A discusión

OPTIMISING ANTI-POVERTY TRANSFERS WITH QUANTILE REGRESSIONS*

Christophe Muller

WP-AD 2006-07

Correspondence to: Christophe Muller. Departamento de Fundamentos del Análisis Económico. Universidad de Alicante. Campus de San Vicente, 03080 Alicante, Spain. Teléfono: 965903614. Fax: 965903898. E-mail: cmuller@merlin.fae.ua.es

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

OPTIMISING ANTI-POVERTY TRANSFERS WITH QUANTILE REGRESSIONS

Christophe Muller

ABSTRACT

Anti-poverty transfer schemes are one of the main way of fighting poverty. Under perfect observation of incomes, designing such scheme boils down to solving an optimisation program under constraints, which can be achieved with well-defined methods. In contrast, when incomes cannot be perfectly observed, the schemes are usually based on predictions of living standards using ancillary regressions and household survey data to predict the unobserved living standards of households. In this paper, we study the poverty minimisation program under imperfect information. We show why using predictions of living standards helps to deal approximately with an otherwise intractable problem. Then, we propose a new approach to the practical optimisation procedure based on improved predictions of living standards in terms of the targeting problem to be solved. Our new empirical methodology to target direct transfers against poverty is based on observable correlates and on estimation methods that can focus on the poor: the quantile regressions. We illustrate our results using data from Tunisia.

Keywords: Poverty minimization, Quantile regressions, Policy targeting.

1. Introduction

Anti-poverty transfer schemes (APTS) are one of the main way of fighting poverty. The aim of such schemes is to minimize poverty under budget and information constraints by transferring positive monetary amounts to the poor. In the absence of the observation of the living standards of all households, they are constructed from predictors of living standards based on ancillary regressions estimated from household survey data. These living standard predictors are used to assess the unobserved living standards of households in the population of interest.

In this paper, we first study the minimisation program of poverty under budget constraint and imperfect information. We show why using predictions of living standards helps to deal approximately with an otherwise intractable problem. Then, we propose a new approach to the practical optimisation procedure based on improved predictions of living standards in terms of the targeting problem to be solved. Our new empirical methodology to target direct transfers against poverty is based on observable correlates and on estimation methods that can *focus on the poor*: the quantile regressions.

We illustrate our method by using data from Tunisia, from which we estimate 'focused' transfer schemes that improve anti-poverty targeting performances. Post-transfer poverty can be substantially reduced with the new estimation method.

How to optimize anti-poverty transfers? A theoretical approach to this question is to solve the minimisation program of poverty under budget constraint. However, availing oneself of such theoretical framework is insufficient. Indeed, people's incomes are generally unobserved. Then, what is needed is a way to implement the APTS when only limited information is available: (1) complete information

on living standards and household characteristics for a sample of households ex-
tracted from a household survey, and (2) observations of individual characteristics
for the whole population, while living standards are not observed.

Another related question is: How to assess the consequences of the proposed
targeting scheme on poverty? This involves estimating poverty, or other welfare
criteria, using the household survey data and applying the transfers formula ob-
tained at the previous stage to calculate post-transfer living standards. Thus, in
all cases the method for determining the optimal transfer is crucial.

There is a growing economic literature on empirical implementation of anti-
poverty transfer schemes, which compensates for the absence of related mathe-
matical literature[1]. In these papers, OLS predictions of incomes and the survey
data are used to derive some numerical solution to the transfer problem, generally
with a quasi-Newton method for solving the first-order conditions of the optimi-
sation problem. In these cases researchers often assimilate the observed sample
to the whole population to simplify the calculus. Thus, the extrapolation issue is
not fully dealt with. Moreover, the obtained targeting efficiency of such transfer
schemes is weaker than desired. In particular, monetary leakages and exclusion
of poor households are important issues, and unpalatable negative transfers may
occur. In these conditions, one could wonder if better methods could not be de-
signed to deal with the anti-poverty targeting problem. The object of this paper
is to expose how such method can be theoretically investigated, and to propose a
better practical solution than the ones currently in force.

Little theoretical investigation of the properties of the associated optimisation
program has been carried out. Besley and Kanbur (1988) study optimal food
subsidies against poverty, characterising the first-order conditions for FGT in-

dicators with parameter α as implying the equalization of FGT indicators with parameter $\alpha - 1$ [2]. Kanbur et al (1991) deal with the optimisation problem by implementing numerical simulations, for the case of nonlinear income taxation. A salient contribution is that of Bourguignon and Fields (1990, 1997), from now B&F, who determine, under perfection information, the optimal formula to allocate anti-poverty budgets by using positive transfers for FGT poverty indicators. Chakraborty and Mukherjee (1998) derive the optimal subsidies to the poor as a function of the density function of incomes and the derivative of the kernel function of the considered poverty indicators, albeit only under a priori conditions on the subsidy function.

