The average treatment effect and average partial effect in nonlinear models

Nguyen Viet, Cuong

28 March 2007

The average treatment effect and average partial effect in nonlinear models

Nguyen Viet Cuong[1]

Abstract

In the literature on program impact evaluation, the popular impact parameters can the average treatment effect, the average treatment effect on the treated, the average partial effect, and the average partial effect on the treated. In empirical studies, these parameters are not always presented and estimated clearly. In addition, when outcome functions are nonlinear, the estimation of these parameters is not straightforward. This paper discusses the estimation of these parameters in nonlinear models of outcomes and illustrates the estimation in an example of a micro-credit program in Vietnam.

Keywords: Average treatment effect, average partial effect, treatment effect, impact evaluation, nonlinear models.

JEL classification: C13; C21; H43; J41

[1] National Economics University, Hanoi, Vietnam; Wageningen University, the Netherlands.
Email: c_nguyenviet@yahoo.com

1. Introduction

There are several widely-used parameters in impact evaluation of a program. When the program variable is a dummy one, parameters of interest can be the average treatment effect (ATE) or the average treatment effect on the treated (ATT). ATE measures program impact on expected outcome, while ATT measures program impact on expected outcome of the program's participants. When the program variable is a continuous one, we can define a parameter of the average partial effect (APE) which measures the change in the expected outcome due to a small change in the program level. Similarly, the average partial effect on the treated (APET) measures the change in the expected outcome of the program participants due to a small change in the program level.

Although the above parameters are discussed thoroughly in the impact evaluation literature, they are not presented clearly in empirical studies. In nonlinear models, they are not directly estimated, and standard errors of the estimated parameters are not calculated by standard statistical software. Meanwhile, coefficients of program variables might not have economic meaning or can have misleading explanation. For example, in evaluation of a micro-credit program in Vietnam that will be presented the fourth section, we run linear regression of log of per capita income on loan size and other explanatory variables. The coefficient of the loan size is estimated at 0.00002. This looks rather small, and a quick interpretation of the estimate is that an increase of 1 VND in loan size leads to an increase of 0.002% in per capita income.[2] This explanation does not give clear economic meaning. Thus after running the log-linear regression, we should estimate the parameters of program impacts for per capita income instead of logarithm of income. In this example, the estimate of APET is equal to 0.06. It means that an increase of 1 VND in loan size leads to an average increase of 0.06 VND in per capita income of the program participants.

In this paper, we will discuss the estimation of ATE, ATT, APE and APET in nonlinear models. The paper is structured in 5 sections. The second section presents the impact evaluation parameters. The third section presents estimation of these parameters. In the next section, the estimation method is illustrated in an empirical study. Finally, the fifth section concludes.

2. Parameters of interest

2.1. Potential outcome framework

[2] VND is Vietnamese currency. 1 USD is approximately equivalent to 16 000 VND (January 2007).

The case of a binary program

The main objective of impact evaluation of a program is to assess the extent to which the program has changed outcomes for subjects. In other words, impact of the program on the subjects is measured by the change in welfare outcome that is attributed only to the program. In the literature on impact evaluation, a broader term “treatment” is sometimes used instead of program/project to refer to intervention whose impact is evaluated.

To make the definition of impact evaluation more explicit, suppose that there is a program assigned to some people in a population *P*. For simplicity, let’s assume that there is a single program, and denote by *D* the binary variable of participation in the program, i.e. $D=1$ if she/he participates in the program, and $D=0$ otherwise. *D* is also called the variable of treatment status. Further let Y denote the observed value of the outcome. This variable can receive two values depending on the participation variable, i.e. $Y=Y_1$ if $D=1$, and $Y=Y_0$ if $D=0$.[3] These outcomes are considered at a point in time or over a period of time after the program is implemented.

The impact of the program on the outcome of person i is measured by:

$$\Delta_i = Y_{i1} - Y_{i0}, \tag{1}$$

which is the difference in outcome between the program state and the no-program state. The problem is that we cannot observe both terms in equation (1) for the same person. For those who participated in the program, we can observe only Y_1, and for those who did not participate in the program we can observe only Y_0.

