3

Chapter Research Design and Methodology

3.1 Introduction to Research Methodology

Research methodology is one of the vital parts of any research. Appropriateness, accuracy and reliability of research result depend on the methodology the researcher has adopted. Present chapter deals with the entire research design and methods adopted to carry out the research systematically. The chapter also covers the methods of selecting the study area, determination of sample size, sampling frame and procedure, tools and techniques of data collection etc. In addition, the present chapter describes the different methods of analysing data according to the objectives of the study, conceptual framework of the model, specification of the models etc.

3.2 Design of the study

The present study is based on the primary data collected through survey method by interviewing the respondents living in Siliguri Municipal Corporation Area (SMCA) with some structured open-ended and some close-ended schedules. Cross-sectional data were collected from a sample of 400 households spreading over Siliguri Municipal Corporation Area (SMCA). The study is confined to the SMCA of the state of West Bengal. The area constitutes some part of plains of Darjeeling district and North- Western part of Jalpaiguri district. Further, SMCA has 47 administrative wards (municipal wards) under Darjeeling district and Jalpaiguri district. But, number of municipal wards was not equally distributed in both districts at the time of accreditation to the corporation city from municipality. So, out of total 47 wards, 33 wards fall under the jurisdiction of Darjeeling district and remaining 14 wards fall under the jurisdiction of Jalpaiguri district.

3.2.1 Selection of the study area

A multi-stage sampling method was used for the present study. In the first stage, to ensure adequate representation of the wards of SMCA under Darjeeling district and Jalpaiguri district, a simple systematic sampling technique is applied for selection of wards. Following the simple systematic sampling technique, 14 wards of SMCA out of total 33 wards falling under the Darjeeling district jurisdiction and 6 wards of SMCA from total of 14 wards falling under the Jalpaiguri district jurisdiction were obtained. Thus, a total of 20 wards from the total of 47 wards of SMCA were selected for the study. In the second stage, probability proportion to population size (PPS) method was used to calculate the sampling unit from the each selected ward of SMCA (under study) falling under the jurisdiction of Darjeeling district area and jurisdiction of Jalpaiguri district area.

3.2.2 Selection of ward number

A simple systematic random sampling technique was applied for the selection of wards under study and it was done in the following way:

Total number of wards in Siliguri Municipal Corporation Area (SMCA) = 47

Total number of wards under Darjeeling district = 33

Total number of wards under Jalpaiguri district = 14

Proportion of wards (Sampling Fraction) (33/14) =2.36

Now, first any ward number between 1 and 2 was chosen randomly, and then the sampling fraction (2.36) was added to the selected ward number. Thus, we obtained the second study ward. Again, to get the third study ward, the sampling fraction (2.36) was added to the

second study ward and the procedure was continued till the last ward number was reached. Following this procedure, all the calculated ward numbers are of fraction in nature. To deal with this problem, each calculated ward number was rounded off at one decimal point. If the ward number having calculated decimal value equal to or above .50 (i.e. decimal value >= .50), absolute value of 1(one) was added to the whole digit number of that ward number. In this way, we identified the ward numbers for the study. For example, we first randomly chose ward number 1 and then sampling fraction (2.36) was added to this ward number 1. Here, we obtained second ward number 3.36 but decimal value is less than 0.50, so we took only 3. Thus, our second study ward was ward number 3.But in the third case, calculated value was 5.72, by rounding off at one decimal point, we got ward number 6 (i.e. third study ward) and procedure was applied till the last number of ward (i.e. ward number 47) was reached. Thus, we obtained 14 wards of SMCA from Darjeeling district area and 6 wards of SMCA from Jalpaiguri district area for the study.

Table 3.1: Ward Number Selection Procedure

Serial Number	Particular (ward number +sampling fraction)	Ward Number (without Round Off)	Ward Number (Rounded Off)	Covering District
1	1*	1	1	Darjeeling
2	1+2.36	3.36	3	Darjeeling
3	3.36 + 2.36	5.72	6	Darjeeling
4	5.72 + 2.36	8.08	8	Darjeeling
5	8.08 + 2.36	10.44	10	Darjeeling
6	10.44 + 2.36	12.80	13	Darjeeling
7	12.80 + 2.36	15.16	15	Darjeeling
8	15.16 + 2.36	17.52	18	Darjeeling
9	17.52 + 2.36	19.88	20	Darjeeling
10	19.88 + 2.36	22.24	22	Darjeeling
11	22.24 + 2.36	24.60	25	Darjeeling
12	24.60 + 2.36	27.96	28	Darjeeling
13	27.96 + 2.36	30.32	30	Darjeeling
14	30.32 + 2.36	32.68	33	Jalpaiguri

15	32.68 + 2.36	35.04	35	Jalpaiguri
16	35.04 + 2.36	37.40	37	Jalpaiguri
17	37.40 + 2.36	39.76	40	Jalpaiguri
18	39.76 + 2.36	42.12	42	Jalpaiguri
19	42.12 + 2.36	44.48	44	Jalpaiguri
20	44.48 + 2.36	46.84	47	Darjeeling

*Selected randomly

From the above table, it is clear that ward number 1, 3, 6, 8, 10, 13, 15, 18, 20, 22, 25, 28, 30 and 47 from Darjeeling district area and ward numbers 33, 35, 37, 40, 42 and 44 from Jalpaiguri district area were selected for the study respectively.

