Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 135 (2018) 105–112

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd International Conference on Computer Science and Computational Intelligence 2018

Comparing Better Environmental Knowledge Based on Education and Income Using the Odds Ratio

Bagus Sumargo*

* Corresponding Author: Bagus Sumargo, Statistics Department Faculty of Mathematics and Natural Science Universitas Negeri Jakarta, Dewi Sartika Building 6st Floor Rawamangun Muka, Jakarta 13220, Indonesia, E-mail: bagussumargo@unj.ac.id, Phone: +62-81316044605

Statistics Department, Faculty of Mathematics and Natural Science, Universitas Negeri jakarta, Jakarta 13220, Indonesia

Abstract

This paper to find of existing better environmental knowledge regarding the demography and economics aspect in DKI Jakarta Provinces. First, hypotheses asserting relationship between environmental knowledge, and demographic and economic variables are reviewed by logistics binary regression. Second, based on better environmental knowledge variables significantly are compared to determine the ratio between the categorical at each that variables. The analysis unit is the DKI Jakarta population aged 15 and older in 2013 at Environment Caring Attitude Survey be conducted by Statistics-Indonesia. Social Environmental factors is a weak point of the cause of sustainable development cannot run smoothly, enhance one of the factors that need improvement is environmental knowledge. Condition of environmental knowledge in DKI Jakarta Provinces in alert category, and only the higher education variables and the middle income of the population with good environmental knowledge. using The Odds Ratio methods, it will be known the amount of population comparison that has better environmental knowledge for each variable of higher education or middle income.

Keywords: Education and Income; The Odds Ratio; Better Environmental Knowledge

1. Introduction

DKI Jakarta Province is the capital city of Indonesia face poor environmental quality and bad category. This is evident from the environmental quality index in 2011 of 36.68 and by 38.43 in 2012. In 2013 and in 2014, the environmental quality index of DKI Jakarta was 36.20 and 36.88 respectively number 50 as the critical value limit.

10.1016/j.procs.2018.08.155

So the environmental quality in Jakarta is categorized as "alert" [1]. It is appropriate that the population who have a high level of formal education is also expected to have awareness and awareness of the environment is also good. The close relationship between man and the environment is expressed [2] that there is a causal relationship between man and the environment and vice versa, so that the two cannot be separated from each other, and even the relationship can be described as a dynamic circularly. The same thing is also stated by Fransson and Garling [3], who said that the badness of environmental awareness is a factor that causes awareness and awareness of the environment is low. That is one of the psychological main to base in order to protect the environment is environmental awareness about. Low levels of knowledge that can cause lack of concern for, environment because the knowledge that the foundations someone in behave. Knowledge will conduct caring environment will encourage the attitude of a person to be friendlier for the environment. Sumargo, B[4] stated that Social Environmental factors is a weak point of the cause of sustainable development cannot run smoothly in Indonesia.

Based on the above explanation by using the case of population in DKI Jakarta, there is anomalous condition of good education and shown by data of enrollment rate of school age 13-15 year counted 99,47 percent, for age of 16-18 as much 70,23 percent and age 19-24 percent 22.52 percent, and above the average percentage of the national figure of Indonesia. In other words, the strength of the formal knowledge level owned by the residents of DKI Jakarta becomes a social capital to adapt or to mitigate the unhealthy environment. But the reality is an anomaly, percentage of households or residents in DKI Jakarta who have environmental caring behavior below the national average percentage [1].

2. Knowledge Environmental

Dunlap and jones [5] mentioned that awareness of environment is the extent to which people realized and know about the environment, efforts and support has done to the problem of the environment and or the initiative of the to contribute personally to find a solution for the environment. Caring environment divided into two a major component namely the environment and concern. Components environment referring to heterogeneity issues and environmental problems while components concern an expression refers to concern to the environment that is response someone above the surroundings around. As we know that The lack of knowledge about the environment is one of the factors causing the low environmental concern [6], such as age and gender [2], income [7], and education level [8] and knowledge about the environment can be obtained from counselling or training [9]. Knowledge of the environment is a very important aspect in discussing. awareness of environment. Knowledge will affect attitudes and behavior someone to its environment. Knowledge of conduct an environmentally friendly constitutes the principal component the internal factor that could lead to a change in attitude somebody to keep more friendly for the environment [1].

