

## **Chapter 3**

### **Methods of Data Collection and Analysis**

#### **3.1 Logic of the Research Design**

The purpose of subjective well-being research from an economic point of view is to seek general patterns within the several variables that are hypothesized to affect this valuation. The literature on subjective well-being paved the way for asking individuals to evaluate their own well-being. For this thesis, the researcher has imported this subjective approach into the basic needs framework in order to allow individuals to play a central role in deciding whether, and to what degree, their basic needs are met. The researcher hopes that by adding the subjective approach to this problem, it will be possible to overcome some of the objective conceptual hurdles involved in identifying the reasons for meeting basic needs.

The analysis of subjective well-being has advantages over the scientific understanding of what affects people's happiness beyond rising income. Subjective well-being refers to individuals' evaluation of the life satisfaction or happiness brought about by experiencing positive and negative effects. Estimating the factors that satisfy basic needs by the individual has great relevance to closing the gap between the concepts and results in the literature. Individuals evaluate their level of subjective well-being depending on their circumstances, but also by comparing themselves with others, past experiences and future expectations (Frey and Stutzer, 2002b). Studying subjective well-being in developing countries is an even wider

concept in that it connects to the debate on the definitions of poverty (Chapter 2), where income or consumption poverty is viewed as an issue of human development or social exclusion (Gough and McGregor, 2007:3). In order to understand well-being in developing countries, results could be imported from empirical studies on developed economies. However, special care should be taken, given the differences in the conditions that influence happiness in emerging economies and developed economies. Nor can it be taken for granted that the negative determinants of poverty and unhappiness are the simple mirror image of the positive determinants of wealth and happiness.

It is very common for the recent literature on subjective well-being to use data asking people directly about their personal well-being with questions like “I am satisfied about everything in my life” and “In general, I consider myself a very happy person” by ranking Strongly Agree, Agree, Neutral, Disagree and Strongly Disagree on a Likert scale. A life satisfaction scale would be obtained in the first question, while a happiness scale would be obtained with implication, as it provides insight about real life aspects of households that are specific to a given region of the study as well as providing information to policymakers about what individuals actually need.

Subjective approaches to well-being have been used to measure the perceived poverty line, thus replacing income based approaches (Kingdom and Knight, 2006; Pradhan and Ravallion, 2000; Rojas, 2008). In the present research, the basic needs approach will be used to measure poverty as it is perceived by a household. This should capture the psychological, demographic and social aspects that are taken into

account in subjective well-being approaches by considering an individual or household extremely, not extremely, or not at all poor.

The scope of basic needs satisfaction is very broad, as it can cover needs from primary education, health care and nutrition to access to water and shelter. Interestingly, according to recent empirical evidence (Sen's Capability Approach, 2005), Sen's capability approach overlaps with the concepts of well-being. The terms above are similar to the capability approach, which provides a more complete approach for measuring poverty by putting the notion of individual freedom at the heart of discussion. In this approach, Sen understands capabilities as the choices that individuals can make according to the characteristics of commodities they have (Sen, 1983, 1985, 1987, 1999). However, he considers that subjective approaches to addressing the well-being of individuals are not very reliable since individuals tend to adapt to each burden so as to overlook the burden itself (Sen, 1984).

### **3.2 Research Design**

The survey research underlying the present thesis was designed to investigate comparisons of the impacts of natural disaster on well-being, social capital and happiness level of differentially-affected Nargis areas. As such, it is more positive than normative. In this study a household is defined as a group of individuals residing together, pooling all or most of the income and resources, and basically sharing the same economic benefit from their livelihoods. The sample characteristics were examined from the data on whole households randomly selected in this study.