3.2.3 Determination of sample size

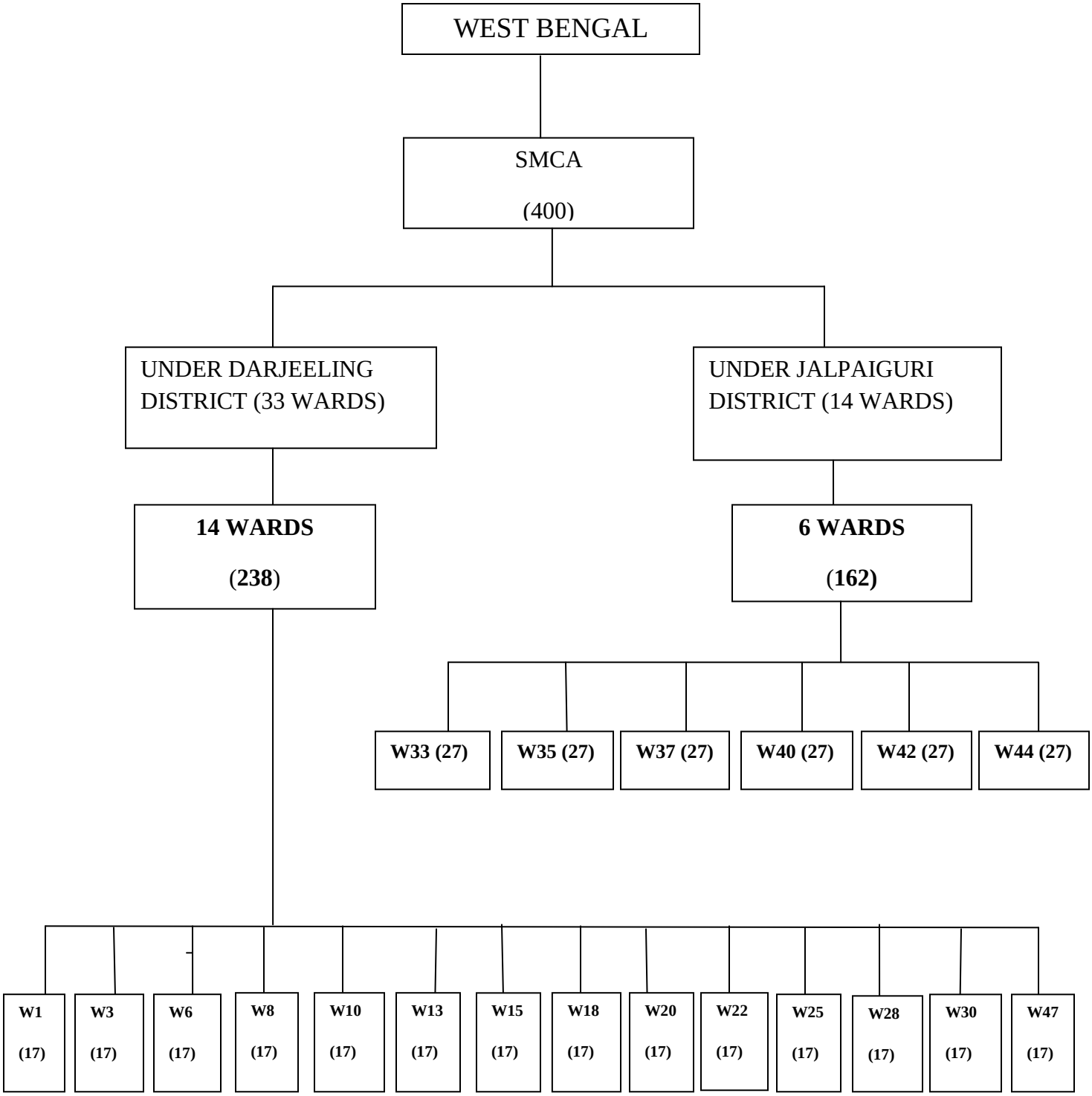
Determination of sample size is vital for any research study because it represents all the characteristics or features of the whole population. Moreover, the coverage, quality, adequacy and accuracy of study results primarily depend on sample size. Besides some non-technical factors, sample size also depends on many technical factors such as purpose of the study, proportion of the population having the characteristics under study, degree of accuracy and level of margin for the study etc. However, in the present study, sample size was calculated using the formula $n= Q / (P*\alpha^2)$; where n = Sample size to be estimated, P = Prevalence rate of disease among the households (without multiplying 1000), Q = (1-P) and α is equated with the standard error of the estimated prevalence rate (i.e. level of margin for the study). The pilot study on 50 households each living in Darjeeling district area and Jalpaiguri district area under SMCA reveals that the annual prevalence rate of disease (proportion of persons who are exposed to utilisation of public health facilities) for the said two districts are 0.288 and 0.372 (without the multiplier 1000) respectively. The average size of a household in SMCA was 4.17 (i.e. 4,72,374 /1, 13,269) according to 2001 Census. This present study considered the level of margin at 5 % (i.e. $\alpha = 0.05$). Putting all these values in the formula, we have obtained n=0.712/ (0.288* 0.0025) ≈989 persons or 989/4.17 ≈ 237 households for Darjeeling district area. But for the sake

of simplification, we considered 238 households to equally divide the total households among the wards of Darjeeling district area. So, we divided the total 238 households of Darjeeling district area by 14 wards (obtained by simple systematic sampling technique). Thus, we obtained 17 (i.e. 238/14) households from each selected wards of Darjeeling district area for the study. On the other hand, by applying the same formula, for Jalpaiguri district area,we obtained n=0.628/ (0.372* 0.0025) ≈ 675 persons or 675/4.17 ≈ 162 households. Again, by dividing total 162 households of Jalpaiguri district area by 6 wards (obtained by simple systematic sampling technique), we obtained 27 (i.e. 162/6) households from each selected wards of Jalpaiguri district area. Thus, for the SMCA as a whole, the required minimum sample size is 1664 (i.e.989+ 675) persons or 400 (i.e. 238+162) households.

3.2.4 Selection of households

For the present study, households were selected by random sampling technique. It was done through several steps. At first, a list of all the houses of the respective ward was prepared as per holding number issued by Siliguri Municipal Corporation. Then each house in the ward was assigned an identification number in the numerical order. The process was applied on all selected wards obtained by simple systematic sampling technique, irrespective of whether character (i.e. disease) presents or not. Thus, sample frame was prepared. Then, randomly twenty (required seventeen plus three extra considering non-response, absence etc.) numbers from each of selected wards of SMCA under Darjeeling district area and thirty ((required twenty-seven plus three extra considering non-response, absence etc.) numbers from each of selected wards of SMCA under Jalpaiguri district area respectively were selected from a random number table (by starting anywhere in the table and then reading either horizontally across the rows or vertically down the columns). In this way, two hundred and eighty households from Darjeeling district area and one hundred and eighty households from Jalpaiguri district area falling under SMCA were selected. However, for SMCA as a whole, total four hundred households were selected from for the study taking 238 from Darjeeling district area and 162 households from Jalpaiguri district area respectively.

Figure 3.1: Multi-stage Sampling



Ward number (Number of Households) is given in parenthesis.