There are several shortcomings in the theoretical results reached so far. First, the fact that transfers should be positive must be respected. Indeed, it is much harder to have people pay than to have them receive money. Nonetheless, many theoretical results correspond to transfers that are allowed to be negative and the crucial positivity constraints are omitted. Second, the issue of statistical extrapolation from an observed sample to the global population is not accounted for. We shall fully treat this aspect. Third, the central issue that only imperfect information is available, and notably incomes are generally not observed, is usually not dealt with in the theoretical literature. What would be needed is an optimisation program accounting for characteristics that can be observed. Fourth, even under perfect information, the optimisation programs are generally presented in discrete form and solved by simple intuitive rules without making explicit the mathematical features of the problem. Allowing for continuous specifications and using the calculus of variations, we shall exhibit the fundamental structure of the poverty minimisation problem under perfect information. Fifth, the link between the theo-

retical optimisation program and the statistical estimates used for the predictions of living standards is missing. Sixth, the passage from the optimisation program under perfect information to the optimisation program under imperfect information is missing. Seventh, the choice of stochastic treatment has not been fully discussed.

We discuss the optimisation problem in Section 2. In Section 3, we present the theoretical solution under imperfect information. In Section 4, we explain the chain of statistical treatment for our new method. In Section 5, we discuss a few illustrative results using Tunisian data. Finally, Section 6 concludes.

2. The Optimisation Problem

2.1. The perfect information case

Most of the poverty indices used in applications are additively decomposable and can be written in the following form

$$\int_0^z f(y, z) d\mu(y) \tag{1}$$

where μ is the probability distribution function of real living standards y, and z is the poverty line. Function f is denoted the 'kernel function' of the poverty index. We consider now the most popular of these indices in applied work[3].

The Foster-Greer-Thorbecke (FGT) poverty indices with α, the poverty aversion parameter of the public planner, is defined as

$$\int_0^z (1 - y/z)^\alpha d\mu(y) \tag{2}$$

Foster, Greer and Thorbecke (1984) discuss the economic properties of these indices. In these formulae, the kernel function, indicated by $(1 - y/z)^{\alpha}$, describes the contribution to aggregate poverty of an household of living standard y for $y < z$ (poor households).

The Watts' poverty index (Watts, 1968) is

$$\int_0^z -\ln(y/z) \ d\mu(y) \tag{3}$$

Let us first consider the case of perfect information where incomes can be observed from the whole population of individuals. In practice, APTS are organised around households rather than around individuals. Also, living standard variables are used instead of income variables. This enables researchers to account on the one hand for differences in household composition, and on the other hand for the heterogeneity of individual and environment characteristics (through price indices for example). It is easy to translate all the methods and results of this paper to the case of households and living standards instead of individuals and incomes. However, to simplify the notations, we present most of them only for individuals and incomes. A theoretical APTS program based on transfers under perfect information can be written as follows

$$\min_{t(.,.)} \ \int_0^z P\left(z, y + t(z,y)\right) dF_{y+t}(y + t(z,y))$$

$$\text{subject to} \quad : \quad \int_0^{+\infty} t(z,y) dF(y) = B \tag{4}$$
$$\text{and } t(z,y) \quad \geq \quad 0 \text{ for all } y \text{ and } z,$$

where P is the kernel function of the poverty index, y is the income variable, $y + t(z, y)$ is the post-transfer income of which cumulative density function (cdf) is $F_{y+t}(y)$, z is the poverty line. Function $t(z, y)$ is the transfer function that defines the value of the monetary transfer to an individual of income y when the poverty line is z. B is the available budget for the APTS.

In the economic literature, this type of problem has been studied with the FGT poverty measures and a finite population of n individuals. In this case the optimisation program is discrete and can be written as follows.

$$\min_{\{t^i\}} \sum_{i=1}^{n} \left\{ \frac{z - y^i - t^i}{z} \right\}^{\alpha} 1_{[y^i + t^i < z]}$$

$$\text{subject to:} \sum_{i=1}^{n} t^i = B \ \text{ and } t^i \geq 0 \text{ for all } i,$$

where index i is the index of the individual and $1_{[y^i + t^i < z]}$ is the dummy variable defining the poor after transfer.

In this framework, B&F show that perfect targeting minimizing the headcount index (FGT index with $\alpha = 0$) would award transfers so as to lift the richest of the poor out of poverty ('r-type transfer'): $t^i = z - y^i$ if $y^i < z, t^i = 0$ otherwise, ranking the individuals to be served from the richest poor to the poorest poor that can be served with the available budget. In contrast, if the aim is to minimize a FGT poverty index with $\alpha > 1$, it is optimal to allocate the anti-poverty budget to the poorest of the poor ('p-type transfer'). A mathematical proof of these results can be obtained by considering the right-hand-side derivative of the objective function with respect to a small positive transfer to an individual i, and noting for which income this right-hand-side derivative is the more negative. In that case,

the APTS would be: $t^i = y_{max} - y^i$ if $y^i < y_{max}; t^i = 0$ otherwise, including the individuals from the poorest poor to the richest poor, where y_{max} is the highest cut-off income allowed by the budget. As the anti-poverty budget rises, y_{max} increases up to the poverty line, z, and perfect targeting would permit to lift all the poor out of poverty if enough funding is available.