The researcher collected the secondary data in the Irrawaddy Division (Cyclone Nargis affected area). The major sources for secondary data were the Myanmar Agriculture Services (MAS), Settlement and Land Record Department (SLRD), Central Statistical Organization (CSO), United Nation Development Program (UNDP), Food and Agricultural Organization (FAO), World Food Program (WFP), United Nation Office for Coordination of Humanitarian Affairs (UNOCHA), United Nations International Children's Emergency Fund (UNICEF), Groupe de recherche et d'échanges technologiques (GRET) and other non-governmental organizations. Secondary data of the township and village profile data, maps, annual progressive reports, project documents and research papers/reports were collected. The sources of data were both local and national.

After collecting secondary data from different sources to reveal the insights into the study area, the researcher selected the sample study sites based on the different levels of damage inflicted by cyclone Nargis.

Moreover, a pilot survey was conducted by the researcher in two randomly selected villages of Bogalay and Pyapon Townships in May 2010. According to the observations gleaned from the pilot survey, the researcher and her team started the main survey to collect data with re-structured questionnaires. For selecting the respondents, the criterion was Cyclone Nargis affected people who engaged in farming before Nargis and the small farmers who are holding under 5 acres (2 ha). Bogalay and Pyapon Townships were selected based on the different level of damages affected by Cyclone Nargis. Interviewees were comprised of village leaders and household heads of random selected villages. The random households were collected

from villages from each township. To get the 150 households from each township, five villages from each township were chosen with the information of land ownership of the members of the community based organization.

Non-probability random sampling was used in each village because of the limitation of time frame and data availability. The opportunity for estimating the whole population or calculating with precision had to be forgone because sample size did not reflect the population size. The information collected was mainly concerned with the monetary expenditures but also with quantities relating to food items purchased or acquired for consumption. In addition to income and expenditure data, other demographic data, socio-economic data such as basic needs, livelihood, social capital and happiness level of the farmers were also gathered through the structured questionnaire.

For the main survey, enumerators who graduated from university and had experience with social survey activities were trained about the main objectives of the study and how to collect data. After training for the data collection, the team primarily conducted the survey in remote selected areas accessible only by waterway. Access to the disaster areas remains risky especially for outsiders who are not familiar with very wide, low-lying areas with small rivers and channels.

### **3.3 Conceptual Framework of the Study**

There are many ways to measure and test the significance of relative and absolute poverty. Poverty is usually defined by monetary measures to establish the number of people living under the poverty line. The common international poverty

line is \$1.25 at 2005 purchasing-power parity (PPP) (The World Bank 2008). However many aspects of poverty do not include money, but rather, other, very important, poverty measures. Since the list of possible factors is almost endless, many observers have suggested a focus on the basic needs. The basic needs generally mean food, health, education, and sanitation. Most unsatisfied basic needs include such indices as access to clean water, quality of housing, and education of household head, food sufficiency and income. In this study, the measurement of poverty is carried out by using the basic need approach. The following table shows the criteria and variables necessary in the calculation of poverty lines under the basic needs approach. It will be noted that the emotional, social, and spiritual dimensions of well-being are omitted from the basic needs approach because human needs in these areas cannot be purchased.

Based on this review of the relevant literature, we may construct an overall conceptual diagram for the essential concepts that will be measured, analyzed and inter-related in this thesis. This is shown as Figure 3.1.

