CHAPTER VI

FACTORS DETERMINING ECONOMIC SECURITY

Apart from the description on economic security situation as presented in Chapter V, here an attempt was made to relate the identified independent variables with the proxy dependent variables of economic security. These variables were derived from the sampled data to follow the basic characteristics of the each type of model. The relevant information are discussed in Chapter III under the section of 3.5.2. In reality, economic security analysis is a version of poverty analysis. Economic security is represented by the income and food sufficiency. Food sufficiency is a very basic and a major measure of economic security.

In the study area, food supply is limited by the very low level production of essential basic foods, such as vegetables, fruits, and livestock products. Rice is the main food obtained through their production path. People use their livelihood income to purchase foods. The amount of food purchased determines the relationship between livelihood income and food sufficiency. Food sufficiency is also determined by the household's consumption behavior and resource endowment. The analysis in this chapter was carried out with respect to each dependent variable and the results were evaluated under each model. The overall results were taken into consideration to give priority to the significant factors. At the end of the chapter, problems related to economic security and households' and institutionally' needed strategies to mitigate economic insecurity problems are discussed. The data generated for this study were derived from cross-sectional household survey. The models presented do not deal with the temporal variation in data or information.

6.1 Descriptive statistics of independent variables

The descriptive statistics viz, mean, standard deviation, minimum and maximum value of each variable included in the model are presented in Table 6.1.

Table 6.1 Household characteristics of sample used in the regression analysis

Variables	Minimum	Maximum	Mean	Std. deviation
Education of household head	0	13	7.31	3.17
Age of Household head	25	72	45.63	9.81
Young is to adult ratio	0	4	0.91	0.85
Diversity index	1	2.88	1.26	0.38
Log of house hold income/ month	2.48	4.70	3.85	0.39
Perception on consumption	0	1.00	0.71	0.46

(Source: Survey data, 2008)

6.2 Testing OLS regression assumptions

The basic assumptions of OLS regression model were diagnosed through proper statistical procedures.

6.2.1 Normal distribution and data transformation to achieve normality

The normal distribution[7] was tested using histogram to test the normality assumption. The variables which only confirm to this condition were included in the multiple regression analysis. The data transformation was carried out for the variables which did not confirm to this assumption and tested for normality after transformation. The variables which did not confirm to this assumption at both stages (before and after transformation) were did not included during the analysis. For example the above testing is shown in Figure 6.1 and 6.2 for the household income dependent variable. Among the different transformation techniques (square root, power, log and inverse), log transformation[8] was end up with expected result of normality (Figure 6.2).

[7] **Note:** Normal distribution condition:

To obey the normality condition, 65% of the population should be fallen between the interval of mean ± one standard deviation and 95% of the population should be fallen between the interval of mean ± two standard deviation of the given data. Normality is restricted due to the population or sample characteristics or variable itself. The reasons for non-normality are the presence of outliers (scores that are extreme relative to the rest of the sample) and the nature of the variable itself. Since, the study population is a poor population; there is more chance some variables have to be outstanding because of the existence of inequality and extremity. Normality distribution is one of the important assumptions before running the multiple regression analysis.

[8] Log data transformations are carried out for improving the normality of variables. Before data transformation, it was confirmed that non-normality is due to a valid reason (real observed data points) and not due to data entry error or non-declared missing values. In cases where there are extremes of range, base 10 is desirable (here the identified variables representing extreme situation Log (base 10) was selected. The logarithm of any negative number or number less than 1 is undefined, if a variable contains values less than 1.0 a constant must be added to move the minimum value of the distribution, preferably to 1.00 (In this study no any data points less than 1 was observed).

The normal distribution of regression standardized residuals (Figure 6.3) and the normal probability plot of standardized residuals (Figure 6.4) are also shown here.