Analysing closely what supports the results of the r-type and p-type transfer characterisations in B&F reveals that they are based on: (1) first-order conditions that are intuitive in the discrete case, and (2) the cumulative total of transfers to such a level so as to exhaust the available budget. We now translate these intuitions in the continuous case by using the calculus of variations, starting from Euler equations, and we extend them to any objective function that can be written as an integral over the income distribution as in eq. (4). The domains of concavity-convexity of the kernel function and its slopes will guide the re-ranking of incomes in order to sequentially implement the transfers. The sequential treatment of transfers will allow us to avoid the issue of the positivity constraints for transfers.

Differentiating the Euler equations is the key to the interpretation. Indeed, looking at what happens marginally to the Euler equations informs us about what the individual to serve first is if one additional currency unit is available. Meanwhile, it is necessary to check when the sum of transfers hits the budget constraint and to stop the transfers at this stage. To simplify the presentation we assume that the considered distributions are continuous with a well defined density function, implying that Riemann integrals can be employed instead of Lebesgue integrals. We therefore place ourselves in the typical context of the calculus of variations.

Let us forget about z and instead consider a general objective function under the form of a Riemann integral:

$$\int_a^b k(y)dF(y),$$

where a is the lower bound of integration, b is the upper bound, k is a derivable kernel function and F is the cdf of the living standards y. The optimisation program is

$$\max_{t(y)} \quad -\int_a^b k(y + t(y))dF(y)$$

subject to: $\int_a^b t(y)dF(y) = B$ and $t(y) \geq 0$, for all y,

Assume that the transfer function is continuously differentiable. The corresponding Euler necessary condition can be calculated. A convexity condition involving on the kernel function and function $t(.)$ can make the Euler conditions sufficient.

Typically, kernel function k is differentiable in the classical calculus of variations, which implies that it must be continuous. In the case of a poverty index, the kernel is the product of a function $P(y, z)$ and the dummy for the poor $1_{[y+t(y)<z]}$. Although, the latter dummy variable is likely to introduce discontinuities in the kernel function (perhaps far from $y = z$ because of the presence of $t(y)$), this is not necessarily the case if P is continuous at $y = z$ (often $P(z, z) = 0$ for usual poverty indices) and if the transfer function satisfies regularity properties (as $t(z) = 0$ and $t(y)$ continuously monotonous for example). If $t(y)$ is not monotonous, it may be possible to examine separately the intervals where it is increasing or decreasing.

In these simple cases the usual results of the calculus of variations apply and we discuss them now. In order to incorporate the budget constraint in the objective function, we use the following change in variables:

Let y be the 'control variable', and $x(y) \equiv \int_a^y t(u)dF(u)$ be the 'state variable', i.e. the budget spend for individuals of income below the level y. Then, $x(a) = 0$ and $x(b) = B$. By derivation we obtain $\dot{x}(y) \equiv t(y)$. The optimisation program becomes

$$\max_{\{x(y)\}} \quad - \int_a^b k(y + \dot{x})dF(y)$$

subject to $x(a) = 0$ and $x(b) = B$ and $\dot{x} \geq 0$.

The necessary Euler conditions are therefore, forgetting for the moment the positivity constraint: $x(a) = 0$, $x(b) = B$ and $-k_x = \frac{d(k_{\dot{x}})}{dt}$.

Since the kernel function does not depend on x, the later equation can be rewritten as $k_{\dot{x}} = c$ a constant, that is $k'(y + \dot{x}) = k'(y + t(y)) = c$.

A condition of convexity or strict-convexity (for strict optimum) of k is all that is needed to ensure that the Euler conditions are necessary and sufficient (since the argument of k is $y + \dot{x}$).

When instead a discrete optimisation program is considered (as in B&F), with a finite population perfectly observed, the problem can be dealt with Kuhn-Tucker conditions and marginal transfers small enough to avoid re-ranking of incomes. This is not true in the continuous case of which distribution modelling suggests

that the researcher already approximates the true distribution, i.e. that something is not perfectly known in the problem. In a sense this can be considered as a first step in the direction of imperfect information problems.

Let us look at the Euler equation when there are only two individuals $i = 1, 2$. We obtain: $k'(y_1 + t(y_1)) = k'(y_2 + t(y_2))$. All the issues are concentrated in this simple equation. This is the shape of function k that allows us to deduce the priority ranking for serving individuals, and this ranking is the inverse of that the $k'(y)$.

The rule to adopt is therefore the following:

(1) One ranks the $k'(y_i)$, for observed $y_i, i = 1, \ldots, n$;

(2) One identifies the individual i, with income y_i corresponding to the highest $k'(y_i)$, and the following individual j corresponding to the next highest $k'(y_j)$;

(3) One transfers positive amounts of money $t(y^i)$ and $t(y^j)$, respectively to individuals i and j, so that $y^i + t(y^i) = y^j + t(y^j)$. The transfer can start with an amount $y_j - y_i$ given to individual i (the first to be served, with the lower income), and then continue by adding marginal monetary transfers until reaching $y^i + t(y^i) = y^j + t(y^j)$;

(4) the procedure can go on in this fashion, sweeping all the individuals until reaching an individual of income level equal to or just below the poverty line, or until the whole budget has been spend. For objective functions that are not poverty indices, the whole population of individuals can be served if there is enough funding.

