Optimal Targeting and Poverty Reduction with fixed budget and under imperfect information

Abdelkrim Araar[1] and Luca Tiberti[2]
[1]Laval University & Partnership for Economic Policy (PEP), aabd@ecn.ulaval.ca
[2] Laval University & Partnership for Economic Policy (PEP), luca.tiberti@ecn.ulaval.ca

Abstract – Targeted anti-poverty programs are preferred to universal interventions because they would allow to spend more resources on the poor. However, they are often criticized because of their inefficiency in including the real poor and excluding the non-poor. In this paper we develop a new numerical algorithm for optimal poverty group targeting when the total budget is fixed and information on individual welfare is unknown, as is the case for most developing countries. Contrarily to other similar algorithms found in the literature, what we propose is applicable to all additive poverty indices. Tested on household data from Burkina Faso, the new approach gives a better targeting performance than the popular proxy-means test (PMT) approach irrespective of the poverty indicator is used, and replicates quite closely the results obtained through a grid approach (our benchmark). Also, finer the partitions of the group per capita transfer, more efficient the targeting algorithm. Finally, the proposed algorithm converges to the global optimum also in the case of headcount poverty index, where there are multiple local optima.

1. INTRODUCTION

Anti-poverty interventions can be universally distributed across the population or targeted to specific groups of individuals. While universal programs reach all the poor but are often unfeasible in terms of budgetary grounds, targeted interventions may take longer and be more expensive to implement, and administratively difficult to manage. In developing countries, targeting is generally based on proxy-means testing, geographical (at regional, provincial, or a lower level such as communes or villages) or demographic (e.g., young or elderly) categorical eligibility, self-selection or a mix of these approaches (Coady, Grosh and Hoddinott, 2004). All targeted methods generate the so-called "two targeting errors" (i.e., inclusion of non-poor and exclusion of poor) (Cornia and Stewart, 1995). Also sophisticated econometric-based proxy-means tests (PMT) are usually prone to exclude many poor individuals (Brown et al., 2018). Targeted schemes may become even more inefficient when combined with high taxes burdening on the poor (Higgins and Lustig, 2016). Over the last 20 years, social protection schemes have progressed substantially, and one or more programs are now available in most of developing countries (Margitic and Ravallion, 2019). Nonetheless, the literature unanimously agrees that no method clearly dominates (e.g., Coady, Grosh and Hoddinott, 2004; Devereux et al., 2017; Hanna and Olken, 2018), and its effectiveness in terms of poverty reduction may depend on various issues such as the depth of poverty, the inequality level within the poor, the poverty indicator which is used, the level of the poverty threshold, the local context, and so on. For example, higher the percentiles to which the poverty line corresponds, lower the need of targeting as the leakage of non-poor decreases.

Although any improved targeting approaches should aim at reducing one or both of the errors mentioned above, the mechanisms proposed so far in the literature have not been set in a way that, subject to a fixed total budget, poverty reduction per dollar spent is maximized, irrespective of the adopted poverty indicator. In order to fill this gap, in this paper we develop a new poverty group targeting algorithm when budget is fixed and information is on individual welfare is partially or fully unknown to the policy maker, as is especially the case in developing countries. Such information is generally available only for sampled households. At the population level, the policy-maker disposes of information on groups of individuals, such as regions or age groups. Under such circumstances, the goal of the policy maker is to find the optimal group transfers that reduces the most the aggregate poverty. Tested on household data from Burkina Faso, and estimated with a newly developed Stata package, we found that our proposed algorithm is more efficient than a PMT-based targeting approach, irrespective of the poverty indicator.

As usual in optimization exercises, the objective function must be strictly quasi-convex in its arguments of interest, which means that the reduction in poverty is larger for increases in incomes of poorer individuals (as postulated by the transfer axiom). Unfortunately, not all popular poverty indices, such as the headcount or poverty gap rates, obey this condition. While Kanbur (1987) focused on the theoretical rules of optimization, Ravallion & Chao (1989) and Elbers, Fujii, Lanjouw, Özler, & Yin (2007) have also proposed numerical algorithms that maximize the reduction in the squared poverty gap index by group transfers, subject to a fixed budget of transfers. Based on the theoretical findings proposed in Kanbur (1987), these numerical approaches show that, to minimize the squared poverty gap, lump-sum transfers should target the population group with the highest poverty gap until when this group reaches the next poorest one, and so on, until the available budget is depleted. Glewwe (1992) proposed a generalization of the Ravallion and Chao's approach to identify poor people through continuous variables (rather than group/binary variables).[1] Unfortunately, these theoretical and empirical works only offer an optimal targeting solution for the case of the squared poverty gap index, which satisfies the set of optimization conditions.

To overcome these limitations in the existing literature, this paper develops a new numerical algorithm for antipoverty targeting. Our approach helps to find the optimal group transfers that allow the largest possible reduction in any additive poverty indices, like the FGT class of poverty indices. Finer group definitions (e.g., smaller groups based on the combination of identifiable socio-economic characteristics or lower geographic units such as those defined through poverty mapping methods) would allow larger poverty reduction. The rest of this paper is organized as follows. In section 2, the group poverty reduction curves are introduced; in section 3, we introduce the new algorithm and present the cases of single and multiple group targeting, we discuss the relation between the number of partitions and the optimization results, and we conclude by validating empirically the algorithm. In section 4, we briefly present various econometric-based targeting approaches (or, proxy-means test), and section 5 compares the poverty targeting performance obtained with the proposed algorithm with respect to the results estimated with the PMT-based approach. Section 6 draws some concluding remarks.

2. GROUP POVERTY REDUCTION CURVES

As indicated earlier, previously developed algorithms for optimal targeting of population groups with a fixed budget have focused on a subclass of poverty indices for which the analytical solution is feasible, such as the squared poverty gap index (Ravallion & Chao, 1989; Elbers, Fujii, Lanjouw, Özler, & Yin, 2007). When an analytical approach is used as in the previous works, the objective function to be minimized must be strictly quasi-convex. Our work aims to extend the existing literature by suggesting a numerical approach that combines the numerical method of optimization of the objective function and some basic theoretical rules which can help to simplify the computation. More importantly, the algorithm we propose here is assumed to be valid for all classes of additive poverty indices, including the headcount and the poverty gap.

