Discussion Papers in Economics

Poverty targeting in public programs: A comparison of alternative nonparametric methods.

Isha Dewan

Rohini Somanathan

June 2004

Discussion Paper 04-16

Indian Statistical Institute, Delhi
Planning Unit
7 S.J.S. Sansanwal Marg, New Delhi 110 016, India

Poverty targeting in public programs: A comparison of alternative nonparametric methods.*

Isha Dewan †and Rohini Somanathan‡

June 1, 2004

Abstract

Very poor households may be excluded from public programs intended for their benefit for a variety of reasons such as lack information, a permanent residence or membership in social networks. We are interested in methods of testing for such exclusion based on independently drawn samples of program participants and non-participants. We discuss three alternative nonparametric procedures; sign tests, tests for stochastic dominance and a test for distribution crossing. In the cases where there is a poverty threshold below which program participation is difficult, our simulation results suggest that the last of these test procedures is the most powerful. We apply this test to data from a microfinance program in India and find evidence that the poorest households in the area were largely outside the program.

*We are grateful to Charlie Brown, Julie Cullen, Roger Gordon, Subhash Kochar, B.L.S. Prakasa Rao and E. Somanathan for very useful discussions and to the Consultative Group to Assist the Poorest at the World Bank for funding our data collection efforts.

†Indian Statistical Institute, New Delhi, isha@isid.ac.in

‡Department of Economics, University of Michigan and Indian Statistical Institute, New Delhi, rohinis@umich.edu.

1 Introduction

Every public program faces the challenge of reaching intended beneficiaries. Documented deficiencies in many older social transfer mechanisms have led governments, non-government organizations and donor institutions to embrace institutions which use innovative methods of transferring resources to poor households. Some of these (such as the Grameen Bank of Bangladesh), provide credit to poor households for micro-enterprises, some (like social funds in Peru), subsidize investments in social and physical infrastructure and others (such as the Employment Guarantee Schemes in India) provide opportunities for employment on local infrastructure projects during periods of food scarcity. Central to evaluating the success of these programs is an assessment of how well they target the poor.

While many of the programs mentioned above have transformed the lives of millions of households, there is some concern that they may not be adequately serving the very poor. These households may be inadequately informed, educated or nourished to take advantage of these programs, they may not possess required documents such as birth certificates or proofs of residence, they may be socially ostracized, or agency problems may lead bureaucrats to direct resources to other groups. Morduch (1998) finds that eligibility rules are often violated in microcredit programs in Bangladesh. There is also empirical evidence from a variety of social programs in both developed and developing countries that information sets differ among those eligible, and that participation rates are sensitive to program design.[1]

Estimates and tests of poverty targeting based on parametric models could be misleading without enough prior information on the set of households that may be vulnerable to exclusion. Social programs to reduce poverty are designed to exclude wealthy households, leading us to expect mean incomes to be lower among participants than non-participants. If these programs also exclude the poorest households, and these are a small fraction of the population, we may not detect such neglect in a

[1] Heckman and Smith (2003) use data from a job training program and shows how information can have significant effects on participation. Atkinson (1995) compares family allowance programs in Western Europe in the post-war period and discusses the role of differences in design.

comparison of means across the two groups. Intutively, this is because popular parametric procedures typically estimate conditional means and differences in the tails of conditional distributions may not be reflected in their means. This is well illustrated in Paxson and Schady (2002) where logit estimates indicate that the benefits from investments in infrastructure are decreasing in income, but nonparametric regressions reveal that the poorest 7% of households are less likely to benefit than the slightly richer ones.

In this paper we describe three nonparametric procedures that could be used to compare differences in the tails of two distributions. We are interested in evaluating the appropriateness of these procedures in testing for whether the poorest households in the population have been neglected by a program.

The simplest procedure we consider is the sign test, based on the number of participants in the sample below a given population quantile. If, for example, the share of participants in the bottom income quartile of the population is significantly lower than 25% we would expect the bottom quarter of our sample (ordered by income) to contain fewer participants than non-participants for equal sample sizes of the two groups. This popular nonparametric test is fairly crude in that it relies only on the number of sample observations below a certain income threshold and is not sensitive to the levels of income corresponding to these observations.

Our second set of procedures are tests for stochastic dominance. If program participants are mainly poor, but the program in inaccessible to those below a certain threshold level of income, x_0, we would expect the distribution of participants and non-participants to cross at some income $x^* > x_0$ with the distribution of participants, $F(x)$, below that of non-participants, $G(x)$, for incomes below x^* with the reverse true above x^*. Suppose we classify households as poor if their incomes lie in some interval $[a, b]$. In this case, there will be no first order stochastic dominance over any such interval if it includes x^*. It may however be the case that $F(x)$ second-order stochastically dominates $G(x)$ over $[a, b]$. This would be evidence of an anti-poor bias in the program. Second-order stochastic dominance is closely related to Lorenz dominance and discussions and tests for both these orderings have now appeared in the literature on inequality and poverty measurement (Foster and Shorrocks, 1988, Bishop *et. al.*, 1992, Anderson, 1996). We discuss this literature more fully in Section 2.

