Optimal Targeting and Poverty Reduction

Abdelkrim Araar[1] and Luca Tiberti[2]
[1]Laval University & PEP, aabd@ecn.ulaval.ca
[2] Laval University & PEP, luca.tiberti@ecn.ulaval.ca

Abstract - *In this note we develop a new algorithm for optimal anti-poverty group targeting when the total budget is fixed. Contrarily to other similar algorithms found in the literature, the proposed algorithm is applicable to all additive poverty indices. Finally, we briefly discuss some popular indicators of targeting performance, as well as the popular proxy-means test (PMT) approach. The notes concludes by comparing the targeting performance of the proposed algorithm and the PMT. Our proposed algorithm for optimal targeting is found to be more efficient than the PMT irrespective of the poverty indicator.*

1- INTRODUCTION

The main aim of this paper is to introduce a new anti-poverty group-targeting algorithm in the case of fixed budget. We assume that the policy-maker does not dispose of perfect information on the welfare level of each individual in a population. Such information is 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. Also, it is assumed that the population is divided into mutually exclusive population groups, for instance, living areas. The goal of the policy maker is to find the optimal group transfers in order to reduce the most aggregate poverty. Thus, the aggregate poverty index becomes the objective function to minimize.

The analytical optimization requires a set of basic conditions. First, as usual in minimizations, 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 the poverty gap ratios, obey this condition. Various theoretical and empirical works have focused on the case of the squared poverty gap index, which satisfies the set of optimization conditions. 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 subject to a fixed budget of transfers.

In this note, a new data-graph algorithm is developed. This algorithm is conceived to find the optimal group transfers that allow the maximum possible reduction in any additive poverty indexes, like the FGT. The rest of this paper is organized as follow. In section 2, the poverty reduction function is introduced, as well as its normalized form by the cost of transfers. In section 3, the data-graph algorithm is discussed. Also, this note presents two Stata modules that can be used for the optimization of poverty reduction.

2- THE GROUP POVERTY REDUCTION CURVES

As indicated earlier, the previously developed algorithms for the optimal targeting of population groups with 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). Indeed, 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 fill this gap by suggesting a numerical approach that combines the numerical form of the optimization 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. In our view these developments will be a valuable contribution to the literature on this field.

The poverty reduction component

Assume that a lump-sum transfer is only attributed to the population group g. The per capita lump-sum transfer is denoted by τ_g. The change in the contribution to total poverty by individual i living in the targeted group g 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:

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

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 $\mathrm{I}[condition]$ 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 is τ_g/φ_g.[1] 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 aggregate poverty when targeting group g as followed[2]:

[1] 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.

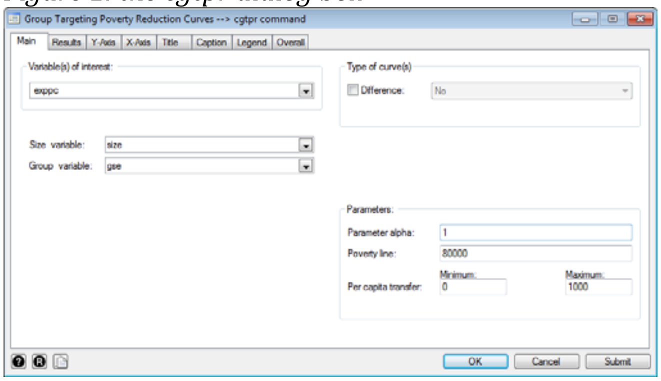
[2] For simplicity, we omit the common denominator $1/z^{\alpha}$ for the class of the FGT poverty indices.

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

The function $PR_g(\tau_g;\alpha)$ is called the *Group Targeting Poverty Reduction* (GTPR). The Stata module ***cgtpr.ado*** can be used to draw the GTPR curves.