The poverty reduction component

Assume that a lump-sum transfer is attributed only to the population group g. Per capita lump-sum transfer is denoted by τ_g and, because of the imperfect information on the individual welfare, is the same for all individuals belonging to g. The change in the contribution to total poverty by individual i living in the targeted group g and having income $y_{g,i}$ is denoted by $d\pi_{g,i}$. When the FGT poverty class is used, for $\alpha = 0$ (which estimates the headcount poverty) we have that:

[1] In addition to the squared poverty gap, any other poverty measure that is concave with respect to the poverty gap (like the Watts index) can be maximized with the analytical approach.

$$d\pi_{g,i}(\tau_g;\alpha=0) = -\frac{1}{n}\mathrm{I}\left[y_{g,i}<z\right]\,\mathrm{I}\left[\left(y_{g,i}+\frac{\tau_g}{\varphi_g}\right)\geq z\right] \qquad (1)$$

where n is the total population (thus $1/n$ represents the population weight represented by individual i), φ_g is the population share of the targeted group g and the indicator $\mathrm{I}[.]$ is equal to one if the condition is satisfied and zero otherwise. Note that, since we target just one group, the per-capita transfer becomes τ_g/φ_g.[2] As from (1), in order to have a reduction in the headcount poverty, two conditions are necessary: (i) individual i must be initially poor (i.e., her initial income $y_{g,i}$ is below the poverty line z); (ii) the transferred amount should be large enough to bring her income at least up to the poverty line. Similarly, for the poverty gap ($\alpha=1$), one can write:

$$d\pi_{g,i}(\tau_g;\alpha=1) = -\frac{1}{n}\mathrm{I}\left[y_{g,i}<z\right]\min\left(\frac{\tau_g}{\varphi_g},\left(z-y_{g,i}\right)\right) \qquad (2)$$

For the squared poverty gap, we have that:

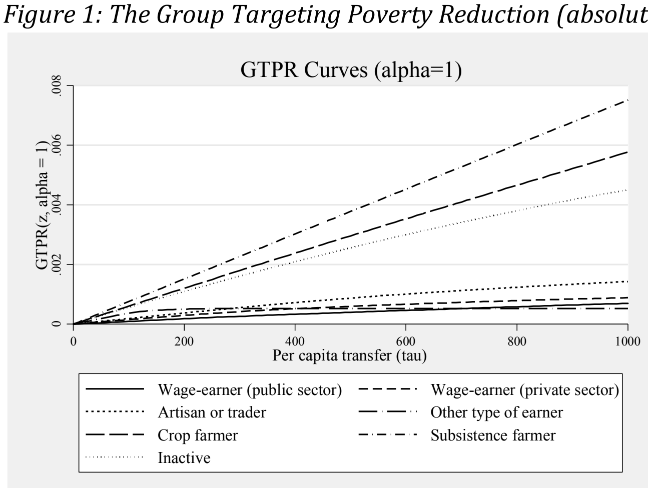
$$d\pi_{g,i}(\tau_g;\alpha=2) = -\frac{1}{n}\mathrm{I}\left[y_{g,i}<z\right]\left(\left(\left(z-y_{g,i}\right)\right)^2-\left(z-min\left(y_{g,i}+\frac{\tau_g}{\varphi_g},z\right)\right)^2\right) \qquad (3)$$

Hence, for the case of additive poverty indices, it is easy to define the reduction in the population poverty when targeting group g as follows[3]:

$$PR_g(\tau_g;\alpha) = \sum_{i=1}^{n_g} d\pi_{g,i}(\tau_g;\alpha) \qquad (4)$$

The function $PR_g(\tau_g;\alpha)$ is called the *Group Targeting Poverty Reduction* (GTPR).[4]

Figure 1: The Group Targeting Poverty Reduction (absolute)



Source: authors' elaboration based on the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso.

[2] Let assume that n=10, that we have two groups whose φ is 0.4 and 0.6 respectively, and that the total budget available for transfers is 1,000. τ_g is then 100. If we target only the first group, then the per-capita transfer rises to 250 (=100/0.4) or to 166.66 (=100/0.6) if only group 2 is targeted.

[3] For simplicity, we omit the common denominator $1/z^\alpha$ for the class of the FGT poverty indices.

[4] The Stata module cgtpr.ado – downloadable from http://dasp.ecn.ulaval.ca/stata_adds/ctgpr.html - can be used to draw the GTPR curves.

Figure 1 illustrates the (absolute) reduction in the poverty gap with 1998 Burkina Faso household data for various per capita transfers ranging from 0 to 1000. In this example, the largest reduction in poverty is obtained by targeting subsistence farmers. The bottom curves show the case of relatively less poor groups for which the transfer, above a certain amount, does not affect poverty (as for "other type of earners") or for which the marginal poverty reduction flattens.

3. OPTIMAL TRANSFERS

3.1 CASE 1: SINGLE GROUP TARGETING

Proposal 1:

For a given per capita transfer τ_g and when only one group is targeted, the optimal targeting is that which prioritises the group with the highest GTPR curve at τ_g.

For instance, based on the results of Figure 1, the optimal targeting prioritizes the *subsistence farmer* group.

3.2 CASE 2: MULTIPLE GROUP TARGETING

Now, assume that the policy-maker aims at targeting more than one group by allocating different levels of lump-sum transfers. Hence, the objective is to find the optimal transfers that reduce the most aggregate poverty, under the transfer constraint $\tau = \sum_{g=1}^{G} \tau_g$ as well as the following inequality $0 \leq \tau_g \leq \iota_g \;\; \forall\, g$, where ι_g denotes the maximum poverty gap within group g. Formally, if we denote the reduction in aggregate poverty by $PR(\boldsymbol{\tau};\alpha) = \sum_{g=1}^{G} PR_g(\tau_g;\alpha)$, the optimization problem can be written as follows:

$$\max_{args:\,\tau_g} \{PR(\boldsymbol{\tau};\alpha)\} \quad s.t. \quad \tau = \sum_{g=1}^{G} \tau_g \text{ and } 0 \leq \tau_g \leq \iota_g \quad \forall g \tag{5}$$

where the vector $\boldsymbol{\tau} = \{\tau_1, \tau_2, \cdots, \tau_G\}$. Based on the analytical approach, the FOCs of maximization are:

$$\frac{\partial PR_g(\tau_g;\alpha)}{\partial \tau_g} - \lambda = 0 \quad \forall g. \tag{6}$$

The SOCs of optimization requires that: $\frac{\partial\partial PR_g(\tau_g;\alpha)}{\partial {\tau_g}^2} \leq 0$. However, for the class of the FGT indices, this condition is only satisfied with $\alpha > 1$.