The last procedure we consider explicitly tests for a crossing of the distributions of participants and non-participants and also estimates the crossing point. These estimated crossing points provides us with an upper bound on the income of the set of households that have been neglected by the program. This approach has the advantage of detecting non-monotonicities in inclusion probabilities, even over small ranges of the income distribution under very general distributional assumptions. It also requires no prior information on the income interval in which such a crossing might occur.

We compare the three tests through simulations on pairs of income distributions which we consider plausible for participants and non-participants in a public program. Each pair is chosen so that the distribution of participants crosses that of non-participants from below, reflecting the relative exclusion of households below a certain minimum income threshold. No first-order stochastic dominance relation therefore exists. We have constructed our examples so that second-order stochastic dominance holds in two out of 3 cases.

We experiment with different sample sizes and find the test of distribution crossing to be the most powerful in rejecting a null of equal distributions in most of our simulations. Simulation results are in Section 3. For all three pairs of distributions we consider, the power of this test rapidly converges to 1 as sample sizes increase. The simulations lead us to the conclusion that if there is reason to believe, *a priori*, that there is a unique threshold below which program participation is difficult, the tests for distribution crossing are the most powerful. In contrast, there are two principal drawbacks to using tests for second-order stochastic dominance in this context. The first is that such dominance will only occur if the crossing point of the two distributions x^* is close to the endpoint b of the interval of interest. These test cannot therefore detect the exclusion of poor households if the fraction excluded is relatively small and in the tail of the income distribution. The second problem relates to the test statistics currently available in the literature. These are mainly union-intersection tests and are therefore usually conservative and have low power unless the distributions being considered are sufficiently different from each other.

In Section 4, we apply the above methods to test for poverty targeting in a rapidly growing microfinance program in India. We collected data on living standards for a sample of households entering the program and randomly chosen non-participant households in the same area. We use this survey data to construct an economic index which can proxy for income and compare the distributions of this index for members of newly formed microcredit groups to randomly chosen non-members in the area. The two empirical distribution functions cross. We estimate the crossing point of these distributions and test for distribution crossing. Our results are statistically significant and suggest that the poorest 5% of households in the area are disproportionately outside the program. Based on their levels of education and their pariticipation in a government program for subsidized foodgrains, it seems that they are also excluded from other programs aimed at poverty alleviation.

Although much of our discussion and our empirical work refers to targeting in poverty alleviation programs, the methods proposed are of more general applicability. They can be used in a variety of situations where the crossing of population distributions is of interest. For example, students in some schools may come from the tails of an income distribution (because the school may admit you either if you are very wealthy or poor and intelligent) while others come from the middle. Some firms may hire some very able managers and low skill workers while others might hire employees of similar ability. Estimates and tests for distribution crossings can be useful in these situations to characterize differences organizational behavior evaluate their economic effects.

2 Methodology

2.1 Some Preliminaries

Almost all the nonparametric procedures we consider are based on empirical estimates of the population distribution functions for each group. These are standard in the literature, but for the sake of completeness, we begin with a definition.

Definition 1 *The empirical distribution function corresponding to a population distribution $H(x)$ is*

$$H_N(x) = \frac{1}{N}\sum_{i=1}^{N} I(X_i \leq x),$$

where $I(A)$ is the indicator function of the set A and N is the number of sample observations.

$H_N(x)$ is a step function with jumps at the order statistics of the sample. The Glivenko-Cantelli theorem (Fisz, 1963), establishes that the empirical distribution function converges uniformly to the population distribution function with probability one.

We denote the distribution of income in the population by $F(x)$ for participants and by $G(x)$ for non-participants. Sample sizes for participants and non-participants are denoted by n and m respectively and the two samples are denoted by $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$.

2.2 Test Statistics

2.2.1 General Tests for Comparing Distributions

We begin by briefly discussing two popular nonparametric tests; the Kolmogorov Smirnov test for the equality of two distributions and the Wilcoxon-Mann-Whitney test for first order stochastic dominance. We point out why it is desirable to go beyond these for the particular problem we are interested in and discuss tests that we believe are appropriate in our context.

The Kolmogorov Smirnov test is used to test the null hypothesis of equal distributions against the very general alternative that the distributions are unequal. The test is based on the maximum difference between the two empirical distribution functions. Large values of the statistic are evidence against equal distributions and lead to the rejection of null hypothesis. The generality of the alternative hypothesis makes it widely applicable. It also means however that rejection of the null is not very informative in terms of ordering the two distributions. Kolmogrov tests are often used for preliminary studies of data since the alternatives involved are very general. As a consequence, the rejection of the null provides us with very little information which can be used to compare the two distributions.

The Wilcoxon-Mann-Whitney test can be used to examine whether one distribution first order stochastically dominates the other. If we find evidence of stochastic dominance of the distribution of non-participants, the program can be judged to be successful in targeting the poor. The test statistic is based on the number of times an X precedes a Y in the combined ordered arrangement of the two independent random samples $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ into a single sequence of $m + n = N$ variables, increasing in magnitude. If a program successfully targets poor households, we would expect the sample of X's to generally precede the Y's and the test would provide support for the stochastic dominance of the distribution of non-participants. The reverse would be true if the program was anti-poor. The test statistic is asymptotically normal and is distribution-free under the null hypothesis. Its power compares very well with parametric tests when the latter are appropriate.[2]

If there is in fact a threshold level of household income below which participation in a program is negligible, the population distributions of participants $F(x)$ will cross that of non participants from below. There will be no first order stochastic dominance of either distribution in this case. These tests will not, in that case provide us with any information about the alternative we are interested in.

