

Cash Transfers and Household Welfare in Albania: A Non-experimental Evaluation accounting for Time-Invariant Unobservables*

Lucia Mangiavacchi[†] and Paolo Verme[‡]

February 24, 2011

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 local linear matching estimator first suggested by Heckman et al. (1997, 1998). Exploiting discontinuities in program design, it is possible to reduce the bias due to unobserved factors with a conditional difference-in-difference estimation. We find the program to have a weak targeting capacity and a non significant impact on welfare.

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

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

*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 Ndhime Ekonomika program. We are specially grateful to Martin Ravallion (The World Bank) and Luc Behaghel (Paris School of Economics) for their valuable suggestions on evaluation strategy.

[†]Department of Applied Economics-University of the Balearic Islands and Microsimula-Paris School of Economics. lucia.mangiavacchi@uib.es

[‡]World Bank. pverme@worldbank.org

1 Introduction

During the socialist period, the transitional economies of Europe and Central Asia relied on a rather complex system of social protection transfers mainly targeted at children, disabled and war veterans. The poor and the unemployed were seldom recognized as vulnerable groups and public policies targeting these groups were the exceptions rather than the rule. The early transition period and the recession that accompanied this period contributed to enlarge quickly the pools of poor and unemployed people and this called for new programs able to target these groups.

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 and restructuring the social assistance systems from categorical to means-tested systems became one of the dominant themes of social protection reforms in transitional economies.

Yet, twenty years into the transition process, countries that implemented comprehensive structural reforms of the social assistance systems are very few and the question of whether social assistance is effective in protecting the poor and improving welfare remains largely unanswered. Studies that addressed this question are not many and the answers provided thus far offer a mixed picture. 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 and 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 and forthcoming) 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 behavioural 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 not many 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. 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 welfare and a higher level of discontent with life with program participants as compared to a control group. More recently, Mangiavacchi et al. (2010) used a collective consumption model to compare the impact of cash and in-kind policies on the welfare of Albanian young children in 2002 and found that, while in-kind transfers are effective in improving poor children's well-being, the *Ndihma Ekonomike* program did not have a positive impact. Seemingly, Giannelli and Mangiavacchi (2010), looking at the factors affecting school participation and attainment of children in Albania, noted that *Ndihma Ekonomike* was unable to prevent school drop outs and school delay of poor children and adolescents.

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 welfare. We consider the 2002 and 2005 surveys separately and exploit targeting failures and a discontinuity in program design occurred during the period to evaluate the impact of these changes on welfare. 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-difference 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 welfare for 2002 and 2005 which is in line with Dabalen et al. (2008) findings on the pooled 2002-2005 sample. When implementing the conditional difference-in-difference matching estimator, we find that a general reduction in the average amount of the benefit (the real value reduced by over 25%) reduces the performance of the program. This means that when we reduce the bias due to unobserved behavioural reaction, we find a positive -but not significant- impact of the program on household welfare. This brings back the empirical results to what the basic economic theory predicts, that is when a family receive a cash transfer, its welfare cannot fall. It also clarifies that the negative impact found in the two years alone is

due to a time-invariant bias originated by unobserved attributes of program participants, such as, for example, a behavioural reaction in the informal labour market.

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

2 The Ndihme Ekonomike Program

2.1 Description, reforms and eligibility

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-tests 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 and the program in recent years decreased in terms of budget and in terms of number of household beneficiaries. Between 2000 and 2005 expenditure on the program declined by about 40%, the number of household beneficiaries by about 20% and the real value of the benefit decreased by over a quarter. By 2005, the program accounted for 0.4% of GDP and covered approximately 120,000 households. According to the World Bank (2006): “*The decline resulted from a number of factors, including a reduction in the overall poverty levels and a decision to reduce the number of beneficiaries through the introduction of a work requirement.*” (p. 169). The average benefit in 2005 was around 2,500 lek per month per household, which was about 15% of the official poverty line and about 8.5% of average household consumption.

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 also the first public service scheme to be decentralized and its administration is now mainly responsibility of municipalities and communes. NE was again revised in early 2005 with the replacement of the means-testing formula and with a few changes on administrative procedures. Between 2004 and 2005, the Government also piloted a scheme that required at least one member of beneficiary households to participate in a work program but it is unclear how vast the effective implementation of this scheme was (World Bank, 2006).

