Impacts of rural non-farm employment on household welfare in Pakistan

Rakhshanda Kousar[1] and Awudu Abdulai[1]
[1]Department of Food Economics and Consumption Studies,
University of Kiel, Germany.

Corresponding Author's E-mail: rkousar@food-econ.uni-kiel.de

Paper prepared for presentation at the 2nd AIEAA Conference
"Between Crisis and Development: which Role for the Bio-Economy"

6-7 June, 2013
Parma, Italy

Summary

This article examines the impact of non-farm work on household welfare, differentiated by female, for rural households in the Punjab province of Pakistan. We employ an endogenous switching regression approach that accounts for selection bias due to observable and unobservable factors to examine the factors that influence the household's decision to participate in non-farm work and the impact of participation on household welfare. Given we find no substantial selection bias on unobservable factors; we also use PSM approach to check the robustness of our results from the ESR estimates. Separate estimates are also provided for male and female to address gender heterogeneity. The empirical results reveal that participation in non-farm work significantly increases per head expenditures and reduces household poverty level. This confirms the potential role of non-farm work in improving rural household welfare and poverty alleviation in rural areas of developing countries.

Keywords: non-farm work, household welfare, impact assessment, Pakistan

JEL classification: J16, J22, Q1

Impacts of rural non-farm employment on household welfare in Pakistan

Rakhshanda Kousar[1] and Awudu Abdulai[1]
[1]Department of Food Economics and Consumption Studies,
University of Kiel, Germany.

1. INTRODUCTION

Developing countries are characterized by the major portion of population living in the rural areas, engaged in agricultural activities. Agriculture sector has been defining the livings of the poor people, giving them employment opportunities in the past few decades. At present, agricultural sector has limited its potential to fulfill the increasing demands of employment and food for the rapidly growing population due to a long list of limiting factors like: the small sized farm holdings, being the leading problem in rural economy in developing countries (Wan and Cheng, 2001; Rusu, 2002), imperfect credit markets coped with missing insurance facilities (Dercon, 2002), limited access to fertililizer (Lamb, 2003; Duflo et al., 2011), unavailability of credit (Odhiambo and Magandini, 2008) and increased cost of production (Tan et al., 2008). So there is need to focus on the diversification of livelihoods of rural people through the generation and promotion of non-farm earning opportunities.

The development of various non-farm activities has been generally recognized to have a potential in raising employment opportunities and stimulating the growth of rural economies. The important share of income is derived from non-farm activities and attracts the large rural labor force. Haggblade et al. (2010), for instance, reported that non-farm income contributed 35-50 percent of rural household income and 33 percent of rural labor force across the developing world. This sector is growing (Lanjouw and Lanjouw, 2001) and the widely quoted empirical evidence from a variety of different locations of developing countries indicates that it increases food production and farm income (Barrett et al., 2001; Babatunde and Qaim, 2010), minimizes the income gap between farm and non-farm households (Mishra et al., 2002; Holden et al., 2004), ensures food security and poverty alleviation in the developing countries (Owusu et al., 2011), relaxes liquidity constraint and purchases inputs for farming (Oseni and Winters, 2009; Chang et al., 2011).

Given the significance of the non-farm in contributing to household welfare, many studies have explained the determinants of participation in non-farm work (Ellis, 2000; Abdulai and CroleRees, 2001; Barrett et al., 2001). However, very few studies have analyzed the impact of participation in non-farm work on household welfare (e.g., Owusu et al., 2011). The studies that have analyzed the impacts did not account for selectivity bias that may occur as a result of unobservable factors. From an econometric standpoint, analyzing the welfare impact of non-farm work may be affected by unobserved factors. Failure to distinguish between the causal effect of participation and the effect of unobserved heterogeneity could lead to biased estimates and misleading policy implications.

The present study acknowledges that the differences in welfare outcome variables between those farm-households that did and did not participate in non-farm work could be due to observed and unobserved factors. As indicated by Heckman et al., (1997), it is likely that the differences between two individuals with or without exposure to program or technology may be more systematic even after conditioning on unobservable or observable factors. In this study, we employ endogenous switching regression approach to examine the factors that influence the household's decision to participate in non-farm work and the impact of participation on household welfare measures such as per capita expenditure and poverty levels. The correlation coefficients between the error terms of selection equation and error terms of outcome equations were not significantly different from zero, even at the 10 percent level, suggesting the absence of any endogenous switch. This indicates that there is no substantial selection on unobservable factors. We therefore employed the propensity matching approach to further examine the impact of participation on household welfare, and to also check the robustness of our findings from the endogenous switching regression model. The study utilizes cross-sectional rural household level data collected in 2010 from a randomly selected sample of 341 households in Punjab province of Pakistan. The study provides separate estimates for males and females to address gender heterogeneity. This examination will be critical particularly in drawing policy recommendation for poverty alleviation in rural areas of developing countries.

The rest of the paper is organized as follows: In section two, overview of the significance of non-farm sector in Pakistan is described. In section three, the conceptual framework and estimation strategy is presented. In section four, description of the survey area and data set are provided. In section five, empirical results are reported. Finally, some concluding remarks are given in the section six.

2. SIGNIFICANCE OF NON-FARM SECTOR IN PAKISTAN

Among the South Asian countries, Pakistan is a lower middle income country. Nearly two third of the population and 80 per cent of the country's poor people are concentrated to the rural parts of the country. Major economic activities in these areas are related directly or indirectly to the agriculture sector, but at present this sector has limited margins for landless poor due many limiting factors. The major limiting factor in rural areas of Pakistan is the skewed distribution of land ownership. As Anwar et al., (2004) reported that in rural areas of Pakistan, 67 percent households were landless and just 0.1 percent households possessed 1 hectare and above landholdings. Moreover, the fast growing rural labor force cannot be much absorbed in the almost employment overcrowded agriculture sector.

Hence, as a result of class stratification, increasing landlessness of small farmers, and population growth, non-farm sector in Pakistan is expanding. This sector in Pakistan, like many in other developing countries, is a heterogeneous sector covering a wide spectrum of activities. Generally it includes all activities in the rural economy that are not pursued on farms. It is of great importance to rural economies for its productive and employment effects. Almost 45-50 percent of the rural population in Pakistan is directly dependent on non-farm income. Contribution of this sector is also critical for food security, poverty alleviation, farm sector competitiveness and productivity. It offers agricultural services and products to the food and fiber system. These products are critical to the dynamics of agriculture (GOP, 2011).

Despite the critical role of the rural non-farm sector in food security, poverty alleviation, farm sector competitiveness and productivity, this sector has received inadequate attention in the debate in Pakistan (Malik, 2008). There is need to focus on the non-farm sector while designing poverty alleviation strategies for rural areas of Pakistan where vicious land-labor ratio limit income earnings opportunities in agriculture.

3. CONCEPTUAL FRAMEWORK AND ESTIMATION STRATEGY

The theoretical basis for the model presented below draws upon the model of time allocation suggested by Huffman (1991), laterally used by Owusu et al., (2011). Farm household is assumed to maximize utility defined over consumption of goods C and leisure N , i.e., $U = U(C, N_i)$.

Household faces time, non-negative, production function and budget constraints. Time constraint of household is:

$$L_i = L_{iA} + L_{iNF} + N_i \qquad where \quad i = 1, Male, i = 2, Female \tag{1}$$

L_i , L_{iA}, L_{iNF} , is the household total time endowment, time allocated to agricultural and non-farm respectively.

Household also faces non-negative, production function and budget constraints:

$$L_{iA} \geq 0 \qquad L_{iNF} \geq 0 \tag{2}$$

$$Y = Y(L_{iA}, H_i, I; X) \tag{3}$$

$$Y_{iA} + Y_{iNF} + Y_{iO} - P_I I = P_C C \tag{4}$$

Putting value of Y from equation 3, the equation 4 can be written as:

$$P_C C = P_y Y - P_I I - W_1 L_{iA} + W_2 L_{iNF} + Y_{iO} \tag{5}$$

where $Y_{iA} / Y_{iNF} / Y_{iO}$ represents agriculture, non-farm and non-labor income respectively, P_I , P_C , P_y represents vector of prices of variable inputs, purchased goods, and farm output respevtively, H_i represents amount of hired labor, W_1, W_2 represents farm and non-farm wages respectively, I represents vector of purchased inputs and Y represents quantity of agricultural production.

The first order condition for optimal time allocation for farm, non-farm and leisure activities is given as:

$$\partial U / \partial L_i = W_i \partial U / \partial C - \partial U / \partial L = 0 \tag{6}$$

Returns to labor from farm and non-farm work can be obtained by rearranging the equation (6)

$$W_i = {}^{\partial U / \partial L}\!\big/_{\partial U / \partial C} \tag{7}$$

When farm household allocate their time to the three activities, the labor supply function for farm and non-farm activities can be specified as:

$$L_{iA} = L_{iA}(W_1, W_2, P_y, P_I; X) \tag{8}$$

$$L_{iNF} = L_{iNF}(W_1, W_2, P_y, P_I, Y_{io}; X) \tag{9}$$

Following the empirical literature on non-farm work decisions of farm households, it is assumed that the participation decision of the individual is influence by a comparison between the reservation wage (W_i^r) and the potential market wage (W_i^m) in the non-farm sector. Participation in non-farm activities $(L_i = 1)$ occur if $W_i^m > W_i^r$ and positive number of non-farm hours will not be observed $(L_i = 0)$ if $W_i^m \leq W_i^r$. However, these differential wages are not observable, but we do observe the decision of participation. Given that we do observe participation or non-participation, Huffman and Lange (1989) note that an index function can be specified with an unobserved variable (L_i^*), such that

$$L_i^* = \alpha X_i' + \mu_i$$

$$L_i = 1 \quad if \;\; L_i^* > 0 \tag{10}$$

$$L_i = 0 \quad if \;\; L_i^* \leq 0$$

where (X_i') is a vector of individual and household characteristics and (μ_i) is an error term. In order to estimate the relationship between participation in non-farm work and household per head expenditures and poverty level, we start with a linear function

$$Y_i = \beta Z_i' + \gamma L_i + \varepsilon_i \tag{11}$$

Where (Y_i) is per head expenditure and poverty level of household, (L_i) is dummy variable representing participation in non-farm sector, (Z_i') is individual, household and locational characteristics, (ε_i) is random error term.