The variable income distribution:

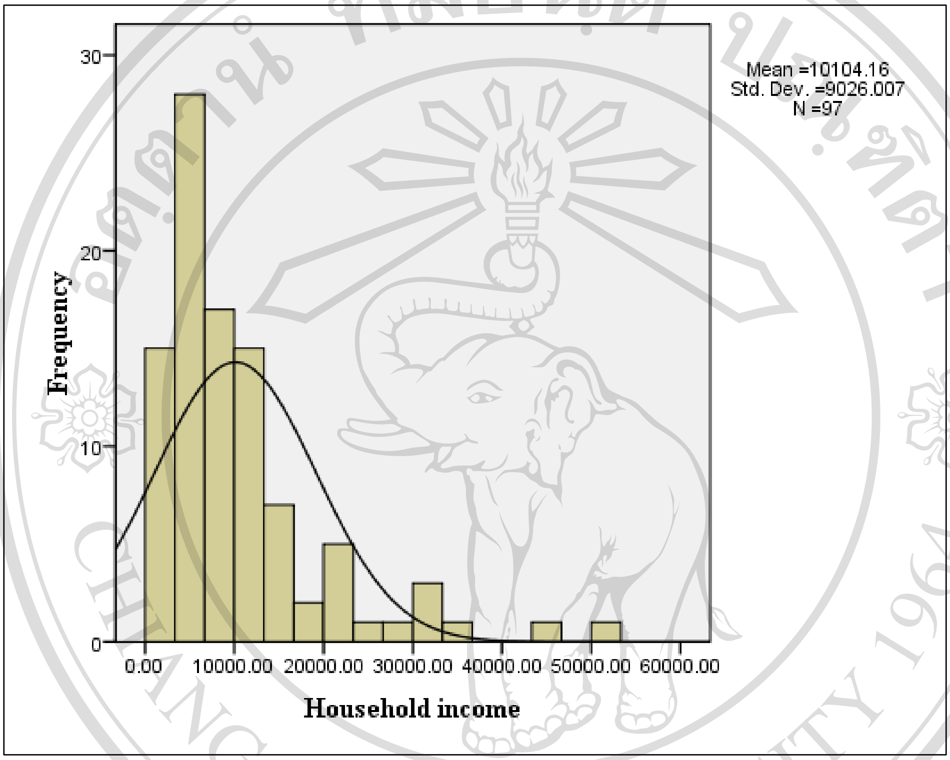


Figure 6.1 Sample households' income distribution before transforming data

The distribution curve after transformed into log:

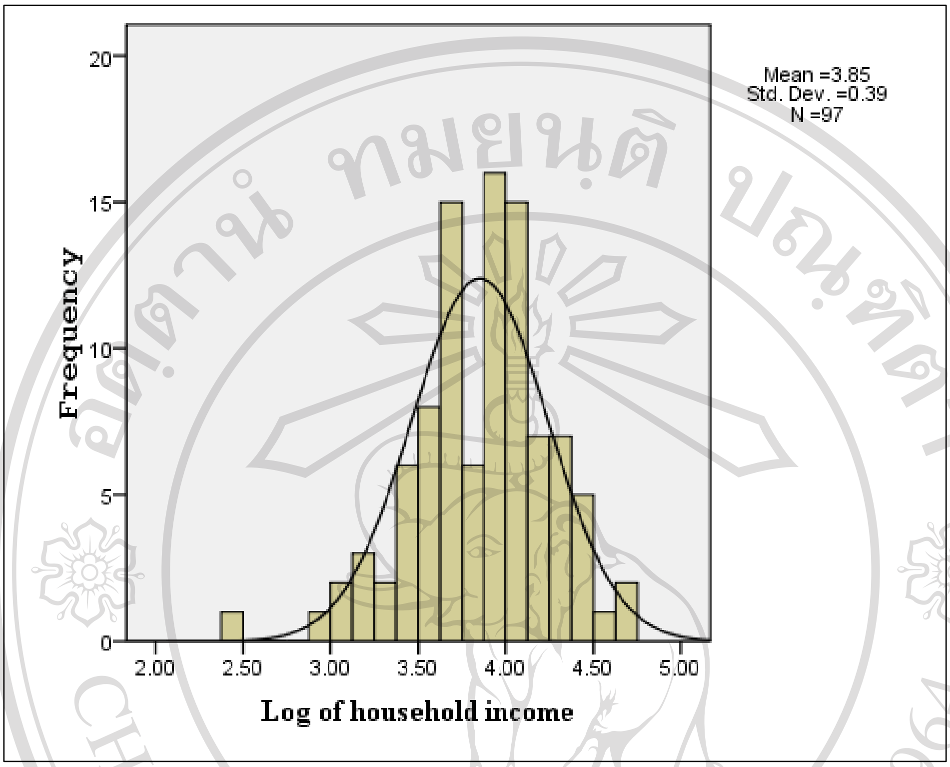


Figure 6.2 Sample households' income (log of household income) distribution after transforming data

