The Empirical Econometrics and Quantitative Economics Letters
ISSN 2286 – 7147
Volume 2, Number 3 (September 2013), pp. 27 - 36.

Policy evaluation of rural labor force training program: Evidence from autonomous minority nationality areas in Southwestern frontier region of China

Jiachun Xie[1], Xingxu Li[2], Songsak Sriboonchitta[3] and Shihti Yu[4]

[1,2] Statistics and Mathematics College,
Yunnan University of Finance and Economics, China.
Email:springchinathai@hotmail.com

[3]Faculty of Economics, Chiang Mai University, Thailand
[4]Department of Quantitative Finance, National Tsing Hua University, Taiwan

ABSTRACT

This paper evaluates the Rural Labor Force Training Program effects in the case of autonomous minority nationality areas in southwestern frontier region of China by using China's Rural Household Survey (RHS) micro data. MTE-based semi-parametric method is used to estimate parameters of interest, i.e., ATE, TT and TUT. The decline trend of MTE value respect to U_T, which represents unobservable characteristics of resistance to participating in training, approves heterogeneity of unobservable characteristics given observed variables of farm households. The results of ATE, TT and TUT shows, households who already participate in training gain most than the randomly selected households if they are participants and those who are not participate in training. It also demonstrates that our Rural Labor Force Training Program is effective in Honghe Prefecture and Dehong Prefecture.

Keywords: Treatment effect, Training program evaluation, MTE-based semi-parametric method, Minority nationality, China

1. Introduction

There is an increasing gap between rural and urban areas, since China has turned into her industrial and urbanization development in the last thirty years. Poverty and backwardness of rural areas caused more and more attentions. In the hope of enhancing farmers' productive skills, vocation and employment ability, in return, increasing their further income and promoting rural labor force migration, the Rural Labor Force Training Program has been initiated by the Chinese government in 2004. Policy makers or scholars may wonder what the effects of the program in place on participants and nonparticipants are. Little literature in training policy evaluation existed since the difficult availability of micro survey data in China. As micro survey data getting into available in recent years, some scholars start to estimate the training policy effects, but still few.

If we want to estimate the training policy effects, we need to use micro data. Heterogeneity and missing counterfactual states are central features of micro data. Due to unobserved heterogeneity, observationally identical people make different choices (Heckman and Li, 2004). Control the observable characteristics, such as income, age, education level, unobservable heterogeneity may affect people acting in different behaviors. Learning ability, challenge courage and adaptive capacity also affect peoples' decision, and they are unobservable. Missing data arises when evaluate problem for social programs. If a farmer chooses to participate in training, we can observe his or her income with the state of participation, but we cannot observe his or her income without participating in training at the same time. We cannot observe the outcomes of all possible choices for the same person at the same time. In order to overcome these problems, we need to construct a counterfactual frame, which is composed of untreated group members who share similar observable characteristics of those who are actually participated in training (LaLonde, 1995; Caliendo and Kopeinig, 2008).

The difficulty of policy effects estimation is how to construct a counterfactual frame in consideration of heterogeneity. This paper uses micro data obtained from China's Rural Household Survey (RHS) of Honghe Prefecture and Dehong Prefecture in Yunnan province, to estimate the Rural Labor Force Training Program effects in the autonomous minority nationality areas in southwestern frontier region of China. Our work builds on previous research by Heckman and Vytlacil (1999, 2005), Heckman (2001), Heckman and Urzua and Vytlacil(2006a, 2006b), Heckman and Li(2004) and Xiang Zhou and Yu Xie(2011), which develops a MTE-based semi-parametric method and their applications.

This paper is organized as: the MTE-based semi-parametric method will be introduced in the next section. Section 3 is data set and empirical results. The last section 4 is concluding remarks.

1. MTE-based semi-parametric Method

Normally, it is easy to find in other literature that the individual person is the unit of studying target. Before the econometric estimation model is being introduced, we have to expatiate that we take the farm-household as our studying target due to the survey data sample characteristics.