For FGT poverty indices with $\alpha > 1$, when the k function is convex ($\alpha > 1$), and therefore the above algorithm is based on the Euler equation, which is sufficient in this case. This is the most relevant case for economists as it gives more importance

to the poorest of the poor. The 'p-type' rule is applied to the case of the head-count index (FGT index with $\alpha = 0$). In that case, 'p-type' transfers just express that it is least costly to start transferring from the poorest of the poor and sweeping up the income distribution, if the aim is merely to reduce the number of the poor.

For the economically most interesting cases (k convex), the above algorithm depends on Euler conditions, including the budget condition. No negative transfer is necessary and the condition of positive transfers is never binding under perfect targeting. We shall show how to implement a similar procedure under imperfect targeting and even with X multivariate.

2.1.1. The imperfect information case

Unfortunately, perfect targeting is not feasible because incomes cannot be perfectly observed. Nevertheless, since the household living standard is correlated with some observable characteristics, it is possible, as in Glewwe (1992), to think about minimizing an expected poverty measure subject to the available budget for transfers and conditioning on these characteristics. In practice, the approach followed in the literature falls short from such a lofty ambition. Practitioners, including Glewwe, design the APTS by merely replacing unobserved living standards with OLS predictions based on observed variables and working with the observed sample as if it was the global population. We shall refine this approach.

Under perfect information, the social planner needs to use observed characteristics X rather than unobserved incomes y to implement the transfers. We therefore consider the following optimisation program.

$$\min_{\{t(X)\}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P\left(z, y + t(X)\right) dF(y, X)$$

$$\text{subject to} \quad : \quad \int_0^{+\infty} t(X)dF_X(X) = B$$

$$\text{and } t(X) \quad \geq \quad 0, \text{ for all } X,$$

where y is the individual income, X is the vector of individual characteristics used in the APTS. P is the kernel function of the objective function, perhaps a poverty index depending on a poverty line z. Function $t(X)$ defines the value of the monetary transfer to an individual of characteristics X. To simplify the notations we eliminate the dependence of this function with z. Typically, the transfers do not directly depend on individual's statements about their income y become such statements are believed to be totally unreliable. Thus, $y+t(X)$ is the post-transfer income. The expectation in the objective to minimise is over the joint distribution of y and X, of which cdf is denoted $F(y, X)$. The marginal pdf of characteristics X is denoted F_X. Three simple but important changes have been performed from the optimisation program characterising perfect targeting: (1) the introduction of correlates X; (2) the incorporation of the poverty line in the kernel function P instead of as an argument of an integral (case $b = z$); (3) the use of the joint cdf of y and X instead of the cdf of y or that of $y + t(y)$.

3.   The Theoretical Solution under Imperfect Targeting

3.1.   The general situation

The previous optimisation problem under imperfect targeting can be transformed by using a change in variable so as to integrate the budget constraint in the objective function. Then, the necessary Euler conditions can be derived. They

lead to integral equations that are implicit in $t(X)$. At this stage several difficulties need to be tackled. First, the transfers must be positive, which potentially implies a large number of inequality constraints in the optimisation program. Second, if several characteristics are used to define vector X, the Euler equations are multidimensional and may be numerically intractable. Finally, to be practically useful the theoretical solution must correspond to a convenient statistical estimation method.

To simplify the notations we dealt with the case where the X is unidimensional and take values from 0 to $+\infty$, as well as y. We start from

$$\max_{\{t(X)\}} \int_0^{+\infty} \int_0^{+\infty} -P(z, y + t(X)) 1_{[y+t(X)<z]} dF(y, X)$$

subject to

$$\int_0^{+\infty} t(X) dF_X(X) = B \text{ and } t(X) \geq 0, \forall X.$$

We now change the variables in order: (1) to adopt the usual notations of the calculus of variations, and (2) to integrate the budget constraint in the objective function. Let $v \equiv X$ be the 'control variable', which is the characteristic used to assess the income level, and $x(v) \equiv \int_0^v t(u) dF_X(u)$ be the 'state variable', here the budget spend for individuals such that the control variable is below the level X. This yields the derivative $\dot{x}(v) = t(v)[= t(X)$, the transfer function with the initial notations].

By conditioning on the control variable, we can define the kernel function in the integral of the objective function as

$f^0(x, \dot{x}, v) \equiv \int_0^{+\infty} -P(z, y + \dot{x}) 1_{[y+\dot{x}<z]} f_{y|v}(y|v) dy . f_v(v)$
$= \int_0^{z-\dot{x}} -P(z, y + \dot{x}) dF_{y|v}(y|v) . f_v(v),$

where $F_{y|v}$ (respectively $f_{y|v}$) is the cdf (respectively density) of y conditioning on v, and f_v is the density of v, assuming the density functions exist and are well defined.

