CHAPTER III

The present chapter discusses sources of the data used, area of the study, sampling technique and methodology. Besides the primary data, secondary data have been used to present the overall picture of migrant labourers of Bihar who have made their destination in two states of NER— Assam and Sikkim. Besides secondary data, the study mostly relies on primary data collected from the field during the period of April to October 2016. The primary survey was done in order to collect information from the Bihari migrant laborers in Assam and Sikkim.

3.1 DATA SOURCE AND SAMPLE

For the secondary data, especially the migration data for 1991 and 2001, information available in the public domain of the organizations like the Registrar General of India (RGI) and the Census Commissioner of India and the National Sample Survey Office (NSSO) have been considered. In the secondary data, the study keeps special focus on interstate migrants and migrants specifically from Bihar in both the selected states. We also try to see patterns, growth rate, migrants by place of birth and by place of last residence. Additionally, profile of the migrants is also constructed on the basis of origin, stream of migration, duration of migration, reasons for migration, etc.

The primary data was collected from the Bihari migrant laborers working in the states of Assam and Sikkim with the help of a pre- structured questionnaire. For better understanding, the migrant labourers of Bihar taken in this study are categorized into three broad groups as— self-employed, regular salary/wage employees and casual labourers. They work basically in the construction sector,

vegetable or fruit vendor, street hawker/vendor, washer man, plumber, load pickers in go-downs, cobblers, domestic helpers, worker/helper at shops, carpenters, mason, etc.

For proceeding further, using well-organized schedules, 350 respondents (sample) were interviewed and collected information relating to the research questions and objectives given in Chapter I. Of the total sample, 200 were collected from Assam and 150 from Sikkim through purposive random sampling method. Since we were in search of respondents/samples in the vast areas we narrowed down to snowball sampling so that the respondents are easily met. The respondents were interviewed at their work places or at their homes, depending on their free time and availability. Depending on the concentration of the population of Bihari migrant labourers, sample size distribution has been done proportionately (refer Figure 3.1). Of the total 350 respondents, 14 of them were taken for special interaction and in-depth interview. Through this in-depth interview and interaction, a few case studies have been included in this study in order to understand their livelihood condition and sustainability in detail.

Figure 3.1: Sample Size Distribution
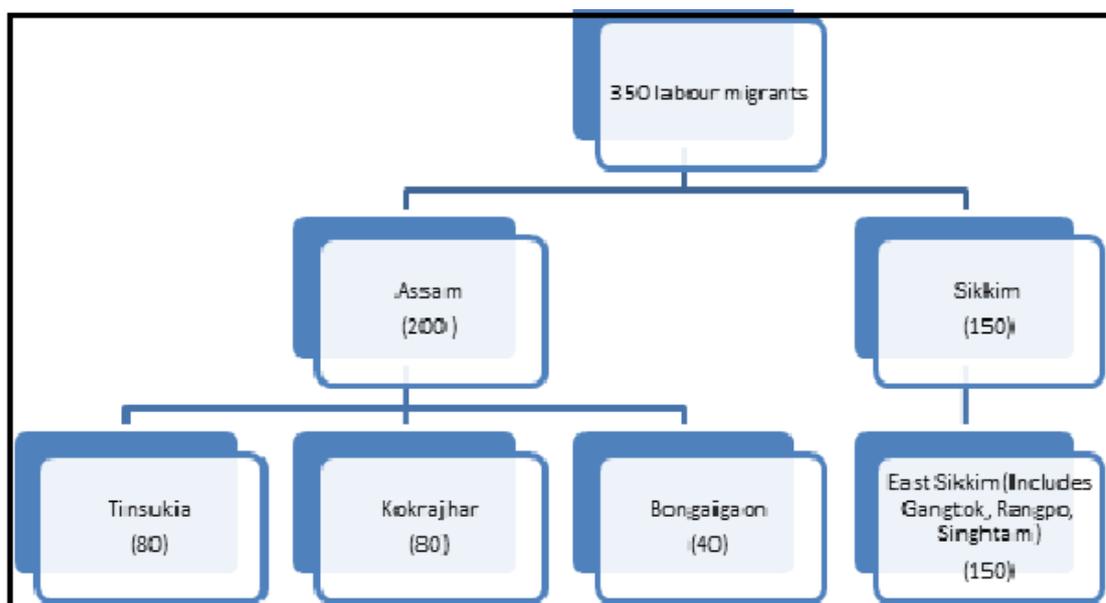
Understandably, the immigrants who have already settled in a particular place for many years become familiar with the local environment and can enjoy most of the benefits (like, ration card, voter ID, passport, etc.) given by the government to the local people. Therefore, for understanding clearly the livelihood and sustainability of the migrants, the Bihari migrant labourers who came to Assam and Sikkim for the period of less than 25 years (time period of migration) are considered for this study. Objectives set in this study to understand the livelihood sustainability of the informal migrant labourers may be quite difficult if we include the migrant labourers living in the study areas for generation/long period.