Suppose that there is a target population N has being studied, let i represent the ith member in N. The choice indicator T denotes the training status, $T = 1$ if a farm household make the decision that participate in training, that is also say a household is treated. Here when a farm household participates in training it means any one family member participated in training. $T = 0$, denotes a household that has not participated in training, that is no family member participates. In employment and training programs, income Y is a typical outcome variable. We further denote Y_1 as the potential income for the participator, Y_0 for the non-participator. $Y_1 - Y_0$ is the *treatment effect* of the training policy. The difficulty for evaluation is, for a given household, that we are observing either Y_1 or Y_0 at the same time, but not both. Then the observable result Y can be written in a general way,

$$Y = TY_1 + (1 - T)Y_0 \qquad (1)$$

It is convenient to establish three equations to demonstrate how to estimate the *treatment effect* and its relative parameters of interest. The first two equations are the potential outcome equations. We assume that:

$$Y_0 = \beta_0 X + U_0 \qquad \text{if} \quad T = 0 \qquad (2)$$

$$Y_1 = \beta_1 X + U_1 + \varphi \qquad \text{if} \quad T = 1 \qquad (3)$$

Vector X represents the observable variables, U_0 and U_1 represents the unobserved variables in the potential outcome equations, and φ represents the benefit associated with treatment ($T = 1$).

Substitute (2) and (3) into (1), we obtain:

$$Y = T\,(\beta_1 X + U_1 + \varphi\,) + (1 - T)\,(\beta_0 X + U_0)$$

$$= \beta_0 X + (\beta_1 X - \beta_0 X + U_1 - U_0 + \varphi)T + U_0 \qquad (4)$$

We can take equation (4) in another expression:

$$Y = \alpha + \beta_i T + \varepsilon \qquad (5)$$

where $\alpha = \beta_0 X$, $\beta_i = Y_1 - Y_0 = \beta_1 X - \beta_0 X + U_1 - U_0 + \varphi$, $\varepsilon = U_0$. Here β_i is the heterogeneous effect of training for household i. When $\beta_1 \neq \beta_0$ or $U_1 \neq U_0$, β_i varies in the population.

Another equation is the selection equation, the farmers decide whether or not participate in the training program ($T = 1$ or $T = 0$) based on the latent variable T^*,

$$T^* = \gamma Z - V$$

$$T = 1 \qquad ,\text{if} \quad T^* > 0 \qquad (6)$$

$$= 0 \qquad ,\text{otherwise}$$

where T^* is a latent variable represents the net benefit of participation. Z is an observed vector of variables and predict the treatment probability and may include some X, V for unobserved variables. Finally, we assume that error term (U_0, U_1, V) have zero means and are jointly independent of X and Z.

We also can rewrite the treatment selection equation (6) in the following form:

$$\widetilde{T^*} = P(Z) - U_T$$

$$T = 1 \quad , \text{if} \quad \widetilde{T^*} > 0 \qquad (7)$$
$$= 0 \quad , \text{otherwise}$$

Where U_T represents unobservable characteristics of resistance to participating in training and $U_T = F_V(V)$. $P(Z) = \Pr(T = 1|Z = z) = F_V(\gamma Z)$ denotes the propensity score of being treated given observable variables of Z. Without lost of generality we assume that $U_T \sim Unif\ [0,1]$ (Heckman and Vytlacil, 1999; Heckman, Urzua and Vytlacil, 2006a, 2006b).

Based on the previous outcome equations and treatment selection equation, the Marginal Treatment Effect (MTE) can be defined as:

$$\text{MTE}(x, u_T) = \text{E}(Y_1 - Y_0|X = x, U_T = u_T) \qquad (8)$$

The MTE is introduced by the evaluation literature by Björklund and Moffitt(1987). It is the average gain to people who are indifferent to participating in training $T = 1$ or nonparticipating $T = 0$ give X and Z.

The other traditional mean treatment parameters are:

$$\text{ATE} = \text{E}(Y_1 - Y_0|X) \qquad (9)$$
$$\text{TT} = \text{E}(Y_1 - Y_0|X, T = 1\ ) \qquad (10)$$
$$\text{TUT} = \text{E}(Y_1 - Y_0|X, T = 0\ ) \qquad (11)$$

The Average Treatment Effect (ATE) is defined for the whole population. ATE evaluate the average difference between a set of members in N that are randomly selected for treatment and another set of members that are randomly selected for control, given a vector of observable variables of X.