Based on the literature review above, the researcher estimates that the variables that influence the knowledge environmental awareness are the classification of residence, sex, age, education, working status, household income, information facility, environmental information source and elucidation/training with the binary logistic regression. Mamady [7] divides the above variables into several aspects of demography and economics, while Arandia [2] mentions an aspect of access to information.

3. Method

The unit of analysis is the DKI Jakarta population aged 15 and older in 2013 at Environment Caring Attitude Survey [1]. The analytical method used was the binary logistic regression analysis. Inferencing analysis is a technique of analyzing sample data whose results are applied to the population. Inferencing analysis in this study aims to determine what variables affect the knowledge about environmental in DKI Jakarta. The method of analysis used is binary logistic regression analysis. Binary logistic regression analysis is used to check whether there is a significant influence between explanatory variables and response variables.

Binary logistic regression[10] analysis was adopted to check for whether a significant effect could be between the explanatory variables and the response variable. Variables used consisted of one response variable (environmental knowledge) and eight explanatory variables were gender, age, education, work status, household income, information facility, information source, and elucidation/training. Binary logistic regression is used when a binary-scale response variable with two categories denoted by $Y = 1$ to declare a success event and $Y = 0$ to declare a failed event. In the regression analysis, the mean value of the response variable is obtained from the value of the explanatory variable. This value is called the conditional mean denoted by $E(Y|x)$ with Y representing the response variable and x denoting the explanatory variable. It is assumed that this average value can be seen as a linear equation in x, which is formulated as follows:

$$E(Y/x) = \beta_0 + \beta_1 x \quad (1)$$

This equation shows that E (Y | x) can be valuable regardless of the value of x in the range -∞ and + ∞. In logistic regression, the conditional mean value must be greater than or equal to zero and less than or equal to 1 $\left[0 \le E(Y/x) \le 1\right]$. E(Y|x) is denoted as π (x) to represent the conditional mean value of Y obtained from the specified x. Logistic regression model used is:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + ... + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + ... + \beta_p x_p)} \quad (2)$$

with β_P is the parameter value j = 0, 1, ..., p, and p are the number of variables. The transformation of π(x) is the core of logistic regression called logit transformation. This transformation is formulated as follows:

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] \quad (3.1)$$

$$g(x) = \ln\left[\exp(\beta_0 + \beta_1 x_1 + ... + \beta_p x_p)\right] \quad (3.2)$$

$$g(x) = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p \quad (3.3)$$

Logit g(x) is a linear function of a continuous parameter and has a value between -∞ and + ∞ depending on the value of x.

The regression model assumes that the value of the response variable can be written as y = E(Y|x) + ε. Value ε is called error and expresses deviation of observation value from conditional mean. It is assumed that it follows the normal distribution with zero mean and constant value variance according to the value of the explanatory variable. Then the conditional distribution of the response variable will follow the normal distribution with the mean E(Y|x) and variance constant. In logistic regression, the value of the response variable determined by x is written as $y = \pi(x) + \varepsilon$. The possibility of a value is two, ie when y = 1 then $\varepsilon = 1 - \pi(x)$ with a probability of $\pi(x)$ and when y = 0 then $\varepsilon = -\pi(x)$ with a chance of $1 - \pi(x)$. Then ε it has a distribution with zero mean and variance equal to $\pi(x)\left[1 - \pi(x)\right]$. Thus the conditional distribution of the response variable follows the binomial distribution with the probability determined by the conditional mean $\pi(x)$.

Discrete explanatory variables must be modified in order to be used in the analysis. Modifications are made using other variables that can describe the level of the discrete variables. This variable is called the dummy variable. When the discrete variable has as many k possible values it will be required as much as k-1 dummy variable. Let the j-explanatory variable (x_j) having as many k categories, will be formed as k_j-1 dummy variable denoted as D_{j1} and the coefficient of this variable denoted β_{j1} by l = 1, 2, ..., k_j-1. Then logit for model with p variable and as much as j the variable is discrete:

$$g(x) = \beta_0 + \beta_1 x_1 + ... + \sum_{j=}^{k_j - 1} \beta_{j1} D_{j1} + ... + \beta_p x_p \quad (4)$$