3.1. Empirical impact evaluation challenges

In non-experimental study, estimation of impact of non-farm work on the welfare of household is very substantial because there is need of information on the counterfactual situation had they not had participated in non-farm work. In experimental studies, information on counterfactual situation is provided by randomly assigning households to treatment and control status, where the welfare outcome observed on the non-participants are statistically representative of what would have occurred without participation for participants.

Moreover, in non-experimental studies, households are not randomly distributed to the two groups (participants and non-participants), rather participation in non-farm activities may be dependent on the benefits from participation. Therefore, participants and non-participants may be systematically different. Thus, possible self-selection occur if unobserved factors influence both the error term (μ_i) of the participation equation (10) and the error term (ε_i) of the outcome equation (11), thus resulting the correlation of the error terms. The implication of this is that the use of standard regression techniques (ordinary least square (OLS)) to estimate the parameters of the equation would result in biased and inconsistent estimates.

Some authors have employed the Heckman selection method or instrumental variable approach (IV) but these two methods assume that outcome function would differ only by unobservable factors between the participating and non-participating households in the non-farm work. According to Heckman et al., (1997), it is likely that the differences between two individuals with or without exposure to program or technology may be more systematic even after conditioning on unobservable or observable factors. So in this study, we employ endogenous switching regression approach to examine the factors that influence the household's decision to participate in non-farm work and the impact of participation on household welfare measures such as per capita expenditure and poverty levels. As we found no endogenous, therefore we also employed the propensity matching approach to further examine the impact of participation on household welfare, and to also check the robustness of our findings from the endogenous switching regression model

3.2. Endogenous switching regression model

We specify the binary decision choice of household to participate in non-farm work conditioned on observed covariates as:

$$L_i^* = \alpha X_i' + \mu_i$$

$$L_i = 1 \quad if \;\; L_i^* > 0 \tag{12}$$

$$L_i = 0 \quad if \;\; L_i^* \leq 0$$

To account for selection biases we adopt an endogenous switching regression model of welfare outcomes, (i.e. per head expenditure and poverty level) where households face two regimes (1) to participate, and (2) not to participate defined as follows:

$$Y_{1i} = \beta_1 Z_i' + \gamma_1 L_{1i} + \varepsilon_{1i} \tag{13a}$$

$$Y_{2i} = \beta_2 Z_i' + \gamma_2 L_{2i} + \varepsilon_{2i} \tag{13b}$$

where Y_i is per head expenditure and poverty level of in regimes 1 and 2, Z_i represent a vector of exogenous variables thought to influence outcome function. γ_1, γ_2 are parameters to be estimated and $\varepsilon_1, \varepsilon_2$ are error terms.

Finally, the error terms are assumed to have a trivariate normal distribution, with zero mean and non-singular covariance matrix expressed as:

$$\operatorname{cov}(\mu_i, \varepsilon_1, \varepsilon_2) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1\mu} \\ \sigma_{12} & \sigma_2^2 & \sigma_{2\mu} \\ \sigma_{1\mu} & \sigma_{2\mu} & \sigma^2 \end{bmatrix} \tag{14}$$

where $\sigma_1^2 = \operatorname{var}(\varepsilon_1); \sigma_2^2 = \operatorname{var}(\varepsilon_2); \sigma^2 = \operatorname{var}(\mu_1); \sigma_{12} = \operatorname{cov}(\varepsilon_1, \varepsilon_2); \sigma_{1\mu} = \operatorname{cov}(\varepsilon_1, \mu_i);; \sigma_{2\mu} = \operatorname{cov}(\varepsilon_2, \mu_i);$ σ^2 represents variance of the error term in the selection equation and σ_1^2, σ_2^2 represents variance of the error term in the outcome equations.

According to Maddala (1983), when there are unobservable factors associated with selection bias, the important implication of the error structure is that because the error term (μ_i) of the selection equation (12) is correlated with the error terms $(\varepsilon_1, \varepsilon_2)$ of the welfare outcome functions (13a) and (13b), the expected values of $\varepsilon_{1i}, \varepsilon_{2i}$ conditional on the sample selection are non-zero:

$$E(\varepsilon_{1i} | L_i = 1) = E(\varepsilon_{1i} | \mu_i > -X_i'\alpha) = \sigma_{1\mu}\left[\frac{\phi(X_i'\alpha/\sigma)}{\Phi(X_i'\alpha/\sigma)}\right] \equiv \alpha_{1\mu}\lambda_1 \tag{15a}$$

$$E(\varepsilon_{2i} | L_i = 0) = E(\varepsilon_{2i} | \mu_i \le -X_i'\alpha) = \sigma_{2\mu}\left[\frac{-\phi(X_i'\alpha/\sigma)}{1-\Phi(X_i'\alpha/\sigma)}\right] \equiv \alpha_{2\mu}\lambda_2 \tag{15b}$$

where ϕ and Φ are the probability density and cumulative distribution functions of the standard normal distribution, respectively. The ratio of ϕ and Φ evaluated at $\alpha X_i'$,represented by λ_1 and λ_2 in equations (15a) and (15b) is referred to as the inverse mills ratio (IMR) which denotes selection bias terms.

Previous studies have used a two-stage method to estimate the endogenous switching model. (e.g Lee, 1978; Feder et al., 1990; Fuglie and Bosch, 1995; Freemanet al., 1998). In the first stage, a probit model of the criterion equation is estimated and the inverse Mills ratios λ_1 and λ_2 are derived according to definitions in equations (15a) and (15b). In the second stage, these predicted variables are added to the appropriate equation in (13a) and (13b) respectively to yield following sets of equations

$$Y_{1i} = \beta_1 Z_i' + \alpha_{1\mu}\lambda_1 + \gamma_1 L_{1i} + \upsilon_1 \tag{16a}$$

$$Y_{2i} = \beta_2 Z_i' + \alpha_{2\mu}\lambda_2 + \gamma_2 L_{2i} + \upsilon_2 \tag{16b}$$

The coefficients of the variables λ_1 and λ_2 provide estimates of the covariance terms $\alpha_{1\mu}$ and $\alpha_{2\mu}$, respectively. Since the variables λ_1 and λ_2 have been estimated, however, the residuals υ_1 and υ_2 cannot be used to calculate the standard errors of the two-stage estimates. While Lee (1978) suggested a procedure to derive consistent standard errors most especially for the two stage approach, Maddala (1983) argue that such procedure require potentially cumbersome and complicated process which most studies using the earlier two stage approach failed to implement. Thus, in the present study, a single stage approach where Full-Information Maximum Likelihood (FIML) method proposed by Lokshin and Sajaia (2004) using the movestay command in the statistical software STATA is employed for the empirical analysis. The FIML simultaneously fit the selection (i.e., equation 12) and outcomes (i.e., equation 13a and 13b) equations in order to yield consistent standard errors, thus, making λ_1 and λ_2 in equations 16a and 16b, respectively homoskedastic.

The FIML's log likelihood Function for switching regression model employed in this study proposed by Lokshin and Sajaia (2004) is described below:

$$\ln L_i = \sum_{i=1}^{N} \left\{ \begin{array}{l} L_i w_i \left[\ln F\left(\frac{(X_i'\alpha + \rho_{1\mu}(Y_{1i} - Z_{1i}'\beta / \alpha_1}{\sqrt{1-\rho_{1\mu}^2}}\right) + \ln(f((Y_{1i} - Z_{1i}'\beta / \alpha_1) \right] \\ + (1-L_i) w_i \left[\begin{array}{l} \frac{\ln(1 - F(X_i'\alpha + \rho_{2\mu}(Y_{2i} - Z_{2i}\beta) / \gamma_2)}{\sqrt{1-\rho_{2\mu}^2}}) \\ + \ln(f((Y_{2i} - Z_{2i}\beta) / \gamma_2) \end{array} \right] \end{array} \right\} \tag{17}$$

The signs of the correlation coefficients $\rho_{1\mu}$ and $\rho_{2\mu}$ have economic interpretations (Fuglie and Bosch 1995). If $\rho_{1\mu}$ and $\rho_{2\mu}$ have alternate signs, then individuals participate in non-farm work on the basis of their comparative advantage: those who participated have above average returns from participation and those who choose not to participate have above-average returns from non-participation. On the other hand, if the coefficients have the same sign, it indicates hierarchical sorting: participants have above-average returns whether they participate or not, but they are better off participating, whereas non-participants have below-average returns in either case, but they are better off not participating.

The ATT of non-farm participation can be calculated as

$$ATT = E(Y_{1i} - Y_{2i} | L_i = 1) = Z_i'(\beta_1 - \beta_2) + (\sigma_{1\mu} - \sigma_{2\mu})\lambda_1 \tag{18}$$

In equation 18, $E(Y_{1i} | L_i = 1) = Z_i'\beta_1 - \sigma_{1\mu}\lambda_1$ represents the expected outcome for households who participated, had they chose to participate in non-farm; $E(Y_{2i} | L_i = 1) = Z_i'\beta_2 - \sigma_{2\mu}\lambda_1$ represents the expected outcome for households who participated, had it been they chose not to participate in non-farm.