The Treatment on the Treated (TT) means to the average difference by treatment status for these people who are actually treated.

The Treatment effect of the Untreated (TUT) refers to the average difference by treatment status for these who are not treated. The MTE can be used to construct ATE, TT and TUT with different weights, see Heckman and Vytlacil (1999, 2000, 2006a, 2006b; Heckman and Li, 2004).

$$\text{ATE} = \int_0^1 \text{MTE}(u_T)\, du_T$$
$$\text{TT} = \int_0^1 \text{MTE}(u_T)\, h_{TT}(u_T) du_T$$
$$\text{TUT} = \int_0^1 \text{MTE}(u_T)\, h_{TUT}(u_T) du_T$$

where the weights are:

$$h_{ATE}(u_T) = 1$$
$$h_{TT}(u_T) = \frac{[\int_{u_T}^1 f_P(P) d_P]}{E(P)}$$
$$h_{TT}(u_T) = \frac{[\int_0^{u_T} f_P(P) d_P]}{1 - E(P)}$$

where f_P is the density of $P(Z)$.

MTE is the essential for semi-parametric method. Before calculate MTE, firstly we have to get the propensity score $P(Z) = \Pr(T = 1|Z = z) = F_V(\gamma Z)$. Normally we can obtain $P(Z)$ through *Probit* or *Logit* model, in our research we choose *Logit* model. Given $X = x$ and $P(Z) = p$, we can get the observable outcome Y:

$$\mathrm{E}(Y|X = x, P(Z) = p)$$
$$= \mathrm{E}(\beta_0 x + (\beta_1 x - \beta_0 x + U_1 - U_0 + \varphi)T + U_0|X = x, P(Z) = p)$$
$$= \mathrm{E}(\beta_0 x + (\beta_1 - \beta_0)xT + (U_1 - U_0)T + \varphi T + U_0|X = x, P(Z) = p)$$
$$= \beta_0 x + (\beta_1 - \beta_0)xp + K(p) \qquad (12)$$

where

$$K(p) = \varphi p + \mathrm{E}(U_1 - U_0|T = 1, P(Z) = p)p + \mathrm{E}(U_0|P(Z) = p) \qquad (13)$$

Before MTE has been introduced, we assumed that the error term (U_0, U_1, V) have zero means and are jointly independent of X and Z. If we relaxing the assumption of normality and then using a polynomial of the propensity score, that can be induced to MTE-based semi-parametric method. Under this method, function $K(p)$ can be expressed by a polynomial of p. Let ϑ denotes the degree of the polynomial, we obtain:

$$\mathrm{E}(Y|X = x, P(Z) = p) = \beta_0 x + (\beta_1 - \beta_0)xp + \varphi p + \sum_{i=1}^{\vartheta} \phi_i p^i \qquad (14)$$

Differentiating the equation (14) with respect to p, we obtain MTE:

$$\mathrm{MTE} = \frac{\partial \mathrm{E}(Y|X=x, P(Z)=p)}{\partial p} = (\beta_1 - \beta_0)x + \varphi + \sum_{i=1}^{\vartheta} i\phi_i p^{i-1} \qquad (15)$$

This expression provides a link between MTE and $\mathrm{E}(Y|X = x, P(Z) = p)$.

Followed by Heckman and Vytlacil (2006b) there are four steps involved in semi-parametric method:

Step1: Establish model:

$$Y = \beta_0 X + (\beta_1 - \beta_0)XP(Z) + \sum_{i=1}^{\vartheta} \psi_i P(Z)^i + \varsigma$$

where we assume $\mathrm{E}(\varsigma|X = x, P(Z) = p) = 0$. Fit local linear regressions of Y, X and XP on P, and extract their residuals R_Y, R_X and R_{XP}.

Step2: Regress R_Y on R_X and R_{XP} using least squares to estimate β_0 and $\beta_1 - \beta_0$, and denote its residuals as R_Y^*.