We now discuss tests that are appropriate when population distributions cross.

2.2.2 Sign Tests

A sign test is a popular procedure to test for population quantiles. The relevant population quantile is the level of household income below which we are interested in comparing the shares of participants and non-participants. In terms of the notation we have been using so far, this population quantile would correspond to the endpoint b of the income interval of interest $[a, b]$, with $a = 0$.

[2]See Gibbons and Chakraborti (1992), chapter 7, for derivations of these test statistics and their distributions and Hettmansperger (1984) for power comparisons.

To elaborate, suppose we wanted to test the null hypothesis that the level of income was irrelevant for program participation of households in the lowest income quartile. We would first find the value of household income in our sample below which 25% of sample households lie. This sample quantile, given by the income of the $\frac{m+n}{4}^{th}$ ordered observation, is a consistent estimate of the corresponding population quantile. This can be used to test the null hypothesis against the alternative that the fraction is less than .25. Rejection of the null would support the claim that households with incomes in the bottom quartile are disproportionately excluded from the program.

The test statistic is based on the number of observations for participants, S_n that are below the relevant sample quantile. S_n has a binomial distribution with parameters n (the size of the sample of participants) and p (the population fraction below the income quantile being used). As long as sample sizes are reasonable, critical values for the test can be based on the normal approximation. In the simulations exercises described below, we use $p = .25$ and $p = .05$ as alternative cutoffs for the sign test.

2.2.3 Tests for second-order Stochastic Dominance

If the distributions of participants and non-participants cross, we cannot rank them in terms of first-order stochastic dominance. We may, however, be able to rank them in terms of higher orders of stochastic dominance.

Definition 2 *We say that the distribution of participants, F(x), dominates that of non-participants, G(x), in second-order iff*

$$\int_{-\infty}^{t} G(x)dx \geq \int_{-\infty}^{t} F(x)dx$$

Higher orders of stochastic dominance are defined by the ordering of higher order integrals of these distributions. Note that stochastic dominance of order r always implies stochastic dominance of order $(r+1)$ but the reverse is not true.

If two distributions cross at an income level close to the end point b of the interval of interest, $F(x)$ is likely to dominate $G(x)$ in second-order. A test which supports such an alternative, against the null of equal distributions, therefore provides evidence supporting the exclusion of households in the interval $[a, b]$. If however, the distributions cross at an income level close to a, such a test would be relatively uninformative. Loosely speaking, the closer the income corresponding to the crossing point is to a, the higher the order of stochastic dominance we need to consider to rank these distributions.

Several test procedures are available for testing second-order stochastic dominance and the related concept of Lorenz dominance. Bishop, Fromby and Thistle (1992) and Beach and Davidson (1983) propose tests for Lorenz dominance based on sample quantiles. Deshpande and Singh (1985) provide a one sample statistical test for second-order stochastic dominance when one distribution is known. Schmid and Trede (1998) consider the dominance of F over G when the supports of F and G are contained in $[0, 1]$ and G is uniform. Since very little in known in practice about the distributions of participants and non-participants in a program, we restrict ourselves to tests which can be used for any (unknown) specification of the two distributions. Moreover, since we are interested in poverty rather than inequality we focus on second-order stochastic dominance rather than Lorenz dominance. Anderson (1996) proposes a test for second-order stochastic dominance based on a comparison of frequencies for the groups in pre-determined income intervals and which does not use any information on the actual levels of income within these intervals. Davidson and Duclos (2000) and Kaur, Prakasa Rao and Singh (1994) propose the same test statistic for testing second-order stochastic dominance at a fixed point in the interval of interest. Kaur, Prakasa Rao and Singh then use the infimum of these statistics to test for second-order stochastic dominance. We favor their approach over the others in the literature since their test is consistent and they have carefully identified all points of change in the statistic over the interval of interest.

A thorough discussion of the usefulness of how higher orders of stochastic dominance can be used for poverty measurement is beyond the scope of this paper. We refer the reader to Chakravarty and Muliere (2003) and Davidson and Duclos (2000) for a review of some of the important literature in this area. We restrict ourselves here to two comments that are relevant to the question we are interested in, namely, testing the exclusion of the very poor from a public program. The first is that most available tests for second-order stochastic dominance that are distribution free are based on the union-intersection principle. This means that the hypothesis being tested is based on several simple hypotheses and the test statistic is a function of the statistics from these simpler tests. In the case of Kaur, Prakasa Rao and Singh (1994) for example, this means that the integrals of the empirical distribution functions for the two groups are compared at several points in the income interval $[a, b]$ which is of interest and the test statistic is based on the smallest difference between the integral estimates in this interval.

