Error of inclusion (H_D)
Punjab	12	14.3	0.20	12	14.3	100	100	0.20
Gujarat	18	26.9	0.25	15	21.5	79.9	82.6	0.22
Haryana	12	12.4	0.26	8	8.7	69.8	72.7	0.23
Tamil Nadu	18	36.8	0.26	12	21.5	58.4	61.3	0.22

Table 5
Comparison of state targeting with equivalent district targeting

Target states	State targeting		Equivalent district targeting				
	Total population covered (millions)	Poor population covered (millions)	Poor population covered (millions)	Additional number of poor covered (millions)	% increase in the number of poor covered	Number of districts covered	Number of districts not in target states
Poorest state	68.7	32.2	35.8	3.6	11.2	35	17
Two poorest states	95.3	43.9	48.2	4.3	9.8	47	26
Three poorest states	145.1	65.3	70.3	5	7.7	91	25

In Table 5 the two alternatives of state targeting and 'equivalent district targeting' are evaluated. In the poorest state of Bihar the size of the population is 68.7 million and 32.2 of them are poor. With 'equivalent district targeting' it is possible to reach 35.8 million poor people and thus increase the coverage by 3.6 million people. Of the 35, districts that will be covered by the 'equivalent district targeting' 17 are not in Bihar but in other states, because the poverty rates in these districts are lower than in the 18 districts of the state of Bihar.

When the program is extended to the two poorest states by also including the state of Orissa, the number of people covered by the state targeting program will be 95.3 million and the number of poor people covered will be 43.9 million. With 'equivalent district targeting' the number of poor people covered will increase to 48.2 million, thus increasing the number of poor people covered by 4.3 million and reducing the leakage to the non-poor by 9.8%. Of the 49 districts that will be covered under the 'equivalent district targeting' 26 are not in Bihar or Orissa but in other states because the incidence of poverty in these districts is higher than in the 24 districts in Orissa and Bihar that are not covered by the 'equivalent district targeting' program.

Although the number of poor people covered with 'equivalent district targeting' rises to 3.6 million with two states and to 4 million with three states, the percentage increase in the number of the poor is gradually reduced from 11.2% with one state to 9.8% with two states and to 7.7% in three states. One reason is the concentration of poor districts in the poorest states. With the expansion of the program to more states, an increasing number of the additional districts that will be covered in the 'equivalent district targeting' will, therefore, be in the poorest states. The other reason is the high incidence of poverty in those districts that will be covered by state targeting; but will not be covered by the 'equivalent district targeting' program. In the particular case of India, the marginal gains from district targeting may, therefore, be smaller as the program is expanded to a large number of people. As noted earlier, the presence of a relatively large number of the poor even in the most affluent states and districts requires more accurate specification of the target groups in two directions: First, narrow the target area, possibly even to the level of the village, and second, add additional targeting criteria that are easy to identify and monitor.

6. Concluding remarks

Narrow geographical targeting of poverty alleviation programs can improve their coverage, reduce the leakage to the non-poor and the program costs. In India, by narrowing the target area from the level of states to districts, the size of the population in each target area is reduced from 40 to 2 million, thus supplying the information for more accurate targeting. The method suggested in this paper for poverty mapping in small geographical areas provides the necessary tool for determining the target areas. Our analysis of the data in India indicates that the gains from district targeting in terms of the savings in costs and the increase in the number of the poor covered will be in the order of 10%. While these gains are by no means negligible, their rather moderate magnitude typically reflects the wide spread of poverty in all the states and districts in India where all too often one finds a slum next to a fancy apartment building.

However, even in India district or even smaller area targeting can, in addition to the financial gains, offer a number of important advantages over state targeting

- The political support of the non-poor is essential for funding poverty alleviation programs. District targeting is likely to generate more political support because it is spread over a larger number of states and thus benefit more diverse populations of various ethnic origins and different religions.
- Even when the Federal program is based on state targeting, the information on the incidence of poverty at the district level will provide additional criteria to the state authorities for distribution within the state.
- District targeting can leave much greater freedom to the district authorities in the design of the program and tailor it to the specific characteristics of the district population.
- Implementation of the program at the district level can give much greater role to local institutions and the NGOs and cater to the specific needs of the target population.