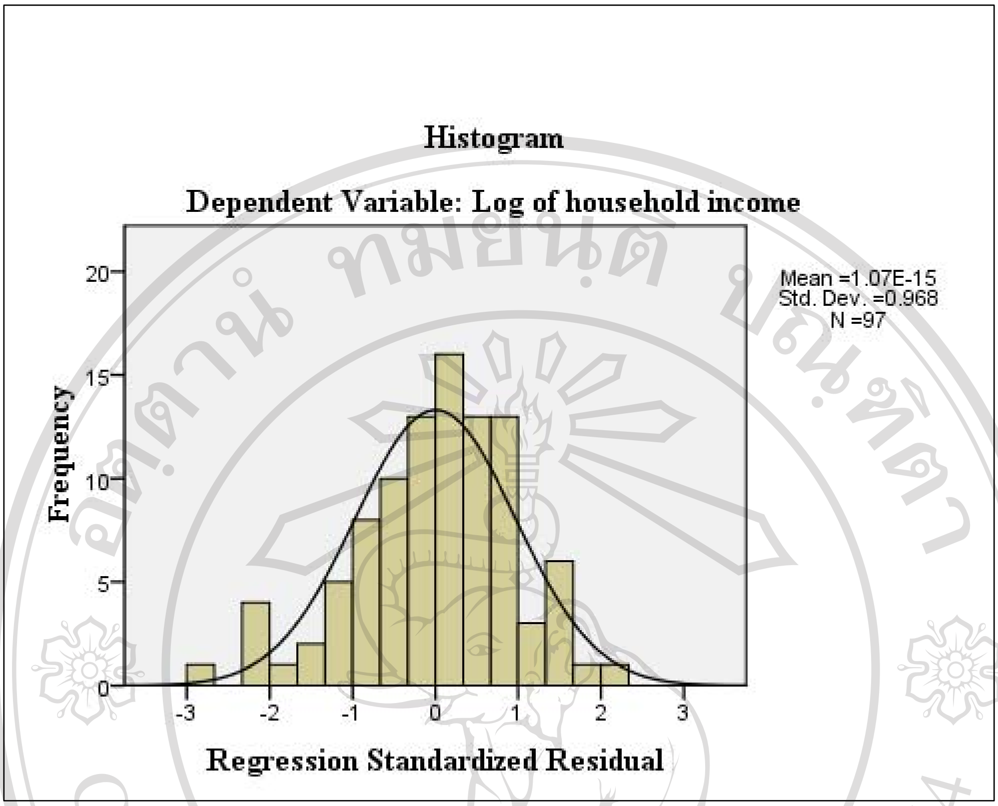
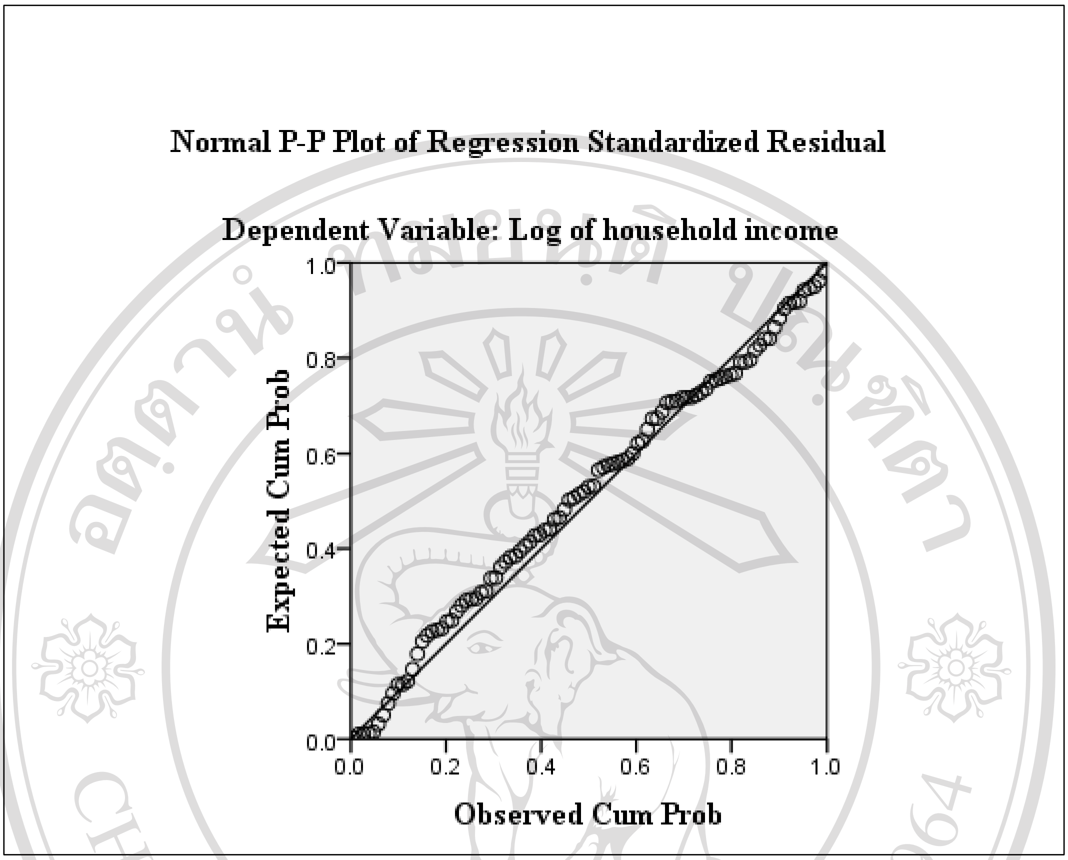