Most tests for second-order stochastic dominance tend to have low power and often yield conflicting results when applied to the same data (Davidson and Duclos (2000)). This is not surprising since these test statistics are sensitive to each of the points in the interval in which the comparison is made and are therefore very sensitive to outliers in the sample. Different statistics are sensitive to different types of outliers and therefore yield conflicting results unless the sample error is small relative to differences in the population distributions. The second observation is that when the crossing of participant and non-participant distributions is close to the bottom of the income interval of interest, we require tests for higher orders of stochastic dominance and very little is known about the properties of these types of tests.

In the simulations below, we use the test for second-order stochastic dominance given by Kaur, Prakasa Rao and Singh (1994) because it is not based on arbitrarily determined quantiles and the behavior of the statistic within the interval of interest is well studied. The calculation of the statistic used in our simulations is reproduced in the appendix for easy reference.

2.2.4 Tests for Distribution Crossing

We now turn to explicit tests for distribution crossing. These tests are appropriate if there is reason to believe that there is a single crossing in our interval of interest. This would be true if income is the main determinant of participation and the program finds it difficult to reach households below a certain income threshold. They have the added advantage of providing an estimate of the crossing point, which is an upper bound on the incomes of the set of households who are relatively neglected by the program.

Hawkins and Kochar (1991) and Chen et al. (2002) have considered point as well as interval estimation of the crossing point x^*. We derive estimates of the crossing point based on the methodology in Chen et al. (2002) which we summarize here.

Suppose that $\lim_{N\to\infty} m/N = \gamma$ for some $\gamma \in (0,1)$. Let $Z_1, \ldots, Z_N$ be the combined sample of $X's$ and $Y's$ and $Z_{(1)} < Z_{(2)} < \ldots, Z_{(N)}$ be the order statistics of this sample. We wish to test

$$H_0 : F(x) = G(x)$$

against the alternative

$$H_A : F(x) < G(x) \text{ when } x < x^* \text{ and } G(x) < F(x) \text{ when } x > x^*.$$

Chen et al. (2002) proposed the following supremum-type criterion function for testing H_0 against H_A:

$$\lambda(x) = \sup_{t\le x}(G(t) - F(t)) + \sup_{x\le t}(F(t) - G(t)) - |F(x) - G(x)|.$$

They prove that under the null hypothesis $\lambda(x) = 0$ and under the alternative hypothesis the crossing point x^* is the unique maximizer of $\lambda(x)$. An estimate of $\lambda(x)$ is given by corresponding values based on the empirical distribution functions $G_m(x)$ and $F_n(x)$:

$$\lambda_N(x) = \sup_{t\le x}(G_m(t) - F_n(t)) + \sup_{x\le t}(F_n(t) - G_m(t)) - |F_n(x) - G_m(x)|.$$

Since empirical distribution functions are step functions with jump points as order statistics, $\lambda_N(x)$ attains its maximum at some point $Z_{(j)}$ Therefore

$$\sup_x \lambda_N(x) = \max_{0\le j\le N} \lambda_N(Z_{(j)})$$

They propose the statistic

$$J_N = \sqrt{\frac{mn}{N}} \max_{0 \le j \le N} \lambda_N(Z_{(j)})$$

for testing H_0 against H_A. Exact critical points for small samples are tabulated in Chen et al. (1998). The asymptotic distribution is not standard and the authors obtain asymptotic critical regions using Monte-Carlo simulations. The relevant critical value is presented with our simulation results below.

Since the empirical distribution functions are only estimates of the population distribution functions, we may encounter multiple crossings in our sample even if the population distributions exhibit a unique crossing point. In the case of multiple estimates of the crossing point, we use the smallest value since this is our most conservative estimate of the upper bound on income of the households neglected by the program.

3 Simulation Results

We now compare the test procedures described above using simulated data from three alternative pairs of distributions for participants and non-participants. In each case, the distributions cross with the participant distribution intersecting that of non-participants from below. We rely mainly on alternative parameterizations of Weibull distributions since these are reasonable approximations for observed income distributions and are easy to work with.

We first consider an example that has been used in both Hawkins and Kochar (1991) and Chen et al (2002). We include it here for purposes of comparison. The income of participant households, $F(x)$ has an exponential distribution and that of non-participants a Weibull distribution:

$$\begin{aligned} F_1(x) &= 1 - e^{-x^2}, \\ G_1(x) &= 1 - e^{-\frac{2x}{\sqrt{\pi}}}. \end{aligned} \tag{1}$$

The two distribution cross at $x = 1.13$ and 72% of the population of each group lies to the left of this point. F second-order stochastically dominates G for all possible intervals in $[0, 1]$.

Our second example has been chosen so that almost the entire population is to the left of the crossing point. Second-order stochastic dominance of F is therefore strongest here. The distributions

$$\begin{aligned} F_2(x) &= 1 - e^{-(\frac{x}{3})^3} \\ G_2(x) &= 1 - e^{-(\frac{x}{2})^2}, \end{aligned} \tag{2}$$

cross at $x = 6.75$.

In our last example,

$$\begin{aligned} F_3(x) &= 1 - e^{-(\frac{x}{2})^3} \\ G_3(x) &= 1 - e^{-(\frac{x}{3})^2}, \end{aligned} \tag{3}$$

the distributions cross at $x = 8/9$. Only 9% of each group has incomes below this point. As expected, there is no second-order stochastic dominance of either distribution.