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

¹Details 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>).

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. This process is time consuming and can potentially hinder program participation. As noted by the World Bank (2006): *“Applicants are required to meet 28 different eligibility criteria and to produce at least nine documents and certificates from seven different ministries and agencies. All evidence has to be produced in written form and much of it has to be notarized. Officially, there is no charge for these documents; however, in many cases a “small payment” will assist the process.”* (p. 170).

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 longer 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. In addition to these changes and as already described, the program decreased in terms of budget and average benefit and a pilot work program was also being introduced during the period considered. These combinations of factors will allow us to compare outcomes of the program in two separate periods when program conditions were very different and use the changes occurred between 2002 and 2005 to get rid of the time-invariant bias generated by unobserved characteristics.

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; 3) Migration episodes within the family and 4) Property different from land. These are observable characteristics to prioritize when considering program participation in the evaluation strategy.

2.2 Targeting

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. If we consider households, 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%. Similar patterns can be observed if we consider individuals. In the absence of behavioural considerations, the *Ndihma Ekonomike* program would appear to have had a positive effect on poverty (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 12.7% in 2005) but undercoverage and leakage rates (Cornia and Steward 1995) were very high in both years considered. In 2002, about three quarters of the poor were not targeted and 57.3% 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.²

The targeting performance over time is mixed. If we compare our results with those of Alderman (2001), which refers to a survey carried out in 1996, we find that targeting has worsened.³ 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.⁴ Coverage and undercoverage rates and the targeting coefficient improved between 2002 and 2005 but this has been accom-

²See below for details on the targeting coefficient.

³The 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.

⁴Consumption for all years is net of NE benefits.

panied by an increase in leakage and a decrease in adequacy (Table 2).⁵ 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 reduced.

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 favour 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 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.

⁵Our results on coverage, leakage and targeting coefficient coincide with those published in World Bank (2007).

Fourth, the targeting errors we have computed may be due to non take-up of the benefit since the household is responsible for the application. The non take-up is due to several reasons and some of these reasons may be deduced from Table 3. For example, most of the households with an income level below the threshold that do not receive NE do not live in urban areas. Living in rural areas can increase the difficulty of reaching the municipal office to apply for NE. Low levels of education may as well play an important role since, to complete the application, it is necessary to present the proper documentation and to understand all the conditions that may be expressed in legal language, difficult to read for most people. Indeed, in Table 3, the average years of education is rather small for the eligible non recipients. Another possible reason for non take-up may reside in the small amount of the benefit. The cost and stigma of applying to NE may be considered larger than the actual welfare improvement that would be reached if the household would receive the benefit. This is especially true if the household already has a source of income which would reduce substantially the amount of the benefit.

Fifth, 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.

3 Evaluation Strategy

The data we use 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 present and retrospective labor participation, migration and household assets, allowing us to identify the NE transfer and also recover some of the variables used for eligibility.⁶

Estimates from the two samples are fully comparable. The 2002 and 2005 surveys covered 3,599 and 3,640 households respectively, employed similar questionnaires and used the same sampling procedure. Both surveys include a community questionnaire with information on local services and socio-economic conditions⁷. This helps us controlling for community fixed effects and determining the behavioral traits of administrators otherwise unobserved. Our unit of ob-

⁶Data 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.

⁷Note that the community questionnaire is not administrated at municipality/communes level, but at a smaller territorial unit such as rural villages or urban blocks.

servation is the household.

Let $D = 1$ define households treated by the program and $D = 0$ households 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 *Average impact of Treatment on the Treated* (ATT):⁸

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

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 has 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 have 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 have 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) 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 in more detail these four conditions and how we address them.

(1) *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 -

⁸See 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.

$(Y_0, Y_1) \perp D | X$.⁹ 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 . These two assumptions are known as the “strong ignorability” assumptions following Rosebaum and Rubin (1983). However, 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. Rosebaum 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 variable ($Y_{(.)} = \beta X_{(.)} + U_{(.)}$), 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 the identification condition became:

$$E(U_0 | P(Z), D = 1) = E(U_0 | P(Z), D = 0) \quad (2)$$

This identification condition is weaker than than the “strong ignorability assumptions” and they can be used to construct alternative matching estimators to address the question of selection on observables. Heckman et al. (1998) also proved that the estimator with this exclusion restriction has lower asymptotic variance so it is better, if possible, to apply this two step procedure. Even if the assumption in (2) is weak with this method we are not perfectly sure to reduce the selection bias due to unobservables. However it is possible to get rid of the time-invariant component of this bias using a conditional DID estimator, as described below.