In the literature on program impact evaluation, two popular parameters are the Average Treatment Effect (ATE), and the Average Treatment Effect on the Treated (ATT) (Heckman, et al., 1999).[4]

ATE is the expected impact of the program on a person who is randomly selected and assigned to the program. It is defined as:

$$ATE = E(\Delta) = E(Y_1 - Y_0) = E(Y_1) - E(Y_0). \tag{2}$$

This is the traditional average partial effect (APE) in econometrics.[5] To see this, let’s write the observed outcome in a switching model (Quandt, 1972):

$$Y = DY_1 + (1-D)Y_0, \tag{3}$$

[3] Y can be a vector of outcomes, but for simplicity let’s consider a single outcome of interest.
[4] Review of impact evaluation literature can be found in other studies, e.g., Heckman and Robb (1985), Blundell and Dias (2002), Imbens (2004).
[5] See Wooldridge (2001).

where Y is observed outcome, which is equal to Y_1 and Y_0 for participants and non-participants, respectively.

Then,

$$APE = E(Y \mid D=1) - E(Y \mid D=0) = E(Y_1) - E(Y_0) = ATE \,. \tag{4}$$

Most programs are targeted to certain subjects. The important question is the program impact on those who participated in the program. The expected treatment effect on the participants is equal to:

$$ATT = E(\Delta \mid D=1) = E(Y_1 - Y_0 \mid D=1) = E(Y_1 \mid D=1) - E(Y_0 \mid D=1) \,. \tag{5}$$

Except for the case of randomized programs that is discussed in section 3, ATE and ATT are, in general, different from each other, since program participation often depends on the potential outcomes, and as a result $E(Y_1) \neq E(Y_1 \mid D=1)$, and $E(Y_0) \neq E(Y_0 \mid D=1)$. To see this, equation (2) can be rewritten as:

$$\begin{aligned} ATE = E(Y_1) - E(Y_0) &= \left[E(Y_1 \mid D=1)Pr(D=1) + E(Y_1 \mid D=0)Pr(D=0)\right] \\ &\quad - \left[E(Y_0 \mid D=1)Pr(D=1) + E(Y_0 \mid D=0)Pr(D=0)\right] \\ &= \left\{\left[E(Y_1 \mid D=1) - E(Y_0 \mid D=1)\right]Pr(D=1)\right\} \\ &\quad + \left\{\left[E(Y_1 \mid D=0) - E(Y_0 \mid D=0)\right]Pr(D=0)\right\}, \end{aligned} \tag{6}$$

where $Pr(D=1)$ and $Pr(D=0)$ are the proportions of participants and non-participants of the program, respectively.

Define the average treatment effect on the non-treated (ATNT) as:

$$ANTT = E(Y_1 \mid D=0) - E(Y_0 \mid D=0) \,. \tag{7}$$

This parameter can be explained as the effect that the non-participants would have gained if they had participated in the program.

Then, ATE can be written as follows:

$$ATE = ATT\ Pr(D=1) + ATNT\ Pr(D=0) \,. \tag{8}$$

The case of a continuous program

In reality, a program can provide different amounts of treatment for people. For example, a micro-credit program provides people with different amount of loan. Denote the amount of treatment that a person receives by *D*. *D* is a continuous variable, which is higher than zero for the recipients. Further let $Y_{(D)}$ denote potential outcome corresponding to the value of *D*. We can be interested in program impact when the program changes by an amount, denoted by δ. More specifically, we want to measure the change in program impact due a change in the amount of remittances from d to $d+\delta$:

$$\Delta_{i(D=d+\delta)} - \Delta_{i(D=d)} = Y_{i(D=d+\delta)} - Y_{i(D=d)} \,. \tag{9}$$

Since we cannot estimate (9) for each household, we are interested in its average:

$$E\lfloor\Delta_{(D=d+\delta)} - \Delta_{(D=d)}\rfloor = E\lfloor Y_{(D=d+\delta)}\rfloor - E\lfloor Y_{(D=d)}\rfloor. \tag{10}$$

Expectation in (10) can be written for those who receive remittances:

$$E\left[\Delta_{(D=d+\delta)} - \Delta_{(D=d)} \middle| D > 0\right] = E\left[Y_{(D=d+\delta)} - Y_{(D=d)} \middle| D > 0\right]. \tag{11}$$

We can divide the right-hand sides of (10) and (11) by δ to obtain parameters called the average partial effect (APE) and the average partial effect on the treated (APET), respectively:

$$APE_{(d,\delta)} = \frac{E\lfloor Y_{(D=d+\delta)} - Y_{(D=d)}\rfloor}{\delta}, \tag{12}$$

$$APET_{(d,\delta)} = \frac{E\left[Y_{(D=d+\delta)} - Y_{(D=d)} \middle| D > 0\right]}{\delta}. \tag{13}$$

$APE_{(d,\delta)}$ measures how the average impact changes due to a change in the program amount d to $d+\delta$. $APET_{(d,\delta)}$ also measures the change in the average impact due to a change in the program amount d to $d+\delta$ but for the program participants.