Step 3: Regress R_Y^* on P using standard nonparametric techniques to estimate $K(p)$. We rewrite:

$$\tilde{Y} = K\big(P(Z)\big) + \tilde{v}$$

where $\tilde{Y} = Y - X'\widehat{\beta_0} - (X'(\widehat{\beta_1 - \beta_0}))P(Z)$ and assume $\mathrm{E}\big(\tilde{v}\big|P(Z)\big) = 0$. $K(p)$ can interpreted as the conditional expectation $\mathrm{E}\big(\tilde{Y}\big|P(Z) = p\big)$.

Then,

$$\frac{\partial \mathrm{E}(\tilde{Y}|P(Z)=p)}{\partial p} = \frac{\partial K(p)}{\partial p}$$

Step4: Finally, construct $\text{MTE}(x, u_T)$,

$$\widehat{\text{MTE}}(X = x, U_T = u_T) = x'\left(\widehat{\beta_1 - \beta_0}\right) + \frac{\partial \widehat{K(p)}}{\partial p}\ |_{p=u_T}$$

In our research, R software of version 2.13.2 is used for our analysis. Library of *Hmisc, survival*, *splines*, *stats* and *sampleSelection* is used.

3. Data set and Empirical results

3.1 Data sources and selection of the sample

The data for our study were obtained from The China's Rural Household Survey (RHS) of Honghe Prefecture and Dehong Prefecture done by local State Statistical Bureau (SSB) offices in Yunnan province. The data appears to be of good quality and bunches of information about rural household income, consumption, production, accumulative and social behaviors. Two-stage sample was selected in each prefecture. The first stage involved the selection of 348 villages from 13 counties of Honghe Prefecture and 5 counties of Dehong Prefecture. In the second stage, it was involved the stochastic sampling of households from the selected villages. There were two main methods have been adopted for collecting data. One is the sampled households fill in a daily diary on expenditures and other relative information. Another one was visiting survey. Sampled households were visited on every month by an interviewer to check the diaries and collect data.

Our sample data is a sequence of cross sections since 2005 and is ongoing. The training information enrolled into this survey only from 2006 to 2008. In this research, baseline variables are measured in 2006 period and income variable also measured in 2007, since the causal inference of training. Finally, totally 2053 households who with positive annual net income and participated in training only in 2006 not in 2007 have been selected in our research.

3.2 Measures

Based on Mincer income model (1974), not only human capital investment has been taken into consideration but also material resources capital investment, farm household characteristics and living village characteristics. We take households who participate in training as the treated group, who not participate in training as the control group. In our research, per capita net income Y of 2007 has been selected as our outcome variable and been taken logarithm. 2053 households with positive net income in 2006 and 2007, then we obtain 570 households in the treated group and 1483 households in the control group. The average rate of training participation is 27.76%. The following Table 1 and Table 2 are variables definition and sample statistical description respectively.

Table 1 Variables' Definition

Variables	Definition
Y	Per capita net income in 2007 and take logarithm
Lin06	Per capita net income in 2006 and take logarithm
Edu*	Highest education levels, 5 for college and above, 4 for special technical school or senior high school, 3 for junior high school, 2 for primary school, 1 for illiteracy
Exp	Average year of work experience = average age - Average years of education-6
Expsq	Square of exp
T	Binary variable for participation in training, 1 indicate participation, 0 otherwise
Avland	Per capita land scale
Lnassets*	Capital assets of farming per capita, take logarithm
Cadre*	Binary variable, 1 indicate household with village official family members
Labor	Number of labor force
Emlabor*	Number of labor force that is migrating into other sectors or other areas
Residents	Number of residents
Lvin06	Average net income at a village level
Burden*	Calculated by (Residents- Labor)/Labor
Children	Number of children under 15 years old
Min*	Binary variable, 1 indicate a minority group village, 0 otherwise
Geo*	Topography indicator, whether the village is on plains, or in hills or mountains
Dis*	Distance between village and county
Vitr06*	Average ratio of the participants at a village level

Note:* indicate variables construct *Z* for treatment selection model.