Conditioning on income correlate X is also interesting in that it naturally introduces the general notion of regression of y on X, embodied in the conditional density $F_{y|X}$. In economic applications, characteristics X are bounded (with upper bound $\bar{X}$), and we denote the corresponding upper bound for v as $\bar{v}$.

Then, the optimisation problem can be rewritten as

$$\max_{\{x(.)\}} \int_0^{\bar{v}} f^0(x, \dot{x}, v) dv \tag{5}$$

$$\text{subject to } x(0) = 0 \text{ and } x(+\infty) = B \text{ (the transfer budget constraint)}$$

$$\text{and } \dot{x}(v) \geq 0, \forall v \text{ (the transfer positivity constraint)}$$

Therefore, the poverty minimising problem can be put in the form of a problem of calculus of variations. In favourable situations, a typical necessary condition is the Euler condition, which is $f_x^0(x, \dot{x}, v) = \frac{d\left[f_{\dot{x}}^0(x(v), \dot{x}(v), v)\right]}{dv}$, for all $v \in [0, \bar{v}]$. In our case, f^0 does not depend on x, and the Euler condition simplifies to

$$f_{\dot{x}}^0 = c$$

where c is a constant.

Several difficulties are to be considered at this stage to be able to obtain this Euler condition and use it for applied work. First, f^0 has to be differentiable with respect to $\dot{x}$. Second, in practice correlates X should be multivariate, which leads to a system of Euler equations to solve. Third, the distribution is assumed to be described by a well defined pdf $f_v(v)$, which does not cancel on the interval

of interest, for future convenient discussion of the integration calculus. This may be an issue when some characteristics X are discrete. Fourth, how to deal with positivity constraints is not obvious. Potentially, the latter issue could correspond to a NP-hard optimisation program if the positivity does not straightforwardly result from the considered functional forms. We avoid these difficulties, assuming them away, to simplify the presentation of the issues and focus on the core of the economic problem at hand. We shall deal later with the issue of positivity constraints.

Reverting to our initial notations, the set of equations to solve are as follows, assuming that we are in the case where $P(z, z) = 0$, an usual situation for the poverty indices used in practice:

$$\int_0^{z-t(X)} \frac{\partial P(z, y + t(X))}{\partial y} f_{y|X}(y) dy = -c/f_X(X)$$

$$\int_0^{\bar{X}} t(X) dF_X(X) = B \text{ and } t(X) \geq 0, \forall X.$$

In general these equations in $t(X)$ can only be solved numerically, sometimes with a lot of difficulty, notably if one wants to extend to cases where X is multivariate. In the multivariate case, the first equation corresponds to a gradient vector of dimension the number of considered income correlates, and the second equation incorporates a multivariate integral. We therefore turn to a practical approach involving a sequence of statistical estimations.

4.   The Chain of Statistical Treatment

4.1.   *The practical approach based on predictions of incomes*

Typically in the applied programs, the transfers are based on 'proxy-means tests' that are supposed to identify the poor by some observable characteristics such as geographical location, household size, type of accommodation. Such proxy-means tests are generally calculated by running OLS (Ordinary Least-Squares) regressions of living standards based on the household survey data in order to investigate the household characteristics correlated with poverty. A prediction of the household living standard is then obtained, based on the regressions, that can be compared with a poverty line to assess how poor the considered household is. Then, some assistance would be delivered to households identified as poor and not to others. However, this empirical approach has always lacked clear theoretical basis. We propose a practical approach to link means tests to the poverty minimisation problem.

The optimisation program used to define the empirical transfers corresponds to using the predictions $\hat{y}$ as if they were the actual incomes and apply the above optimisation procedure to derive the transfers $t(\hat{y})$. Thus, we obtain a simple rule to calculate the transfers. Moreover, using $\hat{y}$ instead of X in the definition of transfers helps us to deal easily with the use of multidimensional X and to avoid multivariate Euler equations. These reasons that we elicit justify the use of proxy-means tests as part of an approximate optimisation technique.

In this framework, the algorithm used for perfect targeting can be employed as well with imperfect targeting and multivariate X, provided the living standards y are replaced in the algorithm by predictions $\hat{y}$. Hopefully, if the predictions

are accurate enough, the result of the algorithm should be close to the optimal solution.

If P is the kernel function of a poverty index, incorporating the dummy function of the post-transfer poor $(1_{[y+t(X)<z]})$, the criterion to minimise under imperfect information, where the y have been replaced by their predictors $\hat{y}$ plus the corresponding transfer can be considered as defining a post-transfer poverty *estimator*, $E_{emp}P(\hat{y}+t(y))$, based on an observed sample of individuals with E_{emp} the empirical expectation operator.

In a second step, in order to estimate the impact of the transfers on poverty, the transfers $t(\hat{y})$ are applied to a sample of individuals characterised by the observed incomes y (rather than using the predictions $\hat{y}$ for approximating their income).

In another practical approach adopted in some empirical papers, the positivity of transfers is not imposed a priori. What is done is just calculating the optimal transfers according to a rule and truncating them to zero to eliminate the possibly negative transfers. This is not valid since the allocated budget is overcome in that case. Else, one should explain where do come from the additional resources, or which poor households are not served to meet the budget. In contrast, when the above sequential allocation method is applied based on predictors $\hat{y}$, all transfers are automatically positive or null.