(2) *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; 3) Estimation of a standard OLS welfare regression using a set of variables T influencing outcome; 4) Estimation of the residuals of the welfare equation and creation of the corresponding variable; 5) Matching treated and non treated groups with a local linear matching estimator based on propensity scores and using residuals as outcome variable; 6) Estimate of the single mean difference in outcomes between treated and matched group.

⁹The symbol ‘ \perp ’ in this paper stands for ‘independence’.

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:

$$\hat{\alpha} = \frac{1}{n_1} \sum_{i \in I_1 \cap S_p} [U_{1i} - \sum_{j \in I_0 \cap S_p} W(i, j) U_{0j}] \quad (3)$$

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):

$$W(i, j) = \frac{G_{ij} \sum_{k \in I_0} G_{ik} (P_k - P_i)^2 - [G_{ij} (P_j - P_i)] [\sum_{k \in I_0} G_{ik} (P_k - P_i)]}{\sum_{j \in I_0} G_{ij} \sum_{k \in I_0} G_{ik} (P_k - P_i)^2 - (\sum_{k \in I_0} G_{ik} (P_k - P_i))^2} \quad (4)$$

A fixed bandwidth of 0.06 and a biweight kernel ($G(\cdot)$) are used for the estimator. We impose a common support condition because S_p needs to be determined to compute α . 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-difference (DID) across the two years to capture the impact of program reduction between the two years. 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 $DID = E(Y_{1t} - Y_{0t'} | X, D = 1) - E(Y_{0t} - Y_{0t'} | X, D = 0)$, where t and t' represent time after and before treatment respectively. In fact, it is sufficient for $E(Y_{0t} - Y_{0t'} | X, D = 1) = E(Y_{0t} - Y_{0t'} | X, D = 0)$ to hold. Under additive separability and index sufficiency, this condition becomes:

$$E(U_{0t} - U_{0t'} | P(Z), D = 1) = E(U_{0t} - U_{0t'} | P(Z), D = 0). \quad (5)$$

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:

$$\hat{\alpha}_{DID} = \left\{ \frac{1}{n_{1t}} \sum_{i \in I_{1t} \cap S_p} \left[U_{1ti} - \sum_{j \in I_{0t} \cap S_p} W(i, j) U_{0tj} \right] \right\} - \left\{ \frac{1}{n_{1t'}} \sum_{i \in I_{1t'} \cap S_p} \left[U_{1t'i} - \sum_{j \in I_{0t'} \cap S_p} W(i, j) U_{0t'j} \right] \right\} \quad (6)$$

We use this estimation to evaluate the impact of the policy change occurred between 2002 and 2005. Note that between 2002 and 2005 Albania experienced rapid growth and poverty reduction. With the DID matching we can also 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. Heckman et. al (1997) show that this estimator is more effective than the cross-sectional method in removing bias when data are contaminated by temporally-invariant components of bias,

(3) *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 basically 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 in section 2) ensure that among the treated and non treated groups we can find comparable households.

(4) *Common labor market.* This condition requires both the treated and control group to reside in the same area and be subject to a common labor market. This is addressed by controlling for local areas using two territorial variables. One is a variable that captures geographic areas

such as coastal and mountainous regions (used for the participation equation) and the second is a variable that captures the district (*rrethe* in Albanian) administrative subdivisions (used for the welfare equation). Albania is a small country of about 28,000 squared kilometers and both these variables can claim to capture local labour markets.¹⁰

Outcome and treatment 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.

Participation equation. The choice of regressors for the participation equation is one of the key choices for the matching procedure. We aimed at including variables that defined selection into the program according to the rules of the programs, variables that captured behavioral traits of local administrators, regional variables and other variables that defined self-selection, including retrospective variables.

To reproduce the decision to apply to the program based on administrative rules, we constructed dummies for eligibility based on the 2002 and 2005 means-test formulae and dummies for two exclusion criteria. The first exclusion criteria is captured by a dummy for employment of household members¹¹ in the formal sector where the formal sector is identified with the variable that indicates individuals who contribute to social security. This proxies the employment exclusion criteria and makes sure that we include only those households whose employment status is likely to be observed by the program administrators. The second exclusion criteria is a dummy for households with properties other than land, which is one of the program criteria to exclude households. From available data, we could also reconstruct a dummy for migration but this variable was never significant in the participation equation and was omitted from the final specification.