3.2 Propensity score matching

To examine the causal effect of non-farm participation on household per head expenditure and poverty level, propensity score matching technique was be used. The basic idea behind the propensity score is that we may reduce the bias if we compare outcomes of treated and control groups which are as similar as possible. It constructs a statistical comparison group by matching every individual observation of participants with an observation with similar characteristics from the group of non-participants. Thus, create the conditions of an experiment in which participants and non-participants are randomly assigned, allowing for the identification of a causal link between the non-farm participants and outcome variables.

The propensity score is the conditional probability of assigning a treatment, given pre-treatment characteristics (Rosenbaum and Rubin, 1983), written as:

$$P(X_i) = \Pr(L_i = 1 | X_i) = E(L_i | X_i) \tag{19}$$

Where $L = \{0,1\}$ is the indicator of exposure to treatment (non-farm participation) and X is the vector of pre-treatment characteristics.

The parameter of interest is the Average Treatment Effect on the Treated (ATT), which can be estimated as:

$$ATT = E\{E[Y_i^1 - Y_i^0 | L_i = 1, P(X_i)]\}$$

$$= E\left\{E\left[Y_i^1 \middle| L_i = 1, P(X_i)\right] - E\left[Y_i^0 \middle| L_i = 0, P(X_i)\right] \middle| L_i = 1\right\} \tag{20}$$

Where Y_i^1 and Y_i^0 are the potential outcome in two counterfactual situations. The propensity score is predicted with probit model. The predicted propensity score is then used to estimate treatment effect.

Conceptually the ATT requires a mean for the unobservable counterfactual, $E\left[Y_i^0 \middle| L_i = 1\right]$ so for the observable quantities in equation (20) to identify the ATT relies on three key conditions introduced into the literature by Rosenbaum and Rubin (1983).

First is "unconfoundedness" $(Y_i^0, Y_i^1 \perp L_i | X_i)$ where $\perp$ denotes independence. According to this, potential outcomes are independent of participation, conditional on the observable covariates, X_i. Given observable covariates, participating group is treated as random and any systematic differences in actual outcomes between participants and non-participants individuals with the same value of the covariates is attributed to the participation in non-farm work.

Second is "common support" where all participated households have a counterpart in the non-participating group for each X_i for which we seek to make a comparison. This condition would appear to create a dimensionality problem when many covariates are matched on; for example, if X_i contains k covariates which are all dichotomous the number of possible matches will be 2^k. However the propensity score reduces the dimensionality of the matching problem because it is possible to match on $P(X_i)$ which is scalar, rather than on the vector of observable variables X_i. This use of $P(X_i)$ is valid so long as the "balancing" property $(prob(X_i | L_i = 1, P(X_i) = P) = prob(X_i | L_i = 0, P(X_i) = P))$ holds (Rosenbaum and Rubin, 1983). In other words, conditional on the propensity score, the means of the covariates should be identical across the treatment and control groups if the balancing property holds.

Since the propensity score is a continuous variable it is unlikely that there are two observations with exactly the same value of $P(X_i)$, so further refinement is needed to estimate equation (20). Various matching estimators have been suggested in the literature. Although all matching estimators normally yield same results but the choice of a matching approach could become important in small sample (Heckman et al., 1997). The most commonly used are nearest neighbor matching (NNM), kernel based matching (KBM), stratified matching, radius matching and Mahalanobis matching methods. The NNM, Radius and KBM methods are employed in this study.

The most straightforward matching estimator is nearest neighbor matching (NNM) which matches each participant with its closest neighbor with similar observed characteristics. It can be done either with replacement or without replacement. Matching with replacement results in bias reduction since each treatment group can be matched to the nearest comparison group as a result of a reduction in the propensity score distance (Smith and Todd, 2005).

Kernel matching (KM) is non-parametric matching estimators that use weighted averages of all individuals in the control group to construct the counterfactual outcome. Kernel matching tends to use more non-participants for each participant, thereby reducing the variance but possibly increasing the bias. To avoid the risk of bad matches by choosing the closest neighbors that are far away, calipers are implemented. This involves imposing a tolerance on the maximum distance in the propensity score allowed. When applying KM one has to choose the kernel function and the bandwidth parameter. The kernel function appears to be relatively unimportant in practice (DiNardo and Tobias, 2001). As noted by Pagan and Ullah (1999) that more important is the choice of the bandwidth parameter with the following trade-off arising: High bandwidth-values yield a smoother estimated density function, therefore leading to a better fit and a decreasing variance between the estimated and the true underlying density function. On the other hand, underlying features may be smoothed away by a large bandwidth leading to a biased estimate. The bandwidth choice is therefore a compromise between a small variance and an unbiased estimate of the true density function.

In radius matching, each treated unit is matched only with the control unit whose propensity score falls in a predefined neighborhood of the propensity score of the treated unit. The benefit of this approach is that it uses only the number of comparison unit available within a predefined radius; thereby allowing for use of extra units when good matches are available and fewer when they are not. One possible drawback is the difficulty of knowing a priori what radius is reasonable (Dehija and Wahba, 2002).

The matching quality depends on the ability of the matching procedure to balance the relevant covariates. The standardized bias approach proposed by Rosenbaum and Rubin (1985) is used to quantify the bias between treated and control groups. Sianesi (2004) has also proposed a comparison of the pseudo-R^2 and p-values of the likelihood ratio test of the joint significance of all the regressors obtained from the probit analysis before and after matching the samples. To ensure that there are no systematic differences in the distribution of the covariates between both groups, the pseudo-R^2 should be fairly low after matching and the joint significance of covariates should be rejected. Sensitivity analysis was also undertaken to check if the influence of an unmeasured variable on the selection process is so strong to undermine the matching procedure. Since it is not possible to estimate the selection bias in practice with non-experimental data, we employed the bounding approach suggested by Rosenbaum (2002) to examine the influence of unmeasured variable on the selection process. The aim of this approach is to determine how strongly an unmeasured variable must influence the selection process to undermine the implication of the matching process.

4. AREA SELECTION AND DATA DESCRIPTION

Data for this study was collected between September 2010 and January 2011 through a cross sectional survey of rural households in Punjab province which is the country's most populous region, constitutes 56 percent of Pakistan's total population. Due to its largest rural society, that province was selected for data collection. There was no large-scale redistribution of agricultural land and asymmetry exists in land ownership. As a result most rural areas are dominated by a small set of land-owning families. It has always contributed the most to the national economy of Pakistan. Its share of Pakistan's GDP was 54.7 percent in 2000 and increased to 59 percent in 2010. It is especially dominant in the service and agriculture sectors of the Pakistan economy. Its contribution is ranging from 52.1 to 64.5 percent in the service sector and 56.1 to 61.5 percent in the agriculture sector. It is also major manpower contributor because it has largest pool of professionals and highly skilled manpower (Pakistan encyclopedia, 2012). It has three broad agro-climatic zones named as lower, central and upper. Two districts from each zone were selected for survey. A stratified random sample of a total of 341 households was selected in each of six districts to ensure representation of all categories of households, which potentially influence the extent and nature of livelihood diversification.

Using a structured questionnaire, these households were interviewed eliciting information on farm and non-farm activities as well as personal, demographic and locational characteristics. Information on agriculture activities included farm size, crop output, price of output, expenditure on variable inputs, family and hired labor, capital assets, own consumption, sale of produce, access to credit. Information on livestock activities included number of animals and poultry birds. Detailed information on the consumption expenditure was fully recorded.

The dependent variable used in the study is a dummy variable that takes the value of one, if the household participated in non-farm work, and the value zero, if no participation was recorded. The outcome variables used in this study are per head expenditure and headcount index as an indicator of household poverty status. The consumption expenditure components include expenditures on food, tobacco, clothing, energy, livestock, health, education, social activities (marriages, deaths, etc.), recreation and other household expenditures over the last year. Headcount index was calculated on the basis of per capita expenditures. Purchasing power parity (PPP) poverty line used in this study is 1.25$ per day per person suggested by World Bank (2008) for Pakistan. Poverty outcome was measured as a binary variable. Since gender plays an important role in the poverty dynamics in Pakistan, the gender stratification was used in the estimation.

The independent variables used in the estimations were based on past research on determinants of participation in non-farm employment (Abdulai and Delgado, 1999; Barrett et al., 2001; Owusu et al., 2011). These variables include household characteristics such as gender of household head, age to capture experience, and education of the household head to present productivity potential, presence of children, household size, household assets (land, livestock) to indicate wealth, access to credit to capture liquidity constraints, village infrastructure development project (road, factory/mill), distance of household from retail shop to indicate employment opportunities, and location characteristics to capture community fixed effects.

Tables 1 and 2 present the descriptive statistics of variable used in estimations and difference in the characteristics of participants and non-participants with their t-values for males and females, respectively. The observed mean difference of 0.50 in the effects of treatment for males (0.59) and females (0.09) is statistically significant at 1% level indicating the presence of gender heterogeneous treatment effects.

Table 1. Descriptive statistics of variables used in estimation for males.

