Dipartimento di Scienze Economiche
Università degli Studi di Firenze

Working Paper Series

# Evaluating Pro-poor Transfers When Targeting is Weak: The Albanian Ndihma Ekonomike Program Revisited

Lucia Mangiavacchi and Paolo Verme

Working Paper N. 08/2009
September 2009

Dipartimento di Scienze Economiche, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze, Italia
www.dse.unifi.it

Stampato in proprio in Firenze dal Dipartimento Scienze Economiche
(Via delle Pandette 9, 50127 Firenze) nel mese di Settembre 2009,
Esemplare Fuori Commercio Per il Deposito Legale
agli effetti della Legge 15 Aprile 2004, N.106

# Evaluating Pro-poor Transfers When Targeting is Weak: The Albanian *Ndihma Ekonomike* Program Revisited*

Lucia Mangiavacchi† and Paolo Verme‡

July 13, 2009

**Abstract**

The Albanian *Ndihma Ekonomike* is one of the first poverty reduction programs launched in transitional economies. Its record has been judged positively during the recession period of the 1990s and negatively during the more recent growth phase. This paper reconsiders the program using a regression-adjusted matching estimator first suggested by Heckman et al. (1997, 1998) and exploiting discontinuities in program design and targeting failures. We find the program to have a weak targeting capacity and a negative and significant impact on welfare. We also find that recent changes introduced to the program have not improved its performance. An analysis of the distributional impact of treatment based on stochastic dominance theory suggests that our results are robust.

Keywords: Social assistance, Poverty, Impact Evaluation, Albania.

JEL: H53; H72; I32; I38; P35; P36

---

* *Acknowledgments:* The authors wish to thank Emil D. Tesliuc (The World Bank), Vilma Kolpeja (The National Albanian Centre for Social Studies) and Mimi Kodheli (Albanian Ministry of Labor and Social Affairs ) for the detailed information provided on the Ndhima Ekonomike program. We are specially grateful to Martin Ravallion (The World Bank) for his valuable suggestions on evaluation strategy, to Barbara Sianesi (Institute for Fiscal Studies) for clarification on the implementation of psmacht2 and to Luc Behaghel (Paris School of Economics) for useful comments on some econometric aspects. Many thanks also to Luca Piccoli (Paris School of Economics) for his advice on non-parametric approach.

†Department of Economics, University of Florence

‡Department of Economics, University of Turin and School of Management, SDA-Bocconi, Milan. *Corresponding Author*: paolo.verme@unito.it; Via Po 53, 10124 Torino, Italy; Tel: +390116704406; Fax: +390116703895

# 1 Introduction

There is a long-standing debate on the relative merits of cash versus in-kind transfers as instruments of redistribution in the developing world. Classical theory on the welfare state suggests that if the objective of a public good program is reducing poverty, the immediate response is providing cash directly to the poor. Cash transfers would be "Pareto-dominant" to public services because individuals would be able to allocate resources more efficiently. On the other hand, in-kind transfers are preferable for their long-term investment properties and the reduced risk of leakage-use of payments in non-desirable commodities.

In the former command economies of Europe and Central Asia, anti-poverty programs launched during the 1990s in response to the transitional recession were very few and built on a complex system of categorical cash transfers heritage of the socialist past. These programs took the form of cash transfers and were initially devised for the poor.

The focus on the poor constituted a break from the past and emerged as a combination of several factors. First, the transitional recession had increased poverty to unprecedented levels and this required a government response. Second, transitional economies acted under a severe budget constraint and the choice of a restricted number of beneficiaries was essential. And third, these countries worked in the framework of international financial assistance and this assistance was largely earmarked to the poor. Targeting the poor with cash transfers was an almost obliged choice for transitional economies.

Were cash transfers for the poor successful in mitigating the negative consequences of transition on poverty? The answer to this question is mixed. Ravallion et al. (1995) found that the safety net in Hungary was able to protect effectively from poverty but did not play an important role in lifting people out of poverty. Okrasa (1999a, 1999b) found for Poland a general positive impact of social transfers on redistribution, a positive but moderate impact on reducing the poverty spell and a positive impact on exiting poverty. Milanovic (2000) found for Latvia a weak pro-poor role of social protection benefits. Lokshin and Ravallion (2000) analyzed the role of the social safety net in protecting the poor from the 1998 Russian financial crisis and concluded that the social safety net in place was largely insufficient to protect the poor. Van de Walle (2004) tested the public safety net in Vietnam and found a very marginal role of the social safety net in protecting people from poverty or promoting an exit from poverty. Verme (2008) looked at social assistance benefits in Moldova using panel data between 2001 and 2004 and found a non-positive impact on welfare.

All these studies emerged in the context of World Bank assistance to transitional economies and share the feature of evaluating bundles of transfers rather than individual programs. This is evidently a limitation given that only a few cash transfers were specifically designed for the poor. Several of the early evaluations also relied on scarce data resulting in incidence rather than impact evaluations with limited or no consideration of behavioral implications. Moreover, only a handful of countries had pro-poor programs in place at the beginning of the 1990s during the deep recession and only some of these countries maintained these programs during the more recent growth phase. As a consequence, evaluations of pro-poor programs during the recent growth phase are scarce and they do not benefit from benchmark evaluations carried out during the 1990s.

One program that received consistent attention during the recession and growth periods is the *Ndihma Ekonomike* (Economic Support) program in Albania. Case (2001) looked at political factors influencing the local budget allocations for the program during the 1990s and found these factors to be relevant. Alderman (2001, 2002) used a 1996 survey to assess the targeting performance and found that a) targeting was rather good as compared to other poverty reduction programs in developing economies; b) local officials use local information to target the poor not easily captured by household surveys and leading to better targeting and c) poorer jurisdictions are better in targeting the poorer than richer jurisdictions. More recently, Dabalen et al. (2008) have looked at the program and tested the poverty implications as compared to the old-age pension program using the pooled 2002 and 2005 living standards surveys. They find a negative impact of Ndihma Ekonomike on poverty and a higher level of discontent with life with program participants as compared to a control group.

In this paper we return to the 2002 and 2005 surveys but follow a different evaluation strategy to assess and validate the impact of Ndihma Ekonomike on poverty. We consider the 2002 and 2005 surveys separately and exploit a discontinuity in program design occurred during the period to evaluate the impact of these changes on poverty. The treatment effect is estimated using a regression-adjusted matching method first proposed by Heckman et al. (1997, 1998). Exploiting a few distinct features of our data, we are able to meet the basic conditions required by the method and estimate single means differences for both years and the difference-in-differences over the period.

In contrast to Alderman (2001), we find the program to have a very poor targeting performance. However, we find great heterogeneity in targeting performance across local administrations supporting both Case (2001) and Alderman (2002) findings in this respect. We also find a negative and significant effect on poverty for 2002 and 2005 which is in line with Dabalen et al. (2008) findings on the pooled 2002-2005 sample. In addition, we find a non-positive effect for the period 2002-2005 indicating that changes in program design have not improved the performance of the program. Results are robust to adjustments in the outcome variable and to an analysis of the treatment distributions based on stochastic dominance theory.

The paper is organized as follows. Section two provides a description of the program. Section three illustrates the evaluation approach and section four presents the results. Section five concludes.