As we intend to see impact of the law and order condition of an area on migrant labourers, the study has chosen two states— Assam and Sikkim on the basis of conflict-ridden and peaceful areas respectively. Within Assam, based on the new reports, we have collected data from the severely affected districts by conflict namely, Kokrajhar, Bongaigaon and Tinsukia. A large number of attacks and atrocities against the Bihari migrants have been made in these districts since the 1980s (Singha 2018). On the other hand, Sikkim is one of the most peaceful states in India, with zero crime rates, zero presence of insurgent group and zero incidences of conflicts against the migrants. Also, a large number of Bihari migrants are found in this state. The data is collected mainly from the East District, especially in and around Gangtok, Rangpo and Singhtam, where a large number of commercial activities are happened and most of the Bihari migrants are found.

3.2 METHODOLOGY

There are numerous factors responsible for migrants to move from one place to another. As given in the first objective of this study and for understanding the major

factors responsible for Bihari migrants to choose Assam and Sikkim, present study begins with the descriptive statistics. With the help of pilot survey, literature and personal justification, altogether 9 (nine) variables/factors, which might influenced Bihari labourers to migrate the most ,have been included in the schedule to fulfill the first objective. The variables are— landlessness, joblessness or unemployment at home, financial crisis at home, inadequate income in the previous occupation, impressed by the city life, social network, pull factors (like good weather, nearby home), push factors (like natural calamities— flood, draught), Family crisis/conflict, conflicts with the neighbors. The respondents were also asked to respond/rank the factors in priority basis from the factors mentioned/provided. Weightage were given accordingly, reasons which got the first priority by the individuals were multiplied by 3, and likewise the reasons which got the second and third priority were multiplied by 2 and 1. After giving weightage to each of the priorities the total weighted scores were calculated, and on the basis of which, final rank orders were drawn. Also the choices selected by the respondents were taken as multiple entries and these responses have been calculated in percentage form.

As given in the second objective, the study tries to identify the major factors that enhanced livelihood sustainability of Bihari migrant labourers in Assam and Sikkim. Besides descriptive statistic, using primary data, correlation and regression analyses have been done to understand livelihood sustainability. Understandably, migrants try to maximize income and in order to strengthen their livelihood at destination they work relentlessly. Their basic objective is to maximize savings and remittances. Hence, it can be said that savings play a motivating factor for migrants. Having understood the importance of savings, we try to understand on what factors the saving depends upon.

To understand the factors responsible for enhancing saving, we do run a regression exercise. Since there seems to have outliers in the variables included in the regression (both dependent and independent), a log transformation is used in this exercise. The regression analysis deals with the intensity of the relationship between the two sets of variables that does not necessarily imply causation. It is used to estimate the mean value of the dependent variable, given the values of the independent variables and also to test the hypothesis (hypothesis suggested by the economic theory) about the nature of dependence.