The distribution functions corresponding to these examples are shown in Figures 1-3.

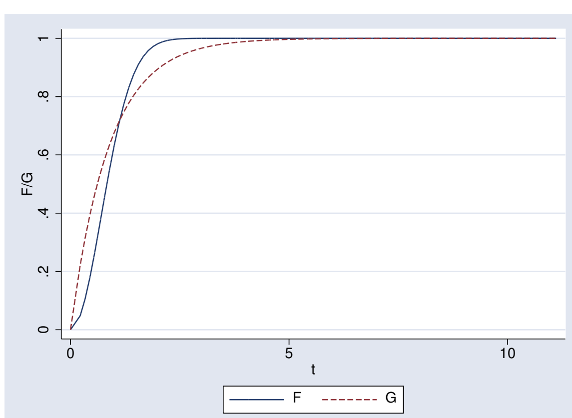

Figure 1: $F_1(x) = 1 - e^{-x^2}, G_1(x) = 1 - e^{-\frac{2x}{\sqrt{\pi}}}$

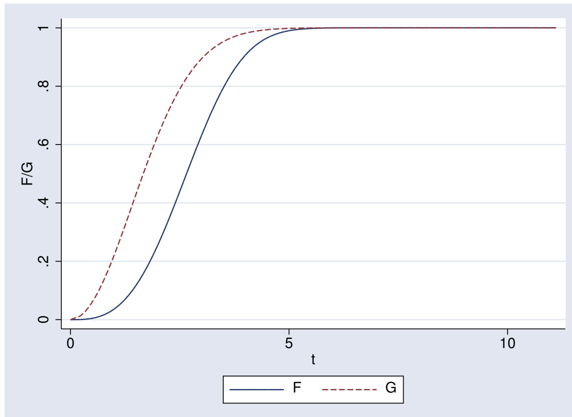

Figure 2 : $F_2(x) = 1 - e^{-(\frac{x}{3})^3}, G_2(x) = 1 - e^{-(\frac{x}{2})^2}$

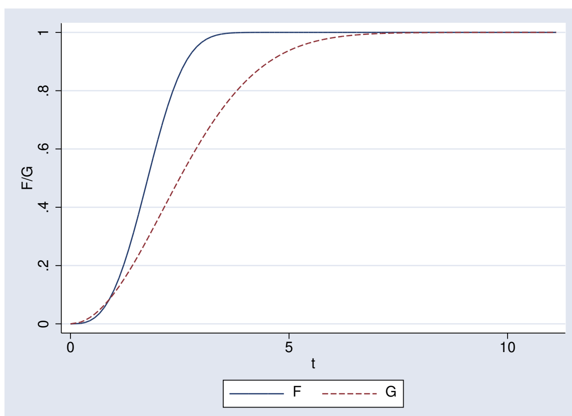

Figure 3: $F_3(x) = 1 - e^{-(\frac{x}{2})^3}, G_3(x) = 1 - e^{-(\frac{x}{3})^2}$

For each of these pairs, we consider 2 sample sizes (i) $n = m = 20$ and (ii) $n = m = 50$ and perform 5,000 iterations. We have used the same sample sizes and number of iterations used by Chen et al. to facilitate comparison between the tests. For the crossing test we use the simulated 5% critical point of 1.529 given in Chen et al.(2002). For the sign test for the first two examples, we count the proportion of observations to the left of the $N/4^{th}$ observation in the combined sample. This is therefore a test for whether the proportion of participants in the bottom income quartile of the population is less than 25%. For the third example, there is an early

Table 1: Power comparisons for three nonparametric tests.

Distributions	Sample sizes	Crossing Test	SSD Test	Sign Test ($p = .25$)	Sign Test ($p = .05$)
F_1, G_1	n=m=20	0.485	0.187	0.185	
	n=m=50	0.899	0.25	0.664	.
F_2, G_2	n=m=20	0.577	0.759	0.312	
	n=m=50	0.960	0.878	0.927	
F_3, G_3	n=m=20	0.623	-		0.003
	n=m=50	0.974	-		0

crossing of the two distributions and we therefore test for proportions in the poorest 5% of the population by counting the proportion of observations to the left of the $N/20^{th}$ observation in the combined sample. We use critical values based on the standard normal distribution.

Computing the power for tests of second-order schochastic dominance is more complicated because even for large samples, we find simulated critical values to be considerably different from those of the standard normal distribution, which has been proposed by the authors for testing purposes. The test statistic proposed by Kaur, Prakasa Rao and Singh (1994) is an infimum of random variables each with a standard normal distribution and they propose using the critical values from the standard normal to reject the null hypothesis. We find the test based on these critical values overly conservative therefore generate simulated critical values under the null hypothesis using the exponential distribution in Example 1 for both groups and 5,000 iterations. We obtained a simulated 5% critical value of .72 with a sample size of 20 for both groups and .96 with a sample size of 50. The power comparisons in Table 1 are based on these critical values.

As seen from Table 1, the test for distribution crossing is more powerful than the others in most cases and its power rapidly improves with sample size. We also performed simulations with bigger samples and found that for m=n=100, the power of this to be 1 in all three examples. We found no similar convergence for the other two tests. Both the sign test and the SSD test do best in the second example when the difference between the two distributions is marked. We do not test for second-order stochastic dominance in the third example since no such dominance holds.