Variable	Description	Participants N=202 (59.24%)		Non-participants N=139 (40.76%)		Difference in means
		Mean	S.d	Mean	S.d	
PerHExpend	Per head expenditures (Rupees)	99738.11	289855.2	72494.24	78173.28	27243.86*
Headcount	Head count index to capture poverty	0.35	0 .48	0 .40	0 .49	- 0.04*
AgeHead	Age of the household head in years	49.63	0.82	46.77	0.95	2.86**
Age2Head	Square of head age	2490.28	1215.56	2286.22	1098.84	204.07*
HeadEdu	Years of education of household head	2.39	1.24	1.73	0.97	0.66***
HHSizeOver14	No.of household members above 14 years	5.49	2.72	4.25	2.67	1.23***
Ch0L05	No. of children under 5 year of age	1.04	1.45	0.95	1.35	0.09
Child14	No. of children between age 6-14 years	6.02	6.78	6.58	5.96	-0.57
livstk	1 if household has livestock,0 otherwise	0.78	0.41	0.88	0.32	-0.98**
TCultiLand	Total cultivated land in acres	18.81	42.58	21.54	39.57	-2.74
BorowMon	1 if household takes credit, 0 otherwise	0.26	0.44	0.24	0.42	0.03*
DProg	1 if village development prog, 0otherwise	0.22	0.42	-0.47	0.72	0.69
M0CasW	Village cash wages of male (Rupees)	266.36	70.75	280.76	78.44	-14.39*
Location1	1 if Lahore district, 0 otherwise	0.14	0.35	0.15	0.36	
Location2	1 if Sahiwal district ,0 otherwise	0.23	0.42	0.14	0.35	
Location3	1 if M.Garh district ,0 otherwise	0.32	0.46	0.26	0.44	
Location4	1 if layyah district, 0 otherwise	0.03	0.17	0.00	0.00	
Location5	1 if Sialkot district,0 otherwise	0.20	0.41	0.32	0.46	
Location6	1 if Khushab district,0 otherwise	0.06	0.23	0.12	0.32	
DRShop	Distance of household to retail shop (km)	0.72	0.45	0.82	0.39	-0.09**

Note: Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.

For poverty calculations Purchasing Power Parity (PPP) US$ 1.25 per person per day is used as poverty line.

Table 2. Descriptive statistics of variables used in estimation for females.

Variable	Description	Participants N=202 (59.24%)		Non-participants N=139 (40.76%)		Difference in means
		Mean	S.d	Mean	S.d	
PerHExpend	Per head expenditures (Rupees)	85552.33	203701.3	64742.62	54142.19	20809.71**
Headcount	Head count index to capture poverty	0.38	0.49	0.41	0.50	-0.03*
AgeHead	Age of the household head in years	53	8.64	47.99	11.71	5.00**
Age2Head	Square of head age	2455.13	966.93	2402.13	1192.49	52.99*
HeadEdu	Years of education of household head	2.41	1.41	2.09	1.15	-0.32*
HHSizeOver14	No.of household members above 14 years	6.75	2.16	4.81	2.76	1.94***
Ch0L05	No. of children under 5 year of age	0.75	0.92	1.03	1.45	-0.28*
Child14	No. of children between age 6-14 years	5.13	5.96	6.37	6.51	-1.24
livstk	1 if household has livestock,0 otherwise	0.68	0.47	0.84	0.36	-0.15**
TCultiLand	Total cultivated land in acres	10.62	11.92	20.89	43.16	-10.27*
BorowMon	1 if household takes credit, 0 otherwise	0.41	0.49	0.24	0.43	0.17**
DProg	1 if village development prog, 0 otherwise	0.44	0.51	0.38	0.49	0.06
F0CasW	Village cash wages of female (Rupees)	138.59	74.17	158.85	76.75	20.26
Location1	1 if Lahore district, 0 otherwise	0.18	0.39	0.15	0.36	
Location2	1 if Sahiwal district, 0 otherwise	0.15	0.37	0.20	0.40	
Location3	1 if M.Garh district, 0 otherwise	0.41	0.49	0.29	0.45	
Location4	1 if layyah district, 0 otherwise	0.03	0.17	0.02	0.13	
Location5	1 if Sialkot district, 0 otherwise	0.19	0.39	0.26	0.44	
Location6	1 if Khushab district, 0 otherwise	0.03	0.17	0.09	0.29	
DRShop	Distance of household to retail shop (km)	0.75	0.44	0.76	0.43	-0.01

Note: Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.

For poverty calculations Purchasing Power Parity (PPP) US$ 1.25 per person per day is used as poverty line.

The difference in the rate of participation between males and females reflects the fact that males in the area are more engaged in non-farm activities, while females are predominantly engaged in household activities. There are cultural and social barriers that prevent women from entering and remaining in the labor force. They are not allowed to go outside for work and confined to only household duties are considered honorable for them. Also presented in the Tables 1 are differences in means of the variables used in the estimations for both male and female participants, alongside their significance levels. The significance levels suggest that there are some differences between participants and non-participants with respect to household and farm-level characteristics.

With regards to the outcome variables, there appear to be statistically significant differences in per head expenditures and poverty level of household between participants and non-participants. Poverty appeared to be lower and per head expenditures was higher among non-farm participants than non-participants for both male and female. When estimating poverty using monetary measures, one may have a choice between using income or consumption as the indicator of wellbeing. Most analysts argue that, provided the information on consumption obtained from a household survey is detailed enough, consumption will be a better indicator of poverty measurement than income (Coudouel et al., 2002). Head count index is used as indicator of poverty. It is the share of the population living in household whose consumption is below the poverty line.

We found a significant difference in term of education level of household head. Higher level of education of household head among participants in non-farm work indicates that, education was found to be a better positioned to mobilize capital through high returning non-farm work. In other words illiteracy serves as an entry barrier into high returning and less risky activities.

There is also significant difference between the female participants and non-participants with respect to access to credit. 41 % of participants and 24 % of non-participants have access to credit, while the corresponding figures for males were 26 % and 24 % respectively, revealing the fact that limited access to credit for households is an entry barrier to their non-farm work participation.

Household composition and characteristics seem to matter for participants and non-participants. Participants have more working age members as compared to non-participants for both male and female. The presence of young children (<5 years) in the household impede the probability of participation of female. So caring for young children appears to be compatible with non-farm work in rural areas of Pakistan. In particular, there appear to be differences in the ownership of land and livestock. Non-participants have higher acreage of land and more number of livestock than participants.

Although the comparison in Table 1 and 2 do reveal some significant differences between participants and non-participants, however mean differences do not account for the effect of other characteristics of rural households and cannot be taken as evidence for the specific effects of participation. Multivariate approaches that account for selection bias arising from the fact that participants and non-participants may be systematically different, are essential in providing sound estimates of the impact of participation on per head expenditure and poverty level of household.

5. EMPIRICAL RESULTS AND DISCUSSION

In order to analyze the driving forces behind rural household's decision to participate in non-farm work, we employed endogenous switching regression that can control for both observable and unobservable selection bias. FIML estimates of the endogenous switching regression model for male and female participants are reported in Tables 3-6. The third column of Tables 3-6 presents the estimated coefficients of the selection equation (12) on non-farm work participating or not whereas the fourth and fifth column present per head expenditures and poverty level of household. The empirical results for the probability of non-farm participation are generally in agreement with prediction from the analytical model. Age variable is positive and statistically significant in both specifications for male and females, which represents general experience that increases the marginal value of time in each activity. The results suggest that an increase in age of household head increases the probability of both male and female participation in non-farm work. In both specifications, number of years of schooling of household head significantly increases the participation decision since most of the activities in non-farm work require a certain level of education. Thus education of household members, represented by the household head's level of education, is a powerful source that leads labor out of agriculture and shifts it into high returning non-farm sector (Timmer, 1988). Therefore households without well educated heads are consequently excluded from non-farm activities.

Household composition and characteristics seems to matter as well on the probability of participation for both male and female. The coefficient of the adult household size is positive and significant, suggesting that the presence of working age men and women in the household which is an indicator of non-nuclearity of household, tends to increase the probability of participation of both male and female in non-farm work. These results are in contrast to the study of Barrett et al. (2008) and in line with the study of Abdauli and Regime (2000) in the case of male labor supply. The presence of children in the household had no significant effect on the probability of participation of both male and female in non-farm work, confirmed to the findings of other studies that showed that non-farm work and child care are not necessarily competing activities in rural areas of developing countries (Rosenzweig, 1980; Sahn and Alderman, 1988; Skoufias, 1994; Abdulai and Delgado, 1999).

Males and females belonging to households endowed with valuable physical capital like farm land or livestock are less likely to participate in non-farm activities. As noted by Weiss (1997) that increase in farm size reduces the probability of participation in off-farm labor market. Perhaps they often capitalize their valuable assets in order to smooth consumption in times of income shortfalls (Fafchamps et al., 1998; Abdulai and CroleRees, 2001; Corral and Reardon, 2001; Lanjouw et al., 2001; Dercon, 2002; Barrett et al., 2005; Verpoorten, 2009). Consequently, valuable agricultural assets can be regarded as some kind of entry barriers, since well endowed households are able to run their farm properly and are not dependent on non-farm activities to generate more income or to spread risk. In the absence of valuable endowments, however, households seem to be forced into non-farm employment since farming would not be successful due to low quality of assets. Since endowment with valuable assets also represents the household’s wealth, these findings also support the theory of decreasing risk aversion. Households endowed with valuable physical capital are less risk-averse and therefore less likely to participate in non-farm employment. The presence of a development project in study area enhanced the probability of participation for both male and female in non-farm earning activities. In both specifications, access to credit decreases the likelihood of non-farm participation for both male and female but this variable is significantly not different from zero.

Table 3. Full information maximum likelihood estimates of the endogenous switching regression.