The saving function is given as:

$$Savings = f(State, occupation, number\ of\ years\ of\ migration \dots \dots \dots).$$

To proceed further and to understand the difference of savings accrued by migrants between the origin of the migrants and the destinations, student's t test has been employed. Later, by using state dummies in the regression we can understand whether the state (conflict-ridden Assam and peaceful state of Sikkim) has an impact on savings or not. This inferential test determines the statistical significant difference between the means of two groups. The null hypothesis for the independent t-test is that the mean savings of the Bihari migrants in Assam and Sikkim are not equal.

$$H_0: \mu_a \neq \mu_s$$

It is intended to reject the null hypothesis and accept the alternative hypothesis, which means the savings of the Bihari migrants in Assam and Sikkim are equal.

$$H_1: \mu_{a=} \ \mu_s$$

Before running the regression analysis, correlation analysis is done in order to find the significant relationship between dependent and independent variables and

also among independent variables. When there is high correlation between the independent variables (multicollinearity), we ought to drop those variables. Nevertheless, we run Variance Inflation Factors (VIF) to evaluate how much the variance of the estimated regression coefficients are inflated as compared to when the predictor variables are not linearly related. Normally, pair-wise correlation is limiting, it is possible that pair-wise correlations are small yet there is linear dependence among two or three more variables. Therefore, we rely on VIF test to detect multicollinearity problem in the regression analysis. Nevertheless, the regression model is given below:

$$LS_i = \beta_i X_i + U_i$$

LS_i Signifies log of monthly saving by an individual migrant labour, $\beta_i X_i$ represents the vector of parameters which explain the variation in the dependent variable,X_i represents the vector of explanatory variables, and U_i represents the vector of the unexplained variables in the model.

Further, logit regression model is also employed to understand the migrant's willingness to stay in the destinations. The migrants were asked about the willingness to continue their stay at the destination. If the migrants are willing to stay that means he is able to sustain his life at destination. So, logit model is used to see what determines or what factors impact the migrants' decision to continue their stay or make their stay continue. As both logit and probit are the types of generalized linear model and categorical outcome, either of this can be used to determine the condition/situation mentioned above. But the logit model is better as it is less complex in interpretation. Here, we used logit model because dependent variable is willingness/desire/plan to stay in future. The model estimate posits that migrants' willingness/desire/plan to stay on the basis of maximizing utility and their decisions

may be influenced by certain characteristics or factors associated with him. Aijia (2009) has analyzed that migrant's intention to continue their stay in destination city depends on human, personal and social capital. Another study by Li Nan (2010) stated that the rural labour migrant wish or willingness to stay depends on several factors like years of schooling, experience, occupational status, income level and urban sense of belongingness. While, Hou Hung Ya et al (2004) found that the lower age and educational level and higher income at the destination of the rural labour migrants in the city make them willing to give up their land in the village and settle in the city. Zhu Yu (2004) also stated that stable job and good income in the city affects migrants' decision to stay in the city. In this manner, the present study also tries to discover the factors responsible for the migrants' willingness to stay at the destination.

The Model can be expressed as:

$$Y_i^* = \beta x_i^* + \varepsilon_i^*$$

$Y_i^* = 1$ if $Y_i^* \geq 0$ if migrant intends to stay, $Y_i^* = 0$ otherwise, if migrant does not intend to stay

Where, β is a set of repressor, x_i^* is a set of attributes determining a potential default ($Y_i^* = 1$), and ε_i^* represent the error term.

Further, ordered logit model is also employed when the dependent variable is in different time periods of migration. More the time a migrant spends at the destination indicates that he is able to sustain his life at the destination. Hence, the ordered logit model is applied in order to see on what factors do the length of the stay of the migrant depends on. This model can be understood as an extension of the logistic model that applies to dichotomous dependent variables, allowing for more

than two (ordered) response categories. Here, the response categories are ordered or ranked. The dependent variable is m, which is an ordered categorical variable ranging from 1 to 5. Based on the data and nature of variables available, responses of the migrants were recorded as: 1= migrants staying less than a year, 2= migrants staying from 1.1 to 3 years, 3= migrants staying from 3.1 to 6 years, 4= migrants staying from 6.1 to 10 years, 5= migrants staying more than 10 years. The ordered logit model can be given as:

$$Y^* = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots \beta_k X_{ki} + U_i$$