In empirical studies, it is practically impossible to estimate $APE_{(d,\delta)}$ and $APET_{(d,\delta)}$ at all values of d and δ, since there are not enough observations. Thus the potential outcome is often assumed to be a parametric function of D. If we denote this function by $f(D)$, the impact parameters can be represented by the derivative of $f(D)$ with respect to D.

$$APE_{(d)} = E\left[\frac{\partial f(D)}{\partial D}\right], \tag{14}$$

$$APET_{(d)} = E\left[\frac{\partial f(D)}{\partial D} \middle| D = 1\right]. \tag{15}$$

2.2. Observed outcome function

In empirical studies, the impact parameters are often estimated from the equation of observed outcome. Denote outcome by Y and suppose it has the following function:

$$Y_i = F(D_i, X_i, \varepsilon_i, \beta), \tag{16}$$

where D is the program variable which can be binary or continuous, X and ε are observed and unobserved variables, β are model parameters, and F is a known function and continuously differential.

Impact of the program participation on outcome of person *i* is measured by the following difference:

$$\Delta_i = Y_{1i} - Y_{0i} = F(D_i = 1, X_i, \varepsilon_i, \beta) - F(D_i = 0, X_i, \varepsilon_i, \beta) . \tag{17}$$

This is the difference between the outcome of the person when she participates in the program and her potential outcome when she does not participate in the program. The parameters, ATE and ATT, are expressed:

$$ATE = E(\Delta_i) = E(Y_{1i} - Y_{0i}) = E[F(D_i = 1, X_i, \varepsilon_i, \beta) - F(D_i = 0, X_i, \varepsilon_i, \beta)]. \tag{18}$$

$$ATT = E(\Delta_i \mid D = 1) = E[F(D_i = 1, X_i, \varepsilon_i, \beta) - F(D_i = 0, X_i, \varepsilon_i, \beta) | D = 1]. \tag{19}$$

When D is continuous, the partial effect of the program on outcome of the person is:

$$\delta_i = \frac{\partial Y_i}{\partial D_i} = \frac{\partial F(D_i, X_i, \varepsilon_i, \beta)}{\partial D_i} , \tag{20}$$

and APE and APET are expressed as follows:

$$APE = E(\delta_i) = E\left[\frac{\partial F(D_i, X_i, \varepsilon_i, \beta)}{\partial D_i}\right], \tag{21}$$

$$APET = E(\delta_i \mid D > 0) = E\left[\frac{\partial F(D_i, X_i, \varepsilon_i, \beta)}{\partial D_i}\middle| D > 0\right]. \tag{22}$$

3. Estimation

3.1. Estimation of ATE and ATT

The estimator of ATT is readily derived from equations (19):

$$A\hat{T}T = \frac{1}{n_T}\sum_{i=1}^{n_T}\left[Y_{1i} - F(D_i = 0, X_i, \hat{\varepsilon}_i, \hat{\beta})\right], \tag{23}$$

where n_T is the number of participants. The above summation is calculated only for the participants. It should be noted that the participants' outcome in the presence of the program is their observed outcome.

ATE is estimated using equation (8). Firstly, we estimate ATNT:

$$A\hat{T}NT = \frac{1}{n_{NT}}\sum_{j=1}^{n_{NT}}\left[F(D_j = 1, X_j, \hat{\varepsilon}_j, \hat{\beta}) - Y_{0j}\right], \tag{24}$$

where n_{NT} is the number of non-participants. The outcome in the state of no-program for the non-participants is their observed outcome. ATE is estimated by:

$$A\hat{T}E = \frac{1}{n_T + n_{NT}}\left\{\sum_{i=1}^{n_T}\left[Y_{1i} - F(D_i = 0, X_i, \hat{\varepsilon}_i, \hat{\beta})\right] + \sum_{j=1}^{n_{NT}}\left[F(D_j = 1, X_j, \hat{\varepsilon}_j, \hat{\beta}) - Y_{0j}\right]\right\} \tag{25}$$

If we further assume that function F is monotonic (and continuously differential), we can estimate the asymptotic variance of the ATT estimator using the Delta method.