Dependent variable: non-farm participation of male and per head expenditure

Variables	Description	FIML Endogenous Switching Regression		
		Participation (1/0)	Participation=1 (participants)	Participation=0 (non-participants)
AgeHead	Age of household head (years)	0.044(0.02)*	1486.615(1032.52)	10850.41(5859.70)*
Age2Head	Square of head age	-0.001(0.00)	- 9.553(9.68)	-113.069(62.17)*
HeadEdu	Years of education of household head	0.089(0.02)***	2006.029(1212.06)*	15799.03(5925.29)***
HHSizeOver14	No.of household members above 14 years	0.101(0.03)***	477.980(1639.10)	-2552.426(10025.73)
Ch0L05	No. of children under 5 year of age	-0.018(0.06)	- 11172.18(3115.13)***	-13555.96(19111.12)
Child14	No. of children between age 6-14 years	-0.016(0.03)	- 2648.617(692.17)***	-8396.397(3770.76)**
Livstk	1 if household has livestock,0 otherwise	-0.496(0.24)**	29170.37(11271.17)***	111341.70(56806.98)**
TCultiLand	Total cultivated land in acres	-0.007(0.01)	629.072(351.20)*	5390.478(3056.95)*
BorowMon	1 if household takes credit, 0 otherwise	-0.126(0.18)	-12316.09(9926.92)*	23771.70(61787.72)
DProg	1 if village development prog, 0 otherwise	0.008(0.00)*	11913.81(11815.76)	249.831(665.87)
M0CasW	Village cash wages of male (Rupees)	-0.001(0.00)	116.95(93.42)	-170.992(259.22)
Location1	1 if Lahore district, 0 otherwise	0.166(0.29)	- 12266.68(17977.18)	61118.18(64830.65)
Location2	1 if Sahiwal district , 0 otherwise	0.503(0.31)	- 15005.54(12192.1)	76184.78(71324.03)
Location3	1 if M.Garh district , 0 otherwise	0.190(0.30)	- 11033.16(14055.02)	16619.42(57644.97)
Location4	1 if Layyah district, 0 otherwise	6.446(0.74)***	- 2614.478(31276.57)	25803(35664.1)
Location5	1 if Sialkot district, 0 otherwise	0.132(0.33)	6467.268(20673.71)	29066.35(57972.38)
DRShop	Distance of household to retail shop (km)	-0.324(0.15)**		
Constant		-1.056(0.81)	-9608.821(40579.14)	-899.378(148525.60)
σ_{ei}			70679.45(1.95) ***	262235(13.64)***
ρ_j			-0.336(0.30)	0.758(0.53)

Note:absolute value of robust standard error in parenthesis.
Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.
σ_{ei} denotes the square root of the variance of the error terms in the outcome equations, ρ_j denotes correlation coefficient between the error term of selection equation and error term of outcome equation.

Table 4. Full information maximum likelihood estimates of the endogenous switching regression.

Dependent variable: non-farm participation of male and per head expenditure

Variables	Description	FIML Endogenous Switching Regression		
		Participation (1/0)	Participation=1 (participants)	Participation=0 (non-participants)
AgeHead	Age of household head (years)	0.057(0.03)**	710.811(657.06)	8864.94(4527.70)**
Age2Head	Square of head age	-0.052(0.03)**	-325.627(708.21)	-10041.23(5216.64)*
HeadEdu	Years of education of household head	0.063(0.02)***	3233.42(879.11)***	2962.276(2825.53)
HHSizeOver14	No.of household members above 14 years	0.115(0.03)***	1123.871(1966.81)	-8814.77(6239.28)
Ch0L05	No. of children under 5 year of age	-0. 051(0.06)	-6185.097(3179.638)*	-29598.36(13595.25)**
Child14	No. of children between age 6-14 years	-0.001(0.01)	-2050.238(692.40)***	-3227.91(1804.83)*
livstk	1 if household has livestock,0 otherwise	-0.423(0.23)*	26923.76(10608.73)***	-10451.13(26556.61)
TCultiLand	Total cultivated land in acres	-0.004(0.00)	216.820(189.01)	3417.42(2015.69)*
BorowMon	1 if household takes credit, 0 otherwise	-0.024(0.18)	-15084.88(9266.77)	45493.47(43596.3)
DProg	1 if village development prog, 0 otherwise	0.013(0.01)**	21355.13(13433.69)	532.333(463.14)
F0CasW	Village cash wages of female (Rupees)	-0.001(0.00)	188.854(111.14)*	-88.223(152.82)
Location1	1 if Lahore district, 0 otherwise	1.458(0.41)***	16096. 98(14509.69)	14043.9(34611.43)
Location2	1 if Sahiwal district, 0 otherwise	1.830(0.40)***	18832.9 (14331.83)	-20906.85(36205.26)
Location3	1 if M.Garh district, 0 otherwise	1.558(0.40)***	22716.35(13875.92)	-14001.92(34369.65)
Location4	1 if Layyah district, 0 otherwise	2.093(0.78)***	8100.969(30827.24)	112396.8(61361.67)*
Location5	1 if Sialkot district, 0 otherwise	0.014(0.46)	47265.47(27991.48)*	3048.95(23646.26)
DRShop	Distance of household to retail shop (km)	-0.586(0.21)***		
Constant		-2.832(0.85)***	-38401.17(44267.32)	-70368.19(99072.37)
σ_{ei}			58999.02(7766.79)***	201434.5(70.54)***
ρ_j			-0.009(0.43)	0.001(0.13)

Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.
σ_{ei} denotes the square root of the variance of the error terms in the outcome equations, ρ_j denotes correlation coefficient between the error term of selection equation and error term of outcome equation.

Table 5. Full information maximum likelihood estimates of the endogenous switching regression.

Dependent variable: non-farm participation of male and per head expenditure

Variables	Description	FIML Endogenous Switching Regression Participation (1/0)	Participation=1 (participants)	Participation=0 (non-participants)
AgeHead	Age of household head (years)	0.045(0.02)*	-0.005(0.01)	-0.044(0.01)***
Age2Head	Square of head age	-0.000(0.00)	-7.77e-06(0.00)	0.001(0.00)***
HeadEdu	Years of education of household head	0.082(0.02)***	-0.030(0.01)***	-0.022(0.01)**
HHSizeOver14	No.of household members above 14 years	0.1077(0.03)***	-0.035(0.02)*	0.026(0.02)
Ch0L05	No. of children under 5 year of age	-0.014(0.07)	0.068(0.03)***	0.080(0.03)***
Child14	No. of children between age 6-14 years	-0.011(0.01)	0.009(0.01)*	0.017(0.01)*
livstk	1 if household has livestock,0 otherwise	-0. 598(0.23)***	-0.186(0.09)***	-0.161(0.13)
TCultiLand	Total cultivated land in acres	-0.003(0.00)	-0.002(0.00)**	-0.002(0.00)*
BorowMon	1 if household takes credit, 0 otherwise	-0.126(0.18)	0.052(0.07)	0.061(0.09)
DProg	1 if village development prog, 0otherwise	0.013(0.01)*	-0.121(0.07)*	0.003(0.00)**
M0CasW	Village cash wages of male (Rupees)	-0.000(0.00)	0.001(0.00)	0.001(0.00)
Location1	1 if Lahore district, 0 otherwise	0.180(0.32)	0.288(0.12)***	0.251(0.12)**
Location2	1 if Sahiwal district ,0 otherwise	0.564(0.31)*	0.207(0.13)	0.443(0.13)***
Location3	1 if M.Garh district ,0 otherwise	0.212(0.31)	0.214(0.11)**	0.303(0.09)***
Location4	1 if Layyah district, 0 otherwise	6.521(0.98)***	0.133(0.25)	0.986(0.72)
Location5	1 if Sialkot district,0 otherwise	0 .093(0.35)	0.021(0.13)	0.020(0.32)***
DRShop	Distance of household to retail shop (km)	-0 .370(0.21)*		
Constant		-1.204(0.79)	0.758(0.43)*	1.102(0.32)***
σ_{ei}			0.418(0.03)***	0.378(0.02)***
ρ_j			-0.229(0.50)	0.081(0.48)

Note:absolute value of robust standard error in parenthesis.
Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.
σ_{ei} denotes the square root of the variance of the error terms in the outcome equations, ρ_j denotes correlation coefficient between the error term of selection equation and error term of outcome equation.

Table 6. Full information maximum likelihood estimates of the endogenous switching regression.

Dependent variable: non-farm participation of male and per head expenditure

Variables	Description	FIML Endogenous Switching Regression Participation (1/0)	Participation=1 (participants)	Participation=0 (non-participants)
AgeHead	Age of household head (years)	0.054(0.03)**	-0.008(0.01)	-0.031(0.01)***
Age2Head	Square of head age	-0.050(0.02)**	0.003(0.01)	0.026(0.01)***
HeadEdu	Years of education of household head	0.063(0.02)***	-0.032(0.01)***	-0.018(0.01)**
HHSizeOver14	No.of household members above 14 years	0.114(0.03)***	-0.042(0.02)**	0.021(0.01)**
Ch0L05	No. of children under 5 year of age	-0.055(0.06)	0.072(0.03)**	0.085(0.02)***
Child14	No. of children between age 6-14 years	-0.000(0.01)	0.006(0.01)	0.007(0.01)*
livstk	1 if household has livestock,0 otherwise	-0.431(0.23)*	-0.286(0.10)***	-0.143(0.09)
TCultiLand	Total cultivated land in acres	-0.004(0.00)	-0.002(0.00)*	-0.002(0.00)***
BorowMon	1 if household takes credit, 0 otherwise	-0.017(0.18)	0.096(0.08)	0.014(0.08)
DProg	1 if village development prog, 0 otherwise	0.012(0.01)*	-0.114(0.09)	0.003(0.00)***
F0CasW	Village cash wages of female (Rupees)	-0.001(0.00)	0.000(0.00)	0.000(0.00)
Location1	1 if Lahore district, 0 otherwise	1.441(0.41)***	0.104(0.34)	0.331(0.12)**
Location2	1 if Sahiwal district, 0 otherwise	1.800(0.40)***	0.076(0.37)	0.588(0.13)***
Location3	1 if M.Garh district, 0 otherwise	1.541(0.40)***	0.034(0.35)	0.341(0.10)***
Location4	1 if Layyah district, 0 otherwise	2.057(0.76)***	-0.081(0.46)	0.210(0.13)
Location5	1 if Sialkot district, 0 otherwise	-0.015(0.47)	0.050(0.32)	0.148(0.09)*
DRShop	Distance of household to retail shop (km)	-0.608(0.214)***		
Constant		-2.720(0.89)***	1.270(0.65)*	1.091(0.26)***
σ_{ei}			0.431(0.03)***	0.385(0.02)***
ρ_j			-0.287(0.39)	0.199(0.22)

Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.
σ_{ei} denotes the square root of the variance of the error terms in the outcome equations, ρ_j denotes correlation coefficient between the error term of selection equation and error term of outcome equation.