Table 3.2: District-wise Distribution of Sampling Unit

Covering District	Siliguri Municipal Corporation Area (SMCA)		Total Number of Households Selected
	Number of Wards	Sampling Units	
Darjeeling	14	17	238
Jalpaiguri	6	27	162
Total	20		400

According to 2001 Census, wards under Darjeeling district area of SMCA constitutes nearly 60 percent and wards under Jalpaiguri district area of SMCA constitutes 40 percent of total population respectively. Keeping this in mind, total sampled households were proportionately distributed according to population size of the area. For the present study, a total of 17 households from each selected 14 wards of SMCA under Darjeeling district area and a total of 27 households from each selected 6 wards of SMCA under Jalpaiguri district area were randomly selected. Thus, total 238 households from Darjeeling district area and 162 households for Jalpaiguri district area were chosen for the study. Finally, total sample size for the study area (i.e. SMCA as a whole), is 400 households, which also satisfies the condition of proportion to population size. In Darjeeling district area, sample households is 238 which is about 60 percent (i.e. 238/400*100) of total sample households (i.e. 400) and In Jalpaiguri district area, sample households is 162 which is about 40 percent (i.e. 162/400*100) of total sample households (i.e. 400).

3.3 Research Methodology

3.3.1 Methods of data collection

The present study used cross-sectional household survey method to investigate the research questions. The information from the households were collected through the pre-designed, pre-tested, structured schedule comprising of some open-ended and some close-ended questions. The schedule was prepared after thorough literature search to meet the objectives of the study. A total of 400 households were interviewed (238 in Darjeeling district area and 162 in Jalpaiguri district area). Here, a household is considered as a group of family members or individuals who live in the same house and who share the common benefits or arrangements like toilet, housing area, source of drinking water, utensils etc. Interview was carried out with the head of the family or with his wife or with any interested member who has the knowledge of every member of that family in their first language. Females were interviewed regarding the type of disease, severity of disease, number of disease episodes experienced by the each individual family members, duration of the disease, cleanliness, sanitation, household asset etc. Males were interviewed, specially, regarding the utilisation pattern of healthcare for different disease episodes of each family member, insurance ownership and claims, other health expenditures incurred during the reference period. More specifically, educated member of the family was requested to give the information regarding a) Identification of the household; b) Household level characteristics; c) Demographic, Economic and Epidemiological profile of each household member; d) Utilisation of healthcare facilities; and e) Different components of OOPHE incurred on the household members during the reference period of twelve months (i.e. one year) preceding the interview.

Table 3.3: Ward-wise distribution of total number of sampled households, number of family members in the sampled households

Covering District	Ward Number	Number of Sampled Households	Number of Family Members in the Sampled Households
Darjeeling	1	17	78
	3	17	72
	6	17	74
	8	17	96
	10	17	79
	13	17	62
	15	17	64
	18	17	73
	20	17	62
	22	17	72
	25	17	61
	28	17	80
	30	17	66
Jalpaiguri	47	17	94
	33	27	105
	35	27	111
	37	27	107
	40	27	112
	42	27	103
	44	27	71
Total SMCA	20	400	1684*

* Our planned minimum sample size was 1664 (according to pilot study), but through the survey, we found 1684 samples or persons from 400 households, which is more than our planned ones. Therefore, we continued our study with the sample size of 1684.

Table 3.4:District-wise distribution of Total Number of Sampled Households, Number of Family Members in the Sampled Households

Covering District	Numbers of Ward	Number of Sampled Households	Number of Family Members in the Sampled Households
Darjeeling	14	238	1033
Jalpaiguri	06	162	651
Total	20	400	1684

Table 3.4 depicts the district-wise distribution of sampled households and sampling units. Out of total selected 20 wards, 14 wards from Darjeeling district area and 6 wards from Jalpaiguri district area were selected for the study. Further, 17 households from each 14 wards from Darjeeling district area and 27 households from each 6 wards from Jalpaiguri district area have chosen respectively. Thus, we have obtained a total of 238 households from Darjeeling district area and a total of 162 households from Jalpaiguri district area. Again, through the field survey, we obtained a total of 1033 persons from 238 households in Darjeeling district area and a total of 651 persons from 162 households in Jalpaiguri district area under SMCA. Thus, for the SMCA as a whole, we obtained a total of 1684 persons from 400 families (sampled households).And, the study is ready for analysis with this sample size of 1684 as it satisfies the criteria of minimum sample size of 1664 obtained through the pilot study conducted.

3.3.2 Methods of analysing data

The household survey explored the households' socio-economic and demographic status, epidemiological profile and households' assets, healthcare utilisation patterns, out-of-pocket healthcare expenditure, healthcare financing in the study area. Therefore, whole study is broadly categorised into three parts, viz. burden of disease or epidemiological profile of the household members, health-seeking behaviour of the household members or utilisation pattern of healthcare

facilities and out-of-pocket healthcare expenditure incurred by the households during the reference period of last twelve months (i.e. one year). Percentage method and standard formula of morbidity prevalence were applied to compare the disease pattern with respect to different demographic and socio-economic characteristics of the respondents or household members of Darjeeling district area, Jalpaiguri district area and combined area (i.e. for SMCA as a whole) respectively.

3.3.2.1 Examination of the epidemiological profile of the people of SMCA

To be familiar with epidemiological profile of the people of any region, it is relevant to study the burden of disease of household members, as it deals with the nature health status of the people caused due to disease, disability and injury. Though several measures such as active life expectancy (ALE) (Katz S et al., 1983), disability-free life expectancy (DFLE), disability-adjusted life years (DALYs) study (Murray CJL and Lopez A, 1996, 1997) are available to quantify the disease burden of the people living in any area, but these measures primarily consider the hospital based records and incidence of disease. However, in the present study, burden of disease is measured in terms of morbidity and disability of the people residing in the study area, as reliable data on mortality is rare or not available in this region. Moreover, number of hospitalisation case is not adequate enough in the study area. Rather, present study considered both the incidence and prevalence of diseases experienced by the household members residing in SMCA during the reference period of twelve months. Further, the study considered all illness episodes including hospitalised (in-patient stays) and non-hospitalised (out-patient visits) cases. In order to get systematic and comprehensive analysis of epidemiological profile or burden of disease of the households in terms of morbidity, sampled data on different diseases, disabilities and injuries were grouped under different disease names according to the modalities of Global Burden of Disease (GBD) 2010 study as recommend by World Health Organization (WHO) (Murray and Lopez, 1996). Further, to analyse the nature and types of disease suffered by the household members during the reference period of last twelve months (i.e. one year), reported diseases have been grouped under three broad categories such as Group I, Group II and Group III. This categorisation of diseases was done on the basis of GBD study 2010. Group I represents the diseases such as communicable, maternal, peri-natal and nutritional conditions; Group II represents the all major non-communicable diseases and Group III represents intentional and