We now discuss the sequence of calculus carried out in the case of an APTS for Tunisia. We discuss the data in Section 5.

4.2. The chain

The chain of treatment applied in the case of Tunisia has four stages:

1. Calculus of living standard indicators:

The living standard indicators are calculated and denoted y^i for household i in the surveyed sample, $i = 1, ..., n^4$. In this paper, we use the per capita nominal total consumption value as our living standard indicator for each household's member.

2. Estimation of the predictions of the living standards: $\hat{y}^i$.

Using the survey data, we replace the observed incomes y^i by statistical predictions based on correlates X^i for the same individual i. Typically OLS are used to generate the predictions. In contrast, we shall also use quantile regressions.

One important ingredient of the practical calculus of the optimal transfer solution is the choice of the statistical prediction method for living standards. The most popular method is that of the Ordinary Least-Squares (OLS) applied to an equation where the living standard variable is a linear combination of correlates. The predictor of living standards is searched in the linear form $X^{i\prime}\beta$ with X^i a vector of correlates for given surveyed household i. This functional form is an additional approximation, although it is little restrictive since polynomials of correlates can be generated. The equation to estimate is therefore $y^i = X^{i\prime}\beta + u^i$, where u^i is a stochastic error term assumed of null expectation, which is ensured by entering a constant term in X^i. The variance of u^i is assumed to be finite. In practice, most of the prediction inaccuracy comes from the unobservable error term u^i. We obtain the following optimisation program that delivers the OLS estimator.

$$\min_{\beta} \ \sum_{i=1}^{n} \left(y^i - X^{i\prime}\beta\right)^2$$

Its solution is easy to calculate: $\hat{\beta} = (X'X)^{-1}X'Y$, where X is the matrix composed of the $X^{i\prime}$ as vectors line, and Y is the vector of which coordinates are the y^i. One problem with the OLS method is that it is likely to predict well the mean living standards (because $E(\hat{y}|X) = E(X\hat{\beta}|X) = \bar{y}$, the population mean of the living standards), but not necessarily the living standards of the poor, nor the living standards of the households close to the poverty line.

Our approach is to use instead quantile regressions for generating the predictions of the living standards. Such method, centered on a given quantile $\theta \in]0, 1[$ ensures that the conditional expectation of the θth quantile of the predictor is $X\beta$ $(E(q_\theta(\hat{y}|X)) = X\beta$, where $q_\theta(.|X)$ is the conditional quantile function centered on the θth quantile. The predictions of living standards around the θth quantile of living standards should be better determined than with OLS. To be able to predict well the living standards of the poor or the near poor, a good choice of parameter θ seems therefore such that the θth quantile of y is close to the poverty line. In that way, the targeting scheme can be said to 'focus on the poor'.

Quantile regression estimates can be obtained as solutions to the following optimisation program.

$$\min_{\beta} \sum_{i=1}^{n} \left(y^i - X^{i\prime}\beta\right) \left[\theta - 1_{[y^i - X^{i\prime}\beta < 0]}\right] \tag{6}$$

where $\theta \in]0, 1[$ is the quantile on which the quantile regression is centered. The numerical solution is obtained by solving a linear programming problem, using the performing algorithms in Barrodale and Robers (1973, 1974). Indeed, the previous optimisation program has a linear programming representation, which is obtained by redefining the predicting regression as: $y^i = \sum_{i=1}^{n} X^{i\prime} \left(\beta^1 - \beta^2\right) + (\varepsilon^i - v^i)$, with

$\beta^1 \geq 0, \beta^2 \geq 0, \varepsilon^i \geq 0, v^i \geq 0$ for all $i = 1, ..., n$:

$$\min_z \; c'z \text{ subject to } Az = y, z \geq 0, \tag{7}$$

where $A = (X, -X, I_n, -I_n), y = (y^1, ..., y^n)', z = (\beta^{1\prime}, \beta^{2\prime}, u', v'),$

$c = (0', 0', \theta l', (1 - \theta)l')', I_n$ is a n-dimensional identity matrix. $0'$ is a $K \times 1$ vector of zeros, and l is an $n \times 1$ vector of ones, K is the number of regressors in X.

The dual problem of the primal problem (7) is approximately the same as the first-order-conditions of the quantile regression optimisation program. It is:

$$\max_w \; w'y \text{ subject to } w'A \leq c' \tag{8}$$

given that matrix X is assumed full column rank, the dual and primal problems have simultaneous feasible solutions.

The numerical solution can be obtained by using simplex iterations after a finite but possibly substantial number of iterations. However, using the improved LP algorithm in Barrodale and Robers (1973), the number of simplex iterations becomes small enough to be useful for typical sample sizes. The algorithm can start with an initial value based on preliminary OLS estimates where the intercept estimate is replaced by the $[n\theta]^{th}$ order statistics of the OLS estimates.