3.1. Parameter Estimation

In the logistic regression model, the parameter estimation used is Maximum Likelihood Estimation (MLE) method. The MLE method generates parameter values by maximizing the odds of the data being observed by using a function called likelihood function. This function denotes the probability of the observed data as a function of the unknown parameter (Hosmer and Lemeshow[11]). If Y is 0 or 1, then π (x) means probability that Y equals 1 condition x denoted as P(Y=1|x). While 1-π (x) means a conditional probability that Y equals 0 when x denoted as P(Y=0|x). For the pair (x_i, y_i), where y_i = 1, the contribution to the likelihood function is π (xi), and for the pair where yi = 0, the contribution to the likelihood function is 1-π (xi), where the value π (xi) of π (x) calculated at xi. The function of the pair (x_i,y_i) follows the following Bernoulli distribution:

$$f(yi, \pi(xi)) = \pi(xi)^{yi}\left[1 - \pi(xi)\right]^{1-yi} \quad (5)$$

To obtain the value of β which maximizes L(β) it is decreased to β_0 and β_1 with the equation equal to zero. The likelihood equation is as follows:

$$\sum_{i=1}^{n}[y_i - \pi(x_i)] = 0 \quad (6)$$

$$\sum_{i=1}^{n} xi[y_i - \pi(x_i)] = 0 \quad (7)$$

The calculation of this iteration is done by using computer programming logistic graduation[12].

3.2. Parameter Significance Testing

The simultaneous test in this research is used to know the effect of all exploratory variables simultaneously to environmental knowledge. The hypothesis used is:

$H_0 : \beta_1 = \beta_2 = \beta_3 = ... = \beta_p = 0$ (there is no significant influence of explanatory variables simultaneously on knowledge of environmental)

H_1 : Minimum there is one $\beta_j \neq 0$ (there is at least one explanatory variable that has a significant effect on knowledge about environmental)

with j = 1, 2, ..., p; p = number of explanatory variables. Test statistics to be used are:

$$G = -2\ln\left[\frac{L_0}{L_1}\right] \quad (8)$$

Information:

L_0 = likelihood value of the model without explanatory variables

L_1 = likelihood value of the model with explanatory variables

The G-test statistic follows the Chi-Square (χ^2) distribution (with degrees of free p, so the basis for decision-making is: (i). If the value of G> $\chi^2_{(p,\alpha)}$ or p-value < α, then decline H_0

(ii). If the value of G $\leq \chi^2_{(p,\alpha)}$ or p-value $\geq$ α, then fail to reject H_0

In this simultaneous test it is expected that the decision obtained is reject Ho in order to conclude that entering the explanatory variable into the model gives a better effect in explaining the response variable.

3.3. Partial Test

If the simultaneous test of the parameters obtained the decision is rejected H_0, then further partial test to determine the explanatory variables that have a significant effect on the response variable. Partial significance test parameters were performed using Wald.

The hypothesis used in this test is:

H0: $\beta_j = 0$ (j explain variable has no significant effect on response variable)

H1: $\beta_j \neq 0$ (j explain variable significant to response variable)

with j = 1, 2, ..., p and p = number of explanatory variables. Test statistics to be used are:

$$Wj = \left(\frac{\hat{\beta}_j}{se(\hat{\beta}_j)}\right)^2 \quad (9)$$

Information:

$\hat{\beta}_j$ = parameter estimator β_j

$se(\hat{\beta}_j)$ = standard error of $\hat{\beta}_j$

Wald test statistics follow the distribution of Chi-Square (χ^2) with a degree of freedom of 1 (Agresti, 2002) .Therefore, the basic decision-making is:

i. If W > $\chi^2_{(1-\alpha)}$ or p-value value <, then reject H_0

ii. If W $\leq \chi^2_{(1-\alpha)}$ or p-value $\geq$ value, then reject H_0

In this partial test is expected to get decision is reject Ho in order to be concluded that explanatory variable have significant effect to explain response variable.

3.4. Goodness of fit Test Model

The next test is the model conformity test to examine the difference between the results obtained from the model and the observed results in the observational data. To check the suitability of the model used Hosmer and Lemeshow test statistics.