unintentional injuries including accidents. The present study primarily considered the self-perceived morbidity based on the perception and reporting of symptoms and impairments by individual (Murray and Chen, 1992), and to some extent, disease names as told by the physicians or the medical store sales persons after hearing the symptoms from the sick person, not on official medical records or clinically diagnosed. The observed distributions were cross tabulated according to different background characteristics of the respondents or household members such as age, gender, caste, occupation education level, economic class, households' monthly income, and place of residence of the households in the study area. Further, the diseases suffered by the household members were classified with regard to category of disease, severity of the disease, nature of disease and number of days of suffered during per disease episode. The study considered all types of illnesses episodes as reported by the respondents which were categorised as below (Hill, 1966):

a) Episodes commencing during the reference period and ending within the reference period.

b) Episodes commencing during the reference period and still existing at the end of the reference period.

c) Episodes existing before the commencement of the reference period and ending within the reference period.

d) Episodes existing before the beginning of the reference period and still existing at the end of the reference period.

In the first two cases incidence rate of disease and in later two cases prevalence rate of disease were measured respectively. Here, the measurement of incidence and prevalence rates of disease of the households during the reference period of 12 months (i.e. one year) was done using the standard formula as recommended by Expert committee on Health Statistics of the World Health Organisation (WHO). According to WHO,

$$\text{Morbidity Incidence Rate} = \frac{\mathrm{I}}{\mathrm{p}} x1000 \ .$$

Here, Incidence rates are useful for the study of acute diseases which last for three months or less and peak severity of symptoms occurs and subsides within day or weeks.

$$\text{Morbidity Prevalence Rate} = \frac{\mathrm{C}}{\mathrm{p}} x1000 \ ,$$

where I = Number of new cases of disease during the 12-month reference period (i.e. one year), C = Total number of spells (in all the four categories)during the same reference period and P =Total number of population exposed to the risk of getting affected by disease during the same reference period.

On the other hand prevalence rates consider the number of new and old cases of a disease in a population in a given period of time. Here, prevalence rates are important for the study of chronic diseases which last longer than three months. In addition, disability is measured considering the percentage of people experiencing deterrence from normal activities due to physical or mental impairments during the reference period of 12 months (i.e. one year).

However, to calculate hospitalisation rate, the study used the formula as:

Annual hospitalisation rate = [number of persons hospitalised in the 12-month reference period (i.e. one year) / total number of persons in the sample households]* 1000.

Finally, to measure the burden of disease in terms of morbidity, prevalence and incidence rates were computed separately for Darjeeling district area, Jalpaiguri district area and combined area (i.e. SMCA) respectively. Further, morbidity prevalence rates were cross tabulated according to different demographic and socio-economic background characteristics of the household members and explained with suitable charts, diagrams etc.

3.3.2.2 Analysis of health seeking behaviour or utilisation of healthcare facility the people of SMCA

Theoretically, utilisation of healthcare facilities is defined (Majumder, 2006a, 2014)in three ways: a) Utilisation of healthcare facilities from 'modern source' where opinions or advices are taken from doctors and medical experts by one group , and utilisation of healthcare facilities from 'traditional source' where treatment is sought from paramedical staff including personnel in chemist's shop, family treatment , or by self-medication or and from any systems of medicine (including ayurveda, yoga, unani and others), excepting allopathy and homeopathy by other group; b) Utilisation of healthcare facilities from 'public', 'private' or other institutions (including NGOs and charitable organisations); c) Utilisation of facilities from 'allopathy', 'homeopathy', or 'other' systems of medicine. Therefore, utilisation is the visit of any healthcare facilities by a person during reference the period of twelve months (i.e. one year) and utilisation

rate refers to the number of times the total population use healthcare facilities during reference the period of twelve months (i.e. one year).In the survey, respondents were asked whether they had visited any healthcare facilities to seek treatment during the reference period of twelve months (i.e. one year) or not. If the response is positive, pattern of healthcare utilisation facilities such as modern and traditional, system of medicine and sources healthcare facilities utilised were recorded in the survey. It was also enquired, whether there was any hospitalisation case among the family members during the reference period or not. If the response is positive, number of hospitalisation cases and number of hospitalisation days experienced were also recoded. In addition, distance to the nearest healthcare facilities and modes of transport used to visit the healthcare facilities were also recorded. On the other hand, if the answer is no, then reasons for not visiting any healthcare facilities were asked and how did they manage themselves to overcome from disease episode, that was also the part of the survey. The responses were coded accordingly. Then, all the observed figures were cross tabulated according to different demographic and socio-economic background characteristics of the household members separately for Darjeeling district area, Jalpaiguri district area and combined area (i.e. SMCA) respectively and were expressed in percentage figures and presented by suitable charts, diagrams etc. Finally, health seeking behaviour or utilisation of healthcare services were analysed according to pattern of utilisation, nature of utilisation (type of visit), sources of care and system of medicine utilised by the sick people of study area. On the other hand, in order to examine the effects of burden of disease on utilisation of healthcare facilities, firstly, utilisation pattern of the people or household members was cross tabulated according to category of disease, severity of disease, nature of disease, number of days of suffering of the individual members of the household during the reference period of twelve months (i.e. one year) separately for Darjeeling district area, Jalpaiguri district area and combined area (i.e. SMCA) respectively and expressed in percentage figure.

Finally, to find out the impact of different components of burden of disease on utilisation of healthcare services of the sick people, five different logistic models were constructed with respect to (1) utilisation of all types of healthcare facilities, (2) utilisation of healthcare facilities from modern care in contrast to traditional care, (3) utilisation of IPD services in contrast to OPD services, and (4) utilisation of allopathy and ayurveda system of medicine in contrast to homeopathy medicine, and (5) utilisation of public and private healthcare facilities in contrast to

other sources of care. In each of the first three models, parameters were estimated separately by using binary multivariate logistic regression models and in the last two models, parameters were estimated by applying multinomial logistic regression models.