The results of the second part of FIML endogenous switching regression model are presented in last two columns of Tables 3-6. Identification of the model requires that there should be at least one variable in the participation equation that does not appear in the per head expenditures and poverty equations. In both specifications, distance of household from retail market variable is used as identifying instrument.

The non-significance of covariance terms $\rho_{1\mu}$ in the case of per head expenditures and poverty, in the lower panel of Tables 3-6, shows the absence of endogenous switching in both cases. Also results show that the covariance terms $(\rho_{1\mu}, \rho_{2\mu})$ have alternate signs with $\rho_{1\mu} < 0$ and $\rho_{2\mu} > 0$, which indicates that non-farm participation is based on its comparative advantage. The non-significance of $\rho_{2\mu}$ indicate that in the absence of non-farm work participation, there would be no significant difference in average behavior of male and female in the two groups, caused by unobserved effects.

The results in Tables 3-6 indicate that education of household head exerts a positive effect on per head expenditure for both non-farm participants and non-participants households, but its effect is not significant in the case of female. On the other hand, the negative and significant coefficient of the variable for schooling suggests that education seems to be a key factor to reduce poverty of both participant and non-participant households. These results indicate that education enhance the welfare of household by increasing the efficiencies of individual activities. As noted by El-Osta (2011) that schooling has significant impact on higher rural household earnings.

Family composition appears to be an important factor in explaining per head expenditure and poverty level differences among participants and non-participants for both male and female. Adult household size tends to have a positive effect on per head expenditures for both male and female participants, and negative effect on per head expenditure for non-participants households, although it does not significantly influence for both cases. In the case of poverty level, adult household size tends to decrease the poverty level of male and female participants and increase the poverty level for non-participants male and female. Presence of children (> 5 years & between 6-14 years) tends to decrease the per head expenditure, while it enhance the poverty level of non-farm work participants and non-participants for both male and female.

Households, endowed with sufficient natural capital (e.g. land) and physical capital (e.g. livestock) are found to influence outcome, although at varying levels. The ownership of livestock has positive impact on per head expenditure of household, albeit its effect is inconclusive for non-participant females. In the case of poverty, livestock has positive and significant effect on poverty reduction for male and female participants but no significant impact for both male and female non-participants. Similarly farm size increases the per head expenditure for both participants and non-participants. The coefficient of farm size is negative and statistically significant for both participants and non-participants in the case of poverty.

Table 7. Impact of non-farm participation on per head expenditure and poverty level of household--------ESR results.

	Outcome Variables	Outcome mean		ATT	t-Statistics
		Participants	non-participants		
Male	Per head expenditure	90560.12	55202.69	35357.43	1.92*
	Head count	1.235	1.504	-0.26	4.37***
Female	Per head expenditure	123000.81	79850.55	43150.26	6.42***
	Head count	1.206	1.501	-0.30	3.56***

Note: Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels

Table 7 presents the results of the impact of non-farm participation of male and female on per head expenditure and poverty using endogenous switching regression method. We find that participation in non-farm work significantly increases consumption expenditures and reduces the poverty status of rural households. Non-farm participation increases per head expenditures by about 35357.43 and 43150.26 point compared to non-participants for male and female respectively. The ATT estimates of 0.26 and 0.30 for poverty reduction suggests that household participation in non-farm work decreases probability of poverty by 0.26 and 0.30 point for male and female respectively, suggesting that non-farm work has significant impact on poverty reduction among rural households in Pakistan. Overall, the estimates show that although the participation rate of females was lower but their participation contributed more to improve welfare and reduce poverty as compared to their male counterparts. Thus, non-farm work appears to be more crucial in the case of female in improving the welfare and reducing the poverty status of farm household.

As we find no endogenous switch so in order to check robustness of our ESR findings and for comparability purposes and, we used propensity score matching (PSM) to determine the impact of non-farm on household welfare. A probit model has been employed to predict the probability of participation in non-farm work. The estimated propensity scores for male and female participation are given in Table A-1 and A-2 in the appendix. As the propensity score only serves as a device to balance the observed distribution of covariates across the participant and non-participant groups (Lee, 2008) so the detailed interpretation of the propensity score estimates is not undertaken here. However, the results indicate that most of the variables included in the estimations have the expected signs.

The effect of non-farm work on per head expenditure and poverty status of the households is estimated with the nearest neighbor matching (NNM), radius matching, and kernel-based matching (KBM) methods. The empirical results of average treatment effect (ATT) are given in Tables 8, while the indicators of matching quality are provided in Table 9 for both male and female. Before turning to the causal effects of non-farm participation of household, we briefly discuss the quality of the matching process. After estimating the propensity scores for the participants and non-participants group we check the common support condition1.

A visual inspection of the density distributions of the estimated propensity scores for the two groups (figure1 in appendix) indicates that the common support condition is imposed and the balancing property was satisfied in all the estimates regression models. The bottom half of the graph shows the propensity scores distribution for the non-participants and the upper half refers to that of participants. The densities of the scores are on the y-axis.

1 In this study, the common support region is implemented, following the example of Leuven and Sianesi (2003), discarding observation from the participants group, whose propensity score is higher than the maximum or less than the minimum propensity score of non-participants

Table 8. Average treatment effects and sensitivity analysis for male and female--------PSM results.

	Matching	Outcome variable	No.of neighbours/ Kernel type	Caliper	ATT	Critical level of hidden bias	No.of treated	No.of controlled
Male	NNM	Per head expenditures	4	0.0011	29678.95**(2.15)	1.4-1.5	196	139
		Headcount	1	0.04	-0.14*(1.79)	1.4-1.5	196	139
	Radius	Per head expenditures	-	0.0011	29678.95**(2.15)	1.4-1.5	196	139
		Headcount	-	0.0015	-0.16*(1.84)	1.3-1.4	196	139
	Kernel	Per head expenditures	tricube	0.0011	27765.51**(2.00)	1.4-1.5	196	139
		Headcount	tricube	0.0011	-0.21**(2.06)	1.5-1.6	196	139
Female	NNM	Per head expenditures	1	-	25452.68**(2.15)	1.4-1.5	32	308
		Headcount	3	-	-0.27**(2.57)	2-2.1	32	308
	Radius	Per head expenditures	-	0.0003	36996.14*(1.82)	1.3-1.4	32	308
		Headcount	-	0.19	-0.17*(1.96)	1.2-1.3	32	308
	Kernel	Per head expenditures	normal	0.0001	32738.56*(1.69)	1.5-1.6	32	308
		Headcount	biweight	0.01	-0.23**(2.03)	1.4-1.5	32	308

Note: t-statistics are in parentheses.
Significance of t-statistics of mean difference is at the *10%, **5% and ***1% level

The matching for all three approaches in Table 8 generally indicate that non-farm work of male and female exerts a positive and significant impact on per head expenditure and a negative and significant impact on poverty status of household. Specifically, the NNM, Radius and KBM casual effect of participation on per head expenditures range between 27765.51 and 29678.95 for male; 25452.68 and 36996.14 for female. The magnitude of coefficient suggests that the average treatment effect of participation in non-farm work increases the individual’s welfare by 27765.51 - 29678.95 for male and 25452.68 - 36996.14 for female.

Non-farm participation of male and female also had significant impact on reducing poverty. The estimated impact of participation on poverty reduction as measured by head count index is estimated to range between -0.14 and -0.21 for male; -0.17 and -0.27 for female, suggesting that poverty is lower for participant by 0.14-0.21 for male and 0.17-0.27 for female.

A comparison of the ATT estimates from the endogenous switching regression and the PSM approach reveals that the estimates from the PSM are slightly lower, suggesting that the PSM may be underestimating the ATT. It is widely accepted that in the presence of hidden bias, PSM normally underestimates the average treatment effects, since matching only controls for observable characteristics.

Also presented in Table 8, the critical levels of gamma (Γ), at which the causal inference of significant participation effect may be questioned. Given that sensitivity analysis for insignificant effects is not meaningful, Rosenbaum bounds were calculated only for treatment effects that are significantly different from zero (Hujer et al., 2004). For example, the value of 1.50 for male participation implies that if households that have the same X-vector differ in their odds of participation by a factor of 50%, the significance of the participation on per head expenditure may be questionable. The lowest critical value of Γ is 1.20, whereas the largest critical value is 2.10. We can therefore conclude that even large amounts of unobserved heterogeneity would not alter the inference about the estimated effects, suggesting that the findings are generally insensitive to hidden bias.