This result corroborates the findings reported in Kanbur (1987) on the minimization of aggregate poverty based on the FGT class index, when (α>1). As he reports, to minimize the FGT poverty class for α>1, the group showing the highest FGT(α−1) should be targeted. For instance, to minimize the squared poverty gap, groups should be ranked by their poverty gap (FGT with α=1) and lump-sum transfers made until the poverty gap of the poorest group becomes equal to that in the next poorest group, and so on, up to the exhaustion of the budget.

Unfortunately, this rule proposed by Kanbur is only valid for classes with α>1, and it fails to cover the other popular indices like the headcount or the poverty gap. Thus, the simple algebraically optimization rules are not valid for the cases of α=0, 1. Indeed, poverty reduction is not always a decreasing function of the marginal increase of transfers. This is also explained by the different levels of the density of the population at different levels of income.

The algorithm proposed here tries to overcome this limitation. Also, it takes into account the importance of the group population sizes and considers the cases where the optimization may indicate to prioritize the groups with small population sizes, even if they are less poor. For instance, assume that we have three population groups. Also, the poverty line is equal to 10 and all poor individuals in a given group have the same income, as shown in Table 1 below. Assume that the available per capita transfer is 0.5.

Table 1: Optimal targeting with different groups

	Population share	Headcount Poverty	Income of the poor
Group A	0.1	0.20	8
Group B	0.6	0.30	6
Group C	0.3	0.25	9

Source: authors' elaboration

If we target only group B, each individual would receive 0.83 and the total reduction in poverty is nil even if the headcount is the highest. If we target only group C, each individual in this group would receive 1.66, which is more than enough to allow all individuals in this group to escape poverty. In such a case, the reduction in the population poverty is 0.25*0.3 = 0.075; also, assuming that the policy maker has not perfect information on individual income, the optimal individual transfer to group C is 1, so the per capita cost is 0.3 (=1*0.3). At this stage, the remaining 0.2 budget would be allocated to group A (and not to group B), because its population share is lower than group B, and this would enable all poor in group A to escape poverty, as their per capita group transfer is 2 (= (0.5 – 0.3)/0.1).

3.3 THE DATA-GRAPH ALGORITHM

In what follows, the three main steps of the new algorithm are introduced. The discussion is provided for the first sequence of optimization. The sequences which follow simply replicate the same logic of the first sequence, and stop when the total budget is exhausted.

Sequence 1

STEP I: Estimate the normalized poverty reduction

The first stage starts by computing the reduction in aggregate poverty (at the population level), for different levels of per capita transfer. This is similar to estimate the function $PR_g(\tau_g;\alpha)$ for different levels of τ_g. For instance, if the fixed budget of the per capita transfer is equal to $\bar{\tau}$ for each group, the reduction in aggregate poverty can be estimated for the transfers: $\frac{\bar{\tau}}{1000}, \frac{2\bar{\tau}}{1000,}, \frac{3\bar{\tau}}{1000,} \cdots, \frac{1000\bar{\tau}}{1000,}$. In such an illustrative example, we used 1,000 partitions. In general, higher the number of partitions, more accurate the results. In general, the number of partitions can affect the degree of precision in the estimations, but after a certain threshold (of partitions) the marginal gain of partitioning becomes negligible. Specifically, a higher number of partitions is suggested for indices such as the headcount where the relationship between income and the poverty reduction index is strongly non-linear. Whereas, for indices whose relationship with income is linear (such as the poverty gap and the severity of poverty), the impact of having a finer partition becomes sooner negligible. After that, we normalize the estimated $PR_g(\tau_g; \alpha)$ by the corresponding per capita transfer (τ_g). We denote the ratio between the reduction in aggregate poverty and the per capita transfer for the group g by:

$$\theta_g(\tau_g; \alpha) = PR_g(\tau_g; \alpha) / \tau_g. \quad (7)$$

Note that, by covering the whole potential levels of $\tau_g \in [0, \overline{\tau}]\ \forall g$, the algorithm seeks a global optimum of the poverty reduction.

STEP II: Rank the normalized aggregate poverty reduction

For each group, we rank the θ_g results in a descending order. Note that, for our maximization problem, such a ranking according to θ_g enables to converge quickly to the global optimum. Also, it circumvents the quasi-convexity condition. This is because, with the highest θ_g and its corresponding transfer τ_g, we cannot reach a bigger poverty reduction with lower transfers for group g or by sharing such amount across other groups. This result will be discussed in more details below. After this step, we have the basic data-graph information, which can be used to identify the optimal attribution of transfers for the first sequence. Basically, the results must be organized as in Table 2 below:

Table 2: The structure of the data-graph table

	Group 1		Group 2		Group G	
Position (p)	$\boldsymbol{\theta_1}$	$\boldsymbol{\tau_1}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\tau_2}$	$\boldsymbol{\theta_G}$	$\boldsymbol{\tau_G}$
1						
2						
.						
.						
1000						

Source: authors' elaboration

In general, the notation on the combination ($\theta_{g.p}(s)$; $\tau_{g.p}(s)$) refers to the normalized poverty reduction and the corresponding per capita transfer for the group g at position p in sequence s, in the table above (*s=1* at the first iteration).

STEP III: seek the optimal transfers

Starting from the first position of table 2, we seek the group with the highest θ (e.g., group g), such that the level of transfer to that group is defined as:

$$\tau_g^*(s=1) = \max\left[\tau_{g,p}(s=1) | \theta_{g,p} \geq \theta_{k,p}\right] \qquad \forall g, k \in G \quad (8)$$

Then, we attribute the corresponding transfer to that group g. Obviously, the transferred amount will implicitly satisfy the different constraints in (5).

STEP IV: update the data

Next, we need to update the income of individual i belonging to group g by adding the attributed transfer τ_g^* (i.e. $y_{g,i} = y_{g,i} + \frac{\tau_g^*(s=1),}{\varphi_g}$). Then, we proceed by updating the remaining budget $\bar{\tau}(s) = \bar{\tau}(s-1) - \tau_g^*(s)$, where s refers to the sequence of computation, which it is equal to one for the first sequence.

After these updates, we move to the next sequence and we repeat the four steps above, until the total budget is depleted.

Proposal 2:

The attribution of transfers based on the data-graph algorithm will converge to the global optimum of poverty reduction.