The hypothesis used in this test is:

H_0: Model fit (no difference in yield between observation and prediction result of the model)

H_1: Model is not fit (there is a difference of result between observation and prediction result of the model)

Test statistics to be used are: $$\hat{C} = \sum_{k=1}^{g} \frac{(o_k - n_k^1 \overline{\pi}_k)^2}{n_k^1 \overline{\pi}_k (1 - \overline{\pi}_k)} \quad (10)$$

g = number of groups

$o_k = \sum_{j=1}^{c_k} y_j$ = number of response variable values in group k

$n_k^1 =$ number of subjects in group k

$\overline{\pi}_k = \sum_{j=1}^{c_k} \frac{m_j \hat{\pi}_j}{n_k^1}$ = average estimated chances in group k

C_k = the number of combinations of explanatory variables in the k^{th} group

m_j = number of subjects with C_k a combination of explanatory variables

Statistics $\hat{C}$ follow the distribution of Chi-Square (χ^2) with degrees of freedom (g-2.) Therefore, the basis of decision-making is:

i. If $\hat{C} > \chi^2_{(g-2)}$ or p-value value < α then reject H_0

ii. If $\hat{C} \leq \chi^2_{(g-2)}$ or p-value value ≥ α then fail to reject H_0

In the fit conformity test model is expected to get the decision failed to deny Ho in order to conclude that the model formed is fit or no difference between the observation and prediction results.

3.5. Odds Ratio

Odds ratio is the size of the association that estimates how the tendency of explanatory variables to the response variable, that is the comparison between two events that fall into success and failure category. Odds when x = 1 is defined as $\frac{\pi(1)}{1-\pi(1)}$ and odds when x = 0 is defined as $\frac{\pi(0)}{1-\pi(0)}$.

Odds ratio denoted by OR is θ the odds ratio for x = 1 and odds for x = 0 expressed in the equation:

$$\theta = \frac{\pi(1)/1-\pi(1)}{\pi(0)/1-\pi(0)} \quad (11)$$

Table 1. Opportunity value in Binary Logistic Regression model

		Explanatory Variable (X)	
		x=1	x=0
Respon Variable (Y)	y=1	$\pi(1) = \frac{\exp(\beta_0 + \beta_j)}{1 + \exp(\beta_0 + \beta_j)}$	$\pi(0) = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)}$
	y=0	$1 - \pi(1) = \frac{1}{1 + \exp(\beta_0 + \beta_j)}$	$1 - \pi(0) = \frac{1}{1 + \exp(\beta_0)}$
Total		1,0	1,0

Source: Hosmer and Lemeshow [11]

In binary logistic regression the explanatory variable consists of two categories, that is encoded with 0 (reference categories) and 1, then the interpretation of the coefficient on that variable depends on the magnitude θ. That is, the

risk of occurrence of event y = 1 in the category x = 1 is $\exp(\beta_j)$ at times the risk of occurrence of event y = 1 in the category x = 0. Meanwhile, if the explanatory variable is continuous scale, it can be interpreted that every increase of unit C in the explanatory variable will causing the risk of occurrence of y = 1 equal $\exp(C,\beta_j)$ to greater times.

3.6. Binary Logistic Regression Model

Logistic regression model used in this study if all explanatory variables significantly influence the environmental knowledge is:

$$g(D) = \beta_0 + \beta_1 D_1 + \beta_2 X_2 + \beta_3 D_3 + \beta_4 D_4 + \beta_5 D_{51} + \beta_6 D_{52} + \beta_7 D_{53} + \beta_7 D_{54} + \beta_8 D_{55} + \beta_9 D_6 + \beta_{10} D_{71} + \beta_{11} D_{72} + \beta_{11} D_8 \quad (12)$$

Notes:

D_1: *dummy* variable of sex
X_2 : age variables
D_3: *dummy* variable of education
D_4: *dummy* variable of Working Statute
D_{51} : the first *dummy* variable of income
D_{52}: the second *dummy* variable of income
D_{53}: the third *dummy* variable of income
D_{54}: the fourth *dummy* variable of income
D_{55} : the fifth *dummy* variable of income
D_6 : *dummy* variable of information facilities
D_{71} : the first *dummy* variable *of* information sources
D_{72} : the second *dummy* variable *of* information sources
D_8 : *dummy* variable of elucidation/training

4. Results

The simultaneous test results shown at the Omnibus Test of Coefficients yield the G statistic value of 65,168 with the p-value less than 0.05 then H_0 is rejected. So it can be concluded that there is at least one explanatory variable that influences environmental knowledge.