Example 1:

Figure 1: the cgtpr dialog box



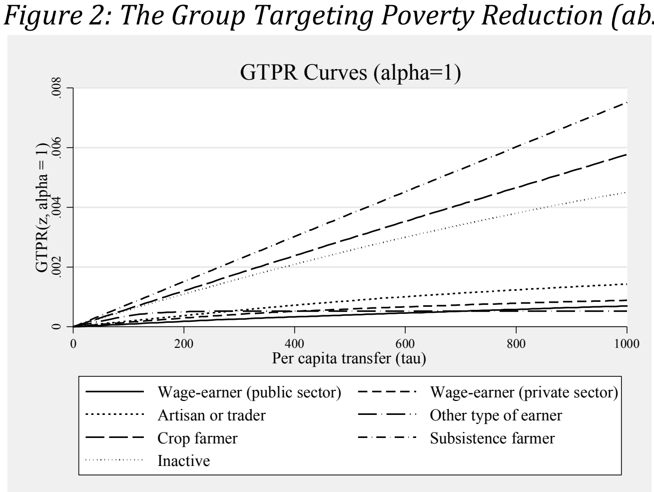
Source: authors' elaboration

The syntax:

cgtpr exppc, alpha(1) hsize(size) hgroup(gse) min(0) max(1000) pline(80000) xline(600)

gives the graph below:

Figure 2: The Group Targeting Poverty Reduction (absolute)



Source: authors' elaboration

Figure 2 illustrates the (absolute) reduction in the poverty gap after different per capita transfers ranging from 0 to 1000. In this example, the largest reduction in poverty is obtained by targeting the group of subsistence farmers. The bottom curves show the case of relatively less poorer groups for which the transfer, after a certain amount, does not affect poverty (as for "other type of earners") or for which the marginal poverty reduction flattens.

3- The optimal transfers

Case 1: Single group targeting

Proposal 1:

For a given per capita transfer τ_g and in the case of a single group targeting constraint (i.e., only one group is targeted), the optimal targeting is the one given to the group with the highest GTPR curve at τ_g .

For instance, based on the results of Figure 2, the optimal targeting prioritizes the *subsistence farmer* group.

Case 2: Multiple group targeting

Now, assume that the policy-maker aims to target more than one group by allocating different levels of lump-sum transfers. The objective becomes to find the optimal transfers that reduces optimally the aggregate poverty, under the transfer constraint $\tau = \sum_{g=1}^{G} \tau_g$, as well as the G range constraints: $0 \leq \tau_g \leq \iota_g \ \forall\, g$, where ι_g denotes the maximum poverty gap within the group g. Formally, if we denote the reduction in aggregate poverty by $PR(\boldsymbol{\tau}; \alpha) = \sum_{g=1}^{G} PR_g\big(\tau_g; \alpha\big)$, 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 that of Kanbur (1987) to minimize the aggregate poverty with FGT(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 cases 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 new suggested algorithm tries to overcome this difficulty. Also, it takes into account the importance of the group population sizes and considers also the cases where the optimization may require 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 the following table. 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 receives 1 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 must receive at least 1 to exit poverty, and the reduction in poverty is 0.25*0.3 = 0.075 for a per capita cost of 2*0.3 = 0.3. At this stage, the remaining 0.2 budget cannot be allocated to group B, but to group A, because its population share is lower, and this enables to have a sufficiently large transfer to escape from poverty (i.e. (0.5 – 0.3)/0.1 = 2).

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 with the computation of 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 effect become negligible. As discussed later, 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 negligible. After that, we normalize the estimated $PR_g(\tau_g;\alpha)$ by the corresponding per capita transfer (τ_g). For simplicity, we denote the ratio between the reduction in aggregate poverty and the per capita transfer for the group g by:

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

Note that, by covering the whole potential levels of $\tau_g \in [0,\bar{\tau}]\ \forall g$, 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 the 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 the Table 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 what follows, the combination ($\theta_{g.p}$; $\tau_{g.p}$) refers to the normalized poverty reduction and the corresponding per capita transfer for the group g at the position p in the table above.

STEP III: seek the optimal transfers

Starting from the first position of the table, we seek the group with the highest $\theta_{g,1}$, and then we attribute the corresponding transfer as a first amount of transfer to that group (for example, the group g). Obviously, the transferred amount implicitly satisfies the different constraints in (5).