Table 2 Sample Statistical Description

	Total Sample		Treated Group		Control Group	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
	Obs=2053		Obs=570		Obs=1483	
Y	7.590	0.737	7.942	0.714	7.455	0.701
Lin06	7.444	0.741	7.758	0.695	7.324	0.723
Edu	2.787	0.789	3.095	0.722	2.668	0.782
Exp	23.369	6.728	22.835	6.511	23.575	6.800
T	0.276	0.448	1.000	0.000	0.000	0.000
Avland	1.358	1.132	1.387	1.142	1.346	1.129
Lnassets	6.185	2.131	6.977	1.993	5.880	2.105
Cadre	0.105	0.306	0.163	0.370	0.082	0.275
Labor	2.926	1.125	2.951	1.074	2.916	1.144
Emlabor	0.122	0.445	0.179	0.503	0.100	0.419
Residents	4.362	1.339	4.186	1.221	4.430	1.376
Lvin06	7.444	0.510	7.465	0.467	7.436	0.526
Burden	0.626	0.603	0.529	0.532	0.663	0.624
Children	0.805	0.845	0.728	0.809	0.835	0.856
Min	0.755	0.430	0.521	0.500	0.845	0.362
Geo	2.576	0.739	2.125	0.859	2.750	0.603
Dis	4.561	0.851	4.328	0.903	4.651	0.813
Vitr06	0.278	0.407	0.695	0.319	0.038	0.099

As is shown in Table 2, there are some lags between the treated group and the control group. Income of participants is higher than the nonparticipants' in 2006 and 2007. Treated group people have higher education level. Material resources capital investment, such as per capita land scale and capital assets of farming per capita in treated group are over to the control group. Differences also appear in family characteristics and village characteristics.

3.3 Empirical results

Before we estimate treatment parameters ATE, TT and TUT, we need to establish a treatment selection model to obtain propensity score, as has mentioned before, we choose *logit* model for our aim and coefficient estimates presented in Table 3. All the variables in Table 3 are the determinants of the probability of participating in training.

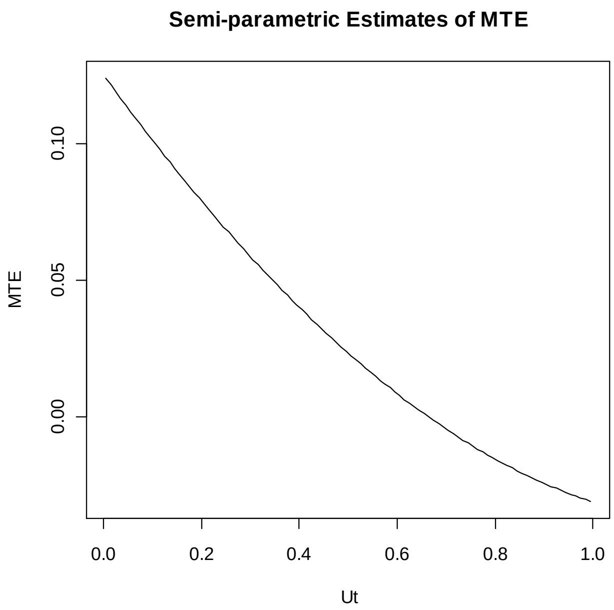
Table 3 Estimated *Logit* Model for Participation in Training

Variables	Coefficient	Std. Error
Edu2	0.61618	0.48639
Edu3	1.18285*	0.47872
Edu4	1.67072***	0.49260
Edu5	2.65236***	0.67487
Lnassets	0.26842***	0.03512
Cadre	0.44422*	0.17809
Emlabor	0.11146	0.13047
Burden	-0.08016	0.10320
Min1	-0.87936***	0.14857
Geo2	-0.19782	0.20451
Geo3	-1.34880***	0.18151
Dis2	2.80532**	0.85673
Dis3	4.31238***	0.83708
Dis4	3.08600***	0.81400
Dis5	3.42785***	0.82124
Vitr06	0.26269.	0.14170

Note: Signif. codes: 0 '***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1 ;' ' 1.

The following Figure 1 shows the semi-parametric estimates of MTE which is a declining trend with respect to u_T, i.e., the unobserved resistance to participating in training. Back to see equation (7), that means the more lower unobservable u_T the more higher probability of participation, i.e., $P(Z)$. The declining trend between MTE and u_T, means that people with higher return of participating in training are more likely participate in training (with lower u_T). Furthermore, the MTE varies significant with respect to the heterogeneity. Returns can vary from 0.15 for low u_T farm households to -0.03 for high u_T farm households.