Proof

For each sequence s of transfer allocation to the group of interest g, the highest poverty reduction per dollar spent is reached $(\theta_g > \theta_k \ \forall k \neq g)$*. Thus, the sequential attribution will converge to the optimal poverty reduction. Formally, at each sequence, based on the data-graph sequence, it is easy to prove the following inequality:*

$$If\ \theta_g(\tau_g^*; \alpha) > \theta_k(\tau_{k,m}; \alpha)\ \forall k \neq g, m\ and\ \theta_g(\tau_g^*; \alpha) > \theta_l(\tau_{l,v}; \alpha)\ \ \forall l \neq g, v; \tag{9}$$

such that:

$$\tau_g^* \leq \tau_{k,m} + \tau_{l,v}$$

then:

$$PR_g(\tau_g^*;\ \alpha)\ > PR_k\big(\tau_{k,m}; \alpha\big) + PR_l(\tau_{l,v}; \alpha) \tag{10}$$

The same inequality rule can be applied for the case of more than two groups. In other terms, at sequence s, by attributing the group per capita amount τ_g^* to group g, any other partition m or v of this amount or lower across the other groups will generate a smaller reduction in poverty per dollar spent.

The following example shows the algorithm clearly. For the sake of simplicity, assume the case of just two population groups (1, 2) with population shares of 75% and of 25% respectively (but it can be easily shown that the inequality above also holds with three or more groups). Also, assume that the total fixed per capita transfer is 50$, and that step II gives the following table:

Sequence 1

Table 3: Data graph in sequence 1.

	Group 1		Group 2	
Position (p)	θ_1	τ_1	θ_2	τ_2
1	0.009	10	**0.010**	**20**
2	0.004	40	0.009	10
3	0.003	20	0.008	40
4	0.0025	30	0.007	30
5	0.002	50	0.006	50

Source: authors' elaboration

According to the results obtained in Table 3, we attribute a first amount of transfer of **20** to group 2. We generate a new per capita vector incomes, where the incomes of the second group increase by 20/0.25 = 80. The remaining population per capita transfer is 30.

Sequence 2

Table 4: Data graph in sequence 2.

	Group 1		Group 2	
Position (p)	θ_1	τ_1	θ_2	τ_2
1	0.008	6	**0.090**	**24**
2	0.007	24	0.008	30
3	0.006	30	0.007	6
4	0.005	12	0.006	12
5	0.004	18	0.005	18

Source: authors' elaboration

In sequence 2 (as illustrated in Table 4), we attribute a second lump-sum of **24** to group 2. We then generate a new per capita vector where the incomes of the second group increase by 24/0.25 = 96. The remaining transfer is 6.

Sequence 3

Table 5: Data graph in sequence 3.

	Group 1		Group 2	
Position (p)	θ_1	τ_1	θ_2	τ_2
1	**0.008**	**6.0**	0.005	3.6
2	0.007	3.6	0.008	2.4
3	0.006	4.8	0.007	1.2
4	0.005	1.2	0.006	4.8
5	0.004	2.4	0.005	6.0

Source: authors' elaboration

According to sequence 3 (shown in Table 5), we attribute a transfer of **6** to group 1 and the incomes of group 1 are increased by 6/0.75 = 8. Since the total budget is now exhausted, the algorithm stops at this sequence.

Thus, the optimal per capita transfers are 6$ for the first group and 44$ for the second group. Equivalently, the optimal *group-per capita* transfers are 8$ (6/0.75) for the first group and 176$ (44/0.25) for the second group. For simplicity, in our illustrative example, the number of positions in each sequence was five. Of course, with real data, it is better to use a significantly larger number of partitions.

Among its advantages, this data-graph algorithm optimizes the reduction in poverty for a fixed budget, it can be applied to any additive poverty index and it considers the corner solution for which some groups can have lower or upper limits of transfer in the optimal solution. The Stata *.ado file "**ogtpr**"[5] can be used to automatically estimate the optimal transfers for the reduction of any FGT indices, including the headcount and the poverty gap. The computation generally takes few seconds using standard household surveys.

[5] "ogtpr" Stata routine is downloadable from http://dasp.ecn.ulaval.ca/stata_adds/otgpr.html

In particular, the *.ado file "**ogtpr**" generates the following outputs: the optimal group and population per capita transfers, the reduction in total poverty and its statistical significance, the quality of the targeting indicator (by comparing the results of our algorithm with respect to the poverty reduction in case of perfect knowledge of individual incomes), and the inclusion and exclusion errors. As an example of the "ogtpr" routine, we use the 1998 *Enquête Prioritaires* (EP II) from Burkina Faso. With seven socio-economic groups and 8478 observations, the execution time is less than ten seconds. Based on the results shown in Table 6, given a population per-capita budget of 4,000 FCFA, the largest reduction in the headcount index is reached when crop farmers and inactive people are targeted. These groups show a population share substantially lower than the largest group (subsistence farmers), but they are likely to be closer to the poverty line. Of course, while this helps to reduce optimally the number of people living in poverty per dollar spent, such a targeting might be socially undesirable and additional analyses and considerations may be needed.

Table 6: Optimal group targeting and poverty reduction

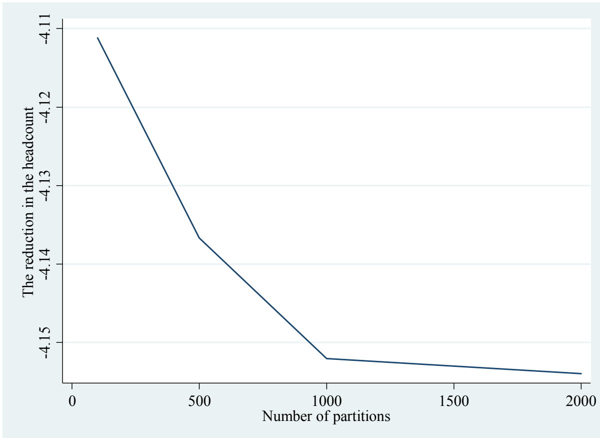
	FGT Index	**Population Share**	**Optimal group per capita Transfer**	**Optimal population per capita Transfer**
Group				
Wage-earner (public)	0.075	4.14	0.00	0.00
Wage-earner (private)	0.139	2.90	0.00	0.00
Artisan or trader	0.155	5.58	0.00	0.00
Other type of earner	0.355	0.57	0.00	0.00
Crop farmer	0.492	16.78	22,971.74	3,854.80
Subsistence farmer	0.606	65.36	0.00	0.00
Inactive	0.463	4.67	3,107.91	145.20
Targeting performance				
a) Reduction with imperfect targeting (in %): -4.152				
b) Reduction with perfect targeting (in %): -28.247				
c) Quality of the targeting indicator (in %) (a/b): 14.699				

Source: authors' elaboration based on the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso. The full results are reported in Annex A

3.4 Number of partitions and optimization

Results and, then, poverty reduction, can be sensitive to the number of the transfer's partitions. Intuitively, a smaller number of partitions reduces the space of the optimum research and, consequently, the level of precision. Using the same data as in the previous example, in Figure 2 we show the relationship between the number of partitions and the total reduction in poverty.