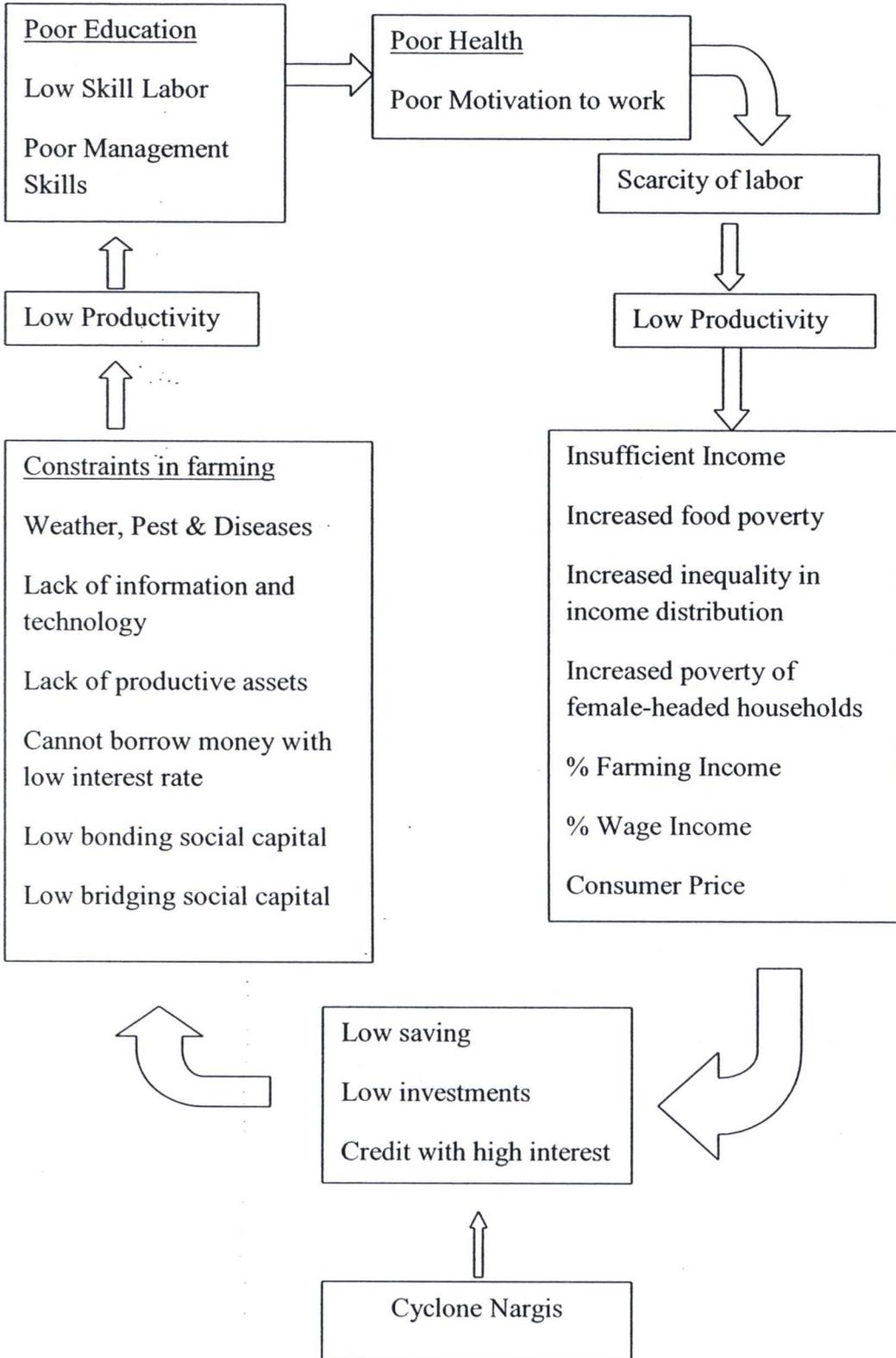


Figure 3.1 Conceptual framework and diagram

Table 3.1 Poverty Maps from Unsatisfied Basic Needs Indicators

Basic Needs	Dimensions	Variables
<b>Food sufficiency</b>	<i>Body:</i> Quality of food and enough	Frequency of having food within seven days according to lifestyle
<b>Housing Condition</b>	<i>Body:</i> Quality of shelter	Construction material used
<b>Sanitation Facilities</b>	<i>Body:</i> i) Availability of clean water ii) System for elimination of waste	i) Water source ii) Type of toilet
<b>Access to health services</b>	<i>Body:</i> Availability of health services	Health service availability
<b>Access to education</b>	<i>Mind:</i> Enrollment of children in schooling age	School attendance of school age
<b>Economic Capacity</b>	<i>Economy:</i> Probability of income sufficiency of household	i) Age of household members ii) Highest educational level reached iii) Education level of household head iv) Condition of economic activity of household head