Figure 6.3 Normal distribution of regression standardized residual



6.4 Normal Probability plot of standardized residual

6.2.2 Multicollinearity

Multicollinearity is another problem observed among livelihood resource variables considered. It was diagnosed by looking at the part and partial correlation and collinearity statistics (tolerance and variance influence factor). Multicollinearity is not considered as a serious problem in a correct functional model. It means that if the model is built with compatible (agreeable) variables then there is no need to leave out any of them (those variables) from the model (in the condition of multicollinearity). The table 6.2 shows the correlation co-efficients between each pair of independent variables included in the model in this study. Among those, all of them were below the multicollinearity criteria of 0.8 (80%).

Table 6.2 Coefficient of correlations among variables

	Hh_edu	Hh_age	YA_ratio	Hh_type	Hh_sex	Divty_index
Hh_edu	1	-0.003	0.015	-0.255	-0.141	0.163
Hh_age	-.003	1	0.541	0.187	0.295	0.049
YA_ratio	0.015	0.541	1	0.075	0.17	0.084
Hh_type	-0.255	0.187	0.075	1	0.059	0.095
Hh_sex	-0.141	0.295	0.17	0.059	1	-0.007
Divty_index	0.163	0.049	0.084	0.095	-0.007	1

6.2.3 Heteroscedasticity

Heteroscedasticity is a common condition existing in cross-sectional micro-level data. It is the condition of error variance increase with the increase in dependent variable (in this study log of income). It is normally examined by plotting regression standardized predicted value (in X axis) against regression standardized residual (in Y axis). If the resulted plot show random array of dots evenly dispersed around zero then the model is free from heteroscedasticity. The data distribution for this studied sample is shown in Figure 6.5. According to this observation, it is free from hereroscedasticity.