This method has not only the advantage of predicting well the living standard around the quantile θ of the living standard distribution, but also to be robust to outliers. Powell (1983, 1986) and Buchinsky and Hahn (1998) discuss the properties of these estimators.

An interest of focused targeting with living standard predictions based on quantile regressions is that it can be related to the theoretically optimal transfer schemes under perfect information in which the transfers should be first implemented for the poorest of the poor, the richest of the poor, or both (Bourguignon and Field, 1997). Since, what need to be accurately determined are the transfers to these sub-populations, focused predictions for the living standards of the poor and near poor may produce more efficient transfer schemes than using OLS predictions.

3. Calculus of the transfers: $t(\hat{y}^i)$.

To calculate the transfers we solve the following minimisation program of FGT poverty index based on the sample for which y^i and X^i have been jointly observed, and predictions $\hat{y}^i$ have been estimated:

$$\min_{\{t(\hat{y}^i)\}} \ \sum_{i=1}^{n} \left\{ \frac{z - \hat{y}^i - t(\hat{y}^i)}{z} \right\}^{\alpha} 1_{[\hat{y}^i + t(\hat{y}^i) < z]} \tag{9}$$

$$\text{subject to} \quad : \quad \sum_{i=1}^{n} t(\hat{y}^i) = B \quad \text{and } t(\hat{y}^i) \geq 0 \text{ for all } i$$

We use the 'r-type transfer rule' for $\alpha = 0$ and 'p-type transfer rule' for $\alpha = 1$

and $\alpha = 2$. Thus, we obtain an approximate rule for the transfers as a function of X. In the case of unidimensional X, this rule could be described as a correlation curve nonparametrically estimated with the observed sample.

4. Estimation of poverty indicators and targeting efficiency criteria:

The final stage consists of assessing the targeting efficiency of the APTS by using, for example, the following Horwitz-Thompson estimator of FGT poverty indicators:

$$\sum_{i=1}^{n} \frac{S_i}{\pi_i} \left\{ \frac{z - y^i - t(\hat{y}^i)}{z} \right\}^{\alpha} 1_{[y^i + t(\hat{y}^i) < z]},$$

where π_i is the inclusion probability of household i in the sample and S_i is its household size. Thus, levels of poverty reached with different APTS based on the same budget can be estimated and compared. This is our main criterion to judge of the quality of the poverty minimisation process with the APTS. Other criteria are also useful. The *leakage* is the proportion of the budget that could be saved since it is spend on non-poor households or to lift poor households *well* above the poverty line instead of just up to the poverty line. The *exclusion* is the percentage of the poor who do not benefit at all from the APTS.

In the next section, we illustrate our theoretical approach by discussing at an APTS proposed for Tunisia (Muller and Bibi, 2005).

5. Illustration for Tunisian Data

We use data from the 1990 Tunisian consumption survey conducted by the INS (National Statistical Institute of Tunisia). This household survey provides information on expenditures for food and non-food items for 7734 households. Other usual information from household surveys is available such as about education, housing, region of residence, demographic information, and economic activities.

We assume that the per capita total consumption expenditure is an adequate living standard indicator of each household member. The APTS we investigate are based on the following predictors: OLS predictor with geographical dummies used as regressors, quantile regression prediction with geographical dummies, OLS predictor with geographical dummies and information on dwelling and demographic

characteristics as regressors, quantile regression prediction with geographical dummies and information on dwelling and demographic characteristics. We use quantile regressions centered on the first and the third deciles.

The considered budget is the one in force for the food price subsidies in Tunisia at the time of the survey, which corresponds to the main anti-poverty policy in Tunisia. A poverty line of TD 280 per capita per year without subsidies is used, consistently with The World Bank (1995). We compare the implementation of the APTS with the main anti-poverty program in Tunisia, which consists of the food price subsidies.

Table 1 presents estimation results of the APTS for (1) two measures of targeting accuracy (leakage and exclusion), and (2) the impacts on poverty. We first look at the estimation results when only regional dummies are used as correlates (Set I). The results show that typical targeting schemes based on OLS can improve on food subsidies in terms of the number of the poor remaining after the implementation of the program. The percentage of the poor shifts from slightly below 13 percent down to slightly above 10 percent. The transfers based on quantile regressions centered on the third decile provide the best scheme among the considered options, if the aim is to reduce the number of the poor (reaching 10.24 percent), while it remains very close to results with OLS (10.73 percent). In contrast, if the aim is to reduce poverty described by the poverty gap (FGT with $\alpha = 1$) or the poverty severity measure (FGT with $\alpha = 2$) , the scheme leading to the smallest poverty level is based on quantile regressions centered on the first decile. For example, in the case $\alpha = 2$, this APTS leads to an adjusted poverty level equal to 0.65, much better that the level of 1.26 obtained with price subsidies. Moreover, leakages and exclusion are smaller with this method too, except for exclusion with

subsidies that is zero since all households consume subsidized products.