For example, per capita income often has the log-linear functional form:

$$\ln(Y_i) = \beta_0 + D_i\beta_1 + X_i\beta_2 + D_iX_i\beta_3 + \varepsilon_i\,, \tag{26}$$

Equation (8) allows the impacts of D to vary across X by using the interaction between X and D. The estimators of ATT, ATNT and ATE are expressed as follows:

$$\begin{aligned} A\hat{T}T &= \frac{1}{n_T}\sum_{i=1}^{n_T}\left[Y_{1i} - \exp(\hat{\beta}_0 + X_i\hat{\beta}_2 + \hat{\varepsilon}_i)\right] \\ &= \frac{1}{n_T}\sum_{i=1}^{n_T}\left\{Y_{1i} - \exp\left[\ln(Y_{1i}) - \hat{\beta}_1 - X_i\hat{\beta}_3\right]\right\} \\ &= \frac{1}{n_T}\sum_{i=1}^{n_T}\left[Y_{1i} - Y_{1i}\exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right)\right] \\ &= \frac{1}{n_T}\sum_{i=1}^{n_T}Y_{1i}\left[1 - \exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right)\right], \end{aligned} \tag{27}$$

$$\begin{aligned} A\hat{T}NT &= \frac{1}{n_T}\sum_{j=1}^{n_{NT}}\left[\exp(\hat{\beta}_0 + \hat{\beta}_1 + X_j\hat{\beta}_2 + X_j\hat{\beta}_3 + \hat{\varepsilon}_j) - Y_{0j}\right] \\ &= \frac{1}{n_{NT}}\sum_{j=1}^{n_{NT}}\left\{\exp\left[\ln\left(Y_{0j}\right) + \hat{\beta}_1 + X_j\hat{\beta}_3\right] - Y_{0j}\right\} \\ &= \frac{1}{n_{NT}}\sum_{j=1}^{n_{NT}}Y_{0j}\left[1 - \exp\left(\hat{\beta}_1 + X_j\hat{\beta}_3\right)\right], \end{aligned} \tag{28}$$

$$A\hat{T}E = \frac{1}{n_T + n_{NT}}\left\{\sum_{i=1}^{n_T}Y_{1i}\left[1 - \exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right)\right] + \sum_{j=1}^{n_{NT}}Y_{0j}\left[1 - \exp\left(\hat{\beta}_1 + X_j\hat{\beta}_3\right)\right]\right\} \tag{29}$$

The asymptotic variances of the estimators can be found using the Delta method. For example, for the estimator of ATT:

$$Var(A\hat{T}T) = T\begin{bmatrix} Var(Y_{1i}) & Cov(Y_{1i}, \hat{\beta}_1) & Cov(Y_{1i}, X_i) & Cov(Y_{1i}, \hat{\beta}_3) \\ Cov(Y_{1i}, \hat{\beta}_1) & Var(\hat{\beta}_1) & Cov(\hat{\beta}_1, X_i) & Cov(\hat{\beta}_1, \hat{\beta}_3) \\ Cov(Y_{1i}, X_i) & Cov(\hat{\beta}_1, X_i) & Var(X_i) & Cov(X_i, \hat{\beta}_3) \\ Cov(Y_{1i}, \hat{\beta}_3) & Cov(\hat{\beta}_1, \hat{\beta}_3) & Cov(X_i, \hat{\beta}_3) & Var(\hat{\beta}_3) \end{bmatrix} T^T\,, \tag{30}$$

where T is the vector of partial derivatives:

$$T = \left[\frac{\partial A\hat{T}T}{\partial Y_{1i}} \quad \frac{\partial A\hat{T}T}{\partial \hat{\beta}_1} \quad \frac{\partial A\hat{T}T}{\partial X_i} \quad \frac{\partial A\hat{T}T}{\partial \hat{\beta}_3} \right], \tag{31}$$

with:

$$\frac{\partial A\hat{T}T}{\partial Y_{1i}} = n_T^{-2} \sum_{i=1}^{n_T} \left[\exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right) - 1 \right], \tag{32}$$

$$\frac{\partial A\hat{T}T}{\partial \hat{\beta}_1} = -n_T^{-2} \sum_{i=1}^{n_T} Y_{1i} \exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right), \tag{33}$$

$$\frac{\partial A\hat{T}T}{\partial X_i} = -n_T^{-2} \hat{\beta}_3 \sum_{i=1}^{n_T} Y_{1i} \exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right), \tag{34}$$

$$\frac{\partial A\hat{T}T}{\partial \hat{\beta}_3} = -n_T^{-2} \sum_{i=1}^{n_T} Y_{1i} X_i \exp\left(-\hat{\beta}_1 - X_i\hat{\beta}_3\right). \tag{35}$$

The complication of the estimator variance depends on the functional form of *F*. Thus, in practice we can use bootstrap techniques to estimate the variance of the ATT estimate.