3.3.2.3 Study of out-of- pocket healthcare expenditure and other related costs

Healthcare expenditure is defined as the total expenditure incurred by any individual or household on restoration of health status from disease during reference period, irrespective of the pattern of utilisation, system of medicine and type of healthcare facility utilised. Total healthcare expenditure is broadly divided into two categories: a) direct healthcare expenditure, and b) indirect healthcare expenditure. Direct healthcare expenditure includes doctors' fee, cost of drugs (medicine), diagnostic test charges, hospital or nursing home charges including surgery etc. On the other hand, indirect healthcare expenditure includes helper costs, total travel costs beginning from first visit to doctor to the end of disease episode during the reference period, special diet costs, expenses of accompanying parson(s), tips, rituals in case of hospitalisation, home care payments, loss of earning of the ill person as well as the accompanying persons during the course of treatment. Sum of these direct expenditures and indirect expenditures is the total healthcare expenditure of the family. But the present study considered out-of-pocket healthcare expenditure (OOPHE) as the total payments made by the individuals or household directly from their own resources to get recovery from illness. It includes any payments to doctors, dentists and other health practitioners; payments for medicines and other medical appliances used covered by co-insurance scheme; balancing bills by physicians; payments for non-covered non-prescription medicines (e.g. self-medication) and hospitalisation charges including surgery and other miscellaneous healthcare goods and other service payments including transportation costs, lodging costs and expenditure on food consumed away from home during health visits for the reference period of last twelve months (i.e. one year). Respondents were asked about the information related to direct expenditure made by him or her or other members of the family during the reference period. The survey also recorded the sources of healthcare expenditure (such as public expenditure, expenditure from own sources or others) used by the household for every member of the family. And how household had financed their healthcare expenditure, it was also the part of survey. It was, further, investigated through the survey whether the household had been protected by any type of health insurance or similar type of benefits or not.

In addition, survey recorded the total sum of money spent by the household on the health or treatment from their own pockets and also the total payments made for health insurance or similar type of benefits during reference period under study. Then, present study calculated the total OOPHE by summing up the payments made by a household for all the members of the family on different components of OOPHE (i.e. public hospital card/ registration fees, transportation costs , doctors'/consultation fees, diagnostic test charges, private health insurance or similar type of premiums, medicine costs, hospital or nursing home charges including surgery not covered by health insurance and other miscellaneous expenditure including the physiotherapy, gym or yoga fees etc.) during the reference period under consideration. Different components of OOPHE were further analysed in order to better understand share of each component in total OOPHE incurred by the households during the reference period. OOPHE of the people or households was cross tabulated according to category of disease, severity of disease, nature of disease, number of days of suffering, hospitalisation or non- hospitalisation cases. Further, OOPHEs were analysed according to demographic and socio-economic characteristics of the sick people to show the variations of healthcare expenditure by each category. Finally, all the observed distributions were cross tabulated and expressed in average and percentage figures and presented by suitable diagrams for the study area (i.e. SMCA).

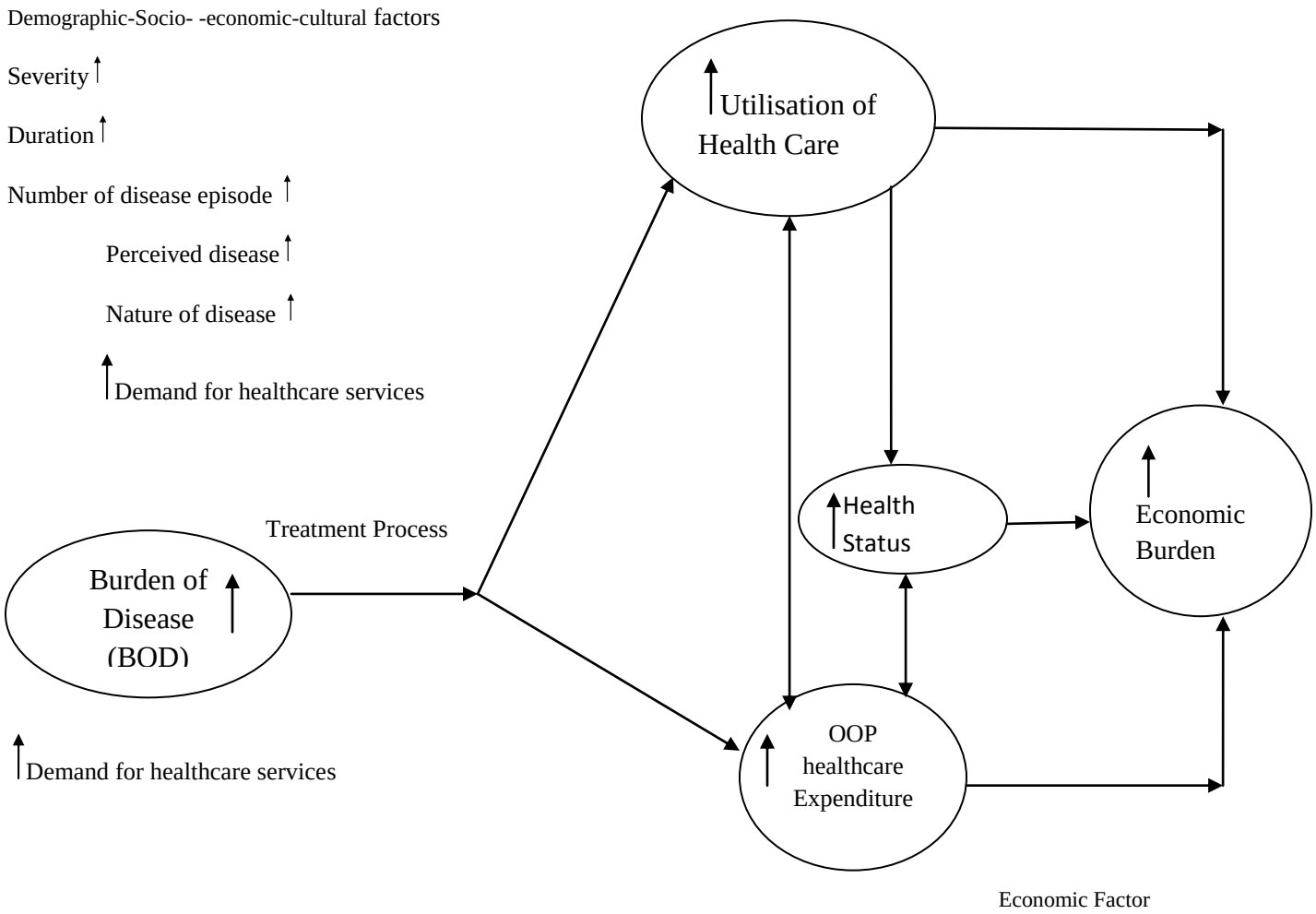
3.3.2.4 Examination of relationship between out-of-pocket healthcare expenditure and utilisation of healthcare facilities of the people of SMCA