Table 2. Simultaneous Testing Results

	Chi-Square	Df	Sig
Model	65,168	13	0,000

Based on partially testing output as in table 2 showed that p-value for each the explanatory variables less than 0.05, so that hypotheses not accepted. The other hand, all of the explanatory variables influence the environmental knowledge significantly.

Table 3. Partially Testing Output

Variable	$\hat{\beta}$	Df	Sig
Sex	-.310	1	.088
Age	-.008	1	.166
Education	.426	1	***.013***
Working Statute	.227	1	.183
Income (earn)		5	.011
Income (1)	-.033	1	.933
Income (2)	-.009	1	.981
Income (3)	.264	1	.501
Income (4)	1.145	1	***.019***
Income (5)	.895	1	.103
Facilities	.866	1	.075
Information Sources		2	.115
Information Sources (1)	-.378	1	.123
Information Sources (2)	-.065	1	.788
Training /Tutorial	.590	1	.135
Constanta	.699	1	.320

The binary logistic regression equations for parameter β:

$$g(D) = 0{,}699 - 0{,}310D_1 - 0{,}008X_2 + \mathbf{0{,}426D^*_3} + 0{,}227D_4 - 0{,}033D_{51} - 0{,}009\,D_{52} + 0{,}264D_{53} + \mathbf{1{,}145D^*_{54}} + 0{,}895D_{55} + 0{,}866D_6 - 0{,}378D_{71} - 0{,}065D_{72} + 0{,}590D_8$$

The results for estimating of parameter showed that: (1). There are positive effect significantly between education and environmental education with regression coefficient 0.426; (2). There are positive effect significantly between income 384–769US$ and environmental knowledge with regression coefficient 1.145.

The goodness of fit be calculated by *output Hosmer and Lemeshow test*. Based on statistical testing get *Chi-*

Square value 7,050 and *p-value* 0,531. Because of *p-value* bigger than 0.05, and we reject hypotheses. In conclusion that with confidence interval 95 percents-that binary logistic regression model is goodness of fit.

Tabel 4. Goodness of Fit Testing

Chi-Square	Df	Sig
13.917	8	0,084

5. Discussion

As we know that the lack of knowledge about the environment is one of the factors causing the low environmental concern [6], such as age and gender [2], income [7], and education level [8] and knowledge about the environment can be obtained from counseling or training [9]. However, as at table 3 showed that the explanatory variables not only sex and gender not influence significantly of environmental knowledge [2], but also information facilities [13]; information sources [14]; and elucidation/training [9]. This research support of: (1) Adugnaw [8] implies that education variables influence environmental knowledge significantly; (2). Mamady, **2016** implies that the income variables influence environmental knowledge significantly (especially at income 384 – 769 US$ level). As a table 5 stated that better educated populations have 1.5 times better environmental knowledge than less educated populations. People with high education have great knowledge, including that about the behavior of environmental concern. It showed at *The Odds Ratio values* for education and income 384 – 769 US$ level. Residents with income 384 – 769 US$ level have 3.1 times better environmental knowledge than others.

Tabel 5. The Odds Ratios Output

Variable	*Odds Ratio*
Education	1.531
Income (4)	3.143

Source: BPS- ECAS Data (processing)

The open problems from this study, we should make fit categorizing of each variables, and the results above just for generalization in DKI Jakarta Provinces.

6. Summary

This study showed that only the high education and income 384–769US$ level variables influence environmental knowledge significantly. The odds ratio indicates that the population with the following characteristics is likely to be knowledgeable about environmental concern high education and income 384–769US$ level. Community with a low economic status are expected to increase their accessibility of information related to the environment. The less educated population should continue to increase their knowledge about the environment, especially through socialization and training. Enhance, the intensity and scope of socialization (elucidation)/training related to the environment should be increased. Classification gender, age, working status, information facilities, information resources, and socialization/training not significantly affect knowledge about environmental concern. It need the best strategy for increasing six variables, and there are have being influence of good environmental knowledge of population or household in DKI Jakarta Provinces.