## 2 The Ndihme Ekonomike Program

*Ndihme Ekonomike* (NE) was introduced in 1993 in response to the economic crisis induced by the transition process and is the only program in Albania targeting specifically the poor.[1] Eligibility to the program is based on means testing and categorical criteria and the program provides cash transfers to selected households on a monthly basis.

When the program was launched it was very large and accounted for about 1.4% of GDP. The economic situation in Albania has improved since but the program continues to be important in size accounting for about 0.4 % of GDP and 10% of government expenditure in 2005 (World Bank 2006) despite a sharp drop in program allocation of about 29% and a subsequent reduction in the number of households beneficiaries from 150,000 in 2002 to 112,000 in 2006. Allocations per household have also decreased between 2002 and 2005 by about 10% (World Bank 2007).

The program design changed on several occasions. NE was originally designed to support urban families without other sources of income and rural families with small land ownership. In 1994 and 1995 the law governing the program was reformed and the program was extended to all poor households. The program was again revised in early 2005 with the replacement of the means-testing formula and a few changes on administrative procedures. NE was the first public service scheme to be decentralized and its administration is now

[1] Details of the program can be found in Kolpeja (2006) and from the Albania Law no. 9355 on Ndihme Ekonomike and social services available from the Albanian Council of Ministers (http://www.mpcs.gov.al/ligje-legjislacioni-social-ligje).

mainly responsibility of municipalities and communes. In this paper we focus on the period 2002-2005 and we are most concerned about the last reform occurred in early 2005.

Application to the program is responsibility of the household. The head of the household files an application form, undergoes an interview at the local NE office and provides a list of documents on the status of the household and its members provided by other state institutions such as the property registry and the employment office. Upon verification of the necessary documentation the household is visited by a social welfare officer who is responsible for drafting a first list of beneficiaries based on personal judgments and on the eligibility criteria established by law.

Eligibility criteria defined by law include categorical "exclusion criteria" and means-tests. Households are excluded from the program if the head of the household is employed or at least one member: 1) owns capital assets with the exception of the living house and agricultural land; 2) is employed or self-employed, except agricultural workers; 3) is unemployed and not registered as job-seeker, with the exception of disabled and agricultural workers; 4) is leaving abroad for any reason except for studying, medical treatment or working for diplomatic offices or international organizations; 5) refuses offers for employment, community work or land if in working age; 6) takes "deliberate actions" aiming to get NE benefit if not eligible. In practice, these criteria aim at excluding those households whose members are likely to have other sources of income and/or exhibit a passive behavior.

The means-testing formula is based on household composition and changed over the period considered. Until 2005, means-tests were based on a formula that computed income thresholds by household as $T = M(0.95H + 0.95E + 0.19W + 0.2375C)$, where $M$ was the national level of unemployment compensation, $H$

referred to the head of household, $E$ was the number of other family members over working age or disabled, $W$ was the number of working age members, and $C$ was the number of household members under working age. In substance, the income threshold was equal to the unemployment benefit per adult equivalent where the equivalence scales were the weights in parenthesis attributed to the different type of household members. An eligible household received a cash transfer equal to the difference between this threshold and actual household income calculated from all sources of income. If the resulting benefit was zero the family was not eligible. The level of the NE benefit was designed to be below incomes generated from unemployment benefit, pension schemes and minimum wage. This was to encourage households to resume work when this became available.

Starting from 2005, a new law regulates program administration. Two major changes have been introduced. The first is that the income threshold is no more linked to unemployment benefit and the second is that the freedom of local officials in granting benefits has been narrowed. The level of benefit that each family can receive now depends on the income threshold computation defined as $T = 2600H + 2600E + 600W + 700C$ where numbers are expressed in local currency (lek). The new law also introduced a lower bound for the transfer at 800 lek, which excludes households previously entitled to a transfer smaller than 800 lek. A maximum transfer of 7000 lek is also established. Moreover, the smaller freedom granted to local officials in assigning benefits reduces *de facto* the capacity of the government to use local information for better targeting, an attractive feature of the program until 2005. Thus, we have an opportunity here to use the discontinuity in program design for evaluating the impact of changes in the means-test and in the freedom of choice granted to local administrators.

In substance and given the characteristics of the program described, we can argue that the key aspects to take into account for selection into the program are: 1) Eligibility based on household income; 2) Employment status of household members; 3) Local heterogeneity in decision making and 4) Urban/rural location (agricultural workers are waived from some of the categorical exclusion criteria for eligibility). These are observable characteristics to prioritize when considering program participation in the evaluation strategy.

The data we dispose of are two rounds of the Albanian Living Standards Measurement Survey (ALSMS), 2002 and 2005. These data contain information on income and cash transfers divided by program as well as sections on labor participation, migration and household assets, allowing us to identify the NE transfer and also recover some of the variables used for eligibility.[2]

Estimates from the two samples are fully comparable. The 2002 and 2005 surveys covered 3,599 and 3,640 households respectively, employed the same questionnaire and the same sampling procedure. Both surveys include a community questionnaire with information on local services and socio-economic conditions[3]. This helps us controlling for community fixed effects and determining the behavioral traits of administrators otherwise unobserved.

# 3 Evaluation Strategy

Let $D = 1$ define individuals treated by the program and $D = 0$ individuals non-treated by the program under study. Let also $Y_1$ be the potential outcome in the treated state and $Y_0$ the potential outcome in the untreated state. We then have two possible potential outcome states for each of the two groups, treated and non-treated. The main parameter of interest in program evaluations is the

[2]Data can be freely downloaded from www.worldbank.org/lsms. The web site also contains information on the questionnaire, variables, sampling procedure and construction of aggregates.

[3]Note that the community questionnaire is not administrated at municipality/communes level, but at a smaller territorial unit such as rural villages or urban blocks.

*Average impact of Treatment on the Treated* (ATT):[4]

$$ATT = E(Y_1 - Y_0 | D = 1) \tag{1}$$

The central problem in program evaluations is that the potential outcomes of the treated $Y_1$ and $Y_0$ cannot be observed simultaneously. We have a missing data problem. We then need an evaluation strategy able to overcome the missing data problem given a set of available data. When the researcher disposes of a random experiment designed *ex-ante*, the treated group can be considered as a representative sample of the population and the estimation of the ATT boils down to the difference between the observed outcome of the treated and the observed outcome of the non-treated in the post-treatment phase.

In our case, we do not dispose of a random experiment and a simple comparison of the post-treatment outcomes of the treated and non treated groups would result in a bias estimate of the ATT. Program participation in NE is based on a number of observable and non observable criteria that self-select into the program only households with certain characteristics and this generates a selection bias. We also do not dispose of a baseline study. The data we have are subsequent to the introduction of the NE program in 1993. In substance, we are confronted with a retrospective evaluation and we need to seek a proper control group before estimating the treatment effect.

As noted by Heckman et al. (1997), critical conditions of non-experimental data are that: (1) Participants and controls have the same distributions of unobserved attributes; (2) The two groups have the same distribution of observed attributes; (3) The same questionnaire is administered to both groups; and

[4]See Rosenbaum and Rubin (1985) or Heckman and Robb (1985). Note that the program evaluation literature has focused mainly on program participants assuming that the indirect effects on non participant are negligible (Todd 2008). This assumption is not always true but generally holds with non-contributive antipoverty program financed by general taxation, which is the case of the NE program.