Figure 2: Headcount Poverty reduction and number of the transfer's partitions



Source: authors' elaboration based on the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso.

From the example above, with a lower number partitions (e.g., 100), the error with respect to the convergence level (where the reduction in poverty practically becomes a parallel line to the x-axis) is less than 1%. We can observe that the convergence is high starting with a partition of 1,000. Furthermore, our tests show that, for $\alpha \geq 1$, the required number of partitions to reach the convergence level is substantially lower than with the headcount poverty.

3.5 VALIDATING THE ALGORITHM

How can we numerically check the validity of the new algorithm and the related Stata routine? As it is well known, among the easiest methods to find the optimal solution is to use the grid approach. We then use the grid approach as a reference tool. Briefly, this approach requires first to compute the reduction of poverty for all potential combinations of transfers, and then to seek the combination that reduces the most of the total poverty. However, this approach is time consuming. For instance, if the number of the partitions is 100 and we have 4 groups, the number of the different combinations is equal to: 103!/(100!*3!) = 176 851.[6]

Using the same data as before, we did a test on three groups among the seven socio-economic groups by using a grid partition of 100 (if all seven groups were used as before, a total of about $1.5*10^{15}$ computations would have been required). The following tables (Table 7 and Table 8) show the optimal transfers for a total per capita budget of 4000.

[6] In general, the number of combinations is equal to:
$(NP + G - 1)!/(NP!)(G - 1)!$, where NP is the number of partitions and G the number of groups.

Table 7: Targeting Headcount index, grid VS data-graph approach

	Approach	
Group	**Grid**	**Data- graph**
	Optimal transfer	
Crop farmer	80.00	56.85
Subsistence farmer	3,920.00	3,943.15
Inactive	0.00	0.00
	Poverty reduction	
Total reduction	**-3.67407%**	**-3.67755%**

Source: authors' elaboration based on the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso.

Table 8: Targeting Poverty Severity Index, grid VS data-graph approach

	Approach	
Group	**Grid**	**Data- graph**
	Optimal transfer	
Other type of earner	240	254.84
Wage-earner-public sec.	0.00	0.00
Inactive	3760.00	3745.16
	Poverty reduction	
Total reduction	**-0.007647**	**-0.007647**

Source: authors' elaboration based on the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso.

The two examples above confirm the validity of our algorithm to find the optimal transfers for any poverty index.

3.6 LOCAL AND GLOBAL OPTIMA

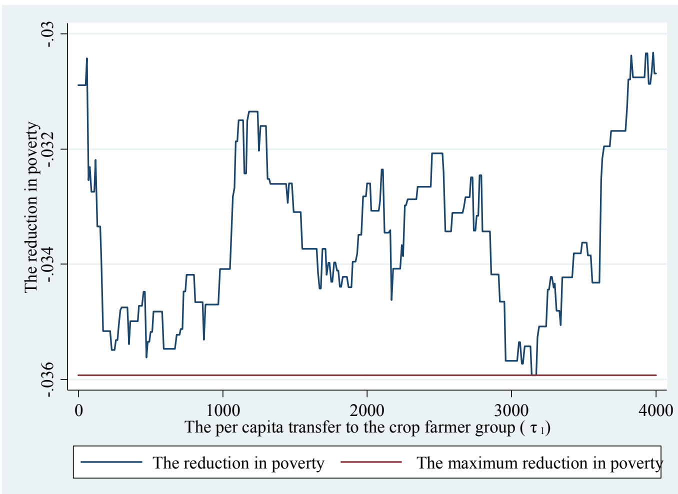
For the case of multiple local optima and non-convexity of the objective function (reduction in total poverty), differently from the Newton-Raphson method, our algorithm allows to reach the global optimum. In fact, the former methodology requires decreasing marginal returns of the objective function. For the headcount, the marginal reduction can be a non-decreasing function, which makes the Newton-Raphson algorithm inefficient. By considering all possible levels of transfer, we try to overcome the non-convexity problem. Specifically, based on the gradual attribution of the amounts as suggested in our algorithm, we try to mimic the convex form of the marginal reduction in the objective function. In other terms, the first attributed amounts of transfers must generate the highest reduction levels in poverty.

The Newton-Raphson algorithm is mainly related to the functional form of the objective function. For instance, with a strictly concave function, starting from the initial values of the arguments of interest, the marginal change in the objective function always decreases (or increases). In fact, the idea is to change the parameters gradually until we get to the optimum and where the marginal change of the objective function converges to 0. Unfortunately, some popular poverty indices are not strictly convex (or quasi-convex). In addition, the set of the budgetary constraints makes the convergence to the optimal solution more tedious.

To better present this idea, below we show the reduction in headcount using the same data of Burkina Faso. Differently from the previous example, and for simplicity of exposition, we only keep the two population groups (Crop farmers and Inactive) which received transfers as shown in Table 6. If the population per capita transfer is 4000, we have that $\tau_2 = 4000 - \tau_1$. Thus, we can seek the maximum reduction in poverty based on τ_1 only.

Figure 3 shows the reduction in poverty for different combinations of (τ_1, τ_2). As we can observe, our objective function is not strictly convex, i.e. for higher levels of τ_1 poverty should not be higher.

Figure 3: Headcount Poverty reduction, local and global optima along the transfer's partitions



Source: authors' elaboration

Assume that we reached the optimal solution (in our example above, it is for τ_1 roughly equal to 3,200); in such a situation it is expected that the average reduction in poverty per dollar spent will be the highest compared to any other combination of transfers. This result is also valid for the case of convex objective functions. Indeed, we can converge to this solution by seeking the groups that generate the maximum average reduction in poverty per dollar spent at each sequence. This is what the algorithm we proposed in this paper does. Hence, differently from the Newton-Raphson algorithm, by considering the different levels of the transfer at each sequence, we are able to overcome the problem of the local optima and we converge quickly to the global optimum.