The Sign test does very badly in case 3 because the two distributions are almost identical initially in this case. This is clear from Figure 3. We conclude that there is a strong case for using the Chen test when we expect a single crossing of the two distributions and when the distributions are fairly similar.

4 An Application to Data from an Indian Microfinance Program.

4.1 Data

Our data is from a microfinance program in the state of Jharkhand in India. Jharkhand is among the poorest of the 27 Indian states, with over half its population below the national poverty line. The program is administered by PRADAN, a non-government organization working in the area to form "self-help groups" of women. These groups of between 10 and 20 women facilitate risk-sharing among their members by saving a pre-determined minimum amount each week and lending accumulated savings to members at terms determined by the group. Most groups eventually establish a savings account at a commercial bank and take loans from the bank for self-employment activities, mainly related to agriculture and livestock. There are currently about 2,000 groups in operation in the Jharkhand.

We surveyed households in villages with newly formed groups and use the above techniques to examine whether the program successfully targeted poor households in the area. The sample consists of 576 households in 24 villages. The survey was conducted over a period of two months starting in August 2002. The villages were chosen from a set of 100 villages in which at least one group was formed during the period April 1st to June 30th, 2002. Very little lending takes place in the months immediately following group formation and a comparison of the characteristics of households in the program with those of other randomly chosen households in the area can therefore be used to evaluate the extent to which the program targeted the poor.

The 100 villages with new microcredit groups were partitioned into 4 geographical clusters and simple random sample of 6 villages was drawn from each cluster. A total of 24 respondents were surveyed from each of these villages- 6 of them members of microcredit groups in the village and the remaining 18 randomly selected non-members from the same village. The ratio of 1:3 for members and non-members were chosen based on the prior belief that the group of non-participants is more heterogeneous and unequal sample sizes would be required to obtain estimates of similar accuracy for both groups.

Survey data were collected on a large number of economic indicators such as the quantity and type of food consumed, the size and condition of the household's main dwelling, land owned and cultivated and the possession of durable goods. In addition, respondents were asked about their contact with the government bureaucracy and about any benefits received from government sponsored programs. Responses to these questions allowed for an assessment of whether the households excluded from the program were also excluded from other official poverty-alleviation programs.

We used a combination of household characteristics for a principal component analysis. The first principal component was used as a proxy for income. Table 2 contains group-wise summary statistics on household size, literacy rates and the variables used to create the income index. Details on survey design and scoring coefficients for the index can be found in Somanathan (2003).

4.2 Results

As can be seen from Table 2, mean values of the index and its constituent components are remarkably similar for the two groups. None of the group-wise differences seen in the table are statistically significant at conventional levels. The empirical distribu-

Table 2: Group-wise Summary Statistics

	Members	Non-members
Household size	6.22	6.15
Literacy rate	.39	.41
Meals consumed in the two days prior to the survey	5.68	5.31
Household foodgrain consumption in normal times (kilograms)	3.82	. 3.97
Rooms in dwelling	3.06	3.35
land owned (hectares)	1.24	1.1
Value of livestock and durables (Indian rupees)	8642	9066
Annual household expenditure on clothing and footwear (Indian rupees)	3275	3428
Economic index	-.021	.01

tions of the income index for participants and non-participants are shown in Figure 4. While these also look similar, it does appear that the very poorest households are predominantly outside the program; 2% of sampled group members and 5% of non-members are below the 5th percentile of the index. We now apply some of the methods discussed in Section 2 to see if these differences are statistically significant.