References

1. Badan Pusat Statistik BPS-Statistics Indonesia; ECAS. 2013a. Indikator Perilaku Peduli Lingkungan Hidup 2013 [Internet]. Jakarta, Indonesia: Jakarta: BPS-Statistics Indonesia; Available from: https://www.bps.go.id/publication/2013/12/27/26e979dbdd579d7b6db6941d/indikator-perilaku-peduli-lingkungan-hidup-2013.html
2. Garcı AP, Iniesta-arandia I, Garcı D, Montes C, Martı B. Factors influencing local ecological knowledge maintenance in Mediterranean watersheds : Insights for environmental policies. 2014;
3. Fransson N. Environmental Concern : Conceptual Definitions , Measurement Methods , and Research Findings. 1999;(November 2017).
4. Sumargo B, Kasuma KAP. social-Environment as a Weak point of Sustainable Development in Indonesia. In: oral presentation ICGRC 2018. 2018.
5. Dunlap RE, Jones RE. Environmental Concern: Conceptual and Measurement Issues. Handb Environ Sociol. 2002;(August):482–524.
6. Mifsud MC. An Investigation on the Environmental Knowledge , Attitudes and Behavior of Maltese Youth. 2011;3:413–22.
7. Mamady K. Factors influencing attitude, safety behavior, and knowledge regarding household waste management in Guinea: A cross-sectional study. J Environ Public Health [Internet]. 2016;2016. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84964904840&partnerID=40&md5=351d35b6b9b88bf82d1e2ccd8ccfe9e0
8. Birhanu A. Environmental Knowledge , Attitude and Participatory Behavior towards Land Degradation in Injibara Secondary and

Preparatory School , Northwestern Ethiopia. 2014;4(17):89–96.
9. Universitesi S, Informa R, Number WR, House M, Street M, Liefl AK, et al. Effectiveness of environmental education on water : connectedness to nature , environmental attitudes and environmental knowledge. 2015;(February):11–4.
10. David W. Hosmer Jr., Stanley Lemeshow RXS. Applied Logistic Regression, 3rd Edition _ Regression Analysis _ General & Introductory Statistics _ Subjects _ Wiley. 2013. p. 528.
11. Hosmer DW, Lemeshow S. Applied Logistic Regression [Internet]. Wiley Series in Probability and Sattistics. 2000. p. 373. Available from: http://pubs.amstat.org/doi/abs/10.1198/tech.2002.s650
12. Landau S, Everitt B. A handbook of statistical analyses using SPSS. [Internet]. Vol. 24, Statistics in Medicine. 2004. 3236-3237 p. Available from: http://doi.wiley.com/10.1002/sim.2134
13. Lyons, E and Glynis M. Factors Predicting Environmental Concern and Indifference in 13- to 16 -year-olds. Environ Behav. 1994;26(2):223–38.
14. Rickinson M. Learners and Learning in Environmental Education : A critical review of the evidence. 2010. 37-41 p.

Appendix 1: The Explanatory Variables

Variable Names	Categorical	Variabel Dummy
SEX	1. Male 2. Female*	SEX=1, if male SEX=0, if female
AGE	Numeric	-
EDUC	1. High 2. Low*	EDUC=1, if > Junior High School EDUC=0, if ≤ Junior High School
WORK	1. Working 2. Un Working*	WORK=1, if Working WORK=0, if un working
EARN	1. < 38.5 US$ * 2. 38.5 - 77 US$ 3. 77 – 192 US$ 4. 192 – 384.5 US$ 5. 384.5 - 769 US$ 6. > 769 US$	• EARN (1)=1, if 38.5 – 77 US$ EARN(1)=0, otherwise • EARN(2)=1, if 77 – 192 US$ EARN(2)=0, otherwise • EARN(3)=1, jika 192–384.5 US$ EARN(3)=0, otherwise • EARN(4)=1, jika 384.5–769 US$ EARN(4)=0, otherwise • EARN(5)=1, if > 769 US$ EARN(5)=0, otherwise
Information Facility	1. Yes 2. No*	FACILITY=1, if Yes FACILITY=0. Otherwise
Information Sources	1. Others Media * 2. Mass Media 3. Combination each others	• SOURCES (1)=1, if Mass Media SOURCES (1)=0, otherwise • SOURCES (2)=1, if Combination SOURCES (2)=0, otherwise
Training	1. Yes 2. No*	TRAINING=1, if Yes TRAINING =0, otherwise

Notes: *) References Categorical