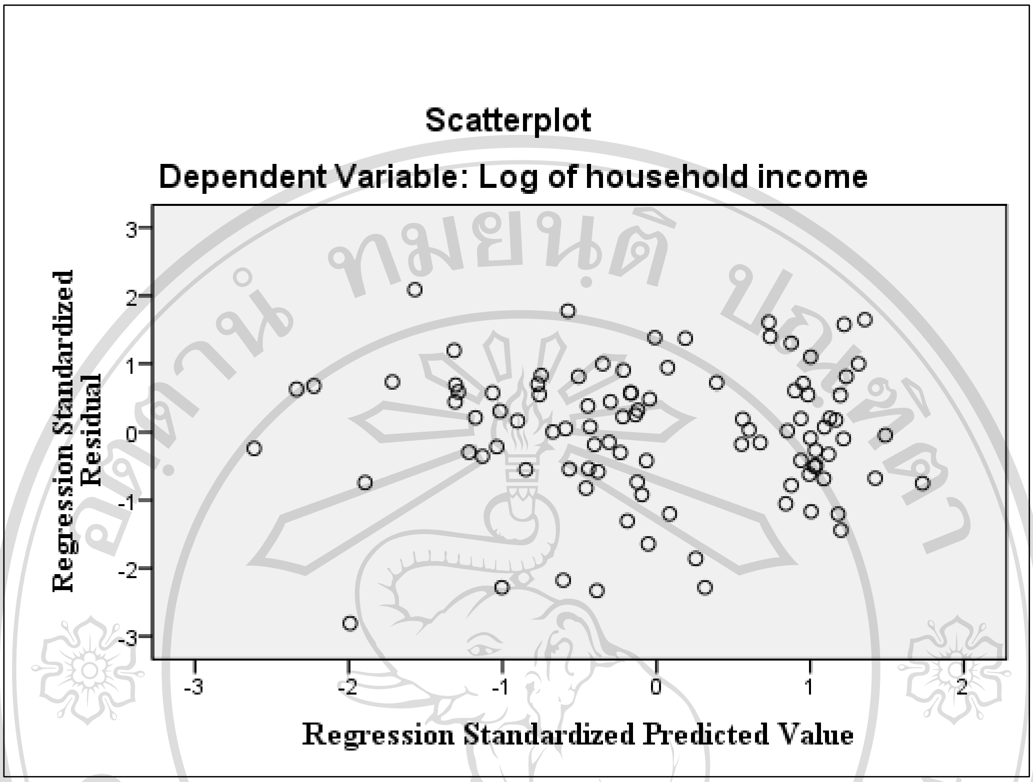


Figure 6.5 Scatter plot showing random array of dots evenly dispersed around zero

6.3 Dummy variables recoding for multiple regression analysis

Dummy variables are variables that take the values of only 0 or 1. Dummy-variable regressors can be used to incorporate qualitative explanatory variables into a linear model, substantially expanding the range of application of regression analysis. Use of dummy variables usually increases model fit (coefficient of determination), but at a cost of fewer degrees of freedom and loss of generality of the model. Too many dummy variables result in a model that does not provide any general conclusions. In this study out of seven variables in the models, only two of them are representing dummy category. They are household types (which shows the nature of partially commercialized and subsistence) and sex of household head. Both of them were qualitative explanatory variables.

When there are dummies in all observations, the constant term has to be excluded. If a constant term is included in the regression, it is important to exclude one of the dummy variables from the regression, making this the base category against which the others are assessed. If all the dummy variables are included, their sum is equal to 1 (which stands for the variable X_0 to the constant term B_0), resulting in perfect multicollinearity. This is referred to as the dummy variable trap. In this study between two different types of observations of included dummies, only one type of observation was included. In the case of sex of household head, male household head was included. In the case of household types, household type 1 was included.

These categorical predictor variables cannot be entered directly into a regression model. Therefore they should be meaningfully recoded to enter into the regression model (using the menu options of Transform →Recode into different variables as a prior process before run the multiple regression models[3]). In this analysis, male and household type 1 observations/characteristics were given the code of 1 and included in the model compare to the base categories of female and household type 2 (coded as 0).

[3] In the case of binary logistic regression model, there is no need to recode them. The model itself creates the base category during the process (by default the lower valued coded type observations treated as base). Further details are in binary logistic regression model on selected dummy variables.