4) Participants and controls are placed in a common economic environment. Condition (1) is the main problem with non-experimental evaluations and will require some assumptions. Condition (2) can be met with a proper matching procedure while conditions (3) and (4) can be met with a proper choice of data.

In this paper, we use a methodology first proposed by Heckman et al. (1997, 1998) to address condition (2) and we exploit two features of our data to address conditions (3) and (4). Heckman et al. (1997) have also shown that if conditions (2), (3) and (4) are met, the remaining bias may not be a major problem. Below, we discuss more in detail these four conditions and how we address them.

*Selection on unobservables.* In non-experimental studies, condition (1) requires the conditional independence assumption where $Y_0$ and $Y_1$ are independent of $D$ conditional on $X$ - $(Y_0, Y_1) \perp D|X$.[5] If this condition is met, the ATT can be estimated simply comparing participants with non participants. Furthermore, with $P(X) = Pr(D = 1|X)$ and $0 < P(X) < 1$ for all $X$, the ATT is defined for all values of $X$ and experimental and non-experimental evaluations can be said to identify the same parameters. These two assumptions are known as the "strong ignorability" assumptions following Rosembaum and Rubin (1983). In fact, if ATT is the only parameter of interest, it is sufficient for $Y_0 \perp D|X$ to hold given that the ATT measures the impact on the treated only.

Rosembaum and Rubin (1983) also showed that the strong ignorability assumptions imply $Y_0 \perp D|(P(X)$ which suggests that matching can be performed on $P(X)$ rather than on $X$. Based on these findings, Heckman et al. (1998) derived that for the estimation of the ATT is sufficient a weaker identifying assumption described as $E(Y_0|P(X), D = 1) = E(Y_0|P(X), D = 0)$. Now, if we partition the $X$ vector of variables into a vector of variables used in program selection $Z$ and a vector of variables used for the outcome equation $T$ and if we consider the econometric specifications of the outcome vari-

[5] The symbol '$\perp$' in this paper stands for 'independence'.

able ($Y_{(.)} = \beta X_{(.)} + U_{(.)}$), we can re-write the basic matching assumptions in terms of residuals as $E(U_0|T, Z, D) = E(U_0|, Z, D)$ and $E(U_0|P(Z), D = 1) = E(U_0|P(Z), D = 0)$ as it is done with similar additively separable models in econometrics. These are weaker assumptions than the strong ignorability assumptions and they can be used to construct alternative matching estimators.

*Selection on observables.* The question of selection on observables is generally addressed with a process of matching where a comparison group for the treated is constructed from a group of non treated based on common observed characteristics. Following from the discussion above, in this paper we use the Regression-Adjusted Matching Estimator (RAME) formally justified in Heckman et al. (1998) and tested in Heckman et al. (1997).

RAME consists of estimating matched outcomes for the treatment group combining a local linear matching on the covariates of eligibility with a regression-adjustment on the covariates of outcome. More in detail, the procedure we follow implies the following steps: 1) Estimation of a probit participation equation using a set of selection variables $Z$; 2) Estimation of the predicted values of participation and creation of the corresponding variable ("pscore"); 3) Estimation of a standard OLS welfare regression using a set of non selection variables $T$; 4) Estimation of the residuals of the welfare equation and creation of the corresponding variable ("res"); 5) Matching treated and non treated groups with a local linear matching estimator and using "res" as outcome variables and "pscore" as propensity scores; 6) Estimate of the single mean difference in outcomes between treated and matched group.

The matching procedure is based on a local linear regression which uses and weighs all the comparison group observations. This procedure has several advantages. It is possible to use more information and achieve a lower variance than methods based on selected observations since all the comparison group observations on common support are included. A local polynomial regression instead of a standard kernel offers a greater robustness to different data design densities and has a faster rate of convergence near boundary points (Fan, 1992). This is a clear advantage given that a large part of our data is concentrated at boundaries. Moreover, according to Caliendo (2008) local linear regression is expected to perform better than kernel estimation when the nonparticipants observations on $P(Z_i)$ fall on one side of the participant observations, which is the case of the propensity score distribution estimated by our participation equation. Finally, nonparametric methods characterize better than traditional matching methods the form of evaluation bias, since they estimate more precisely the function of the dependent variable.

The local linear matching estimator is defined as:

$$\hat{\alpha} = \frac{1}{n_1} \sum_{i \epsilon I_1 \cap S_p} [U_{1i} - \sum_{j \epsilon I_0 \cap S_p} W(i,j) U_{0j}] \tag{2}$$

where $I_1$ is the set of participants, $I_0$ the set of non-participants, $S_p$ is the region of common support and $n_1$ is the number of individuals in the set $I_1 \cap S_p$. The match of each participant is constructed as a weighted average over the outcomes of non-participants where $W(i,j)$ is computed by a local linear weighting function on the distance between $P_i$ and $P_j$ (see also Todd, 2008):

$$W(i,j) = \frac{G_{ij} \sum_{k \epsilon I_0} G_{ik}(P_k - P_i)^2 - [G_{ij}(P_j - P_i)][\sum_{k \epsilon I_0} G_{ik}(P_k - P_i)]}{\sum_{j \epsilon I_0} G_{ij} \sum G_{ij}(P_k - P_i)^2 - (\sum_{k \epsilon I_0} G_{ik}(P_k - P_i))^2} \tag{3}$$

A fixed bandwith of 0.06 and a biweight kernel ($G(.)$) are used for the estimator. We impose a common support condition because $S_p$ needs to be determined to compute . Moreover, to ensure that the propensity score density under the common support is strictly positive, we apply a trimming procedure excluding any $P$ point for which the estimated density is zero and the two percent of the remaining $P$ points for which the estimated density is positive but relatively small.

Based on the $ATT$ estimations for 2002 and 2005, we can then estimate the difference-in-differences (DID) across the two years to capture the impact of changes in program design. Heckman et al. (1997 and 1998) have shown that with panel or repeated cross-section data it is possible to adopt weaker conditional independence assumptions using a difference-in-differences estimator of the type $DID = E(Y_{1t} - Y_{0t'}|X, T = 1) - E(Y_{0t} - Y_{0t'}|X, T = 0)$, where $t$ and $t'$ represent time after and before treatment respectively. In fact, it is sufficient for $E(Y_{0t} - Y_{0t'}|X, T = 1) = E(Y_{0t} - Y_{0t'}|X, T = 0)$ to hold. Under additive separability and index sufficiency, this condition becomes $E(U_{0t} - U_{0t'}|P(Z), T = 1) = E(U_{0t} - U_{0t'}|P(Z), T = 0)$. In substance, the DID estimator does not require $E(U_0|X, D = 0)$ and allows for unobservable but time-invariant differences in outcomes between matched NE beneficiaries and non-beneficiaries. The DID is estimated as:

$$\hat{\alpha}_{DID} = \left\{ \frac{1}{n_{1t}} \sum_{i \epsilon I_{1t} \cap S_p} \left[ U_{1ti} - \sum_{j \epsilon I_{0t} \cap S_p} W(i,j) U_{0tj} \right] \right\} - \left\{ \frac{1}{n_{1t'}} \sum_{i \epsilon I_{1t'} \cap S_p} \left[ U_{1t'i} - \sum_{j \epsilon I_{0t'} \cap S_p} W(i,j) U_{0tj} \right] \right\} \quad (4)$$

We use this estimation to evaluate the marginal impact of the policy intervention occurred between 2002 and 2005. Note that between 2002 and 2005 Albania experienced rapid growth and poverty reduction. With the DID matching we can isolate the impact of the program from the impact of growth since we will perform a matching for both years, comparing individuals equally affected by economic growth.