Source: Adapted and modified from Poverty maps from unsatisfied basic needs indicators in Latin America by United Nation's Economic Commission for Latin America (Feres and Mancero, 2001)

For each of the variables in the table, criteria are chosen that indicate whether a household is considered lacking or not in a basic need. For example, roofing materials such as branches and leaves or temporary shelter may be considered lacking the basic need for adequate shelter. Households that need to spend the whole day to get to the nearest clinic, dispensary or hospital may be considered lacking in the basic need of health services.

### **3.4 Methods of Data Analysis**

Both descriptive and inferential statistical methods are used in the data analysis of the current thesis. Lorenz curves are used to explain the income distribution in these study areas. Analytical tools are extended in the calculation of Gini coefficients and Foster-Greer-Thorbecke indices because of the poverty head count index, the poverty gap and the intensity of poverty for the research area.

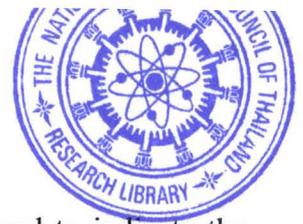
#### **3.4.1 Descriptive Statistics**

A table of descriptive statistics was generated in SPSS 17 to check the data. The total observations, means, maximum and minimum values and standard deviation are included in this table. The number of observations gives the researcher a chance to know whether the variable will be a useful explanatory factor. Maximum and minimum values in the table give researchers a way to conclude whether this is feasible or not to improve the formulation of hypothesis equations. Correlation matrices are also in use to show the dependency of variables and the directions (positively or negatively correlation) of variables among each other.

### 3.4.2 Lorenz Curve and Gini Coefficient

When economists measure income distribution or inequality in society, they typically measure it with respect to income and expenditures. This opens the possibility of observing negative incomes. In theory, income is equal to expenditure plus savings but in some developing countries where financial markets are neither accessible nor mature, income and expenditures plus savings can be unbalanced at times. By definition, negative savings can clearly produce negative incomes. There can be numerous difficulties in determining incomes through research including;

- (i) Seasonal effects where farmers need to spend more than normal when they cannot get food supply from nature or their own farms
- (ii) Periods where farmers have to invest for the next season
- (iii) People forget, particularly when asked in a single interview, about items they may have sold, or money they may have received.
- (iv) People may be reluctant to report income, especially if it is earned illegally
- (v) Some parts of the income are difficult to calculate, for example at this study some family members work at the seasonal informal sectors and do not have regular income.
- (vi) Income levels may be affected by short term fluctuations in the seasonal pattern of home agricultural production. Since the time series data and secondary data for the previous time period were not accessible in this study and the prices and income of a base year were unknown, real income and expenditures could not be calculated.



Nominal income and expenditure were therefore used to indicate the state of distribution.

Therefore, the negative income is unavoidable for some people at certain times in some economies. Moreover, informal or illegal economy data are sometimes inaccessible because of the legal or social prohibitions or even danger to the researcher.

In Chapter 4 and 5, we shall present some possible methodological correlations to the data from income, happiness and well-being surveys conducted in two Nargis affected areas. The first of these is the Lorenz curve, which presents both the actual and a perfectly equal distribution of income or happiness. It shows how far away the actual distribution in society lies from perfect distribution. A perfectly equal distribution in the society means that each proportion is getting equal share of a welfare input or outcome.

In either case, perfect equality means that each proportion of the cumulative population will add up to the equal amount of increase in cumulating the welfare-good. The line of perfect equality, the 45 degree line in the graph means that the X axis, the percentage of population, and the Y axis, the proportion of the welfare is increasing in the same proportion. The slope of separation of the Lorenz curve can be analyzed for a population or regional sub group that lies furthest away from perfect distribution; it is those sub groups which we normally focus our policy on in order to restore greater equality in the aggregate population.