STEP IV: update the data

Next, we need to update the income by adding the attributed transfer to group g (i.e. $y_{g,i} = y_{g,i} + \frac{\tau_{g,1}}{\varphi_g}$). Then, we proceed by updating the remaining budget to be distributed $\bar{\tau}_s = \bar{\tau}_{s-1} - \tau_{g,1}$, where s refers to the sequence of computation.

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

Proposal 2:

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

Proof

At each sequence of attribution of the transfer to the group of interest, the highest poverty reduction per dollar spent is reached ($\theta_{g.1} > \theta_{k.1} \quad \forall k \neq g$). *Thus, the sequential attribution will converge to the optimal reduction in poverty. Formally, based on the organization of the data-graph table, it is easy to prove the following inequality:*

if: $\frac{PR_g(\tau_{g,1};\alpha)}{\tau_{g,1}} > \frac{PR_l(\tau_{l,r};\alpha)}{\tau_{l,r}}$ *and* $\frac{PR_g(\tau_{g,1};\alpha)}{\tau_{g,1}} > \frac{PR_m(\tau_{m,s};\alpha)}{\tau_{m,s}}$;
such that:

$\tau_{g,1} \leq \tau_{l,r} + \tau_{m,s}$

then:

$$PR_g(\tau_{g,p};\alpha) > PR_l\big(\tau_{l,r};\alpha\big) + PR_m(\tau_{m,s};\alpha).$$

The same inequality rule can be applied for the case of more than two groups. In other terms, at sequence s, by attributing the per capita amount $\tau_{g,1}$ *to group* g, *any other partition of this amount across the other groups will generate a lower reduction in poverty per dollar spent.*

The following example shows the algorithm clearly. Assume the case of two population groups (1, 2) with population shares of 75% and of 25% respectively. 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_1	τ_1
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 2, 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 = 100. The remaining transfer is 30.

Sequence 2

Table 4: Data graph in sequence 2.

	Group 1		Group 2	
Position (p)	θ_1	τ_1	θ_1	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

We attribute a second amount of transfer 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 = 116. 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_1	τ_2
1	**0.008**	**6**	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

We attribute an amount of transfer of **6** to group 1 and the incomes of group 1 are increased by 6/0.75 = 8. Since the total budget is exhausted, the algorithm ends at this sequence.

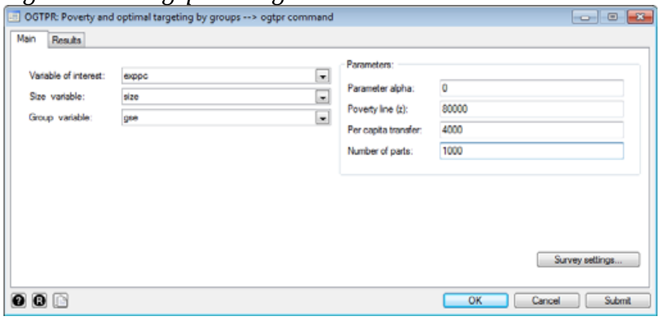
Thus, the optimal per capita transfers are 6$ for the first group and 44$ for the second group. Equivalently, the optimal group-per capita transfers are 8$ (6/0.75) for the first group and 176$ (44/0.25) for the second group. For simplicity, in our illustrative example, the number of positions in each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

The virtues of this data-graph algorithm:

- It optimizes the reduction in poverty for a fixed budget of transfers;
- It can be applied to any additive poverty index.
- 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**" can be used to automatically estimate the optimal transfers for the reduction in poverty measured by any of the FGT indices, including the headcount and the poverty gap. The computation generally takes few seconds using standard household surveys:

Figure 3: the ogtpr dialog box



Source: authors' elaboration

Which gives the output below:

```
. ogtpr exppc, hgroup(gse) hsize(size) alpha(0) pline(80000) trans(4000) ered(1) part(1000)
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 Targeting of Groups for Poverty Reduction
    Number of observations      :8478
    Time of computation        : 8.79      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
```

Group	Fgt Index	Population Share	Optimal G.P.C. Transfer	Optimal P.C. 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	---