Table 9. Indicators of matching quality before and after matching --------PSM results.

	Matching	Outcome variable	Pseudo-R2 (unmatched)	Pseudo-R2 (matched)	p-Value* (unmatched)	p-Value* (matched)	Median absolute bias (matched)	Median absolute bias (unmatched)	% bias reduction
Male	NNM	Per head expenditures	0.112	0.054	0.000	0.888	14.3	6.1	57.34
		Headcount	0.112	0.010	0.000	0.997	14.3	3.3	76.92
	Radius	Per head expenditures	0.112	0.054	0.000	0.888	14.3	6.1	57.34
		Headcount	0.112	0.029	0.000	0.982	14.3	5.1	64.34
	KBM	Per head expenditures	0.112	0.049	0.000	0.929	14.3	6.3	55.94
		Headcount	0.112	0.049	0.000	0.929	14.3	6.3	55.94
Female	NNM	Per head expenditures	0.299	0.197	0.000	0.189	26.9	12.2	54.6
		Headcount	0.299	0.081	0.000	0.995	26.9	7.7	71.43
	Radius	Per head expenditures	0.299	0.176	0.000	0.186	26.9	16.9	37.17
		Headcount	0.299	0.082	0.000	0.995	26.9	6.4	76.34
	KBM	Per head expenditures	0.299	0.157	0.000	0.26	26.9	13.3	50.56
		Headcount	0.299	0.143	0.000	0.947	26.9	15.5	59.09

Note:* p-Value of likelihood ratio test $(pr > \chi^2)$

The fourth and fifth columns in Table 9 present the pseudo-R2 from the propensity score estimation and from the re-estimation of the propensity score after matching on the matched samples for both male and female. The likelihood-ratio test of the joint significance of all the regressors in the probit model of propensity score estimation before and after matching and their corresponding p-values are presented in the sixth and seventh columns of the Table 9 for both male and female. The corresponding p-values of the likelihood-ratio test show that the joint significance of regressors on treatment status could always be rejected after matching. It was, however, never rejected before matching. The relatively low pseudo-R2 after matching and the p-values of the likelihood-ratio test of joint significance of the regressors imply that there is no systematic difference in the distribution of covariates between participants and non-participants after matching.

However, as indicated earlier, the main purpose of the propensity score estimation is not to obtain a precise prediction of selection into treatment but rather to balance the distributions of relevant variables in both groups. The balancing powers of the estimations are ascertained by considering the reduction in the median absolute standardized bias between the matched and the unmatched models. These median absolute standardized bias before and after matching are shown in the eighth and ninth columns of Table 9 for male and female, and the tenth column reports the total bias reduction obtained by the matching procedure. The estimates show substantial bias reductions for both male and female. Rosenbaum and Rubin (1985) suggested that a remaining standardized bias of 20% would be advisable.

6. CONCLUSIONS

This study evaluates the impact of non-farm work on consumption expenditure and poverty status in rural Pakistan. The study utilizes cross-sectional rural household level data collected in 2010 from a randomly selected sample of 341 households in Punjab province of Pakistan. The causal impact of non-farm work participation is estimated by utilizing propensity score matching and switching regression methods to assess robustness of the results. This helps in estimating the true welfare effect of non-farm work by controlling selection problem that normally occurs when observable and unobservable factors influence both on participation in non-farm work and outcomes such as per head expenditures and poverty status of household. The study provides separate estimates for males and females to address gender heterogeneity. The estimates show that although the participation rate of females was lower but their participation contributed more to improve welfare and reduce poverty as compared to their male counterparts.

Both the propensity score matching and switching regression results suggest that participants of non-farm work have significantly higher consumption expenditure and lower poverty than non- participants even after controlling for all confounding factors. The results from this study generally confirm the potential direct role of non-farm sector in improving rural household welfare and alleviating poverty in rural areas of developing countries.

The policy makers should be worried about substantial evidence of the inability of the poor to overcome existing economic and social entry barriers of non-farm activities. Particularly, women face more entry barriers to participation in non-farm work, so policy measures should target them to lower these barriers. Increased and stable household income through non-farm participation in turn smoothes consumption and reduces poverty. Government poverty reduction strategies should address the poor people, especially females to encourage them to engage in non-farm work in order to reduce poverty.

7. REFERENCES

Abdulai, A. and CroleRees, A. (2001). Determinants of income diversification among rural households in southern Mali. *Food Policy* 26: 437-452.

Abdulai, A. and Delgado, C. L. (1999). Determinants of time spent in non-farm employment by farmers in northern Ghana. American Jornal of Agricultural Economics 81: 117-130.

Abdulai, A. and Regmi, P. P. (2000). Estimating labor supply of farm households under nonseparability: Empirical evidence from Nepal. American Jornal of Agricultural Economics 22: 309-320.

Anwar, T., Qureshi, S. K. and Ali, H. (2004). Landlessness and rural poverty in Pakistan. The Pakistan Development Review 43: 855-874.

Babatunde, R. and Qaim, M. (2010). Impacts of off-farm income on food security and nutrition in Nigeria. Food Policy 35: 303-311.

Barrett, C. B., Reardon, T. and Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: Concepts, dynamics and policy implications. Food Policy 26: 315-331.

Barrett, C. B., Bezuneh, M., Clay, D. and Reardon, T. (2005). Heterogeneous constraints, incentives and income diversification strategies in rural Africa. Quarterly Journal of International Agriculture 44: 37-62.

Barrett, C., Sherlund, S. and Adesina, A. (2008). Shadow wages, allocative inefficiency, and labor supply in smallholder agriculture. Agricultural Economics 38: 21-34.

Chang, H., Mishra, A. and Livingston, L. (2011). Agricultural policy and its impact on fuel usage: Empirical evidence from farm household analysis. Applied Energy 88: 348-353.

Corral, L. and Reardon, T. (2001). Rural non-farm incomes in Nicaragua. World Development 29: 427-442.

Coudouel, A., J. Hentschel, Q. Wodon. (2002). Poverty measurement and analysis. In J. Klugman, editor, A Sourcebook for Poverty Reduction Strategies, Volume 1: Core Techniques and Cross-Cutting Issues. The World Bank, Washington, D. C.

De Janvry, A., Fafchamps, M., Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. The Economic Journal 101. PP 1400-1417.

Dehejia, R. H. and Wahba, S. (2002). Propensity score matching methods for non-experimental causal studies. The Review of Economics and Statistics 84: 151-161.

Dercon, S. (2002). Income risk, coping strategies and safety nets. The World Bank Research Observer 17: 141-166.

DiNardo, J. and Tobias, J. (2001). Nonparametric density and regression estimation. Journal of Economic Perspectives 15: 11-28.

Duflo, E., M. Kremer. and J. Robinson. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. American Economic Review, American Economic Association 101: 2350- 2390.

El-Osta, H. (2011). The impact of human capital on farm operator household income. Agricultural and Resource Economics Review 40: 95-115.

Ellis, F. (2000). Rural livelihoods and diversity in developing countries. Oxford University Press.

Fafchamps, M., Udry C. and Czukas, K. (1998). Drought and saving in west Africa: Are livestock a buffer stock? Journal of Development Economics 55: 273-305.

Feder, G., Lau. L. J., Lin, J. Y. and Luo, X. (1990). The relationship between credit and productivity in Chinese agriculture: A Micro-economic model of disequilibrium. American Journal of Agricultural Economics 72: 1151-1157.

Freeman, H. A,. Ehui, S. K. and Jabbar, M. A. (1998). Credit constraints and smallholder dairy production in the east African high lands: Application of a switching regression model. Agricultural Economics 19: 33-44.

Fuglie, K. O. and Bosch, D. J. (1995). Implications of soil nitrogen testing: A switching regression analysis. American Journal of Agricultural Economics 77: 891-900.

Government of Pakistan. (2009). Economic Survey 2008-09. Finance Division, GOP, Islamabad.

Haggblade, S., Hazell, P. and Reardon, T. (2010). The rural nonfarm economy: Prospects for growth and poverty reduction. World Development 38: 1429-1441.

Heckman, J., H. Ichiwura. and P. E. Todd. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. Review of Economic studies 64: 605-654.

Holden, S., Shiferaw, B. and Pender, J. (2004). Non-farm income, household welfare, and sustainable land management in a less-favoured area in the Ethiopian highlands. Food Policy 29: 369–392.

Huffman, W. E. (1991). Agricultural household models: Survey and critique," in M.C. Hallberg, J.L. Findeis, and D. A. Lass. (eds), Multiple Job-Holding among Farm Families. Iowa State University Press, 79-111.

Hujer, R., Caliendo, M. and Thomsen, S. L. (2004). New evidence on the Job creation schemes in Germany –A matching approach with threefold heterogeneity. Research in Economics 58: 257-302.

Kurosaki, T. and Humayun, K. (2004). Effects of Human Capital on Farm and Non-farm Pro-ductivity and Occupational Stratification in Rural Pakistan. COE Discussion Paper 46, Institute of Economic Research, Hitotsubashi University, Tokyo.

(http://hi-stat.ier.hit-u.ac.jp/english/research/discussion/index.html).

Lamb, R. (2003). Fertilizer use, risk and off-farm labor markets in the semi-arid tropics of India. American Journal of Agricultural Economics 85: 359-371.

Lanjouw, J. and Lanjouw, P. (2001). The rural non-farm sector: Issues and evidence from developing countries. Agricultural Economics 26: 1-23.