3.2. Estimation of APE and APET

The estimator of APE and APET are based on equation (21) and (22), respectively:

$$A\hat{P}E = \frac{1}{n_T} \sum_{i=1}^{n_T} \frac{\partial F(D_i, X_i, \hat{\varepsilon}_i, \hat{\beta})}{\partial D_i}, \tag{36}$$

where *n* is the number of observations in sampled data.

$$A\hat{P}ET = \frac{1}{n_T} \sum_{i=1}^{n_T} \frac{\partial F(D_i, X_i, \hat{\varepsilon}_i, \hat{\beta})}{\partial D_i}, \tag{37}$$

where n_T is again the number of program participants, and this summation is calculated only for the participants, i.e., with positive values of *D*.

Since *F* is a nonlinear function, there can be the *D* variable in the estimator of APET. APET can be varied across *D*, and we can estimate APET at a certain level of D. In example of the log-linear functional form given by (26), if *D* is continuous, the APET estimator is:

$$A\hat{P}E = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{\beta}_1 + X_i\hat{\beta}_3\right) \exp\left(\hat{\beta}_0 + D_i\hat{\beta}_1 + X_i\hat{\beta}_2 + D_i X_i\hat{\beta}_3 + \hat{\varepsilon}_i\right) = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{\beta}_1 + X_i\hat{\beta}_3\right) Y_i, \tag{38}$$

$$A\hat{P}ET = \frac{1}{n_T}\sum_{i=1}^{n_T}\left(\hat{\beta}_1 + X_i\hat{\beta}_3\right)\exp\left(\hat{\beta}_0 + D_i\hat{\beta}_1 + X_i\hat{\beta}_2 + D_iX_i\hat{\beta}_3 + \hat{\varepsilon}_i\right) = \frac{1}{n_T}\sum_{i=1}^{n_T}\left(\hat{\beta}_1 + X_i\hat{\beta}_3\right)Y_i . \quad (39)$$

Similar to ATT, the asymptotic variance of the APET estimator can be estimated using the Delta method. However, the variance estimator can have a complicated expression. Thus the asymptotic variance can be estimated using the bootstrap techniques.

4. An empirical study

In this section, we estimate impact of a micro-credit program in Vietnam on the participants. Since 2003, the Vietnam Bank for Social Policies (VBSP) has been established by the government to provide the poor with micro-credit without collateral at low interest rates. According to Vietnam Household Living Standard Survey (VHLSS), the program covered 8% of all households in 2004. The average loan size per borrowing household was around 3540 thousand VND.

The data used for impact measurement are from VHLSS in 2002 and 2004. These surveys were conducted by General Statistical Office of Vietnam with technical support from World Bank. These surveys set up a panel data set, which is representative for the national, rural and urban levels. We focus on the impact of the program in rural areas, since there is no data on communes in urban areas. The number of households of the panel data used for this analysis is 3099.

We use the log-linear model of per capita income which is similar to equation (26):

$$\ln(Y_i) = \beta_0 + X_i\beta_1 + D_i\beta_2 + D_iX_i\beta_3 + \varepsilon_i \quad (40)$$

where Y is per capita income, D is the program variable, X are explanatory variables which can be grouped into 5 categories: (i) household composition, (ii) regional variables, (iii) human assets, (iv) physical asset, (v) commune characteristics (Glewwe, 1991). To measure ATT, D is a dummy variable indicating the program participation. To measure APET, D is loan size which is continuous.

The main problem in estimation of (40) is correlation between the program variable and the error term. There can be unobserved variables such as business skills or production capacity that affect both income and the program variable. Thus, we run fixed-effect regression using the panel data of VHLSS 2002 and 2004 to remove time-invariant unobserved variables.