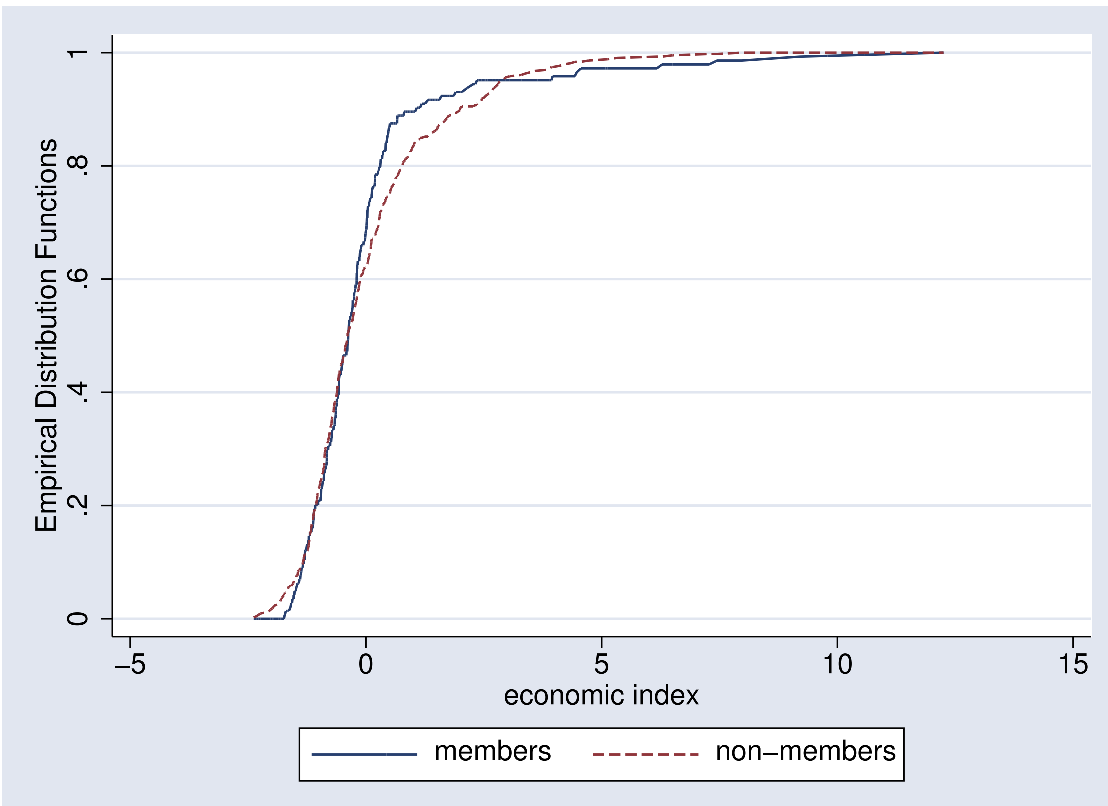

Figure 4: Empirical distribution functions for members and non-members of microcredit groups in Jharkhand, India.

Table 3 contains computed test statistics and p-values for the Kolmogorov Smirnov and Wilcoxon Mann Whitney tests for equality of these distributions. We use a two-sided alternative in both these cases and find that neither test rejects the null hypothesis at the 5% level of significance, although the Kolmogorov Smirnov test does reject equality at the 10% level.

Table 3: Preliminary Nonparametric Tests for Equal Distributions

	Kolmogorov Smirnov	Wilcoxon Mann-Whitney
Value of the Statistic	.125	.684
p-value	.068	.49
Reject H_0 at 5%?	No	No

Table 4 reports the critical values and test statistics from the sign test and the test for distribution crossing. For the crossing point test we use the simulated 5

The distribution crossing test rejects the null of equal distributions against the alternative of their crossing. Using the methodology outlined in Section 2, we estimate of the crossing point x^*. The value of the income index at the crossing point is -1.19 , which is in the 15th percentile of the distribution of the index for the whole sample. The sign test does not reject the null hypothesis of equal shares of members and non-members for the poorest 5% if the population at the 5% level of significance although it does reject it at the 10% level.

4.3 Characteristics of Excluded Households

How poor are the households neglected by the program? Converting our income index back into household characteristics allows us to compare lifestyles of these households with others in the area. Table 5 presents means of selected variables for households

Table 4: Tests for the Exclusion of Poor Households

	Crossing Test	Sign Test
Value of the Statistic	1.756	-1.61
Critical Value	1.529	-1.645
Reject H_0 at 5%?	Yes	No

below the 5th percentile of the income index and other households. The poorest households have dramatically different lifestyles and consumption levels from the rest of the sampled households. They live in smaller dwellings and eat fewer meals. Their expenditure on clothing and footwear is about one-half of the mean for other households, their food grain consumption (by weight) is lower by about one-third. They spend less time in the village, although the differences here are small. The households that are difficult to involve in the microcredit program also seem to be excluded from other public programs and the political process more generally. Only 7% of these households had ever approached a government official compared to 28% of other households. Perhaps the most striking observation is that the fraction of households receiving subsidized foodgrains from a government anti-poverty program was smaller for the poorest 5% than for the rest. The exclusion of these households seems to extend far beyond microfinance programs!

Table 5: Selected Household Characteristics: They Very Poor and the Rest

	Poorest 5%	**Other Households**
Meals consumed during the two prior to the survey	3.79	5.49
Number of rooms in dwelling	1.21	3.29
Foodgrain consumption per day in normal times (kilograms)	.5	.67
Annual expenditure on clothing and footwear (rupees)	243	562
Land owned (hectares)	.29	1.17
Average months spent in the household over the past year	10.8	11.09
Fraction ever approached government official	.07	.28
Fraction received goverment subsidized foodgrains	.46	.52

5 Conclusions

This paper evaluates available nonparametric methods to test for whether the poorest households in a population have been excluded from a program. Parametric tests, which usually rely on some function of differnces in the conditional means of two groups, may not successfully detect these households if there is little prior information on the income level of the households vulnerable to exclusion. We discuss three alternative nonparametric methods that may be appropriate in this context and use simulations with three different pairs of income distributions to compare their power. If the distributions of participants and non-participants exhibit a single crossing, we find that a test of distribution crossing, which explicitly uses this information, is usually more powerful than competing procedures. This is especially true when the distributions of the two groups are similar.

We apply some of the methods to a non-government microcredit program in India. We find evidence that the population distributions of program participants and non-participants cross, with the poorest households largely outside the program. These households also appear to have limited access to public programs which are, in principle, designed for their benefit: they are no more likely to be on official poverty lists and government-sponsored social programs.