Lanjouw, P., Quizon, J. and Sparrow, R. (2001). Non-agricultural earnings in peri-urban areas of Tanzania: Evidence from household survey data. Food Policy 26: 385-403.

Lee, L. F. (1978). Unionism and wage rates: A simulation equation model with qualitative and limited dependent variables. International Economic Review 19: 415-433.

Lee, W. (2008). Propensity Score Matching and Variations on the Balancing Test. In: 3rd Conference on Policy Evaluation, ZEW, Mannheim (Germany), 27-28 October.

Leuven, E. and Sianesi, B. (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. Last accessed 12 March 2009.

Lokshin, M. and Z. Sajaia. (2004). Maximum likelihood estimation of endogenous switching regression models. Stata Journal 4: 282-289.

Maddala, G. S. (1983). Limited Dependent and Qualitative Variables in Econometrics. Cambridge University Press, New York.

Malik, S. J. (2008). Rethinking development strategy-The importance of the rural non farm economy in growthand poverty reduction in Pakistan. The Lahore Journal Of Economics, Special edition: 189-204.

Mishra, A., Morehart, M., El-Osta, H., Johnson, J. and Hopkins, J. (2002). Income, Wealth, and Well-being of Farm Operator Households. Agricultural Economics Report 812, US Department of Agriculture, Economic Research Service, Washington.

Odhiambo, J. and Magandini, N. (2008). An assessment of the use of mineral and organic fertilizers by smallholder farmers in Vhembe district, Limpopo province, south Africa. African Journal of Agricultural Research 3: 357-362.

Oseni, G.and Winter, P. (2009). Rural nonfarm activities and agricultural crop production in Nigeria. Agricultural Economics 40: 189-201.

Owusu, V., Abdulai, A. and Rahman, S. (2011). Non-farm work and food security among farm households in northern Ghana. Food Policy 36: 108-118.

Pagan, A. and Ullah, A. (1999). Nonparametric Econometrics. Cambridge University Press, Cambridge.

Pakistan Encyclopedia. (2012). http://en.wikipedia.org/wiki/Punjab,_Pakistan

Rosenbaum, P. R. (2002). Observational Studies. Springer, New York.

Rosenbaum, P. R. and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika 70: 41-45.

Rosenbaum, P. R. and Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. The American Statistican 39: 33-38.

Rosenzweig, M. R. (1980). Neoclassical theory and the optimizing peasant: An econometric analysis of market family labor supply in a developing country. Quarterly Journal of Economics 94: 31-55.

Rusu, M. (2002). Land Fragmentation and Land Consolidation in Romania, Paper presented at the International CEEC Symposium, 25–28 February, Munich.

Sahn, D. E. and Alderman, H. (1988). The effects of human capital on wages, and the determinants of labor supply in a developing country. Journal of Development Economics 29: 157-83.

Sianesi, B.(2004). Evaluation of the active labor market programmes in Sweden. The Review of Economics and Statistics 86: 133-155.

Singh, I., Squire, L., Strauss, J. (1986). The basic model: Theory, empirical results and policy conclusions. In Singh, I., Squire, L and Strauss, J.(eds), Agricultural Households Models. Baltimore: Johns Hopkins University Press.

Skoufias, E. (1994). Using shadow wages to estimate labor supply of agriculture households, American Journal of Agricultural Economics 76: 215-227.

Smith, J. and Todd, P. (2005). Does matching overcome lalonde's critique of nonexperimental estimators? Journal of Econometrics 125: 305-353.

Tan, S., Heerink, N. and Kruseman, G. (2008). Do fragmented landholdings have higher production costs? Evidence from rice farmers in northeastern Jiangxi province, P.R. China. China Economic Review 19: 347-358.

Timmer, P. (1988). The agricultural transformation. In Chenery, H. and Srinivasan, T. N. (eds.), The Handbook of Development Economics 1. North Holland Press, Amsterdam.

Verpoorten, M. (2009). Household coping in war- and peacetime: Cattle sales in Rwanda, 1991-2001. Journal of Development Economics 88: 67-86.

Wan, G. and Cheng, E. (2001). Effects of land fragmentation and returns to scale in the Chinese farming sector. Applied Economics 33: 183-194.

Weiss, C.R. (1997). Do they come back again? The symmetry and reversibility of off-farm employment European Review of Agricultural Economics 24: 65-84.

Appendix

Table A-1. Probit estimates of propensity score for male's non-farm employment participation.

Variable	Coefficient	Standard error	Z-value
Head	-0.566	0.203	-2.79**
AgeHead	0.006	0.007	0.92
HHSizeOver14	0.127	0.033	3.81***
Ch0L05	-0.004	0.062	-0.06
Child14	-0.006	0.013	-0.45
AvEdMale	0.128	0.053	2.43**
livstk	-0.477	0.224	-2.13**
TCultiLand	-0.001	0.002	-0.27
Dis0vill	-0.000	0.008	-0.03
BorowMon	-0.29	0.177	-0.16
UppCaste	-0.222	0.179	-1.24
LowCaste	0.015	0.251	0.06
Fac0Mil	-0.127	0.167	-0.76
Location1	0.167	0.332	0.50
Location2	0.738	0.323	2.29**
Location3	0.436	0.369	1.18
Location4	-0.550	1.009	-0.54
Location5	0.212	0.356	0.64
Constant	-0.272	0.559	-0.49
Pseudo-R2	0.1118		
Log likelihood	-201.91614		

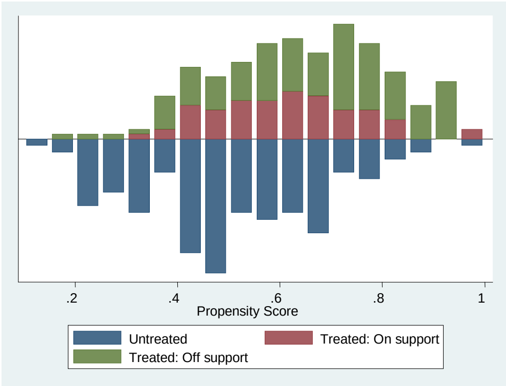
Note: Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.

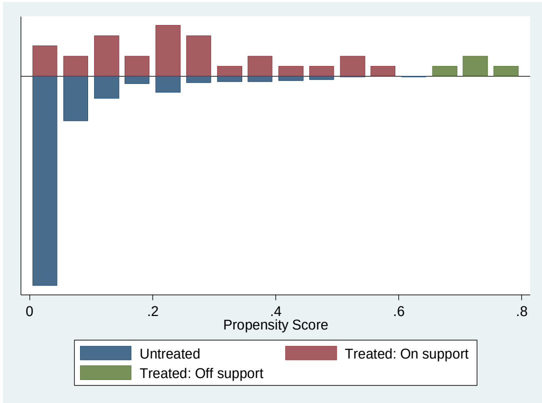
Table A-2. Probit estimates of propensity score for female's non-farm employment participation.

Variable	Coefficient	Standard error	Z-value
Head	-0.350	0.304	-1.15
AgeHead	0.024	0.011	2.09**
HHSizeOver14	0.222	0.054	4.14***
Ch0L05	-0.217	0.113	-1.91*
Child14	-0.003	0.022	-0.14
AvEdFemale	0.042	0.028	1.45
livstk	-0.268	0.318	-0.84
TCultiLand	-0.011	0.008	-1.34
Dis0vill	-0.070	0.023	-3.08***
BorowMon	0.309	0.257	1.20
UppCaste	0.323	0.317	1.02
LowCaste	0.638	0.407	1.57
Fac0Mil	0.328	0.253	1.30
Location1	0.821	0.766	1.07
Location2	0.835	0.750	1.11
Location3	0.265	0.685	0.39
Location4	-0.550	1.009	-0.54
Location5	0.105	0.691	0.15
Constant	-3.005	1.036	-2.90***
Pseudo-R2	0.2979		
Log likelihood	-74.471862		

Note: Significance of t-statistics of mean difference is at the *10%, **5% and ***1% levels.

Male's non-farm work participation

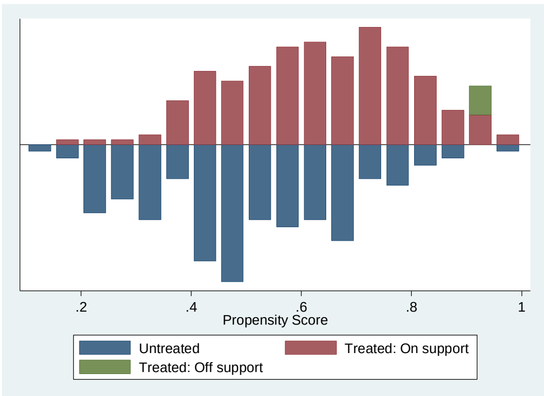


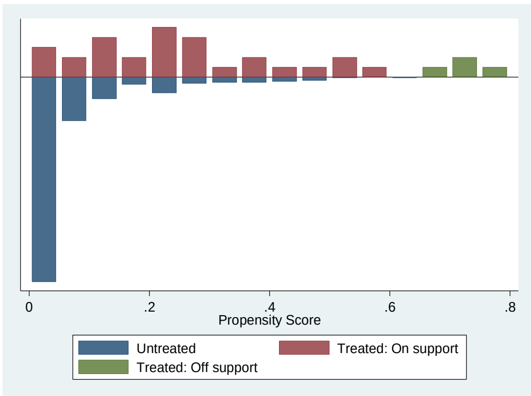


Impact on per head expenditure

Impact on poverty

Female's non-farm work participation





Impact on per head expenditure

Impact on poverty

Figure 1: Propensity score distribution and common support for propensity score estimation.