Table 1 presents results of fixed-effect regression. We use two models. In Model 1, there is no interaction between the program variable and explanatory variables. In Model 2, there are interaction between the program variables and some explanatory variables. All coefficient estimates of the program variables are statistically significant, but the magnitude of the estimates is very different across the models.

Table 1: Fixed-effect regression of log of per capita income

Explanatory variables	The program variable is the program participation (dummy)		The program variable is the loan size (continuous)	
	Model 1	Model 2	Model 1	Model 2
Program variable	0.07457***	0.11944***	0.00002**	0.00011***
	[0.03501]	[0.04245]	[0.00001]	[0.00003]
Age of household head	0.02310**	0.02253**	0.02369**	0.02334**
	[0.01061]	[0.01064]	[0.01062]	[0.01068]
Age of household head squared	-0.00016	-0.00015	-0.00017	-0.00016
	[0.00010]	[0.00010]	[0.00010]	[0.00011]
Head are ethnic minorities	-0.0199	-0.02612	-0.01974	-0.02183
	[0.07584]	[0.07713]	[0.07524]	[0.07542]
Head professionals/technicians	0.18617*	0.18120*	0.18454*	0.18052*
	[0.10180]	[0.10127]	[0.10176]	[0.10116]
Head clerks/service workers	0.16291*	0.17360*	0.16538*	0.17637*
	[0.09730]	[0.09869]	[0.09741]	[0.09890]
Head agriculture/forestry/fishery	-0.02592	-0.02652	-0.0218	-0.02787
	[0.09482]	[0.09363]	[0.09491]	[0.09369]
Head skilled workers/machine operators	0.18218*	0.18337*	0.18508*	0.18192*
	[0.09820]	[0.09759]	[0.09828]	[0.09756]
Head unskilled workers	0.07661	0.07834	0.08074	0.08119
	[0.09493]	[0.09418]	[0.09503]	[0.09418]
Head not working	-0.05617	-0.08036	-0.05153	-0.07587
	[0.09974]	[0.09912]	[0.09976]	[0.09891]
Ratio of members younger than 16	-0.27408***	-0.26410***	-0.27964***	-0.26905***
	[0.08319]	[0.08255]	[0.08314]	[0.08289]
Ratio of members who older than 60	-0.24094**	-0.21016**	-0.24091**	-0.22177**
	[0.09407]	[0.09643]	[0.09401]	[0.09508]
Household size	-0.14760***	-0.15274***	-0.14924***	-0.15401***
	[0.02736]	[0.02766]	[0.02706]	[0.02690]
Household size squared	0.00555***	0.00595***	0.00571***	0.00610***
	[0.00205]	[0.00211]	[0.00202]	[0.00203]
Ratio of members with lower secondary school	0.25812***	0.25886***	0.25768***	0.25714***
	[0.05698]	[0.05607]	[0.05689]	[0.05602]
Ratio of members with upper secondary school	0.51615***	0.52712***	0.51876***	0.52515***
	[0.09538]	[0.09520]	[0.09519]	[0.09479]
Ratio of members with technical degree	0.81194***	0.84668***	0.81819***	0.81555***
	[0.10470]	[0.10593]	[0.10470]	[0.10402]
Ratio of members with post secondary school	0.97186***	0.97138***	0.97773***	0.96757***
	[0.20472]	[0.20574]	[0.20475]	[0.20463]
Area of annual crop land (m2)	0.00001***	0.00001***	0.00001***	0.00001***
	[0.00000]	[0.00000]	[0.00000]	[0.00000]
Area of perennial crop land (m2)	0	0.00001***	0	0.00001***
	[0.00000]	[0.00000]	[0.00000]	[0.00000]
Area of aquaculture water surface (m2)	0.00002**	0.00001**	0.00002**	0.00001**
	[0.00001]	[0.00001]	[0.00001]	[0.00001]