Figure 1 Marginal Treatment Effects (MTE) with unobservable u_T



Finally, three treatment parameters ATE, TT and TUT are presented in Table 4.

Table 4 Estimation of ATE, TT and TUT (Unit: %)

	ATE	TT	TUT
Estimation	3.11	7.88	1.67

The treatment effect of participants is 7.88% while average treatment effect for a randomly selected household is 3.11% which demonstrates that our rural labor force training program do really improve their income. Turn to review TUT, lower than TT, which indicate that nonparticipants gain less than those who participated in training.

This conclusion is against to the result of Aakvik, Heckman and Vytlacil, 2005. In their research, they estimated the Norwegian Vocational Rehabilitation training programs effect by using Latent Variable Model and Factor Structure Model, and finally got a negative estimated average effect of the training, i.e., ATE=-0.014, TT=-0.11.

4. Concluding Remarks

This paper uses micro data obtained from China's Rural Household Survey (RHS) of Honghe Prefecture and Dehong Prefecture in Yunnan province, to estimate the Rural Labor Force Training Program effects in the autonomous minority nationality areas in southwestern frontier region of China. Based on the MTE-based semi-parametric method, we get a decline trend of MTE respect to u_T. This varies decline trend approves heterogeneity of unobservable characteristics given observed variables of farm households. The main aim of MTE-based semi-parametric method application is to answer which groups of farm households benefit most from participation in training. Another one is does our Rural Labor Force Training Program really effective? The three treatment parameters ATE, TT and TUT shows, households who participate in training gain most than the randomly selected households if they are participants and those who are not participate in training. It also demonstrates that our Rural Labor Force Training Program is effective in Honghe Prefecture and Dehong Prefecture.

ACKNOWLEDGMENT

This research was supported by The National Natural Science fund to the project of "Research of Western Minority Nationality Areas Rural Labor Force Training Program Policy Effects Evaluation and Policy Optimization: Evidence from Yunnan Province of China"(NO.71263055) and Yunnan University of Finance and Economics Research Foundation Project to the project of "Farmers' Income Differential and Anti-Poverty Policy Reach in Autonomous Minority Nationality Areas in Southwestern Frontier Region of China "(NO.YC10A001)

REFERENCES

Arild Aakvik, Heckman, James J. and Edward Vytlacil. 2005. Estimation treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs. Journal of Econometrics, 125:15-51.

Heckman, James J. and Xuesong Li. 2004. Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China. Pacific Economic Reviw 9:155-71.

Heckman, James J. and Edward Vytlacil. 1999. Local Instrumental Variable and Latent Variable Models for Identifying and Bounding Treatment Effects. Proceedings of the National Academy of Sciences,96:4730-4734.36.

Heckman, James J. and Edward Vytlacil. 2005. Structural Equation, Treatment Effects, and Economic Policy Evaluation. Econometrica, Vol.73, No.3, 669-738.

Heckman, James J. and Edward Vytlacil. 2000. The Relationship Between Treatment Parameters within a Latent Variable Framework. Economics Letter, 66:1,33-39.

Heckman, James J..2001. Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture. Journal of Political Economy 109:673-748.

Heckman, James J., Sergio Urzua and Edward Vytlacil. 2006a. Understanding Instrumental Variables in Models with Essential Heterogeneity. The Review of Economics and Statistics 88:389-432.

Heckman, James J., Sergio Urzua and Edward Vytlacil. 2006b. Estimation of Treatment Effects under Essential Heterogeneity. Retrieved October 12, 2010(http://jenni.uchicago.edu/underiv/documentation_2006_03_20.pdf)

Marco Cliendo and Sabine Kopeinig. 2008. SOME PRACTICAL GUIDANCE FOR THE IMPLEMENTATION OF PROPENSITY SCORE MATCHING. Journal of Economic Survey, Vol.22, No.1,pp.31-72

Robert J. LaLonde. 1995. The Promise of Public Sector Sponsored Training Programs. Journal of Economic Perspectives 9,148-168.

Xiang Zhou and Yu Xie. 2011. Propensity-Score-Based Method versus MTE-Based Methods in Causal Inference. Population Studies Center Research Report.