6.4 Multiple regression model results

The tested multiple regression model was significant at 5% significance level with adjusted R-Square of 0.3(even though this value was low, the results were further confirmed with the following models in terms of economic security). The following variables were significant in explaining the dependent variable of log of household income: household head education, household type 1, and male household head. The positive and significant sign on household head education indicated that while keeping other variables in the model constant, one unit increase in education caused 0.022 unit increases in log of household income. De Haan and Dubey (2003) also found income poverty correlates with illiteracy at the level of Orissa's districts of India.

The co-efficient of male household head measured the relative difference in the log of household income for male and female headed. The log of income differential for male headed was 24.1 percent higher than that for female headed household.

The co-efficient of household type 1 (partially commercialized) measures the relative difference in the log of household income for household type 1 (partially commercialized) and household type 2 (subsistence). The log of income differential for household type 1 (partially commercialized) is 29.8 percent higher than for household type 2 (subsistence) (Table 6. 3). Partially commercialized households (in terms of higher money value of job related equipments, higher additive borrowing capacity and continuous employment pattern) have the relatively better position in income security than the subsistence one.

The other variables considered age of household head, young is to adult ratio and diversity index did not show significant influence on log of income at 5% significant level. The affects of all these variables on the other considered dependant variables (dichotomous food sufficiency and dichotomous income sufficiency variables) are discussed under each model in the following pages.

Table 6.3 Determinants of household income (economic security) in study population: OLS regression results (with constant)

Independent variables	Unstandardized Coefficients	Standard error	t-value	Sig.
Education of household head	.022*	.011	1.992	.049
Age of household head	-.003	.004	-.638	.525
Young is to adult ratio	-.063	.047	-1.337	.184
Household type	.298*	.071	4.190	.000
Diversity index	.013	.022	.608	.545
Sex of household	.241*	.079	3.069	.003
Constant	3.515*	.262	13.42	.000

Number of observations: 97, R-Square: .344 (Prob>F), Adjusted R-Square: .300

Regression mean square > residual mean square (.838 > .107)

Dependent variable: log household income

*Significant at 5% level

6.5 Binary logistic regression model

Since the studied population is comprised of identified rural poor households, most of the characteristics are skewed in nature (poor households representing slightly same characteristics except for a few households who had better ownership features). Binary logistic regression is a suitable measure to build a model to differentiate basic characteristic levels in relation to economic security.

In this study, binary logistic regression was utilized to depict the dichotomous variables i) probability of having food sufficiency over probability of not having food sufficiency and ii) probability of having income sufficiency over probability of not having income sufficiency in relation to selected independent variables.

Logistic regression can be used to predict a dependent variable on the basis of continuous and/or categorical independents and to determine the percent of variance in the dependent variable explained by the independent variable; to rank the relative importance of independent variable; to assess interaction effects; and to understand the impact of covariate control variables. The impact of predictor variables is usually explained in terms of odd ratios.

Logistic regression estimates the odds of a certain event occurring. Note that logistic regression calculates changes in the log odds of the dependent variable, not changes in the dependent variable itself as OLS regression does. Logistic regression has many analogies to OLS regression: logit coefficients correspond to b coefficients in the logistic regression equation, the standardized logit coefficients correspond to beta weights, and a pseudo R^2 statistic is available to summarize the strength of the relationship.

Unlike OLS regression, however, logistic regression does not assume linearity of relationship between the independent variables and the dependent variable, does not require normally distributed variables, does not assume homoscedasticity, and in general has less stringent requirements. It does, however, require that observations be independent and that the independent variables be linearly related to the logit of the dependent. The predictive success of the logistic regression can be assessed by looking at the classification table, showing correct and incorrect classifications of the dichotomous. Goodness-of-fit tests such as the likelihood ratio test are available as indicators of model appropriateness, as is the Wald statistic to test the significance of individual independent variables.

It is important to be careful to specify the desired reference category, which should be meaningful. Binomial logistic regression by default predicts the higher of the two categories of the dependent (usually 1), using the lower (usually 0) as the reference category. In SPSS binomial logistic regression, categorical independent variables must be declared by clicking on the "Categorical" button in the Logistic Regression dialog box. In this analysis considered dependent (food sufficiency and income sufficiency) and independent categorical variables (household sex and household types) were predicted by default set up as follows: i) Dependent variables coding: a) food sufficiency: have food sufficiency-1, not have food sufficiency-0, b) income sufficiency: have income sufficiency-1, not have income sufficiency-0 ii) independent variables coding: a) household sex: male-1, female -0, b) household types: household type1(partially commercialized)-1, and household type 2(subsistence)-0.