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 (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 the first quantile as regressor in the selection equation.

We also added in the participation equation regional dummy variables including coastal areas,

¹⁰We used the geographic variable in the participation equation because it improved significantly the balancing tests.

¹¹With the exception of self-employed in rural areas.

central areas, mountainous areas and Tirana (with coastal areas as base category). This has two major advantages. First, we expect residents in those different areas to have different information and opportunities about the NE program. Secondly and as already discussed, the matching should be with treated and controls residing in the same local labor market.¹²

Moreover, it is well-established that information on past income is essential to use for matching in means-tested programs as this improves the quality of matching (Heckman et al. 1997) and addresses the Ashenfelter’s dip problem (Ashenfelter, 1978). Since information about past outcomes was not available, we included in the treatment equation a dummy variable for the subjective assessment of the financial and life situation during the past three years. The variable codes ‘1’ those individuals that replied “somewhat deteriorated” or “deteriorated a lot” to both these questions: “Do you feel that your financial situation in the past 3 years has:” and “Do you feel that your life in general in the past 3 years has:”. Other controls in the treatment equation included: The age of the household head, a dummy indicating university education of the household head and the share of workers among the household members.

In substance, we are able to capture all major factors that determine program participation and the decision to apply to the programme. 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.

Welfare equation. 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 household’s economic situation), community variables (presence of educational, health and financial institutions) and district dummies. Note that employment status variables are included into the participation equation and also in the outcome equation as suggested by Heckman et al. (1997). In the outcome equation we control also for female labour force participation and the share of peoples self-employed in farm activities. As already explained, matching is based on the residuals of this equation which improves on efficiency as compared to more standard matching methods based on outcomes (Heckman et al., 1997 and 1998).

4 Results

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, however, the performance of the program worsened over the period mainly because the average amount of the benefit was reduced. This means that the overall impact of the program, once controlled for time-invariant unobserved attributes, is positive even if small and not statistically different from zero.

¹²The use of the district variable used further in the welfare equation was probably a better choice but the number of observations in each district was small and it was not possible to reach perfect balancing of covariates with this variable.

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 average treatment effect for 2002 and 2005 and the difference-in-difference 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 4) shows that all variables selected contribute significantly to selection into the program with the exception of two variables in 2002 (dummy for household with no land properties and share of employed).¹³ As expected, the dummy that capture deterioration of household conditions during the past three years and the dummy indicating whether households are below the means-tests' thresholds are the most relevant variables for selection, both with a positive sign and both with similar coefficients for the two 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. Higher education and higher age of the head of the household predict lower participation. The variable capturing the weak district ability to target households is significant and with the expected sign in both years. 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 median or good targeting record, other selection criteria being equal. The regional dummies are also relevant with a positive and significant coefficient for central and mountainous regions for both years.

The participation prediction capacity of the probit models based on the hit or miss method are 84.38% for 2002 and 85.45% for 2005. These are rather good scores considering that not all eligibility criteria could be used. Also, the set of variables selected is identical for the two years and ensured that treated and controls observations were fully balanced in both years.¹⁴

The OLS model (Table 5) has a fairly good explanatory power as compared to models of this kind. The model explains about 31.1% of the variance of welfare in 2002 and about 28.7% in 2005. Significant variables in both years are health and higher education of the head of the household (both with positive signs), share of employed in household size, the number of migrants in the household, households living in single houses and the number and age of children in the family (always negative).

In Table 6 we report the estimations of single and double differences. Program treatment seems to have a negative effect on welfare.¹⁵ Both single differences for 2002 and 2005 show negative and significant values. The average treatment effect for 2002 is estimated at about 15.5% of the poverty line. This negative effect rises to 25.1% in 2005 resulting in a negative

¹³The the dummy for the capital city Tirana is non significant but as compared to Coastal areas

¹⁴The predicted probabilities of participation are estimated with the Stata module 'pscore' by Becker and Ichino (2002) and the balancing tests are those performed by this module.

¹⁵Single 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.