*Common questionnaire.* We will estimate counterfactual outcomes from the comparison group of non treated individuals found within the same survey used to observe the treated group. This ensures that the questionnaire administered to both groups is the same, which satisfies condition (3). Note also that the questionnaire is the same for the two years considered.

The problem of this choice is that finding good matches of the treated in the pool of non-treated may be difficult due to self-selection. However, a combination of factors specific to our data ensures that this is not the case. Among the pool of non treated individuals it is common to find eligible households who did not apply to the program and eligible households who applied to the program but were rejected. According to Kolpeja (2006): *"The number of applicants for NE is much higher than those who receive the benefit. Some estimations indicate that about 30-35 percent of applications are rejected. The reasons for the refusal of NE benefit are: a) incompatibility with (eligibility) criteria (about 5 percent), insufficient funds (15-20 percent), and c) provision of false information (10 percent)."* We also find in the pool of treated non eligible households who were selected. In substance, program leakage and under-coverage (documented further in the paper) ensure that among the treated and non treated groups we can find comparable households. Indeed, we will see that our matching procedure will achieve full common support.

*Common labor market.* This condition is addressed by controlling for local areas using a territorial dummy variable which ensures that matching takes into account the local economic environment. The territorial variable selected is the district (*rrethe* in Albanian). Albania is a small country of about 28,000 squared kilometers subdivided into 36 districts. We judged the average district to be of reasonable size to represent local labour markets. Smaller territorial units were also difficult to use in the regressions due to sample size. In addition, the participation equation includes a dummy for urban and rural areas capturing the different features of urban and rural labour markets.

*Key variables.* Our objective is to measure the welfare improving capacity of the NE program and our outcome variable is a measure of welfare. We opted to use household expenditure per capita normalized by an absolute poverty line, which is a standard practice in similar studies (Ravallion et al. 1995; van de Walle 2003). The consumption aggregate we use has been elaborated by the World Bank, includes food, clothings, household articles, utilities, education and durables and is computed in the same way for the two years considered.

The treatment group $D = 1$ is identified with a treatment indicator variable for households receiving benefits (the survey reports the last NE payment received and the referring period). The comparison group includes all non treated households on common support weighted with the matching procedure already described.

To reproduce the assignment process and define the $Z$ vector of variables, we constructed dummies for eligibility based on the 2002 and 2005 means-test formulae and dummies for the exclusion criteria already described. We were able to reconstruct from data four of the six exclusion criteria and two of these have been retained in the final specification of the selection equations. The first variable is employment of any household member[6] in the formal sector where the formal sector is identified with the variable that captures individuals who contribute to social security. This proxies the employment exclusions criteria and makes sure that we capture only those households whose employment status is likely to be observed by the program administrators. The second variable captures households with at least one member unemployed and not seeking work.

To take into account the freedom of choice attributed to local administrators in selecting participants, we constructed a targeting coefficient for each of the 36 Albanian districts following a methodology proposed by Galasso and Ravallion

[6] With the exception of self-employed in rural areas.

(2005). The targeting coefficient measures the difference between the proportions of the poor and non-poor households receiving the transfer and varies between '1' (perfect targeting) and '-1' (perfect leakage). We split this variable into three quantiles and used dummies for two of these quantiles as regressors in the selection equation.

We also add in the participation equation a dummy variable for urban and rural areas. This has two major advantages. First, we expect urban and rural residents to have different information and opportunities about the NE program. And we also know from the categorical exclusion criteria that rural residents are not covered by some of these criteria. These two factors imply that urban and rural residency can be considered as an important determinant of participation. We preferred this option to splitting the sample into urban and rural areas to rely on as many observations as possible in the matching procedure and reduce the problem of dimensionality.

In substance, we are able to capture all four major factors that determine program participation as described in section 2. The capacity to predict participation of the probit models is estimated with the hit or miss method. The method classifies observations as '1' if the estimated propensity score is larger than the sample proportion of the treated and '0' otherwise.

The $T$ vector of variables selected for the outcome equation includes characteristics of the head of the household (age, health and education), household characteristics (dummies for number of children according to age) and community variables (presence of educational, health and financial institutions). Note that employment status variables are included into the participation equation and are excluded from the outcome equation.

# 4 Results

If we limit our analysis to the comparison of welfare with and without treatment, we find that the incidence of Ndihma Ekonomike on poverty is relevant. Table 1 shows that the poverty headcount index and the poverty gap index in 2002 would have been 1% and 0.6% higher respectively in the absence of the program. Such incidence increases in 2005 for the poverty headcount ratio to about 1.2% and decreases for the poverty gap ratio to about 0.4%. In the absence of behavioral considerations, the Ndihma Ekonomike program would appear to have a positive effect on poverty (Table 1).

[Table 1]

The overall targeting capacity of the poor is weak (Table 2). The program covers a considerable share of the population (11% in 2002 and 12% in 2005) but undercoverage and leakage rates (Cornia and Steward 1995) have been very high in both years considered. In 2002, about three quarters of the poor were not targeted and 57% of the households treated by the program were non poor. The Galasso and Ravallion (2005) targeting coefficient also indicates that the targeting capacity of the program is very low.

The targeting performance over time is mixed. If we compare our results with those of Alderman (2001), which refer to a survey carried out in 1996, we find that targeting has worsened.[7] Figure 1 shows that the targeting curve by decile was steeper in 1996 as compared to 2002 and 2005 indicating that the share of NE expenditure going to lower deciles was higher than the share going to upper deciles in 1996 as compared to subsequent periods.[8] Coverage and undercoverage rates and the targeting coefficient improved between 2002 and

[7] The survey used by Alderman (2001) is a different survey from those we use but both sets of surveys are nationally representative and we have reconstructed the same consumption indicator used by Alderman.

[8] Consumption for all years is net of NE benefits.

2005 but this has been accompanied by an increase in leakage and a decrease in adequacy (Table 2).[9] Figure 1 also shows that the share of NE expenditure going to the poor has marginally decreased between 2002 and 2005 especially for the third decile. In section two we noted that NE expenditure during this last period has declined by about 29% and here we find that this decline has not been pro-poor. In other words, between 2002 and 2005 improvements in coverage have been achieved at the expenses of leakage and adequacy. The program has been able to capture more poor households but expenditure per capita has got thiner overall and marginally thiner for poor households.

The targeting performance of the program may be explained in terms of several factors. First, funds may be misallocated with insufficient funds reaching poor areas and excessive funds reaching rich areas. The central NE budget allocation mechanism to local administrations determines *ex-ante* the funds available for local areas. Case (2001) found that political constituencies were an important factor in explaining budget allocations and Kolpeja (2006) has noticed that 15-20% of applications rejected are because of lack of funds. These two findings could explain a bias allocation of funds in favor of richer areas. Such problems are generally difficult to address but can be improved if the design of the budget allocation criteria are demanded to an independent body.