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

Targeting by transfers and poverty status

Poor	Targeted No	Targeted Yes	Total
No	37.17	11.03	48.19
Yes	41.38	10.43	51.81
Total	78.55	21.45	100.00

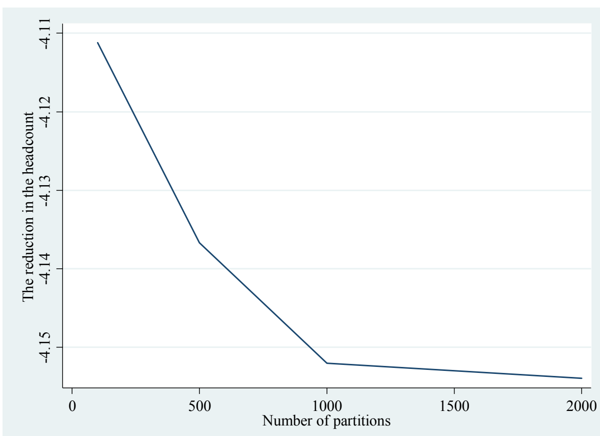
In particular, the *.ado file "**ogtpr**" generates the following outputs: the optimal group and population per capita transfer, the reduction in total poverty and its statistical significance, the quality of the targeting indicator and the inclusion and exclusion errors.

For the example above, we used the Burkina Faso household survey of 1998. For the computation of the optimal transfers for seven socio-economic groups and 8478 observations, the execution time is less than ten seconds. Based on the results shown above, for the optimal reduction of the headcount index, we are required to target the groups of crop farmers and inactive people. 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. If these groups are prioritized, poverty reduces faster per dollar spent. 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.

NUMBER OF PARTITIONS AND OPTIMIZATION

Let now discuss about the impact of the choice of the number of the transfer's partitions and its implication on the level of the precision of convergence. Intuitively, it is clear that by selecting a lower number of partitions we reduce the space of optimum research, and consequently the level of precision. Using the same data as in the previous example, we show in the next Figure the relationship between the number of partitions and the total reduction in poverty.

Figure 4: Poverty reduction and number of the transfer's partitions



Source: authors' elaboration

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 1000. Furthermore, our tests show that, for $\alpha \geq 1$, the required number of partitions can be low to reach the convergence level.

VALIDATING THE NEW SUGGESTED ALGORITHM

How can we numerically check the relevance of the new algorithm and the related developed Stata routine? As it is well known, among the easiest methods to find the optimal solution is to use the grid approach. 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 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.[3]

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 show the optimal transfers for a total per capita budget of 4000.

Table 6: Targeting Headcount index, grid VS data-graph approach

	Approach	
Group	**Grid**	**Data- graph**
	Optimal transfer	
Crop farmer	80.00	56.85
Subsistence farmer	3920.00	3943.15
Inactive	0.00	0.00
	Poverty reduction	
Total reduction	**-3.67407%**	**-3.67755%**

Source: authors' elaboration

Table 7: 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**	**-.0076469**

Source: authors' elaboration

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

BASIC FOUNDATION OF THE SUGGESTED ALGORITHM

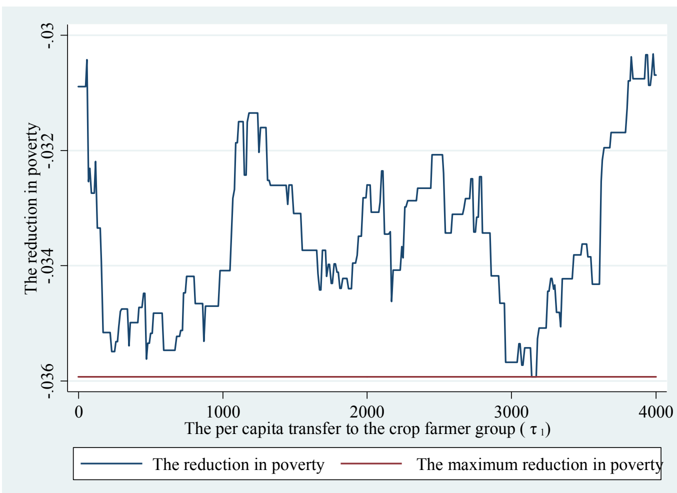
For the case of multiple local optimums and non-convexity of the objective function (reduction in total poverty), differently from the Newton-Raphson method, our algorithm allows to reach the global optimum. Indeed, the former methodology requires decreasing marginal returns of the objective function. 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. Indeed, based on the gradual attribution of the amounts with the suggested algorithm, we try to mimic the convex form of the marginal reduction in the objective function (total poverty). In other terms, the first attributed amounts of transfers must generate the highest reduction levels in poverty.