Explanatory variables	The program variable is the program participation (dummy)		The program variable is the loan size (continuous)	
	Model 1	Model 2	Model 1	Model 2
Commune variables				
Have non-farm enterprises	-0.04084	-0.03896	-0.04256	-0.04267
	[0.02728]	[0.02738]	[0.02727]	[0.02736]
Distance to nearest agr. extension center (km)	0.00518***	0.00520***	0.00517***	0.00522***
	[0.00130]	[0.00129]	[0.00130]	[0.00129]
Have national electricity network	-0.03069	-0.04618	-0.0275	-0.04412
	[0.03766]	[0.03242]	[0.03771]	[0.03289]
Have car road	-0.05024	-0.03948	-0.05183	-0.04032
	[0.03953]	[0.03496]	[0.03977]	[0.03558]
Distance to nearest town (km)	-0.00019	-0.00027	-0.00012	-0.00021
	[0.00147]	[0.00146]	[0.00147]	[0.00146]
Distance to nearest daily market (km)	0.002	0.00208	0.00205	0.00207
	[0.00163]	[0.00165]	[0.00162]	[0.00165]
Distance to nearest periodic market (km)	-0.00245	-0.00218	-0.00253*	-0.00227
	[0.00151]	[0.00151]	[0.00152]	[0.00151]
Interaction terms				
Program variable * Head professionals/technicians		-0.80912***		-0.00014***
		[0.13031]		[0.00003]
Program variable * Head clerks/service workers		0.54442***		0.00014***
		[0.20166]		[0.00005]
Program variable * Ratio of members who older than 60		-0.39889		
		[0.25328]		
Program variable * Ratio of members with technical degree		-0.55146***		
		[0.20759]		
Program variable * Area of annual crop land (m2)		0.00001		0
		[0.00000]		[0.00000]
Program variable * Area of perennial crop land (m2)		-0.00002***		-0.00000***
		[0.00000]		[0.00000]
Program variable * Head unskilled workers				-0.00007***
				[0.00002]
Program variable * Head not working				-0.00003
				[0.00003]
Program variable * Age of household head				-0.00000***
				[0.00000]
Constant	7.98330***	8.00094***	7.97149***	7.98318***
	[0.28759]	[0.28572]	[0.28836]	[0.28794]
Observations	6198	6198	6198	6198
R-squared	0.83	0.83	0.83	0.83

Robust standard errors in brackets. Standard errors are corrected for sampling weights.
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Estimation from VHLSS 2002-2004

Table 2 estimates ATE, ATT, APE and APET. All the estimates are statistically significant at the 5% level. For example, Model 1 shows that the program can increase income per capita by 356 thousand VND. The program increases per capita income by 276 thousand VND for the borrowing households. Also according to Model 1, an increase of 1 VND at the mean of the loan size leads to an increase of around 0.06 VND in per capita income for the borrowers (APET).

Table 2: Estimation of ATT and APET

	Model 1	Model 2
ATE (thousand VND)	356.1**	342.6**
	[172.1]	[171.3]
ATT (thousand VND)	276.1**	252.6**
	[127.2]	[122.3]
APE	0.073**	0.075**
	[0.034]	[0.032]
APET	0.057**	0.063**
	[0.025]	[0.029]

Figures in brackets are standard errors. Standard errors are corrected for sampling weights and estimated using bootstrap (non-parametric) with 500 replications.
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Estimation from VHLSS 2002-200

5. Conclusion

In impact evaluation of a program, one is often interested in ATE, ATT, APE and APET. Although the above parameters are discussed thoroughly in the impact evaluation literature, they are not presented clearly in empirical studies. In nonlinear models, they are not directly estimated, and standard errors of the estimated parameters are not calculated by standard statistical software. Coefficients of program variable in nonlinear models can have misleading economic meaning. If the function of outcome is known, we can get the estimates of ATT and APET easily from estimates of function parameters and the observed outcome. The standard error of the estimates can be also calculated using the Delta method. If the derivation is complicated, the standard error can be estimated using the bootstrap method.

References

Blundell, R., and Dias M. C. (2002), "Alternative Approaches to Evaluation in Empirical Microeconomics", *Portuguese Economic Journal*, Vol. 1, No. 2, 91-115.

Glewwe, P (1991), "Investigating the Determinants of Household Welfare in Cote d'Ivoire", *Journal of Development Economics* 35: 307-37.

Heckman, J., and R. Robb (1985), "Alternative Methods for Evaluating the Impact of Interventions - An Overview", *Journal of Econometrics*, 30, 239-267.

Heckman, J., Lalonde, R., and Smith, J. (1999), The Economics and Econometrics of Active Labor Market Programs, in: Ashenfelter, A. and D. Card, eds., Handbook of Labor Economics, Vol. 3.

Imbens, G. (2004), "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review", *The Review of Economics and Statistics*, 86(1), 4-29.

Quandt, R. (1972). Methods for estimating switching regressions. *Journal of the American Statistical Association,* 67(338):306-310.

Wooldridge J. M. (2001). *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, Cambridge, Massachusetts London, England.