This analysis was primarily illustrative, however, and the focus on district targeting reflects primarily our limited access to more detailed data that could allow us to estimate the spatial distribution of poverty in smaller geographical areas. The outcome of this analysis highlights, the potential gains from a more detailed analysis that can provide guidelines for narrower geographical targeting, possibly even at the village level. At the district level, the analysis can be extended by dividing the rural population into sub-groups according to specific and easily identifiable socio-economic characteristics. An obvious example is the size of landholdings. By dividing the rural population into three groups of landless rural households, small land holder (e.g. >1 ha) and large landholders we can, in principle, divide the entire rural population into nearly 1400 groups of households [(466 districts) × (3 landholding groups)] and achieve much better targeting. The analysis in this paper should, therefore, be regarded as a first step in a sequence of more refined analyses that can identify the most effective levels of targeting, the most effective programs at these levels and the potential gains from targeting.

Appendix A

The data in the Statistical Abstract include:

- Cropped area (cropping pattern: area under different crops)
- Crop yields (all crops)

- Irrigated area
- Area under modern varieties
- Rainfall
- Agricultural wages
- Farm harvest prices.

The ‘Fertilizer statistics’ published by the Fertilizer Association of India provides data on fertilizer use at the district level. In addition, The Center for Monitoring the Indian Economy has data in its publication ‘Profiles of Districts’ (November 1993). The following is the list of variables compiled in this source:

- Total area
- Area under forests
- Population density
- Proportion in urban areas and the total population in the district
- Workers as a percentage of the total population
- Share of the main occupation categories in the total workforce
- Average landholding under cultivation
- Value of output of major crops (per hectare, per capita)
- Density of roads (road length per 100 km^2)
- Railway routes (length per 100 km^2)
- Post offices density (number of brunches per 1000 persons)
- Telephone density (number of telephone instruments per 1000 persons)
- Bank credit to agriculture (in rupee per hectare, per capita).

References

Baker, J., & Grosh, M. (1994). *Measuring the effects of geographical targeting on poverty reduction*. LSMS Working Paper No. 99. Washington, DC: The World Bank.

Besley, T., & Kanbur, R. (1993). The principles of targeting. In M. Lipton & J. Van der Gaag (Eds.), *Including the poor*. Washington, DC: The World Bank.

Bigman, D., & Deichmann, U. (2000). Geographical targeting: a review of different methods and approaches. In D. Bigman & H. Fofack (Eds.), *Geographical targeting for poverty alleviation*. The World Bank, Regional and Sectoral Studies.

Bigman, D., & Fofack, H. (2000). *Geographical targeting for poverty alleviation*. The World Bank Sectoral and Sectoral Studies.

Bigman, D., Dercon, S., Guillaume, D., & Lambotte, M. (2000). Community targeting for poverty reduction in Burkina Faso. *The World Bank Economic Review*, *14*(1), 167–194.

Datt, G., & Ravallion, M. (1993). Regional disparities, targeting, and poverty in India. In M. Lipton & J. Van der Gaag (Eds.), *Including the poor*. Washington, DC: The World Bank.

Dreze, J., & Srinivasan, P. V. (1996). *Poverty in India: regional estimates, 1987–1988*. Discussion Paper No. 129. Mumbai: IGIDR.

Glewwe, P., & Kanaan, O. (1989). *Targeting assistance to the poor using household survey data, policy, planning and research*. Working Papers WPS 225. Washington, DC: The World Bank.

Hentschel, J., Lanjouw, J., Lanjouw, P., & Poggi, J. (2000). Combining census and survey data to study spatial dimensions of poverty: a case study for Ecuador. *The World Bank Economic Review*, *14*(1), 147–166.

Jalan, J., & Ravallion, M. (1998). Are there dynamic gains from poor-area development program. *Journal of Public Economics*, *67*(1), 65–85.

Kanbur, R. (1987). Transfers, targeting and poverty. *Economic Policy*, (4), 141–147.

Murthi, M., Srinivasan, P. V., & Subramanian, S. V. (1999). *Linking the Indian census with the National Sample Survey*. Working Paper. Cambridge: Centre for History and Economics, King's College.

Ravallion, M. (1993). Poverty alleviation through regional targeting: a case study for Indonesia. In K. Hoff, A. Braverman, & J. Stiglitz (Eds.), *The economics of rural organization*. Oxford: Oxford University Press.

Ravallion, M. (1998). Poor areas. In A. Ullah & D. Giles (Eds.), *Handbook of applied economic statistics*. New York: Marcel Dekkar.

Ravallion, M., & Chao, K. (1989). Targeted policies for poverty alleviation under imperfect information: algorithms and applications. *Journal of Policy Modeling*, (11), 213–224.

Ravallion, M., & Wodon, Q. (1997). *Poor areas or only poor people*? LSMS Working Paper No.1798. Washington, DC: The World Bank.