4. An Alternative targeting method: The Proxy Means Test (PMT)

It is useful to compare the effectiveness of our algorithm to a common targeting approach such as the proxy-means test. Before moving to this comparison, it can be worthy to quickly remind a few things on how the PMT works.

In the real life, targeting indicators X_h for household h can be several, such as the age, the education level, the region, etc. Also, the indicators can be in discrete or continuous form. Assume that the proposed application $A(X)$ is used to predict the level of income starting from a set of indicators X such that:

$$y_h = A(X_h) + \epsilon_{A,h} \qquad (11)$$

The error $\epsilon_{A,h}$ depends, *inter alias*, on the application (functional form and estimation method), as well as the quality of the set of indicators X_h. In a developing country context household income is not generally observed. As originally suggested by Glewwe (1992), we can use household survey datasets reporting household income and a set of indicators, and easily estimate the household income by an OLS estimation. The estimated coefficients are then used to determine the eligibility of the applicants to any anti-poverty program. Other empirical works have suggested different econometric models in which, for example, more importance is given to a specific income percentile. For instance, we can give more weight to the poor group or to those that are close to the poverty line (Muller and Bibi, 2010).

As well known, imperfect targeting generates two types of errors: type I error occurs when we incorrectly predict as non-poor some individuals who are actually poor (exclusion error); type II error is when we include as poor some individuals who are actually not poor (inclusion error). Weighted regressions giving more weight to one group like the poorest decile or those just below the poverty line (to the detriment of another group) only reduces one type of error, but it increases the other one. To better explain this issue, assume that we use a weight that is equal to one for the poor and zero for the non-poor. Obviously, the estimated coefficients are lower compared to those using the whole population. Thus, we underestimate incomes and it is expected that the error of type I converges to zero (all non-poor's incomes are underestimated, thus we have a greater probability to identify the real poor as poor). However, the trade-off is that we incorrectly predict as poor those people who are actually not poor. As for any program evaluation, social efficiency requires a redistributive efficiency (i.e., reduction of type I error by including the largest possible number of poor into the program), and the economic efficiency (i.e., reduction of type II error, and then the cost of wrongly targeting the non-poor).

In what follows, we use a true subsample of 1,000 observations and we select the level of the poverty line to have a sufficiently large group of poor (about 45.7 %). We regress the log of the per capita expenditures on a set of six indicators (household size, sex, age and level of education of the household head, living area and regions). The OLS model gives an R2 of about 0.54. The first line in Table 9 shows that, using the OLS method, 33.28% of the population that are truly poor are correctly identified as poor. When only the observations of the poor are used (by assigning a weight equal to 1 for all those below the poverty line, and 0 otherwise), the error of type I is largely reduced to the detriment of the error of type II. The inverse is observed when the estimation focuses on the non-poor group. The last line of the table shows the results obtained through a quantile regression model at a percentile that is equal to the headcount ratio. Even with the quantile regression estimation, the sum of the two errors is higher than what found with the simple OLS.

Table 9: True status Versus Estimated status, by different methods

Method	**(0,0)**	**(1,0) ERROR_I**	**(0,1) ERROR_II**	**(1,1)**
OLS	*42.93*	12.42	11.38	*33.28*
w=1 if $y_h < z$	*5.99*	0.47	48.32	*45.22*
w=1 if $y_h \geq z$	*53.41*	37.5	0.9	*8.2*
QREG(H0)	*42.23*	11.93	12.08	*33.76*

Source: authors' elaboration based on 1,000 observations sampled from the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso.

As a useful exercise, we estimate the contribution by each indicator to the decrease in the sum of the two errors. This can be used to select pertinent indicators. To do so (1) we estimate the OLS model with all of the explanatory variables except the explanatory variable of interest, and then we compute the sum of the two errors of targeting (SER1); (2) we estimate the OLS model with all of the explanatory variables including the explanatory variable of interest, and then we compute the sum of the two errors of targeting (SER2); (3) the contribution of the explanatory variable refers to the decrease in the sum of the two errors after including the explanatory variable of interest (Contribution = SER2-SER1). The results of this exercise for the OLS model are reported in Table 10.

Table 10: Contribution of each indicator to the decrease in the sum of the two errors of targeting

Indicator	contribution
Region	-1.74
Sex	-1.33
Education	-2.45
Age	-0.48
Area	-2.56
HH size	-5.53

Source: authors' elaboration based on 1,000 observations sampled from the 1998 *Enquête Prioritaire* (EP2) from Burkina Faso.

5. POVERTY TARGETING MEASURES AND COMPARISON BETWEEN OGT AND PMT

5.1 THE QUALITY OF A TARGETING INDICATOR

The quality of targeting indicators (like the population group indicator in our case) depends on the form of the poverty index of interest, on the distribution of incomes and on the level of the budget allocated to cash transfers. Denote by x the household type to be targeted (e.g., according to the region of residence or the age groups of its members), or targeting indicator. Let $PR^*(y,\tau;\alpha)$ be the maximum possible reduction in poverty with per capita transfer τ and and assuming that the policy maker has perfect information on the individual welfare y. Also, let $PR^*(x,\tau;\alpha)$ be the maximum possible reduction in poverty with per capita transfer τ when individuals are targeted based on the indicator x. The quality of the targeting indicator x is:

$$\rho(x,\tau;\alpha) = \frac{PR^*(x,\tau;\alpha)}{PR^*(y;\tau;\alpha)} \qquad (12)$$

The $PR^*(y,\tau;\alpha)$ can be estimated using standard household surveys. The quality index $\rho(x,\tau;\alpha)$ helps to select the appropriate targeting indicators. In the example provided above with the graph-based algorithm, the quality of the targeting indicator "economic groups" is estimated at 14.699% (see the results reported in Table 6 – bottom panel).