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. Our interpretation for this results is the following. The most basic economic theory says that if you give some money to an household, it is not possible that its welfare reduces. In this perspective, the negative ATT estimated for 2002 and 2005 would not make sense, since the purpose of the estimator is to compare households with the same observed and unobserved attributes.

In Table 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 can see that the matching procedure almost eliminates the bias on observables. In 2002, the percentage in bias reduction is in between 82.1% and 99.7% depending on the variable considered. In 2005, these values vary in between 59% and 98.8%. Moreover, for none of the two years the means tests between treated and controls are significant after matching.

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). However, the negative estimation of the ATT parameter suggests that conditional independence assumption is not verified. RAME gives a biased estimation of the program impact since it is not able to account for some unobserved attributes that are driving this negative results. As a final cross-sectional analysis, 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 groups¹⁶. 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. 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, treatment has always a negative and significant effect on consumption and this negative effect increases over the period.

The estimation of the conditional Difference-in-Difference, brings back the results on the track provided by the economic theory, respecting identification condition and eliminating the time-invariant component of the bias caused by unobserved attributes. This allows us to infer on the Average Treatment Effects on the households covered by the program, which is found to be positive, but small and not significant. It also allows us to derive some conclusions on behavioural responses of treated households not observed in the data. For instance the behavioural effect on the informal labour market. Informal labour market is characterized by very low wages and even the small amount provided by the NE is sufficient to discourage people from supply labour force

¹⁶See Foster and Shorrocks (1988) and Abadie (2002).

on this market¹⁷.

5 Concluding Remarks

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 (Aderman, 2001 and 2002). More recently, the program was found to have a negative effect on poverty and life satisfaction (Dabaleni 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 semi-parametric matching estimator first proposed by Heckman et al. (1997, 1998), we found *Ndihma Ekonomike* to have a negative and significant effect on household welfare in 2002 and 2005. 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. When cash transfers decrease rather than increase welfare, behavioral unobserved changes due to program participation must be at work. These changes may be explained by labor supply distortions produced by this kind of safety nets, as suggested by Kanbur et al. (1994) when discussing the behavioural effects of social security on labour supply in developing countries. Considering the relevance of the informal economy in Albania, one can say that *ceteris paribus* the NE transfer limits the supply of informal labour and hence reduces welfare in the long run. Additionally, given the high level of corruption of the Albanian institutions, it possible that bribes requested during the NE application representing an other unobserved factor in the program evaluation, inducing distortion in the program's effectiveness.

All these possible omitted characteristics may lead conditional independence assumptions be too strong for our evaluation, while the weaker assumptions that justify conditional DID are more likely to be not rejected since all these factors are probably time-invariant. Indeed, DID estimation using repeated cross-sections brings back the results on the track provided by the economic theory, showing a positive but not significantly different from zero impact on household welfare.

The natural implications of all these findings is that *Ndihma Ekonomike* should be further

¹⁷Informal works may also be likely unsafe and dangerous.

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, may be conditional cash transfers of use the money to provide in-kind services more effective in order to reduce household vulnerability to poverty in the long run.

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.

Ashenfelter, Orley C. (1978). Estimating the Effect of Training Programs on Earnings. *The Review of Economics and Statistics*, MIT Press, vol. 60(1), pages 47-57, February.

Becker, S. and Ichino, A. (2002) Estimation of average treatment effects based on propensity scores. *The Stata Journal* Vol.2, No.4, pp. 358-377

Caliendo, M. and S. Kopeinig (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22, 3172.

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 D. van de Walle and Nead, K. (Eds.), *Public spending and the poor*. Johns Hopkins University Press.

Dabalén, A., Kilic, T. and Wane, W. (2008). Social transfers, labor supply and poverty reduction. The case of Albania. *World Bank Policy Research Working Paper* 4783.

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, 1731-1777.

Galasso, E. and M. Ravallion (2005). Decentralized targeting of an anti-poverty program. *Journal of Public Economics* 89, 705-727.

Giannelli, G. C. and Mangiavacchi, L. (2010). Children's Schooling and Parental Migration: Empirical Evidence on the "Left Behind" Generation in Albania", *LABOUR: Review of Labour*

Economics and Industrial Relations, 24 (s1), pp.76-92.

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, 605-654.

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

Hochman, H.M., and J.D. Rodgers (1969). Pareto optimal redistribution. *The American Economic Review* 59(4), 542-557.