On the basis of Lorenz curves for each sub group, a second measure, Gini coefficients, may be calculated to reflect the situation of economic welfare in the Nargis slightly-affected area and heavily-affected area. The Gini coefficient is the area between the Lorenz curves which bends beneath the perfect equality line and the 45 degree straight line. It indicates how the distributional pattern in the community is away from the perfect equality line. The graph shows that the Gini is equal to the area marked 'A' divided by the sum of the areas marked 'A' and 'B' (that is,  $Gini = A/(A+B)$ ).

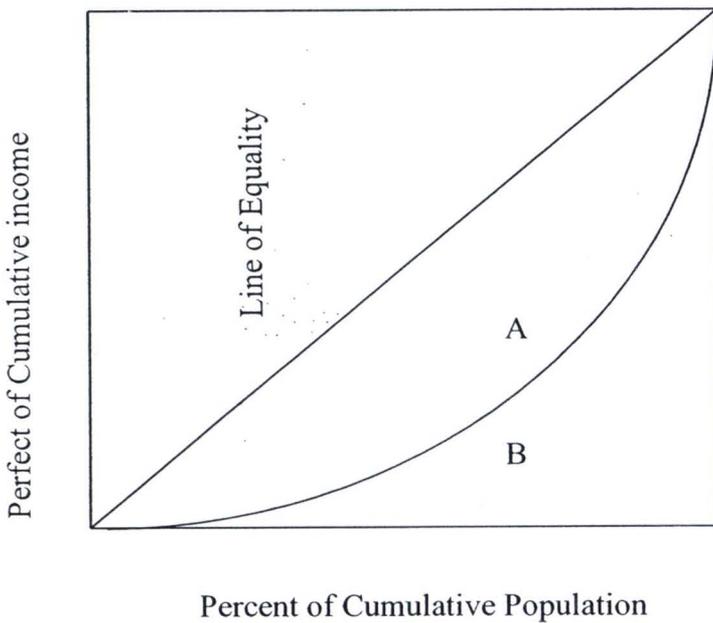


Figure 3.2 Lorenz Curve and Gini Coefficient of Income

The range of Gini coefficient is between zero and one. The perfect equality line is the situation at which the distribution of social good is increasing in the same proportion with the increase in cumulative population. That means, good consumption

increases one unit with the one unit increase of population. The area under the perfect equality line is one unit square.

If the distribution of income is approaching absolutely unequal, the area between absolute equality line and actual distribution will approach to one and if the distribution is equal, the Lorenz Curve will approach to the perfect equality line and the area between them will approach to zero.

The weakness of the Gini coefficient is the effect that occurs if a negative value in the data is neglected. In the real world, there could be negative income in the primary data collection because of the effect of periodical shocks or seasonal cyclical pattern. Sometimes, it is not convenient to check only the income inequality to compare the income of the two places. Income tends to be understated for many reasons:

(1) People forget, particularly when asked in a single interview, about items they may have sold, or money they may have received.

(2) People may be reluctant to report income, especially if it is earned illegally.

(3) Some parts of the income are difficult to calculate, for example some family members work in seasonal informal sectors and do not have regular income.

(4) Income levels may be affected by short term fluctuations for the seasonal pattern of home agricultural production. Since the time series data and secondary data for the previous time period were not accessible in this study and the prices and income of a base year were unknown, real income and expenditures could not be

calculated. Nominal income and expenditure were therefore used to indicate the state of distribution. A separate happiness score was also developed to measure subjective well being (SWB).