5.2 PERFECT TARGETING, POVERTY REDUCTION AND PERFORMANCE OF TARGETING METHODS

Define the $PR^*(y,\tau;\alpha)$ function. Given a continuous distribution of incomes, the maximum reduction in headcount is equal to:

$$PR^*(y,\tau;\alpha=0) = -\int_{z-\vartheta(\tau)}^{z} dF(y) \quad (13)$$

where $\vartheta(\tau)$ is the corresponding income such that:

$$\int_{z-\vartheta(\tau)}^{z} (z-y)dF(y) = \tau \quad (14)$$

Thus, for the headcount index, the transfer will be equal to the distance between the poverty line and the income, and the transfer should be targeted in priority to those with incomes close to the poverty line. When $\alpha = 1$, we have that:

$$PR^*(y,\tau;\alpha=1) = -\int_{0}^{\eta(\tau)} \frac{(\eta(\tau)-y)_+}{z} dFy \quad (15)$$

where $\eta(\tau)$ is the corresponding income such that:

$$\int_{\mathbf{0}}^{\boldsymbol{\eta}(\boldsymbol{\tau})} (\boldsymbol{\eta}(\boldsymbol{\tau}) - \boldsymbol{y})_+ \boldsymbol{dF}(\boldsymbol{y}) = \boldsymbol{\tau} \ \ \textit{and}\ \boldsymbol{\eta}(\boldsymbol{\tau}) \leq \boldsymbol{z} \quad (16)$$

Thus, to maximize the reduction in the poverty gap, transfers should be prioritized to individuals with the largest poverty gaps. This would avoid transfer loss for those with incomes close to the poverty line and that would go well beyond this threshold after the transfer. The optimization with $\alpha = 1$ is also valid for poverty indices of higher orders.

Illustrative example: as shown in Table 11, assume that we have a population of six poor individuals. The poverty line is assumed to be equal to 11, and the per capita transfer is 1$ (the total budget of transfers of 6$).

Table 11: Optimal transfer with $\boldsymbol{\alpha} = \mathbf{0}$ and $\boldsymbol{\alpha} \geq \mathbf{1}$

#	**Income**	**Poverty gap**	**Optimal transfer**	
			$\boldsymbol{\alpha} = \mathbf{0}$	$\boldsymbol{\alpha} \geq \mathbf{1}$
1	0	10	0	4
2	2	8	0	2
3	3	7	0	0
4	7	3	3	0
5	8	2	2	0
6	9	1	1	0

Source: authors' elaboration

For this example, we find that: $\vartheta(\tau = 1) = 7$ and $\eta(\tau = 1) = 3$.

Let now denotes by m a given method for poverty targeting optimisation and by m^* the method that generates the optimal reduction in poverty. Generalizing what we had in (7) for the PMT method, given the indicator x (for instance, the proxy means test method or the optimal group targeting), the performance of method (m) in terms of optimizing poverty reduction is:

$$\vartheta(m,x,\tau;\alpha) = \frac{PR^{m}(x,\tau;\alpha)}{PR^{m^*}(x,\tau;\alpha)} \quad (17)$$

The performance index $\vartheta(m,x,\tau;\alpha)$ can be used to validate the pertinence of the method m or to show its limits.

5.3 OTHER POPULAR MEASUREMENTS OF TARGETING PERFORMANCE

Other measurements of targeting performance are:

Inclusion Error Rate (IER): it is equal to the proportion of non-poor that are wrongly targeted:

$$IER = \frac{P(y \geq z|\hat{y} < z)}{P(\hat{y}<z)} \tag{18}$$

Such index helps in monitoring the economic efficiency of the program.

Exclusion Error Rate (EER): it is equal to the poor that we fail to target.

$$EER = \frac{P(\hat{y}\geq z|y<z)}{P(y<z)} \tag{19}$$

This index helps to monitor the redistribution efficiency of the program.

Normalized Targeting Differential (NTG):

It is defined as the average transfer within the poor group minus that of the non-poor

$$NTG = \frac{E[\tau|y<z] - E[\tau|y\geq z]}{E[\tau]} \tag{20}$$

In the case of programs with a constant lump-sum transfer, we have that:

$$NTG = p(\hat{y} < z|y < z) - p(\hat{y} < z|y \geq z)) \tag{21}$$

The NTG ranges between -1 (imperfect targeting) and 1 (perfect targeting), depending on the probability p of estimating correctly or incorrectly the welfare variable

5.4 COMPARING THE PERFORMANCE OF THE OGT AND PMT METHODS.

It is now interesting to compare the PMT approach versus the optimal group targeting (OGT). Recently, this issue was largely investigated by Brown et al. (2018). This work finds that their OGT approach shows basically the same levels of efficiency compared to the PMT approach. However, their results are applicable only to convex indicators such as the Watts poverty index. In what follows, we compare the performance of targeting of our proposed OGT method compared with the PMT approach.

Briefly, the advantage of the OGT is the possibility of overcoming all modelling-related issues (functional form, weights, etc.) and of using nonlinear optimization algorithms in order to find the optimal targeting (see, for instance, Elbers et al., 2007). Its inconvenience may reside in the limited number of indicators that can be used and on the computational burden to find the optimum. On the other side, the PMT approach is a simple method that allows to use a relatively large number of indicators. However, its limitation is related to its linear form. Obviously, the choice between the different methods may depend on different aspects, like the form of the poverty index to be reduced, the design and the objectives of the targeting program, etc.

In addition to technical considerations, in the real life, the selection of the targeting indicators can depend on various issues, such as social desirability of using such indicators, cost and feasibility of collecting them.

In what follows, we compare the results of the OGT with that of the PMT. For this end, we use again the true subsample which was already used in the earlier examples. Let first assume that the policy-maker disposes of only two categorical variables, which are the level of education of the household head (8 modalities) and the living area (2 modalities). For the OGT approach, we then construct a group variable with the possible combinations of the two indicators. This generates a group variable with 16 modalities which can be used to estimate the optimal targeting and the subsequent reduction in total poverty. For the PMT approach, we start by estimating the semi-log model with the two categorical variables, and then predict the per capita expenditures.

Table 12: Total reduction in poverty (in %)

Method	α =0	α =1	α =2
OGT	9.14	10.55	17.87
PMT	3.92	7.60	13.06
$\frac{\vartheta(m = PMT, x, \tau; \alpha)}{\vartheta(m = OGT, x, \tau; \alpha)}$	42.89	72.04	73.08

Source: authors' elaboration

As it is shown in table 12, the OGT approach is a more efficient method compared to the PMT because it allows a larger reduction in poverty irrespective of the value of α. Given that the administrative costs for implementing the PMT-based transfer are expected to be higher than the OGT-based approach, the OGT method would be even more efficient.