Second, the targeting mechanism in place may not be able to target the poor efficiently, even if perfectly implemented. Means-testing is only one of the criteria used to select households, selection is based on income rather than consumption and the program has no proxy-means tests in place. Program administrators do not have the same information available in surveys to measure poverty and this may partly explain the targeting ratios which we estimate with surveys data on consumption. This problem can be addressed by introducing

[9] Our results on coverage, leakage and targeting coefficient coincide with those published in World Bank (2007).

proxy-means tests based on household surveys to complement or replace the means-test formula.

Third, administrators may not be able to apply the targeting mechanism properly. This may be due to supply side reasons such as difficulties in administrative procedures, collection of documents or misbehavior on the part of administrators or demand side reasons such as fraudulent behavior or lack of information on the part of clients. Alderman (2002) found that the information available to local administrators improved the targeting capacity of the program. World Bank (2007) decomposed the targeting coefficient reported in Table 2 into intra-commune and inter-commune components and found that two thirds of the targeting coefficient is explained by the intra-commune component. The performance of program administrators within communes seems to be more relevant than differences across communes partly explained by factors such as different funding levels. The 2005 program reform reduced the freedom of choice of local administrators. This may be a good or bad factor depending on how good local administrators were in the first place. Our results indicate an improvement in the targeting coefficient between 2002 and 2005 together with a growth in leakage and a reduction in adequacy, a rather mixed picture. Nevertheless, the targeting capacity of administrators can be improved with a combination of training, public information campaigns and anti-corruption measures.

Fourth, targeting during a recession phase may be different from targeting during a growth phase. During a recession public resources are scarcer while poverty is widespread. With more poor it is easier to catch the poor although transfers may be low. Different is the outlook during a growth phase. With more money and less poverty it is easier to spread money around increasing coverage and leakage at the same time. Albania acted counter-cyclically with a 29%

drop in NE program allocations in real terms between 2000 and 2006 (World Bank, 2007) and achieved higher coverage and leakage by reducing average transfers per household. The expenditure reduction may be partly explained by a reduction in needs and applications to the program during the growth phase but the reduction in expenditure per household is hardly a pro-poor policy. This is another aspect of the program that can be improved.

[Figure 1 and Table 2]

Despite the weak targeting performance, was the program able to improve on the living conditions of those targeted? Our results suggests that the program had a negative effect on welfare in 2002 and 2005 and that the performance of the program worsened over the period. In what follows we discuss first the building blocks of the RAME method proposed including the probit participation equation, the OLS outcome equation and the ability of the matching procedure to reduce selection bias on observables. We then report the single means estimates of the treatment effect for 2002 and 2005 and the difference-in-differences estimate for the period 2002-2005. Last, we test the robustness of our results using a different outcome indicator and assessing the distribution of the treatment effect based on stochastic dominance theory.

The probit selection equation (Table 3) shows that means-tests and some selected categorical criteria contribute significantly to selection into the program. As expected, the coefficient for the dummy variable constructed for those households with an income below the income threshold determined by law is positive and significant for both years.

The employment exclusion restriction is negative and significant as we should expect. Households with at least one household member employed or self-employed are less likely to participate to the program. The dummy for households with members unemployed and not job seeking is instead non significant in both years. This is perhaps due to the fact that it may be difficult for program administrators to observe this household attribute with accuracy. This variable was nevertheless maintained in the final specification because contributes to keep matching results much more stable.

The variables capturing the district ability to target households are both significant and with the expected sign. Households living in districts with a bad targeting record are less likely to be selected into the program than households living in districts with a good targeting record, other selection criteria being equal. The dummy for urban areas is also significant but with a positive sign in 2002 and a negative sign in 2005. We expected this variable to be significant but to have a consistent sign over the period. This is not the case which would suggest that changes in program design have been in favor of rural households. Indeed, World Bank (2007) found that improvements in coverage observed between 2002 and 2005 are almost entirely explained by improvements in coverage in rural areas. Central and local administrators seem to have put a major effort in improving conditions in rural areas and this has shifted the balance between urban and rural areas.

The participation prediction capacity of the probit models based on the hit or miss method are around 76% for both years, which are rather good scores considering that not all eligibility criteria could be used.

[Table 3]

The OLS model (Table 4) has a fairly good explanatory power as compared to models of this kind, also considering that the program eligibility variables are excluded. The model explains about 31% of the variance of welfare in 2002 and about 27% in 2005. Significant variables in both years are health and higher education of the head of the household (both with positive signs) and the number and age of children in the family (always negative).

[Table 4]

In Table 5 we report the estimations of single and double differences. Program treatment seems to have a negative effect on welfare.[10] Both single differences for 2002 and 2005 show negative and significant values. The average treatment effect for 2002 is estimated at about 16.6% of the poverty line. This negative effect rises to 25.7% in 2005 resulting in a negative effect also for the period 2002-2005. Single differences are significant in both years at the 1% level while the double difference is non significant.

Table 5 also includes results using the OECD equivalence scale.[11] Our previous results are based on an outcome variable that measures consumption per capita relative to the poverty line. Poverty studies often use consumption per adult equivalent to take into account household composition in addition to household size and poverty ratios are known to be very sensitive to the use or non use of equivalence scales. However, when we use our measure of welfare per adult equivalent we find that the treatment effect is still negative and significant at the 1% level for both years considered. The difference-in-differences is also negative but non significant as before. As compared to the use of consumption per capita, single and double differences for the adult equivalent measure show much higher values, -23.3% of the poverty line in 2002 and -38.2% in 2005.

[Table 5]

In Tables 6 and 7 we test the capacity of the matching procedure described to reduce the bias between treated and control groups based on the observed participation variables $Z$ used in the probit selection equation. For both years, we obtain full common support with no observations falling out and the matching

[10] Single means difference and respective standard errors are estimated with the Stata module psmatch2 (Leuven and Sianesi, 2003). Bootstrapped standard errors were also estimated but the difference with the standard errors reported in the table is negligible.

[11] We use the OECD original scale attributing a weight of one to the first adult in the household, 0.7 to other adults and 0.5 to children.

procedure almost eliminates the bias on observables. In 2002, the percentage in bias reduction is in between 75.7% and 99.5% depending on the variable considered. In 2005, these values vary in between 75% and 99.7%. For none of the two years the means tests between treated and controls are significant.

In substance, we have been able to reduce very significantly the bias arising from non-overlapping support and the bias arising from differences in observables. Given the use of a common questionnaire for treated and untreated groups and considering the use of local fixed effects, the remaining bias arising from differences in unobservables should be small (as the experiment in Heckman et al.,1997, would suggest).

[Tables 5 and 6]

As a final test, we exploit stochastic dominance theory to assess the distributional impact of treatment. Stochastic dominance of first degree can be assessed by comparing the cumulative distribution functions (CDFs) of the outcome variable for the treated and control groups.[12]. This is equivalent to test our results for all reasonable poverty lines. In Figure 2, we compare the CDFs for both years using consumption per capita and consumption per adult equivalent as we did in Table 5. As it can be seen, the CDFs for the control groups always dominate the CDFs for the treated groups in all four quadrants. It is also evident that, for both outcome variables used, dominance of the control group increases over the period. Overall, irrespective of the poverty line and of equivalence scales, treatment has always a negative and significant effect on consumption and this negative effect increases over the period.

[Figure 2]

[12]See Foster and Shorrocks (1988) and Abadie (2002).