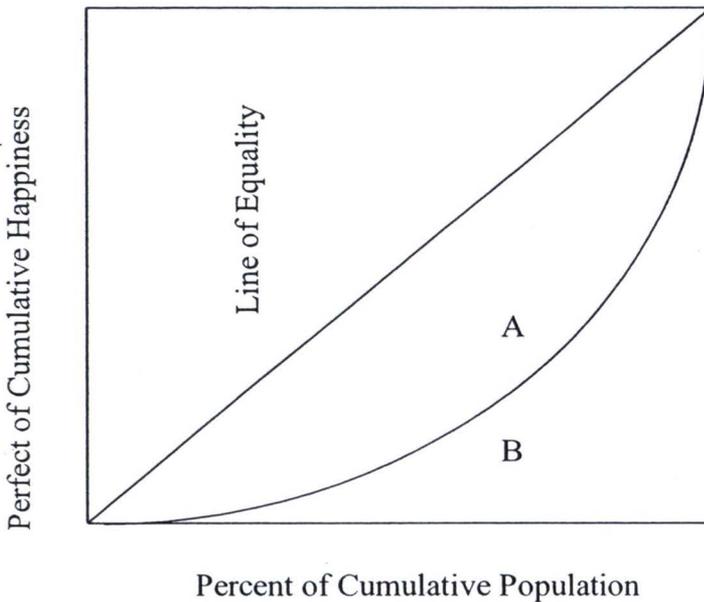


Figure 3.3 Lorenz Curve and Gini Coefficient of Happiness

### 3.4.3 Incidence, Depth and Intensity of Poverty Foster\_Greer\_Thorbecke-Schoch

$$FGT_{\alpha} = \frac{1}{N} \sum_{i=1}^H \left( \frac{z - y_i}{z} \right)^{\alpha}$$

N= number of people in the sample population

H= number of poor people

z = poverty line (1095 Myanmar Kyats)

$y_i$  = actual expenditure on food consumption per adult male equivalent

$\alpha = 0$ , incidence of poverty, the Headcount ratio, or the fraction of the population which lives below the poverty line.

$\alpha = 1$ , depth of poverty, how much government must give, on average, to each poor household to eradicate poverty.

$\alpha = 2$ , intensity of poverty, the index combines both poverty and income inequality among the poor.

$\alpha = 3$ , urgency of poverty, i.e., if we do not support them soon, they will likely starve.

Absolute poverty is measured in Foster\_Greer\_Thorbecke-Schoch coefficients calculation by arranging the income in ascending order. But the information of how poor or how much poor is needed cannot be measured. The income by the cost of basic needs method is used in the measurement of relative poverty. The incidence or headcount index of poverty shows the percentage of households in the population that receive less income or welfare than a given poverty threshold. The incidence, depth and intensity of poverty give us a chance to compare the two cyclone Nargis affected areas. The above formula was also adapted to measure the incidence, depth and intensity of unhappiness, where  $z$  = Likert scale of 1 to 5.

The income gap or depth of poverty is an indicator of how much people are deviating from the standard income on average. If we are interested in poverty reduction, the income gap tells us which population will need support. It will also tell us how much this population will need to climb out of poverty.

While the poverty gap indicator might indicate how much, on average, the whole population needs, the intensity of the gap is not yet counted. In the population of poor, some people are further away but some are not so highly deviated so when the gap is summed up and taken on average, the pattern of distribution is blended. The intensity can tell us which poor population is more unequal when we calculate the intensity of poverty. This is done by giving more weight to people who are further away from the poverty line. Intensity of poverty is an indicator that we can use to give the greatest weight to the person with the income farthest away from the standard poverty line. The distribution pattern of income of the poor is weighted by taking the square of the income gap. The intensity of poverty of a population is calculated by squaring the gap takes average on it.

One difficulty for this study is that Myanmar does not have an official poverty line. Expenditure on food items as a percentage of total household expenditures is a widely used indicator to measure household access to food. According to the Integrated Household Living Conditions Survey in Myanmar (UNDP, 2007), 69 percent of all household expenditures in Myanmar, including health, are spent on food nationally, which is very high. In all states and divisions the percentage exceeds 60 percent, which indicates the broad scope of household vulnerability towards food insecurity. Although average consumption expenditures of non-poor households are nearly twice that of poor households, the percentage spent on food is also high, indicating that vulnerability to food insecurity is not limited to the poor in Myanmar. The average cost of a monthly food basket at the poverty line is approximately MMK 118 492 (USD 99 at the unofficial exchange rate). The poverty line is K162, 136 per

adult per year. This poverty line was integrated by United Nation Development Programme in 2007 from the Integrated Household Living Survey.