Kanbur, R., M. Keen, and M. Tuomala (1994). Labor supply and targeting in poverty alleviation programs. *The World Bank Economic Review* 8(2), 191.

Kolpeja, V. (2006). Program implementation matters for targeting performance: Evidence and lessons from eastern and central europe. mimeo, 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. Wired at: <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.

Mangiavacchi, L., Perali, F. and Piccoli, L. (2010). Child Welfare in Albania Using a Collective Approach. *DEA Working Papers 41*, *Universitat de les Illes Balears, Departament d'Economia Aplicada*.

Milanovic, B. (2000). Social transfers and social assistance: An empirical analysis using latvian household survey data. *World Bank Policy Research Working Paper* 2328.

Oates, W.E. (1972). *Fiscal federalism*. Harcourt.

Okrasa, W., (1999). The dynamics of poverty and the effectiveness of Poland's safety net (1993-1996). *World Bank Policy Research Working Paper* 2221.

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* 2218.

Ravallion, M. (2008). Evaluating anti-poverty programs. In T. P. Schultz and J. Strauss (Eds.), *Handbook of development economics* Vol.4. North Holland.

Ravallion, M. and B. Bidani (1994). How robust is a poverty prole? *The World Bank Economic Review* 8 (1), 75-102.

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*, 4, Schultz, Paul and Strauss, John (Eds.), Elsevier.

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. (forthcoming) The Poverty Reduction Capacity of Public and Private Transfers in Transition, *Journal of Development Studies*.

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 (2006). Albania: restructuring public expenditure to sustain growth. Report 36543 - AL.

World Bank (2007) Albania: Urban Growth, Migration and Poverty Reduction. A Poverty Assessment. Report No. 40071- AL.

Table 1: Poverty

	2002		2005	
	With NE	Without NE	With NE	Without NE
Households				
Headcount Ratio	19.1	20.1	14	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: Targeting

Index	Construction	2002	2005
1 Coverage	Households treated/Total households	11.0	12.7
2 Adequacy	Average treatment/Average consumption	9.2	8.3
3 Undercoverage	Poor not treated/Total poor	75.4	67.5
4 Leakage	Non poor treated/Total treated	57.3	64.2
5 Targeting	(Poor treated/Total poor)-(Non poor treated/Total non poor)	0.17	0.23