6.5.1 Food sufficiency binary logistic regression model results

Hosmer and Lemeshow (chi-square test)[4] was used to test the overall fit of the food sufficiency binary logistic regression model developed here with the real observed raw data. Overall fit is meant by the data used for this model is really represented the predicted model or not. If the data used is really represented the model then, the significance of the test is more than 0.05. In this test the significance value is .342 (Table 6.4), this is more than 0.05, and then the following null hypothesis was accepted[5]. It means that there was no significant difference between observed data and the model predicted. The model can better explain the observed data.

Null hypothesis (HO): there is no difference between observed data used (to build the model) and the food sufficiency model predicted

Alternative hypothesis (H1): there is difference between observed data used to build the model (all observations entered under each variable in the structure or form of binary logistic regression using SPSS) and the model predicted (out put).

[4] This test is requested by checking “Hosmer and Lameshow goodness of fit” under the options button in SPSS

[5] If the significance of the test is small (i.e., less than 0.05) then the model does not adequately fit the data. At this point of less than 0.05 alternative hypothesis telling the truth of the model predicted does not adequately fit the data was accepted. According to these conditions (criterias), in this model the null hypothesis was accepted and the alternative hypothesis was rejected.

This model was significant at 5% significant level. The model correctly predicted 78.4 percentages of both food sufficient and food deficit households. Two variables were significant in explaining the probability of having food sufficiency compare to the probability of not having food sufficiency (dependent variable in odd ratio). By looking at Exp (B) the variables effect on odds can be expressed as follows: while holding other variables constant 1 unit increase in household head education cause 1.269 unit increase in odds (probability) of having food sufficiency compared with not having food sufficiency. Odds of having food sufficiency is more than 6.1 times higher for the households which are in household type 1 than those in household type 2 (reference group is household type 2) (see Table 6.5). Women's role in household food security is considered crucial, and it is widely accepted that despite the daily household chores, women activities are mostly revolved around the household welfare through production, processing and acquisition of food in norms. As women labour force is totally dedicated in food related, they did not significantly differ from male head.

Table 6.4 Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9.006	8	.342

Table 6.5 Determinants of household food sufficiency (economic security) in study population: binary logistic regression results

Independent variables	B	Standard error	Exp(B)	Sig.
Household head education	.238*	.089	1.269	.008
Age of household head	-.024	.015	.976	.104
Young is to adult ratio	-.444	.297	.642	.136
Household type	1.808*	.569	6.100	.001
Diversity index	-.186	.293	.830	.524
Household sex	.332	.553	1.394	.548

-2 log likelihood: 86.366, Number of observations: 97, Correctly predicted: 78.4%
*significant at 5% level

6.5.2 Results of income sufficiency binary logistic regression model

Hosmer and Lemeshow (chi-square test) was used to test the overall fit of the income sufficiency binary logistic regression model developed here with the real observed raw data used for this purpose. Overall fit is meant by the data used for this model is really represented the predicted model or not. If the data used is really represented the model then, the significance of the test is more than 0.05. In this test analysis, the significance value is .686 (Table 6.6), this is more than 0.05, then the following null hypothesis was accepted[6]. It means that there was no significant difference between observed data and the model predicted. The model can better explain the observed data.

[6] If the significance of the test is small (i.e., less than 0.05) then the model does not adequately fit the data. At this point of less than 0.05 alternative hypothesis telling the truth of the model predicted does not adequately fit the data was accepted. According to these conditions (criterias), in this model the null hypothesis was accepted and the alternative hypothesis was rejected.

Null hypothesis (HO): there is no difference between observed data used (to build the model) and the income sufficiency model predicted

Alternative hypothesis (H1): there is difference between observed data used to build the model (all observations entered under each variable in the structure or form of binary logistic regression using SPSS) and the model predicted (out put).