In order to examine the relationship between OOPHE and utilisation of healthcare facilities, OOPHE of the sick people were cross tabulated according to pattern of care (modern or traditional), different sources of treatment (public, private or others), type of visit (hospitalisation or non- hospitalisation case) and systems of medicine (allopathy, homeopathy, ayurveda, and others) utilised for the treatment of different of illness episodes experienced by any member(s) in the family during the same reference period. Finally, by applying econometric approach, a log linear multivariate regression model in log-log form was developed to measure the impact of burden of disease and other related factors on OOPHE incurred by the people of SMCA through the utilisation of healthcare services. Finally, co-linearity test is done to check whether there is any correlation among one predictor variable and the remaining predictor variables.

3.4 Conceptual Framework of the model

Diagrammatical presentation depicted below shows how burden of disease impacts the utilisation of healthcare services and OOPHE by the households through the process of treatment or health seeking behaviour, which is influenced by socio-economic, cultural and demographic factors of the individuals or the household directly or indirectly. When health status of the people decline or when more people suffer from any kind of physical or mental illness, burden of disease in the society increases. Increase in burden of disease causes an increase in the demand for healthcare services. Further, increase in the demand for healthcare services results into an increase in both the utilisation of healthcare services and OOPHE by the sick people or the household members.

Figure 3.2: Impact of BOD on Utilisation of Healthcare service and OOPHE



Thus, the more will be the burden of disease, the more will be the demand for healthcare services, and this in turn leads to an increase in utilisation of healthcare services. Further, the more will be utilisation of healthcare services, the more will be OOPHE. On the other hand, the more will be OOPHE, the more will be utilisation of healthcare services. OOPHE by the household can be a good reflection of demand for healthcare services. OOPHE of the households is backed by affordability or economic profile of the household. Individual or group health insurance coverage plays vital role in this regard. Health insurance coverage by the household may influence them to utilise more healthcare services with the expectation of reimbursement. It may lead to decrease in out-of-pocket health expenditures by the households, which, in turn, improves the health status and reduces the burden of diseases. Thus, it is expected that high income-group household and insured household has high capacity to bear high OOPHE than the others. On the other hand, utilisation of healthcare services is demand as well as supply driven phenomenon. Therefore, to improve the health status of the people or to reduce the burden of disease, both the utilisation of healthcare services and OOPHE should be increased accordingly.

3.4.1 Explanation of the model

When a person perceives himself or herself as sick or when a person's normalcy is disturbed due to physical, biological or psychological reason, the person feels the need or demand for healthcare services to get relief from sickness. Similarly, if the family member or others feel the health disorder of small children or aged persons who cannot express their own difficulties, then also medical care or health service care is sought. Perception of illness is influenced by various socio-cultural, behavioural factors. Different questions come to the mind of the sick person or other members of the family such as what type of disease it is, how severe it is and associated risks, where to go for treatment, how much would be the treatment costs etc. First, categorisation of disease or disease type such as communicable, non-communicable or others needs to be perceived which is influenced by several factors such as past experience, factual knowledge, awareness and other personal factors like education, consciousness etc. After perceiving the severity and associated risks of disease, it is important to take decision of whether to visit any healthcare service provider or not (i.e. utilise the healthcare facilities or not). It is expected that change in type of disease, severity of disease, nature of disease and duration of sufferings of the

ill person lead to a change in utilisation of healthcare facilities. On the other hand, the longer will be duration of disease the more will be utilisation of healthcare services and vice versa. Thus, high disease burden leads to a high utilisation of healthcare services subject to the availability of healthcare service in that area and affordability of the household. Further, socio-economic, cultural and demographic factors of the household may influence to take decision regarding the choice of a particular pattern of care, sources of care, system of medicines etc. However, it is perceived that educated and urban people utilise the modern healthcare services and people living in the backward areas or rural areas utilise the traditional healthcare services. It may be due to lack of awareness, on-availability of modern healthcare facilities or others. Further, healthcare expenditure may also depend on the choice of healthcare service providers. It is also expected that hospitalisation or outdoor patient (OPD) visit at formal private healthcare provider much is costlier than those at public healthcare and other charitable healthcare institutions. On the other side, poor health has significant effect on OOPHE incurred by the individual member or the households. As disease type, degree of severity, number of disease episodes and duration of disease episodes changes, OOPHE by the household also changes in the same direction. Further, an increase in OOPHE leads to a less access to healthcare services which increases the disease burden. However, it is believed that good health improves the productivity, increases income and this, in turn, leads to an improve the economic well-being as well as the living standard of the individual. On the contrary, if more expenditure is made to increase the utilisation of healthcare services with the expectation of getting better health or reducing the burden of disease, household may have to face financial crisis which would affect the economic well-being adversely and thereby, would reduce the living standard of the people. Therefore, model portrays that burden of disease has considerable impact on utilisation of healthcare services and OOPHE incurred by the household.