Figure 1: Share of NE Expenditure by Quintile

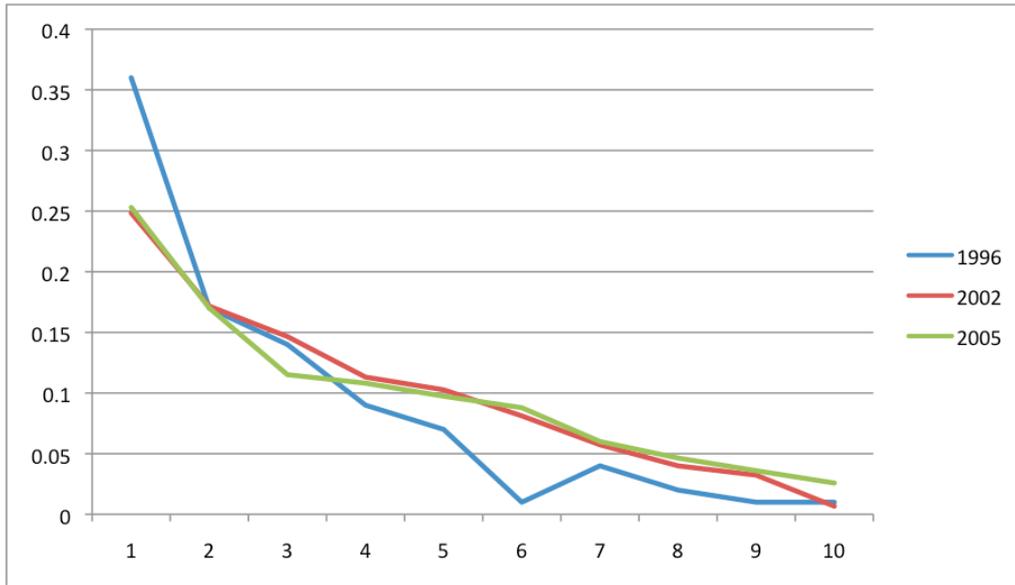


Table 3: Eligibility

	2002		2005	
	Eligible without NE	Non Eligible with NE	Eligible without NE	Non Eligible with NE
Share of hh with a member employed	0.01	0.16	0.18	0.19
Share of hh in urban area	0.21	0.62	0.27	0.32
Average level of education	6.68	8.46	6.80	8.32
Number of hh (sample)	626	177	147	388
Number of hh (population)	129605	31619	30304	67171
Share of hh with child 0	31.33	13.13	34.13	11.91
Share of hh with child 1	18.08	27.08	17.99	21.86
Share of hh with child 2	23.75	34.24	22.94	35.12
Share of hh with child 3	16.66	14.29	17.85	19.76
Share of hh with child 4	8.11	7.31	5.16	7.45
Share of hh with child 5	1.48	1.89	1.09	2.99
Share of hh with child 6	0.53	1.60	0.84	0.60
Share of hh with child 7	0.07	0.45	0.00	0.32

Table 4: Selection equations (Probit)

VARIABLES	2002	2005
Financial and life situation deteriorated in the past 3 years	0.336*** (0.0741)	0.366*** (0.0706)
Dummy for means testing	0.726*** (0.0736)	0.767*** (0.0905)
Dummy for employment	-1.176*** (0.159)	-0.935*** (0.123)
Dummy for hh with properties different from land	-0.163 (0.130)	-0.186** (0.0867)
Household head has completed university	-0.761*** (0.248)	-0.682*** (0.207)
Share of employed in household size	-0.110 (0.224)	-0.527*** (0.198)
Age of the household head	-0.0269*** (0.00226)	-0.0232*** (0.00232)
District's targeting coefficient (bad)	-0.506*** (0.101)	-0.444*** (0.0846)
Central regions (1)	0.445*** (0.0899)	0.654*** (0.107)
Mountain regions (1)	0.734*** (0.0906)	1.052*** (0.102)
Capital city Tirana (1)	0.136 (0.163)	-0.183 (0.149)
Constant	-0.236* (0.141)	-0.280* (0.159)
Observations	3599	3640
Participation prediction capacity (hit or miss method, %)	84.38	85.45

Dep. Var.: Dummy for treatment. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. (1) Base category: Coastal areas.

Table 5: Welfare equations (OLS)

VARIABLES	2002	2005
Urban area	-0.0512 (0.0725)	0.176* (0.0934)
Age of the hh head	0.00248 (0.00803)	-0.0212* (0.0123)
Age of the hh head (squared)	-3.92e-05 (7.51e-05)	0.000170 (0.000108)
HH head is in good health	0.106*** (0.0385)	0.107* (0.0596)
HH mean education	-0.0338* (0.0195)	0.0165 (0.0252)
Head has completed university	0.557*** (0.0658)	0.697*** (0.0910)
Share of employed in household size	0.763*** (0.142)	0.534*** (0.155)
Share of urban family members working in farm	0.128 (0.520)	1.469*** (0.261)
HH share of female employed	-0.0583 (0.214)	0.253 (0.235)
HH number of migrants	0.0662* (0.0401)	-0.221*** (0.0420)
HH owns dwelling	0.108* (0.0641)	0.0487 (0.0900)
HH lives in single house	-0.0860** (0.0435)	-0.157*** (0.0548)
HH has 1 under five child	-0.396*** (0.0382)	-0.385*** (0.0588)
HH has 2 under five children	-0.746*** (0.0533)	-0.827*** (0.0627)
HH has 3 or more under five children	-0.952*** (0.155)	-0.893*** (0.125)
HH has 1 child (6-18)	-0.409*** (0.0447)	-0.290*** (0.0510)
HH has 2 children (6-18)	-0.628*** (0.0477)	-0.488*** (0.0563)
HH has 3 children (6-18)	-0.866*** (0.0581)	-0.646*** (0.0680)
HH has 4 or more children (6-18)	-1.043*** (0.0664)	-0.974*** (0.0709)
Pre-school exists in the community	-0.0287 (0.0527)	0.186 (0.117)
Primary school exists in the community	0.0322 (0.0534)	-0.0749 (0.0776)
Secondary school exists in the community	0.0147 (0.0507)	-0.104* (0.0626)
Ambulatory exists in the community	0.120** (0.0515)	0.0643 (0.0854)
Hospital exists in the community	0.0234 (0.0522)	-0.216*** (0.0741)
Government or private bank exists in community	0.116** (0.0530)	-0.00284 (0.0472)
Credit cooperative exists in the community	-0.0159 (0.0934)	-0.144 (0.224)
Dummies for districts	yes	yes
Constant	2.171*** (0.255)	2.670*** (0.428)
Observations	3599	3638
R-squared	0.311	0.287