[3] 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.

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 gradually the parameters until arriving at the optimum and where the marginal change of the objective function converges to 0. Unfortunately, some popular poverty indices are not strictly convex. In addition, the set of the budgetary constraints makes the convergence to the optimal solution more tedious.

To better present this idea, in the following graph, we show the reduction in headcount using the same data of Burkina Faso but, for simplicity, we only kept two population groups (Crop farmers and Inactive). If per capita budget of transfer is 4000, we have that: $\tau_2 = 4000 - \tau_1$. Thus, with this basic constraint, we can seek the maximum reduction in poverty based on τ_1. Obviously, the constraint reduces our degree of freedom by one.

Figure 4: Poverty reduction and number of the transfer's partitions



Source: authors' elaboration

The Figure above shows the reduction in poverty for different combinations of (τ_1, τ_2), with the sum of the two transfers always equal to 4000 (so, $\tau_2 = 4000 - \tau_1$). As we can observe, this objective function is not strictly convex, i.e. for higher levels of $\tau_1$poverty should not be higher.

Assume that we are at the optimal solution; in such a situation it is expected that the average reduction in poverty per dollar will be the highest compared to any other combination of transfers. This result is also valid for the case of convex objective functions. Based on this, we can converge to this solution by seeking the groups that generate the maximum average reduction in poverty per dollar at each sequence. This is what the algorithm we proposed in this paper does. Furthermore, 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.

OTHER SUGGESTED POVERTY TARGETING NOTIONS

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 for transfers. Let denote by x the indicator to be targeted (like regions and age groups). Let $PR^*(y,\tau;\alpha)$ be the maximum possible reduction in poverty with per capita transfer τ and 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:

$$\rho(x,\tau;\alpha) = \frac{PR^*(x,\tau;\alpha)}{PR^*(y;\tau;\alpha)} \tag{8}$$

The $PR^*(y,\tau;\alpha)$ can be estimated using the household survey. The quality index $\rho(x,\tau;\alpha)$ helps to select the appropriate targeting indicators. In the example provided above, the quality of the targeting indicator “economic groups” is estimated at 14.699% (see the output results below Figure 3).

Perfect targeting and poverty reduction

It may be helpful to define the $PR^*(y,\tau;\alpha)$ function. Under the continuous form of the distribution of incomes, the maximum reduction in headcount is equal to:

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

where $\vartheta(\tau)$ is a corresponding income such as:

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

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

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

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

$$\boldsymbol{\int_0^{\eta(\tau)} (\eta(\tau) - y)_+ dF(y) = \tau}$$

Thus, among the easiest solutions is to attribute the transfers to the poorest individual with larger poverty gaps. This will avoid transfer loss for those with incomes close to the poverty line. It follows that the proposed optimization with $\alpha = 1$ is also valid for the case of the highest values of α.[4]

[4] The last range of α that discussed shortly is that where $\alpha \in \,]0,1[$. Within this range, an increase in inequality within the poor group will decrease the poverty. Thus, the optimal transfers is this that targeted the middle class of the poor group.

Illustrative example: Assume that we have a population of six poor individuals. The poverty line is assumed to be equal to 10, and the per capita transfer is 1$ (the total budget of transfers of 6$).