# 5 Conclusion

The paper evaluated the poverty reduction capacity of the Ndihma Ekonomike program in Albania. The program is one of the earliest poverty reduction program implemented in transitional economies and had a positive record in terms of targeting during the 1990s (Adelrman, 2001 and 2002). More recently, the program was found to have a negative effect on poverty and life satisfaction (Dabalen et. al., 2008).

We find the targeting performance of the program to be weak and to have worsened as compared to the 1990s. Between 2002 and 2005 coverage has improved, especially in rural areas, but the average benefit per household has decreased (especially for the poor) together with an increase in leakage. This explains a decline in the overall budget share reaching the poor. Both undercoverage and leakage rates remain very high by any standard. Weak targeting may be explained by various factors including central budget allocation mechanisms, the design of the targeting methodology, the behavior of clients and administrators and the business cycle. All these factors are probably at work.

Making use of a regression-adjusted matching estimator first proposed by Heckman et al. (1997, 1998), we find Ndihma Ekonomike to have a negative and significant effect on household welfare in 2002 and 2005. Changes in program design between 2002 and 2005 seem to have worked in favor of rural households but, overall, the negative impact has increased. The estimated difference-in-differences between 2002 and 2005 is also negative, although non significant. Results seem to be robust. Using per adult equivalent welfare instead of per capita welfare increases marginally the negative impact. Testing stochastic dominance of first degree comparing the cumulative distribution functions of the outcome variables for the treated and control groups shows that the control groups dominate invariably the treated group all along the curves.

The natural implications of these findings is that Ndihma Ekonomike should be further revised. Possible reforms include the shift of the budget allocation decisions to an independent body, the redesign of the targeting mechanism with the introduction of proxy-means test and anti-corruption measures combined with public information campaigns and training. A viable option would be to discontinue the program and replace it with a new program. This would allow to redesign the program altogether and to evaluate its performance with a randomized experiment. This paper exploited the poor targeting performance to the advantage of the matching procedure with *ex-post* data but this is a second best solution to the evaluation of a properly targeted program with a randomized experiment.

# References


Abadie, A. 2002. "Bootstrap tests for distributional treatment effects in instrumental variable models." *Journal of the American Statistical Association* 97: 284-292.

Alderman, H. 2001."Multi-tier targeting of social assistance: The role of inter-governmental transfers." *The World Bank Economic Review* 15: 33-53.

Alderman, H. 2002. "Do local officials know something we don't? Decentralization of targeted transfers in Albania." *Journal of Public Economics* 83: 375-404.

Caliendo, M. and S. Kopeinig. 2008. "Some practical guidance for the implementation of propensity score matching." *Journal of Economic Surveys* 22: 31-72.

Case, A. 2001. "Election goals and income redistribution: Recent evidence from Albania." *European Economic Review* 45: 405-423.

Cornia, G. and F. Steward. 1995. "Two errors of targeting." *In Public spending and the poor*, ed. D. van de Walle and K. Nead. Johns Hopkins University Press.

Dabalen, A., Kilic, T. and Wane, W. 2008. "Social transfers, labor supply and poverty reduction. The case of albania." World Bank Policy Research Working Paper no. 4783, World Bank, Washington, DC.

Fan, J. 1992. "Local linear regression smoothers and their minimax efficiencies." *The Annals of Statistics* 21: 196-216.

Foster, J. and A. Shorrocks. 1988. "Poverty orderings". *Econometrica* 56: 173-177.

Galasso, E. and M. Ravallion. 2005. "Decentralized targeting of an antipoverty program." *Journal of Public Economics* 89: 705-727.

Heckman, J., H. Ichimura, and P. Todd. 1998. "Matching as an econometric evaluation estimator." *Review of Economic studies* 65: 261-294.

Heckman, J., H. Ichimura, and P. Todd. 1997. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *Review of Economic Studies* 64: 605-54.

Heckman, J. and Robb, R. 1985. "Alternative Methods For Evaluating The Impact of Interventions", in *Longitudinal Analysis of Labor Market Data*, ed. Heckman, J. and Singer, B. New York: Wiley.

Kolpeja, V. 2006. "Program implementation matters for targeting performance: Evidence and lessons from eastern and central europe." Unpublished manuscript, World Bank.

Leuven, E. and B. Sianesi. 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." http://ideas.repec.org/c/boc/bocode/s432001.html.

Lokshin, M., and M. Ravallion. 2000. "Welfare impacts of the 1998 financial crisis in Russia and the response of the public safety net." *The Economics of Transition* 8: 269-295.

Milanovic, B. 2000. "Social transfers and social assistance: An empirical analysis using Latvian household survey data." World Bank Policy Research Working Paper no. 2328, World Bank, Washington, DC.

Okrasa, W. 1999. "The dynamics of poverty and the effectiveness of Poland's safety net (1993-1996)." World Bank Policy Research Working Paper no. 2221, World Bank, Washington, DC.

Okrasa, W. 1999b. "Who avoids and who escapes from poverty during the transition? Evidence from polish panel data, 1993-96." World Bank Policy Research Working Paper no. 2218, World Bank, Washington, DC.

Ravallion, M., D. van de Walle, and M. Gautam. 1995. "Testing a social safety net." *Journal of Public Economics* 57 (2): 175-199.

Rosenbaum, P and D. B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70: 41-55.

Rosenbaum, P and D. B. Rubin. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *American Statistician* 39: 38-39.

Todd, P. 2008. "Evaluating social programs with endogenous program placement and selection of the treated." In *Handbook of development economics Vol.4.*, ed. T. P. Schultz and J. Strauss. North Holland.

van de Walle, D. 2003. "Are returns to investment lower for the poor? Human and physical capital interactions in rural Vietnam." *Review of Development Economics* 7: 636-653.

Verme, P. 2008. "Social Assistance and Poverty Reduction in Moldova 2001-2004. An Impact Evaluation." World Bank Policy Research Working Paper no. 4658, World Bank, Washington, DC.

World Bank. 2006. "Albania: restructuring public expenditure to sustain growth." Report no. 36543 - AL, World Bank, Washington, DC.

World Bank. 2007. "Albania: Urban Growth, Migration and Poverty Reduction. A Poverty Assessment." Report no. 40071- AL, World Bank, Washington, DC.