Dep. Var.: Consumption per capita. *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses

Table 6: Average Treatment Effects and Difference in Difference

Variable	Mean Treated	Mean Controls	Difference	S.E.	T-stat	sig. level
ATT 2002	-0.173	-0.018	-0.155	0.044	-3.49	***
ATT 2005	-0.249	0.002	-0.251	0.057	-4.44	***
Difference in Difference			-0.096	0.061	-1.57	

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Means Tests

Variable	Sample	2002				2005				t-test	t	p>t
		Mean	Control	%bias	%reduct bias	Mean	Control	%bias	%reduct bias			
Financial and life situation deteriorated in the past 3 years	Unmatched	0.29	0.17	28.50	6.48	0.00	0.34	0.18	35.60	8.26	0.00	0.00
	Matched	0.27	0.30	-5.10	-0.75	0.45	0.32	0.32	0.40	98.80	0.06	0.95
Dummy for means testing	Unmatched	0.66	0.20	104.50	23.50	0.00	0.29	0.05	67.40	19.47	0.00	0.00
	Matched	0.66	0.66	-0.70	-0.10	0.92	0.27	0.27	1.10	98.40	0.14	0.89
Dummy for employment	Unmatched	0.02	0.22	-65.80	-11.07	0.00	0.04	0.23	-57.60	-10.26	0.00	0.00
	Matched	0.02	0.02	-2.20	-0.77	0.44	0.04	0.04	-1.40	-0.38	0.71	0.71
Dummy for hh with properties different from land	Unmatched	0.04	0.10	-20.30	-3.86	0.00	0.11	0.21	-25.50	-5.08	0.00	0.00
	Matched	0.04	0.05	-1.40	-0.26	0.79	0.12	0.13	-2.90	-0.53	0.60	0.60
Household head has completed university	Unmatched	0.01	0.12	-49.50	-8.20	0.00	0.01	0.12	-45.40	-7.75	0.00	0.00
	Matched	0.01	0.01	-3.00	-1.19	0.23	0.01	0.02	-2.70	-0.93	0.35	0.35
Share of employed in household size	Unmatched	0.07	0.20	-70.80	-12.99	0.00	0.09	0.22	-65.10	-12.12	0.00	0.00
	Matched	0.07	0.08	-4.70	-0.92	0.36	0.09	0.10	-1.00	-0.20	0.84	0.84
Age of the household head	Unmatched	44.21	51.91	-57.70	-11.91	0.00	46.87	52.79	-46.40	-9.77	0.00	0.00
	Matched	44.55	44.53	0.20	0.03	0.98	47.07	46.59	3.70	0.60	0.55	0.55
District's targeting coefficient (bad)	Unmatched	0.10	0.40	-73.20	-13.52	0.00	0.14	0.34	-47.50	-9.30	0.00	0.00
	Matched	0.10	0.11	-2.40	-0.50	0.62	0.14	0.15	-1.90	-0.36	0.72	0.72
Central regions	Unmatched	0.26	0.28	-5.40	-1.13	0.26	0.34	0.26	15.60	3.43	0.00	0.00
	Matched	0.26	0.26	0.30	0.05	0.96	0.34	0.31	6.40	1.02	0.31	0.31
Mountain regions	Unmatched	0.59	0.22	80.20	18.07	0.00	0.55	0.23	71.00	16.27	0.00	0.00
	Matched	0.58	0.58	0.60	0.09	0.93	0.55	0.56	-3.70	-0.56	0.58	0.58
Capital city Tirana	Unmatched	0.03	0.19	-51.70	-9.03	0.00	0.03	0.20	-53.20	-9.46	0.00	0.00
	Matched	0.03	0.04	-0.80	-0.21	0.84	0.04	0.04	-2.80	-0.73	0.46	0.46

Figure 2: Stochastic Dominance (CDFs)