This model was significant at 5% significant level. The model correctly predicted 75.3 percentage of both income secure (not under official poverty line) and income insecure (under official poverty line) households. Four variables were significant in explaining the probability of having income secure compared to the probability of not having income secure (dependent variable in odd ratio). By looking at Exp (B) the variables effect on odds can be expressed as follows: while holding other variables constant, 1 unit increase in household head education cause 1.368 unit increase in odds (probability) of having income sufficiency compared with not having income sufficiency. While holding other variables constant 1 unit increase in household head age cause 0.944 unit decrease in odds (probability) of having income sufficiency compared with not having income sufficiency. While holding other variables constant 1 unit increase in young is to adult ratio cause 0.192 unit decrease in odds (probability) of having income sufficiency compared with not having income sufficiency. Odds of having income sufficiency is more than 3.003 times higher for the households which are in household type 1 (which are partially commercialized) than those in household type 2 (which are subsistence and treated as reference group during analysis) (see Table 6.7).

Table 6.6 Hosmer and Lemeshow test of income sufficiency binary logistic regression model

Step	Chi-square	df	Sig.
1	5.654	8	.686

Table 6.7 Determinants of Household income sufficiency (economic security) in study population: Binary Logistic regression results

Independent variables	B	Standard error	Exp(B)	Sig.
Education of household head	.314*	.107	1.368	.003
Age of household head	-.058*	.021	.944	.006
Young is to adult ratio	-1.650*	.524	.192	.002
Household type (1)	1.100*	.521	3.003	.035
Household sex (1)	.464	.612	1.590	.448
Diversity index	-.158	.629	.854	.801

-2 log likelihood: 94.034, Number of observations: 97, Correctly predicted: 75.3 %
*significant at 5% level

Effects of variables on considered dependent variables among the models are discussed here. Most of the results from regression analysis presented in the above models were consistent with the hypothesized assumption.

In the multiple regression model, household head education, male household head and household type 1 (partially commercialized) showed positive significant association with log of income.

In the binary logistic food sufficiency regression model, household head education and household type 1 (partially commercialized) showed positive association with log odds of food sufficiency.

In the binary logistic income sufficiency regression model, household head education, household type 1 (partially commercialized) showed positive association while age of household head and young is to adult ratio showed negative relationships with the odds of income sufficiency.

Diversity index did not show any significant relationship in any of the model. It reflects economic specialization is more important than economic diversification. This result further confirms how it is important that the people should be educated in their primary occupational categories.

Household head education and household types are the variables commonly cause significant effect on the specified dependent variables in all three models. Household type altogether says productive tools related to category of occupation are very important to develop an economically active population. While it says seasonality in employment affects on economic security. It means that during a year period peoples' economic situations differ. The qualitative analysis results presented in Chapter V says the frequency of economic vulnerability is high during the months October, November, and December. Another important variable included in household type is additive borrowing capacity; therefore additive borrowing capacity should be increased by increasing production capacity in terms of productive tools. It jointly suggests the importance of skills and full education with the provision of productive tools or borrowing capacity for investment (to purchase productive tools). In the binary logistic income sufficiency regression model, female headed households did not significantly differ from male headed households. This may be due to the effects of adult members (sons or daughters) in female headed family significantly contributes in maintaining the household welfare.

Educational status of household's head plays important role in the household resource management, technology adoption and even in the household food consumption behavior. The empirical results from the regression analysis showed a positive effect of education on economic security which might be due to the overwhelming effect of education in successful handling of scarce resources.

6.6 Summary of this chapter

The above chapter identified the factors influencing different aspects of economic security. For this purpose log of income, logit of food sufficiency (log odds of food sufficiency) and logit of income sufficiency (log odds of income sufficiency) were used as dependent variables. Among the three models education of household head and dummy variable household type 1 (partially commercialized) were significantly explain the nature of economic security. In the multiple regression models, the dummy male headed was further significant in explaining log of income. In the income sufficiency binary logistic regression model, logit of income sufficiency was further determined by age of household head and young is to adult ratio. These factors were taken into evaluation during proposing reasonable recommendations.