6 Appendix

Test statistic for second-order stochastic dominance

Let X_1, X_2, X_n be a random sample from F and Y_1, Y_2, Y_m be a random sample from G. For a fixed $x \in [a, b]$ and for $i = 1, 2, \ldots, n$,

$$U_i(x) = \begin{cases} x - X_i & \text{if } X_i < x, \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

and for $j = 1, 2, \ldots, m$,

$$V_j(x) = \begin{cases} x - Y_j & \text{if } Y_j < x, \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Let

$$\bar{U}(x) = \frac{1}{n}\sum_{i=1}^{n} U_i(x) = \int_{-\infty}^{x} F_n(y)dy,$$

and

$$\bar{V}(x) = \frac{1}{m}\sum_{i=1}^{m} V_i(x) = \int_{-\infty}^{x} G_m(y)dy.$$

Let

$$Z_{n,m}(x) = \frac{\bar{U}(x) - \bar{V}(x)}{\sqrt{\frac{1}{n}s^2_{n,U}(x) + \frac{1}{m}s^2_{m,V}(x)}}$$

where

$$s^2_{n,U}(x) = \frac{1}{n}\sum_{i=1}^{n}[U_i(x) - \bar{U}(x)]^2$$

and

$$s^2_{m,V}(x) = \frac{1}{m}\sum_{j=1}^{m}[V_i(x) - \bar{V}(x)]^2$$

For large sample sizes their test rejects H_0 if and only if

$$Z_{n,m} = \inf_{a \le x \le b} Z_{n,m}(x) > z_\alpha,$$

where z_α is the upper α point of the standard normal distribution.

References

[1] Anderson, Gordon, “Nonparametric Tests of Stochastic Dominance in Income Distributions ”, *Econometrica*, 64(5), 1996, 1183-1193.

[2] Atkinson, Anthony B. “On Targeting Social Security: Theory and Western Experience with Family Benefits ” in van-de-Walle, Dominique and Kimberly Nead, ed. *Public Spending and the Poor: Theory and Evidence*, John Hopkins Press for the World Bank, 1995.

[3] Beach, Charles M. and Russell Davidson, “Distribution-Free Statistical Inference with Lorenz Curves and Income Shares. ”, *The Review of Economic Studies*, 50(4), 1983, 723-735.

[4] Bishop, John A., John P. Fromby and Paul D. Thistle, “Convergence of the South and Non-South Income Distributions, 1969-1979. ”, *American Economic Review*, 82(1), 1992, 262-272.

[5] Chakravarty, S.R. and Muliere, P., “Welfare indicators: A review and new perspectives ”, Metron , 61, 2003, 1-41.

[6] Chen, G.J., J.H. Chen and Y.M. Chen, *Statistical inference on comparing two distribution functions with a possible crossing point*, Technical Report, Department of Mathematics, Anhui University, 1998.

[7] Chen, Guijing, Jiahua Chen and Yuming Chen, “Statistical inference on comparing two distribution functions with a possible crossing point”, Statistics and Probability Letters, 60, 2002, 329-341.

[8] Davidson, Russell and Jean-Yves Duclos, “Statistical Inference for Stochastic Domainance and for the Measurement of Poverty and Inequality”, *Econometrica*, 68(6), 2000, 1435-1464.

[9] Deshpande, J.V. and Singh, H. “Testing for second-order stochastic dominance ”, Communications in Statist.,Theory and Methods, 14, 1985, 887-893.

[10] Fisz, Marek, *Probability Theory and Mathematical Statistics*, New York, Wiley, 1963.

[11] Foster, James E. and Anthony Shorrocks, "Poverty Orderings ", *Econometrica*, 56(1), 1988, 173-177.

[12] Gibbons, Jean Dickinson and Subhabrata Chakraborti, *Nonparametric Statistical Inference*, M. Dekker, New York, 1992.

[13] Hawkins, D.L and Subhash Kochar, "Inference for the Crossing Point of Two Continuous CDF's", *The Annals of Statistics*, 19(3), 1991, 1626-1638.

[14] Heckman, James J. and Jeffery A. Smith, "The Determinants of Participation in a Social Program: Evidence from a Prototypical Job Training Program", *NBER Working Paper* number 9818, 2003.

[15] Hettmansperger, T.P. *Statistical Inference Based on Ranks*, John Wiley, New York, 1984.

[16] Kaur, A. , Prakasa Rao, B.L.S. and Singh, H. , "Testing for second-order stochastic dominance of two distributions ", *Econometric Theory*, 10, 1994,849-866.

[17] Morduch, Jonathan, *Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh*, working paper available at http://www.wws.princeton.edu/%7Erpds/Downloads/morduch_microfinance_poor.pdf, 1998.

[18] Paxson, Christina and Norbert R. Schady, "The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru", *World Bank Economic Review,* 16(2), 2002, 297-319.

[19] Somanathan, Rohini, "Poverty Targeting in PRADAN's Microfinance Program: A Study Conducted on Behalf of CGAP ", available from the Consultative Group to assist the Poorest at the World Bank, Washington D.C., April, 2003.

[20] Schmid, F. and Trede, M. "A Kolmogorov-type test for second-order stochastic dominance", Statistics and Probability Letters, 37, 1998, 183-193.