#### **3.4.4 Tests of Means**

An independent sample t test was conducted by village, wealth status, Nargis affected status, gender and so on. The test was used to determine whether or not the mean values were significantly different between groups. Comparisons of means also give the researcher further opportunity to study the nature of two Nargis affected areas comparing similarities and differences. This allows for further refinement of the hypotheses by choosing promising variables and adapting the assumptions from these tables.

#### **3.4.5 Correlation Matrix Analysis**

For the correlation matrix analysis, each category of variables was subdivided into a set of specific nominal, ordinal and cardinal indicators in order to choose the most significant for hypothesis testing. Correlation matrices help find out the possibility of multicollinearity and highly associated variables in the hypotheses equations. It also indicates the direction of the relationship between variables and can help to identify the most suitable variable for our regression equations. If the absolute correlation coefficient that is higher than 0.5, it could lead to multicollinearity. Some variables were left out after the test of means and correlation matrices because of their low levels of significance.

### 3.4.6 Regression Analysis and Final Hypothesis Testing

Linear regression analysis was used for the hypotheses testing. The situation of the two different areas was determined by the other socio-economic factors, the Nargis affected status. Linear regressions were estimated using SPSS 17 to discover the most significant relationship between happiness and other socio-economic factors.

#### 3.4.6.1 Adjusted R Square

Adjusted R square is used to show how well future outcomes are likely to be predicted by the model. If adjusted R square is at least 0.5, then the variables in the model are strongly explained.

$$\text{Adjusted R square} = (\text{TSS} - \text{RSS}) / \text{degree of freedom}$$

Total sum square (TSS) means the sum of the square of each actual observation minus the true mean. The residual sum of squares (RSS) means the total sum square minus the variations explained by the regression itself. The degree of freedom is the number of observation deducted by the number of variables in the model including the intercept.

#### 3.4.6.2 F statistics and Level of Significance

F statistics is used to indicate whether the model is acceptable or not. It shows the external validity of each variable in the model. The model is acceptable when the F statistic is higher than 4.2.



### 3.4.6.3. Student T-statistics of the Regressors

In order for a given variable to be considered significant, its T-statistic must be greater than or equal to 2.58 for the 99% level of significance, 1.96 for the 95% level, or 1.67 for the 90% level. It is common practice to mark these three levels with \*\*\*, \*\*, and \* respectively. They also correspond to the probability of making an error in assuming the variable is meaningful of .01, .05, and .10 respectively.

### 3.4.6.4 Logic of the Signs

The sign of beta coefficient indicates the direction of the relation between hypothesized variables. If the beta coefficient is positive, the variable is directly related to the independent variable and if it is negative, the variable is inversely related to the dependent variables in the model.

## 3.5 The Variables

In order to design the questionnaire, key respondents in the study area were asked about the factors or variables that could be significant in satisfying their basic needs. Several sets of variables have been introduced in order to identify the factors that determine basic needs fulfillment. In this section, the research used the selected satisfiers of the household and adapted them to the perception of subjective well-being of respondents.

First, the research focused on characteristics of household income, livelihood and social variables. The variables can be described as:

- (1) The income per year of the household is calculated as the sum of profits from agriculture, wages of all members of the family, seasonal income from casual work and others sources of income. In this case, income from agriculture is calculated from the sum of income from crop per season. And also income from remittances.
- (2) Livelihood variables such as respondent time in his/her own field (hours/day), time in a secondary occupation (hours/day), quantity of land they own (measured in acres), and number of jobs in the family.
- (3) Social variables included age, education level of all members of the family, occupation of household members, dummy variable indicating the household head sex, number of household members and sex of members.