6. CONCLUDING REMARKS

Especially in developing countries, the policy maker does not observe the individual welfare of its population, so anti-poverty interventions are generally disappointing in terms of inclusion of the real poor and exclusion of the non-poor. Also, while universal programs reach all the poor but are often unfeasible from a budgetary point of view, means-tested targeted interventions may be expensive and long to implement, and administratively difficult to manage. In this paper, we propose a new anti-poverty group-targeting algorithm in the case of fixed budget and imperfect information on household welfare. The policy-maker generally disposes of information on groups of individuals, such as regions or age groups. Under such circumstances, the goal of the policy maker is to find the optimal group transfers that reduce the most the aggregate poverty. The existing literature on antipoverty optimal targeting design was limited to the case of quasi-convex indices such as the squared poverty gap. This paper expands this literature by developing a new numerical algorithm (along with a newly developed Stata package) for estimating the optimal group transfers which reduce the most any additive poverty indices, like the FGT class of poverty indices. Tested on household data from Burkina Faso, we found that our proposed algorithm is fairly more efficient than a PMT-based targeting approach, irrespective of the poverty indicator. Also, the efficiency of our algorithm generally increases with finer partitions of the group per capita transfer. Finally, the algorithm proposed here replicates quite closely the performance obtained through the grid approach (our benchmark tool) and is able to converge to the global optimum even when, like in the headcount poverty index, there are multiple possible local optima. The empirical exercise presented in this paper is just illustrative and serves to show the performance of the new algorithm in terms of optimal poverty targeting. Appropriate group identification should be defined together with local policy-makers on a case-by-case basis.

REFERENCES

Brown, C., Ravallion, M., & van de Walle, D. (2018). A poor means test? On econometric targeting in Africa. *Journal of Development Economics*, 134 (2018) 109-124

Coady, D., Grosh, M., & Hoddinott, J. (2004). *Targeting of transfers in developing countries: Review of lessons and experience*. The World Bank.

Cornia, G. A., & Stewart, F. (1993). Two errors of targeting. *Journal of International Development*, *5*(5), 459-496.

Devereux, S., Masset, E., Sabates-Wheeler, R., Samson, M., Rivas, A. M., & Te Lintelo, D. (2017). The targeting effectiveness of social transfers. *Journal of Development Effectiveness*, *9*(2), 162-211.

Elbers, C., Fujii, T., Lanjouw, P., Özler, B., & Yin, W. (2007). Poverty alleviation through geographic targeting: How much does disaggregation help? *Journal of Development Economics*, *83*(1), 198–213. https://doi.org/10.1016/j.jdeveco.2006.02.001

Glewwe, P. (1992). Targeting Assistance to the Poor: Efficient Allocation of Transfers when Household Income Is Not Observed. *Journal of Development Economics*, *38*(2), 297–321.

Hanna, R., & Olken, B. A. (2018). Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives*, *32*(4), 201-26.

Higgins, S., & Lustig, N. (2016). Can a poverty-reducing and progressive tax and transfer system hurt the poor?. *Journal of Development Economics*, 122, 63-75.

Kanbur, R. (1987). *Budgetary Rules for Poverty Alleviation*. IMF Staff Papers

Margitic, J., & Ravallion, M. (2019). Lifting the floor? Economic development, social protection and the developing World's poorest. Forthcoming in *Journal of Development Economics*.

Muller, C., & Bibi, S. (2010). Refining targeting against poverty evidence from Tunisia. *Oxford Bulletin of Economics and Statistics*, *72*(3), 381-410.

Ravallion, M., & Chao, K. (1989). Targeted policies for poverty alleviation under imperfect information: Algorithms and applications. *Journal of Policy Modeling*, *11*(2), 213–224. https://doi.org/10.1016/0161-8938(89)90015-X

Annex A

The Stata command

```
ogtpr exppc, hgroup(gse) hsize(size) alpha(0) pline(80000) trans(4000) ered(1) part(1000)
```

The displayed results:

```
Main Information
Number of observations    :  8478
Time of computation       :  8.20      second(s)
Per capita transfer       :  4000.00
Used per capita transfer  :  4000.00
Household size            :  size
Sampling weight           :  weight
Group variable            :  gse
Parameter alpha           :  0.00

Algorithm sequences
Sequence ...1:   Remaining p.c. budget   3952.326 over   4000.000
Sequence ...2:   Remaining p.c. budget   3878.941 over   4000.000
Sequence ...3:   Remaining p.c. budget   3807.127 over   4000.000
Sequence ...4:   Remaining p.c. budget      0.000 over   4000.000


Optimal Group Transfers
  +----------------------------------------------------------------------------------------------+
  |                  Group    |     Fgt Index     Population    Optimal G.P.C.     Optimal P.C.|
  |                           |                      Share          Transfer          Transfer |
  |---------------------------+------------------------------------------------------------------|
  |Wage-earner (public sector) |         0.075          4.140          0.000             0.000|
  |Wage-earner (private sector)|         0.139          2.904          0.000             0.000|
  |Artisan or trader          |         0.155          5.580          0.000             0.000|
  |Other type of earner       |         0.355          0.569          0.000             0.000|
  |Crop farmer                |         0.492         16.781      22971.742          3854.801|
  |Subsistence farmer         |         0.606         65.355          0.000             0.000|
  |Inactive                   |         0.463          4.672       3107.906           145.199|
  +----------------------------------------------------------------------------------------------+

Total Poverty Reduction
------------------------------------------------------------------------------------------
Variable |   Estimate   Std. Err.       t      P>|t|       [95% Conf. interval]  Pov. line
---------+--------------------------------------------------------------------------------
   exppc |  .5180611    .0109517   47.3042   0.0000       .4965334    .5395888     80000
exppc_tr |  .4765408    .0112242   42.4565   0.0000       .4544774    .4986042     80000
---------+--------------------------------------------------------------------------------
    diff.| -.0415202    .0046131   -9.0005   0.0000      -.0505882   -.0324522      ---
------------------------------------------------------------------------------------------

Targeting quality
    - Reduction with imperfect targeting (in %)      :   -4.152
    - Reduction with  perfect targeting (in %)       :  -28.247
    - The quality of the targeting indicator (in %) :   14.699

Targeting by transfers and poverty status
                    |        Targeted
               Poor |        No         Yes |      Total
        -----------+----------------------+----------
                 No |     37.17       11.03 |      48.19
                Yes |     41.38       10.43 |      51.81
        -----------+----------------------+----------
              Total |     78.55       21.45 |     100.00
```