3.4.2 Specification of the model

To justify the title of the thesis, two separate econometric models were presented. The first model deals with the impact of burden of disease on the utilisation of healthcare services and the second model shows the impact of burden of disease on OOPHE incurred by the household. In the present study, burden of disease is expressed in terms of morbidity, which is composed of

category of disease, severity of disease and nature of disease experienced by the household members during the reference period of last twelve months (i.e. one year). These diseases are based on the self-perception and reporting of symptoms and impairments by individuals (Murray and Chen, 1992) and to some extent, disease names as told by the physicians or the medical store sales persons after hearing the symptoms from the sick person. Here, severity of disease indicates the degree of harshness of the attack of illness. Severity of disease has also been classified into three categories such as low, medium and high. Low severity represents a situation in which a person performs all normal activities with symptoms, medium severity represents the impairment of some activities due to attack of disease and high severity represents a situation in which a person is bedridden with chronic condition for more than 7 days per disease episode during the reference period. When a person visits any type of healthcare services with more than one symptom or illness, the disease which is more severe, was included in the study and choice of care for more severe disease was considered. Lastly, nature of disease is proxied by duration of illness episodes. If person suffers from disease for less than or equal to 30 days, it is considered as acute disease and when the illness episode continues for more than 30 days, it is considered as chronic disease.

3.4.3 Modelling the impact of burden of disease on the utilisation of healthcare services

The first model tries to find out the explanation of how utilisation behaviour or health seeking behaviour of the household members changes with respect to change in different components of the burden of disease such as category of disease, severity of disease, nature of disease. In this context, when a person perceives himself or herself as sick or if normalcy is disturbed due to physical or psychological reasons, the first question revolves around whether healthcare facilities will be utilised or not. If he or she convinces himself or herself or convinced by other member (s) of the family or by others to utilise the healthcare facilities, now question arrives at which type of healthcare facilities will be utilised, whether modern care or traditional care .Modern care includes opinions or advices of the doctors and other medical experts; and traditional care means the treatment which is sought by a person from paramedical staff including personnel in

chemist's shop, family treatment, or by self-medication and from any system of medicine (including ayurveda, yoga, unani and others), excepting allopathy and homeopathy by other group. After deciding the pattern of healthcare facilities, the person faces with the choices of which system of medicines to be followed to get cured from illness episode. The choices are many such as allopathy, homeopathy, ayurveda, yoga, physiotherapy, unani etc. Finally, the person decides the sources of care (viz. public, private, charitable healthcare organisation or others) for the present disease episode. However, decision of whether healthcare services will be utilised or not, is of dichotomous nature. Here, utilisation of healthcare facilities is considered as an event (Béland, 1988**)** and the event of utilisation of healthcare services is categorical and binary (dichotomous) in nature. Since**,** the predicted responses are dichotomous, i.e. the response is 'yes' if character is present, coded as 1 and the response is 'no', if the character is not present, coded as 0. Present study considered the burden of disease as the explanatory variable (independent variable) and utilisation of healthcare services as dependent variable. As logistic regression model measures the relationship between categorical dependent variable and one or more independent variables by estimating probabilities using logistic function, a binary-multivariate logistic regression model and three multinomial logistic regression models were estimated to find out the impact of different components of burden of disease on utilisation of healthcare services.

3.4.3.1 Binary multivariate logistic regression model

The merit of using logistic regression function is that it can take any value for the independent variable from the range of -∞ (infinity) to + ∞ (infinity), but the dependent variable always takes the values between 0 (zero) to 1 (one**)** and hence, it can easily be interpreted as a probability distribution function. Logistic regression function is defined as follows:

$$P\mathrm{i} = \frac{1}{1+e^{-Zi}}$$

$$= \frac{1}{1+\frac{1}{e^{Zi}}}$$

$$Pi = \frac{e^{Zi}}{1+e^{Zi}}, \qquad \text{(i)}$$

where$Zi = \beta o + \beta i\ X\ i.$

Here, Zi is a linear function of the explanatory variable of X (or a linear combination of explanatory variables). In our model, if Pi is the probability of utilising the healthcare services by a person, given that the person is sick, then (1- Pi) is the probability of not utilising the healthcare services.

Thus, $1 - \mathrm{Pi} = 1 - \frac{e^{Zi}}{1+e^{Zi}}$

$$\mathrm{Pi} = \frac{1}{1+e^{Zi}}. \qquad \text{(ii)}$$

Dividing equation (i) by equation (ii), we get

$$\frac{\mathrm{Pi}}{1-\mathrm{Pi}} = e^{Zi}. \qquad \text{(iii)}$$

Here, $\frac{Pi}{1-Pi}$is the odds ratio in favour of utilising the healthcare services in contrast to not to utilise the same. It indicates the ratio of the probability that a person or household member will utilisethe healthcare facility to the probability that a person or household member will not utilise the same, keeping other factors as constant or controlling the other factors.

If we take the logarithm on the both sides of the equation (iii), we get

$$\mathrm{Log}\frac{\mathrm{Pi}}{1-\mathrm{Pi}} = \mathrm{Log}e^{Zi}$$

or, $$\mathrm{Log}\frac{\mathrm{Pi}}{1-\mathrm{Pi}} = \mathrm{Zi}. \qquad \text{(iv)}$$

Now, by replacing $Log\frac{Pi}{1-Pi}$ byUi, we get

$$\mathrm{Ui} = \mathrm{Zi}\,. \qquad \text{(v)}$$

Substituting the value ofZi in equation (v), we get

or, $$\mathrm{Zi} = \beta o + \beta i\ X\ i. \qquad \text{(vi)}$$

Here, U is the log of the odds ratio, which is linear in explanatory or independent variable (Xi) and linear in parameters (i.e. β_0 and β_i).Therefore, it is expected that utilisation of healthcare services is linearly related to burden of disease. Instead of taking only one predicted variable, predictor variable can take a linear combination of explanatory variables as follows:

$$Ui = \beta o + \sum \beta i \, X \, i. \qquad \text{(vii)}$$

The above equation (vii), can be rewritten as

$$Ui = \beta_0 + \beta 1X1 + \beta 2 \, X2 + \beta 3 \, X3 + - - - - \beta k \, Xk. \qquad \text{(viii)}$$

This is a multiple regression equation with k explanatory variables; the parameters β_i for all j=0, 1, 2, 3….k are all estimated. In our study, assuming U is the probability of utilising healthcare services, which is dependent variable and $X_i$is a burden of disease which is a set of independent variables, such as, category of disease, severity of disease, nature of disease. Thus, our regression equation will be as follows:

$$Ui = \beta 0 + \beta 1X1 + \beta 2 \, X2 + \beta 3 \, X3 + \beta 4 \, X4 + \ldots + Є. \qquad \text{(IX)}$$

Where, U_i = Probability of utilising any healthcare services by the individual i during the reference period of twelve months.