For the food section, respondents were asked about the food groups they consumed within a seven day period, the average number of meals they are in a day, a dummy variable if food based on whether the food they ate within a seven day period is enough or not, and coping strategies used if food was not enough.

For water and sanitation, the data collected covered the main sources of drinking water before and after Nargis for the rainy season and dry season, the number of minutes it takes to collect that drinking water for the household, a dummy variable if they treat water for safety, the way they treat the water, the number of water containers in a household, toilet facilities and ways to eliminate the waste and the household heads' habit of using soap.

In terms of the health situation of the research area, the respondents were asked the about the health services of the research area, how they get to the health

care center, how many times they use the health care centers services, the availability of medicines and equipment, and how long they have wait to see the health care center staff.

Finally, the questionnaire investigated the housing conditions of the respondents. It asked about the major construction of the wall, roof and floor before and after Nargis, the type of dwelling the household lived in before and after Nargis and other housing-related topics. Dummy variables were used to indicate whether the respondent considered their dwelling to be strong and sturdy or not. In this section we estimate the selected satisfier of the household influence the perception of basic needs satisfaction. The variables used in this section are used to explain the influences on the basic needs. Although their importance may vary between households, the research intention is to find the general pattern. For the welfare section, the respondents were asked about their leisure, religion and social networking. For social capital, length of stay in the area and their level of satisfaction in their own land and neighborhoods were asked to gauge respondents' perception on their social surroundings.

Some indicators in the questionnaire were straightforward. However, there were other indicators such as social networking and happiness that cannot be measured with a simple scale since they are more abstract. The problem is solved by dictating the different facts of the single problem and then combining all the simple scales to make one composite score. To add up all the information, the scale has to be a whole number so that the observations will not be blended. The list of composite

variables is displayed below and there is also an explanation of how the composite scores are derived.

Table 3.2 Calculation Formulae of the Composite Variables

Variable Name	Formula
Social Capital	types of neighborhoods + agree or disagree that this neighborhood of people with different backgrounds get on well together + contact organization to solve the problem affecting people in the respondents' local area + share information to others if there is anyone the respondents could ask for help + involvement in the groups
Physical Happiness	I look attractive + I feel fresh and full of energy + I have great physical well-being at this time + I will probably be free from diseases
Mental happiness	I love to challenge my mind to create something new and different + I often lose myself in the flow of my work or hobbies + I feel secure + I do have a happy memories + I have no worries.
Emotional Happiness	I don't regret the past + I can easily/quickly calm down when I feel bad/stressed + I find myself smiling when no one else is around + I feel peace in my heart
Spiritual Happiness	I can laugh even in the face of difficulty + I know myself and exactly what I want/need + I respect myself + I have clear faith or religion + I feel great peace in my soul + I willingly accept what happens to me + I accept that others may have different beliefs + I feel my life is very bright + I can be myself however I am or wherever I go + I am satisfied about everything in my live
Social Happiness	I love to share with other people + I feel harmony with other people + I am not afraid of what others may think + My own freedom is more important than what others think
Lyubomirsky Subjective Happiness Scale	In general, I consider myself a very happy person + Compared to most my peers, I consider myself happier + Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. To what extent does this characterization describe you? + Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. To what extent does this characterization describe you

### 3.6. Questionnaire

The questionnaire for this study was designed to inquire about the socio-economic situation, the use of economic infrastructures, social capital and happiness level of the research areas. The time frame of household interviews was set up to last between forty-five minutes to one hour.

A pilot survey was conducted in May 2010 in Pyapon Township. The pilot survey data revealed the local norms, gaps of information, and the feasibility of the data analysis questions. Because of the pilot survey, the researcher also discovered some assumptions that had to be adjusted before the main survey was disseminated to the research areas. These insights were used to adapt and refine the questionnaire for the research project.