**Figure 1 - NE Expenditure Share by Consumption Decile**

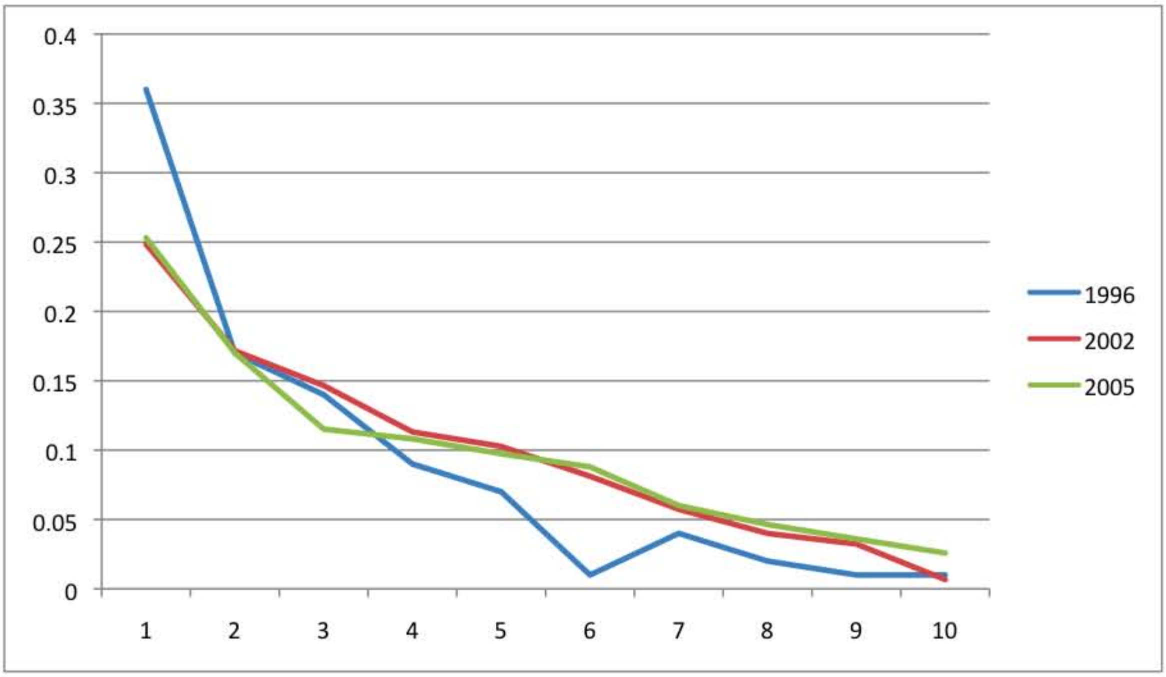

**Figure 2 - Stochastic Dominance of First Degree (CDFs)**

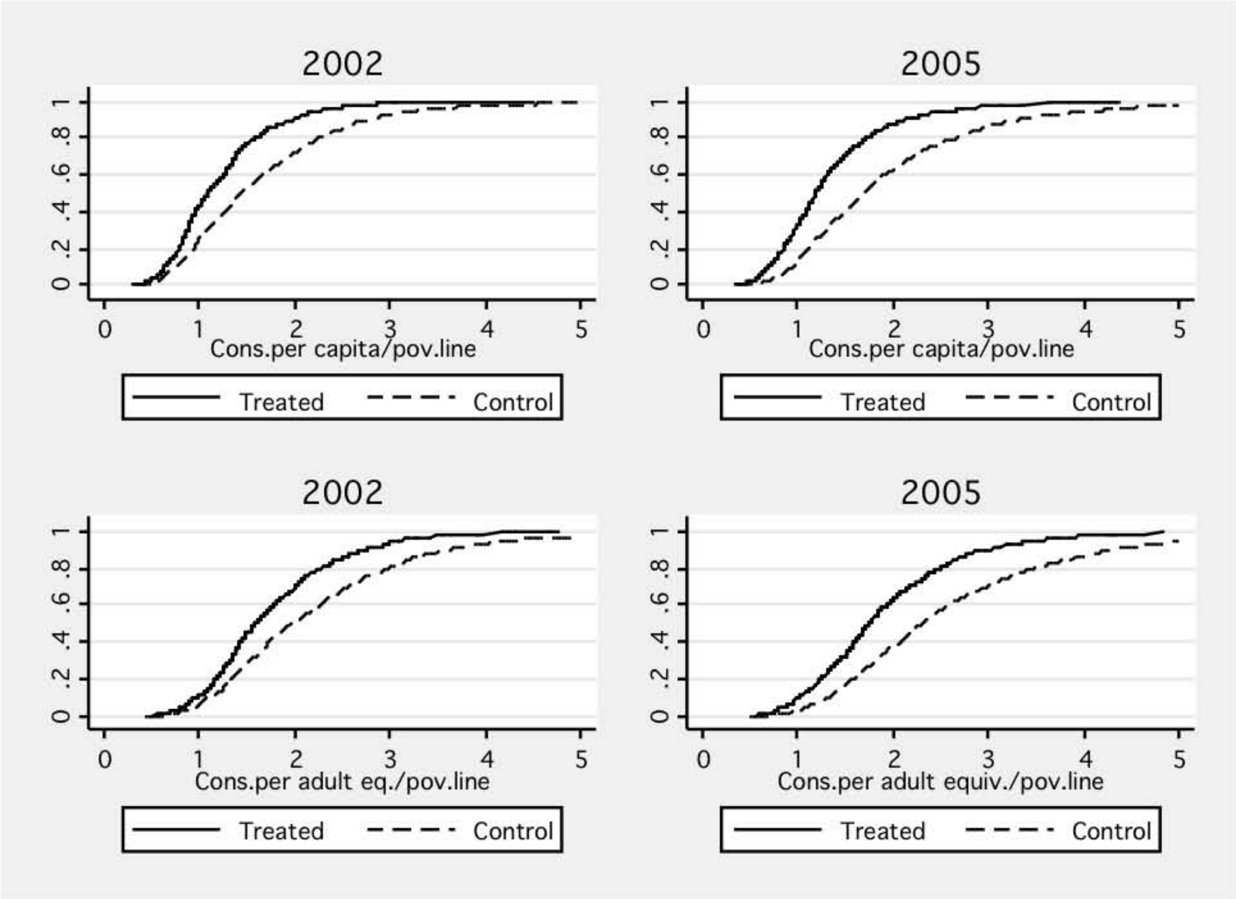

**Table 1**

**Poverty Incidence (%)**

| | 2002 | | 2005 | |
|---|---|---|---|---|
| | **With NE** | **Without NE** | **With NE** | **Without NE** |
| **Households** | | | | |
| Headcount Ratio | 19.1 | 20.1 | 14.0 | 15.2 |
| Poverty Gap | 4.0 | 4.6 | 3.0 | 3.4 |
| | | | | |
| **Individuals** | | | | |
| Headcount Ratio | 24.4 | 25.4 | 17.7 | 19.0 |
| Poverty Gap | 5.4 | 6.0 | 3.8 | 4.3 |

**Table 2**

**Coverage and Targeting**

| | | 2002 | 2005 |
|---|---|---|---|
| 1 | Coverage | 11.0 | 12.7 |
| 2 | Adequacy | 10.3 | 9.3 |
| 3 | Undercoverage | 75.4 | 67.5 |
| 4 | Leakage | 57.3 | 64.2 |
| 5 | Targeting Coefficients | 0.17 | 0.23 |

1 hh treated/population
2 av.transfer/av.consumption
3 poor not treated/tot. poor
4 non poor treated/tot. treated
5 (poor treated/tot. poor)-(non poor treated/tot. non poor)

**Table 3**

**Probit Regression for Program Participation**

| | 2002 | 2005 |
|---|---|---|
| Dummy for hh with income below means-testing threshold | 0.891*** | 1.070*** |
| | (0.0765) | (0.0947) |
| Dummy for hh with members employed or self-employed | -1.173*** | -0.903*** |
| | (0.147) | (0.131) |
| Dummy for hh with members unemployed and not seeking work | 0.0286 | 0.108 |
| | (0.0946) | (0.0792) |
| District targeting coefficient (lower quantile, poor targeting)* | -0.478*** | -0.605*** |
| | (0.102) | (0.0905) |
| District targeting coefficient (upper quantile, good targeting)* | 0.436*** | 0.308*** |
| | (0.0848) | (0.0748) |
| Dummy for urban areas | 0.447*** | -0.292*** |
| | (0.0759) | (0.0675) |
| Constant | -1.704*** | -1.062*** |
| | (0.0893) | (0.0700) |
| Observations | 3599 | 3638 |
| Prediction capacity (hit or miss method, %) | 76.95 | 75.87 |

(*) Base category is intermediate targeting coefficient (middle quantile of three quantiles)

Robust standard error in parentheses. (*) Significant at 10%; (**) Significant at 5%; (***) Significant at 1%.