This picture of the APTS efficiency slightly varies when the set of regressors used in the prediction equations is extended (Set II). By taking advantage of information on dwelling and demographic characteristics, substantial improvements can be reached whether in terms of poverty statistics, leakage or exclusion. The quantile regressions centered on the first quantile remain the best approach for reducing FGT with $\alpha = 2$ and exclusion (except for subsidies). For the poverty severity index, the APTS based on OLS regressions reduces poverty down to 0.39, while the APTS based on quantile regressions centered at the first decile yields a lower poverty level of 0.31, a notable improvement. The improvement on exclusion is still more substantial, passing from 21 percent of excluded poor with OLS, with only 10 percent with quantile regressions centered at the first decile.

As it happens, these two criteria may often be considered as the most important ones for poverty specialists. In particular, FGT with $\alpha = 2$ gives a stronger weight to the poorest of the poor, a generally admitted requirement for normatively valid poverty measures. Moreover, exclusion is related to critical political conditions. Indeed, policies leaving aside a large proportion of the poor are unlikely to be politically and socially implementable in Tunisia. However, if the aim is merely to diminish the number of the poor, OLS based transfers would provide better results, while if the aim is to reduce FGT with $\alpha = 1$ or leakage, the quantile regressions centered on the third decile would be preferable.

6. Conclusion

In this paper, we discuss the theoretical difficulties of designing optimal anti-poverty schemes. We first describe the mathematical bases of the optimisation

program defining an anti-poverty scheme. Then, we propose an approximative solution to this problem based on the use of quantile regressions which allow us to 'focus' the estimation of such schemes on the poor and near poor, consistently with theoretical insight stemming from the poverty minimisation program. An illustration based on Tunisian data shows the efficiency gain obtained with such refining of the poverty minimisation procedure.

[Insert Table 1]

REFERENCES

Atkinson, A. B., 1995, "On Targeting Social Security: Theory and Western Experience with Family Benefits." in Dominique van de Walle and Kimberly Nead (eds.), *Public Spending and the Poor, Theory and Evidence.* Pp. 25-68. The Johns Hopkins University Press for the World Bank.

Barrodale, I. and F.D.K. Roberts (1973), "An Improved Algorithm for Discrete l_1 Linear Approximation," *SIAM Journal of Numerical Analysis*, 10:839-48.

Barrodale, I. and F.D.K. Roberts (1974), "Algorithm 478: Solution of an Overdetermined System of Equations in the L_1 Norm," *Communications of the Association for Computing Machinery*, 17, 319-320.

Besley, T. and R. Kanbur, 1988, "Food Subsidies and Poverty alleviation, " *The Economic Journal*, vol. 98: 701-719.

Bigman, D. and P.V. Srinivasan, 2002, "Geographical Targeting of Poverty Alleviation Programs: Methodology and Application in Rural India," *Journal of Policy Modelling*, 24, 237-255.

Bourguignon, F. and G. Fields, "Poverty Measures and Anti-Poverty Policy," *Recherches Economiques de Louvain*, 56 (3-4), 1990

Bourguignon, F. and G. Fields, 1997, "Discontinuous Losses from Poverty, General Ra Measures, and Optimal Transfers to the Poor," *Journal of Public Economics*, vol. 63: 155-175.

Buchinsky, M. and J. Hahn, 1998, "An Alternative Estimator for the Censored Quantile Regression Model," *Econometrica*, Vol. 66, No. 3, 653-671, May.

Chakravaty, S. R. et D. Mukherjee, 1998, "Optimal Subsidy for the Poor," *Economics Letters*, vol. 61: 313-319. North-Holland.

Foster, J., J. Greer, and E. Thorbecke, 1984, "A Class of Decomposable Poverty Measures," *Econometrica*, vol. 52: 761-765.

Glewwe, P., 1992, "Targeting Assistance to the Poor, Efficient Allocation of Transfers when Household Income is not Observed," *Journal of Development Economics*, vol. 38: 297-321.

Glewwe, P. et O. Kanan, 1989, "Targeting Assistance to the Poor: A Multivariate Approach Using Household Survey Data," Policy Planning and Research Working Paper, # 86, Washington, D.C.

Muller, C. and S. Bibi, 2005, "Focused Targeting against Poverty. Evidence from Tunisia," Working paper IVIE.

Parks, A., S. Wang and G. Wu, 2002, "Regional Poverty Targeting in China," *Journal of Public Economics*, 86, 123-153.

Powell, J.L., 1983, "Least Absolute Deviation Estimation for the Censored Regression Model," *Journal of Econometrics*, 51.

Powell, J. L., 1986, "Censored Regression Quantiles," *Journal of Econometrics*, vol 32.

Ravallion, M. and K. Chao, 1989, "Targeted Policies for Poverty Alleviation under Imperfect Information: Algorithms and Applications," *Journal of Policy*

Modeling, vol. 11: 213-224.

Schady, N. R., 2002, "Picking the Poor: Indicators for Geographic Tarteging in Peru," *Review of Income and Wealth*, Serie 48, No. 3, September.

The World Bank, 2000, "A Profile of Poverty in Tunisia 1995," Washington D.C.

Watts, H.W., 1968, "An Economic Definition of Poverty" in: D.P. Moynihan (ed.), *On Understanding Poverty*, 316-29, Basic Book, New York.