$X1$ = Category of disease

$X2$ = Severity of disease

$X3$ = Nature of disease

All other factors (such as X4, X5, X6, X7……..etc.) represent the demographic and socio-economic characteristics of the sick person. The odds ratios indicate the probability of a person utilising any healthcare services compared to reference category during the reference period, provided that the person is sick. As predicted response is dichotomous in nature, we have assigned the value of one, if it was reported that a person had utilised any healthcare services during the reference period of twelve months (i.e. one year) and zero for opposite.

3.4.3.2 Multinomial logistic regression model

As our estimates are for the measuring of probability of choosing a given option out of three or more options available by a person or household, so two different multinomial logistic regression models were constructed with respect to system of medicine (allopathy, homeopathy, ayurveda and others) and sources of healthcare services (public, private, and other charitable organisations).

3.4.4 Modelling the impact of burden of disease on OOPHE by the people of SMCA

Like any other goods or services, healthcare services are also demanded and it takes the form of demand function. But, due to introduction of user fees in the public healthcare sector and low health insurance coverage, household has to depend on private healthcare sector, which leads to high OOPHE. Thus, to a large extent, households' OOPHE can be a good reflection of healthcare demand. In the present study, OOPHE is the total payment made by the individuals or household to get disease free and healthy life from their own financial resources. Considering theoretical model, it can be assumed that OOPHE (demand for healthcare services) of the household is affected by following factors:

1) **Burden of disease:** The study measured the burden of disease in terms of different components such as category of disease, severity of disease, number of days suffering, nature of disease experienced by the sick people of SMCA during the reference period. It is expected that increase in all these components of burden of disease leads to an increase in OOPHE incurred by the household. In fact, the more is the burden of disease, the more is the demand for healthcare services, which in turn, increases OOPHE of the households.

2) **Choice of care:** Household faces with number of choices to get recovery from disease. It comprises of pattern of care (modern care and traditional care), system of medicines (allopathy, homeopathy, ayurveda, yoga, physiotherapy, unani and others), sources of care (public, private, charitable healthcare organisations) and nature of utilisation (OPD and IPD). Present study considered both the outpatient visits and inpatient (hospitalisation) cases in all sources of care (such as public, private and charitable healthcare organisations). It seems reasonable to assume

that treatment with allopathy medicine is costlier than other systems of medicine. On the other hand, hospitalisation is costlier phenomenon than non-hospitalisation cases, so household has to bear more OOPHE for hospitalisation cases than in outpatient visits, irrespective of choice of care. Further, it is expected that OOPHE in private healthcare facilities is much expensive than in public or charitable healthcare organisations.

3) **Socio-economic conditions of household:** Socio-economic conditions of sick person are represented by affordability of household, education level and the place of residence of the sick person. However, affordability of household is measured in terms of monthly household income. It is assumed that the higher is the affordability of the household, the higher is the chance of incurring OOPHE. On the other hand, higher educated people are more conscious about their health and they are less likely to be fallen sick than the other categories, so probability of incurring OOPHE by them is low than the others. Further, place of residence may also an important bearing on OOPHE, as place of residence of the people changes, living standard, availability and accessibility of services including healthcare facilities also changes, this may lead to change in OOPHE.

4) **Demographic composition:** Healthcare expenditure may be affected by the demographic composition of household. It is reasonable to assume that households having more number of children and elderly persons would make higher OOPHE as they are vulnerable to different types of diseases. It is also expected that the more is the household member the more is the probability of falling sick in the family, which, in turn, increases the OOPHE of the household. There may be the chance of gender biasness with regard to OOPHE, if male earning member becomes ill, probability of incurring OOPHE for him is comparatively higher than the other.

3.4.5 Log-linear regression model (log-log form)

The following model is developed using econometric approach to measure the impact of burden of disease and its associated factors on OOPHE incurred by the household (equation form of the Model):

$\mathrm{Ln}(\mathrm{OOPHE}) = \mathrm{f}(\mathrm{BOD}, \mathrm{COC}, \mathrm{SEC}, \mathrm{DEMO}, \mathrm{Ui})$[Functional form of the Model]

$$Ln(OOPHEi) = \alpha i + \beta 1\ Ln\ (CODi) + \beta 2Ln\ (SODi) + \beta 3Ln\ (NOD\ i) + \beta 4Ln\ (NDAYi) + \beta 5\ Ln\ (POUi) + \beta 6\ Ln\ (TOCi) + \beta 7Ln\ (SOMi) + \beta 8Ln(SOC\ i) + \beta 9Ln(AFFORDi) + \beta 10Ln\ (EDUi) + \beta 11Ln\ (POR\ i) + \beta 12Ln(Agei) + \beta 13Ln(GENi) + \beta 14Ln(MSi) + \beta 15Ln\ (FS\ i) + ui$$

Where the sub index i represents the household and Ln denotes the natural logarithm. And

Ln OOPHEi= natural logarithm of OOPHE made by the household i during the reference period.

α_i= constant term for each household which takes into account the individual characteristics

Here, the dependent variable is natural logarithm of OOPHE made by the household i during the reference period, as proxied by direct expenditure incurred by household on healthcare and the independent variables are natural logarithm of Burden of disease (Ln BOD_i), natural logarithm of Choice of care (Ln COC_i), natural logarithm of Socio-economic conditions (Ln SEC_i), natural logarithm of Demographic composition (LnDEMO $_i$). Since the model is specified in log-log form, the co-efficient estimates are elasticities and therefore, it facilitates us to interpret the relationship between OOPHE and different dimensions of burden of disease through the process of utilisation of healthcare services. Thus, all the coefficients will give the elasticity of OOPHE with respect to explanatory variables stated. The model is also appropriate as the sampled cross-sectional units were drawn from a large population. Further, this multivariate linear regression model expressed in terms of log-log functional forms examines the impact of burden of disease on OOPHE incurred by the households of SMCA. The model provides us a better insight into the variations of OOPHE by households with respect to different dimensions of burden of disease such as types of disease, severity of disease, nature of disease, number of days of suffering and also the associated demographic and socio-economic conditions of the people and different factors related to utilisation of healthcare services etc. The contribution of individual independent variables is evaluated through computation of β_s and is tested for significance using t-test at normal level of significance and multi-co linearity is checked by proper method.