**Table 4**
**OLS Welfare Equations**

| | 2002 | 2005 |
|---|---|---|
| Age of the hh head | 0.00391 | -0.0144 |
| | (0.00799) | (0.0126) |
| Age of the hh head (squared) | -0.000054 | 0.000105 |
| | (0.000075) | (0.000112) |
| Head (1) is in good health | 0.103*** | 0.118** |
| | (0.0377) | (0.0592) |
| HH head-primary school | -0.00137 | -0.171* |
| | (0.0631) | (0.0907) |
| HH head-two years vocational | 0.310*** | 0.250** |
| | (0.0881) | (0.108) |
| HH head-five years vocational | 0.0840 | 0.0370 |
| | (0.104) | (0.124) |
| HH head-general secondary | 0.283*** | 0.283*** |
| | (0.0727) | (0.108) |
| HH head-university degree | 0.781*** | 0.955*** |
| | (0.0869) | (0.128) |
| HH head-postgraduate | 0.974*** | 1.783*** |
| | (0.250) | (0.476) |
| HH has 1 under five child | -0.413*** | -0.419*** |
| | (0.0357) | (0.0594) |
| HH has 2 under five children | -0.767*** | -0.876*** |
| | (0.0535) | (0.0621) |
| HH has 3 or more under five children | -0.964*** | -0.922*** |
| | (0.163) | (0.116) |
| HH has 1 child (6-18) | -0.413*** | -0.312*** |
| | (0.0445) | (0.0525) |
| HH has 2 children (6-18) | -0.657*** | -0.536*** |
| | (0.0472) | (0.0574) |
| HH has 3 children (6-18) | -0.917*** | -0.768*** |
| | (0.0561) | (0.0684) |
| HH has 4 or more children (6-18) | -1.092*** | -1.012*** |
| | (0.0652) | (0.0693) |
| Pre-school exists in the community | -0.00717 | 0.216* |
| | (0.0512) | (0.119) |
| Primary school exists in the community | 0.0400 | -0.0274 |
| | (0.0529) | (0.0792) |
| Secondary school exists in the community | -0.0434 | 0.0238 |
| | (0.0463) | (0.0572) |
| Ambulatory exists in the community | 0.0918* | -0.00847 |
| | (0.0517) | (0.0836) |
| Hospital exists in the community | 0.00269 | -0.0114 |
| | (0.0449) | (0.0549) |
| Bank exists in the community | 0.132*** | 0.0433 |
| | (0.0507) | (0.0481) |
| Credit cooperative exists in the community | -0.0284 | -0.146 |
| | (0.0885) | (0.229) |
| District fixed effects (36 districts-coefficients omitted) | Yes | Yes |
| Constant | 2.062*** | 2.668*** |
| | (0.230) | (0.373) |
| Observations | 3599 | 3638 |
| R squared | 0.309 | 0.268 |

**Table 5**

**Average Treatment Effects**

| | Cons. per capita/poverty line | | | Cons. per adult equiv./poverty line | | |
|---|---|---|---|---|---|---|
| | **2002** | **2005** | **2002-2005** | **2002** | **2005** | **2002-2005** |
| Treated | -0.187 | -0.232 | -0.045 | -0.276 | -0.348 | -0.072 |
| Controls | -0.021 | 0.025 | 0.046 | -0.043 | 0.034 | 0.077 |
| **Difference** | **-0.166** | **-0.257** | **-0.091** | **-0.233** | **-0.382** | **-0.148** |
| *S.E.* | *0.038* | *0.051* | *0.063* | *0.050* | *0.068* | *0.084* |
| *T-stat* | *-4.400* | *-5.060* | *-1.435* | *-4.680* | *-5.610* | *-1.759* |

**Table 6**
**Means Tests 2002**

| Variable | Sample | Mean of Treated | Mean of Controls | % Bias | % Bias Reduction | t-statistic | p-value |
|---|---|---|---|---|---|---|---|
| Dummy for means-testing | Unmatched | 0.663 | 0.204 | 104.5 | | 23.5 | 0.000 |
| | Matched | 0.663 | 0.658 | 1.1 | 98.9 | 0.2 | 0.864 |
| Dummy for hh employment | Unmatched | 0.017 | 0.219 | -65.8 | | -11.1 | 0.000 |
| | Matched | 0.017 | 0.018 | -0.3 | 99.5 | -0.1 | 0.897 |
| Dummy for hh member not seeking | Unmatched | 0.194 | 0.085 | 31.8 | | 7.7 | 0.000 |
| | Matched | 0.194 | 0.197 | -0.7 | 97.8 | -0.1 | 0.923 |
| District targeting coefficient (bad) | Unmatched | 0.103 | 0.401 | -73.2 | | -13.5 | 0.000 |
| | Matched | 0.103 | 0.108 | -1.3 | 98.3 | -0.3 | 0.787 |
| District targeting coefficient (good) | Unmatched | 0.724 | 0.353 | 80.2 | | 16.6 | 0.000 |
| | Matched | 0.724 | 0.705 | 4.1 | 94.9 | 0.7 | 0.501 |
| Urban areas | Unmatched | 0.442 | 0.562 | -24.1 | | -5.1 | 0.000 |
| | Matched | 0.442 | 0.471 | -5.9 | 75.7 | -1.0 | 0.344 |

**Table 7**
**Means Tests 2005**

| Variable | Sample | Mean of Treated | Mean of Controls | % Bias | % Bias Reduction | t-statistic | p-value |
|---|---|---|---|---|---|---|---|
| Dummy for means-testing | Unmatched | 0.285 | 0.048 | 67.4 | | 19.5 | 0.000 |
| | Matched | 0.285 | 0.307 | -6.1 | 91.0 | -0.8 | 0.440 |
| Dummy for hh employment | Unmatched | 0.039 | 0.226 | -57.6 | | -10.3 | 0.000 |
| | Matched | 0.039 | 0.064 | -7.8 | 86.5 | -1.9 | 0.058 |
| Dummy for hh member not seeking | Unmatched | 0.278 | 0.185 | 22.2 | | 5.0 | 0.000 |
| | Matched | 0.278 | 0.261 | 4.1 | 81.4 | 0.6 | 0.520 |
| District targeting coefficient (bad)* | Unmatched | 0.142 | 0.340 | -47.5 | | -9.3 | 0.000 |
| | Matched | 0.142 | 0.142 | -0.1 | 99.7 | 0.0 | 0.980 |
| District targeting coefficient (good)* | Unmatched | 0.436 | 0.282 | 32.6 | | 7.3 | 0.000 |
| | Matched | 0.436 | 0.457 | -4.3 | 86.7 | -0.7 | 0.497 |
| Urban areas | Unmatched | 0.355 | 0.584 | -46.9 | | -10.0 | 0.000 |
| | Matched | 0.355 | 0.412 | -11.7 | 75.0 | -1.9 | 0.053 |