Table 8: Optimal transfer with $\boldsymbol{\alpha} = \mathbf{0}$ and $\boldsymbol{\alpha} \geq \mathbf{1}$

#	Income	Poverty gap	Optimal transfer	
			$\boldsymbol{\alpha = 0}$	$\boldsymbol{\alpha \geq 1}$
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.

Performance of targeting methods

Let m denotes a given method for poverty targeting optimisation and m^* identifies the method that generates the optimal reduction in poverty. The level of performance of method (m) for the optimization of the reduction in poverty with the indicator x (for instance, the proxy means test method or the optimal group targeting) is defined as followed:

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

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

The Proxy Means Test (PMT)

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:

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

The error $\epsilon_{A,h}$ will depend, *inter alias*, on the application (functional form and estimation method), as well as the quality of the set of indicators X_h. Obviously, in the real life and especially in a developing country context, the income of each household y_h is not observed.

As suggested by Glewwe (1992), the easiest way to model the income with a set of indicators is by using the OLS estimation. The estimated coefficients can be 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 to a given segment of incomes is given. For instance, we can give more weight to the poor group (Mapa & Albis, 2013) or for instance those that are close to the poverty line.

As well known, imperfect targeting generates two types of errors: the error of type I occurs when we incorrectly predict as non-poor some individuals which are actually poor; the error of type II is when we include as poor some individuals which are actually not poor. It follows that the weighted regression where more importance to one group is given to the detriment of another one 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 will be low compared to those using the whole population. Thus, we will underestimate incomes and it is expected that the error of type I will converge to zero (all non-poor's incomes are underestimated, thus we have a great probability to identify the real poor as poor). However, its trade-off is that we incorrectly predict as poor those people which 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 1000 observations and we select the level of the poverty line to have a large enough 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 region). 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

The following table reports the contribution of each indicator to the decrease in the sum of the two errors. This norm can be used to select pertinent indicators. To estimate the contribution of each explanatory variable, we do what follows:

1. We estimate the OLS model with all of the explanatory variables except the explanatory variable of interest. 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. 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).

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

OTHER POPULAR MEASUREMENTS OF TARGETING PERFORMANCE

At this stage, it may be helpful to recall briefly some of the measurements of the targeting performance.

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

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

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.

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

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

Normalized Targeting Differential (NTG):

It equals to the average transfer within the poor group minus that of the non-poor

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

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

$$NTG = P(\hat{y} < z|y < z) - P(\hat{y} < z|y \geq z))$$

The NTG ranges between -1 (imperfect targeting) and 1 (perfect targeting).

Among the criticisms that can be addressed to these popular measurements of targeting performance is their focus on counting indices. For instance, assume that our poverty index of interest is the poverty gap, and also that the transfer depends on the estimated income. These indicators do not take into account of the cost of the transfer neither of the reduction in the poverty gap and the poverty severity.

MAPPING, POPULATION GROUPS 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, Ravallion, & Walle (2016). This work finds that the OGT approach shows basically the same levels of efficiency compared to the PMT approach.

It may be helpful to rethink on the technical inconveniences and advantages of each of these two approaches. The advantage of the OGT is the possibility of overcoming all modeling-related issues (functional form, weights, etc.) and of using nonlinear optimization algorithms in order to find the optimal scores (see for instance Elbers et al., 2007). Its inconvenience may reside in the limited number of indicators that one can use and on computational burden to find the optimum. On the other side, the PMT-OLS 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 different 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. In an extreme case, one can suppose that the policy maker can predict the per capita expenditures for each individual using the estimated coefficient of the model, and then attribute the appropriate transfer according to the estimated gap, the poverty index of interest and the total budget of transfers. This turns to be a perfect targeting with the estimated per capita expenditures.

Table 11: Total reduction in poverty (in %)

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

Source: authors' elaboration

As it is shown in the table above, 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.

References

Brown, C., Ravallion, M., & Walle, D. Van De. (2016). A poor means test? econometric targeting in Africa.

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.

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

Mapa, D. S., & Albis, M. L. F. (2013). New Proxy Means Test (PMT) Models: Improving Targeting of the Poor For Social Protection. In *National Convention on Statistics (NCS) EDSA Shangri-La Hotel.*

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