$$= \sum \beta_n X_{ni} + U_i \text{ (n= 1 to k)}$$

$Y_i = 1$, if Y_i^* less that equal to μ_1

$Y_i = 2$, if μ_1 is less than equal to Y_i^* less than equal to μ_2

$Y_i = 3$, if μ_2 is less than equal to Y_i^* less than equal to μ_3

$Y_i = 4$, if μ_3 is less than equal to Y_i^* less than equal to μ_4

$Y_i = 5$, if μ_4 is less than equal to Y_i^*

Where, $\mu1<\mu2<\mu3<\mu4$

This study also tries to understand income difference between at the origin and destination of the migrants. Of course, migrants generally move to new places for better opportunities. Which factor(s) makes this income difference is discussed in this study. With the help of primary data, migrants were investigated about their income back at home and at the destination. Of course, some of the illiterate migrants who have moved to the present destination may not be remembering the wage rate at the origin before he had moved to the present destination. For this, average daily wage

rates of male agricultural field labourers of Bihar from 1990 to 2015 have been extracted from the publications of the Department of Agriculture, Government of Bihar. Also, the past income cannot be compared with the present income directly. For this, first we try to find out the present value of income with the stated past income as the base value and appropriate rate of depreciation. Suppose a migrant was working on an agricultural field and was earning 'X' amount of money annually and we suppose that this was his income 7 years back. It is understood that his income would not be the same in the present time. Hence, in order to find the worth of Rs X we intend to calculate the present value of the past income. To calculate the present value of the past income, we follow:

1. Past Income (Rs X)= The amount of money earned by a migrant at the time of departure

2. Time = Number of years of migration (7 years)

3. Rate of Interest = Bank rate of interest on savings which is decided by the RBI (Reserve Bank of India) annually.

4. An average of 7 years savings deposit rate is taken in present case.

Table 3.1: Method for Discounting Income

Years	Savings rate of interest (r)	Discounted income
2011	r_1	$X_1 = \dfrac{X}{(1+\bar{r})^1}$
2012	r_2	$X_2 = \dfrac{X}{(1+\bar{r})^2}$
2013	r_3	$X_3 = \dfrac{X}{(1+\bar{r})^3}$
2014	r_4	$X_4 = \dfrac{X}{(1+\bar{r})^4}$
2015	r_5	$X_5 = \dfrac{X}{(1+\bar{r})^5}$
2016	r_6	$X_6 = \dfrac{X}{(1+\bar{r})^6}$
2017	r_7	$X_7 = \dfrac{X}{(1+\bar{r})^7}$
2018 (present)	$\bar{r} = \dfrac{r_1+r_2+r_3+r_4+r_5+r_6+r_7}{7}$	$\bar{X} = \dfrac{X_1+X_2+X_3+X_4+X_5+X_6+X_7}{7}$

Hence, $\bar{X}$ is defined as an estimated present income. To calculate, we have taken the monthly income and present income for each migrant is calculated individually. Further, the difference in income is calculated by subtracting the past income from the present income at the destination (refer Table 3.1). This income difference is dependent on several factors and the equation may be written as:

$$InDiff = f(State, Age, Years\ of\ migration, Occupation \ldots)$$

$$InDiff_i = \beta_i X_i + U_i$$

β_i represents the vector of parameters which explain the variation in the dependent variable. X_i represents the vector of explanatory variables and U_i represents the vector of the unexplained variables in the model.

As given in the third objective, in order to assess quality of life of the migrant labourers in the two states, we construct a composite index using different parameters.

Further, a t-test is done to see whether there is significant difference in expenditures on different items in both the states (Assam and Sikkim). Migrants spend time on different activities apart from the work. We compare and contrast times spend on activities in both the states. Spending time with family and recreational activities indirectly indicates psychological well-being of a person. In order to construct a composite index, the procedure of Human Development index (HDI) has been followed religiously. The only difference is arithmetic mean is used instead for geometric mean since the sample data do not have major variations like what is found among countries while calculating HDI. Measurement and descriptive statistics used in index construction under every dimension is given as:

$$Individual\ Index = \frac{Actual\ value - Minimum\ value}{Maximum\ value - Minimum\ value}$$

Individual index values have been normalised by using minimum and maximum values of each indicator and transformed to index values between 0 and 1. Finally, a composite index is calculated by taking the arithmetic mean of all for dimensions.

$$Composite\ Index = \frac{MEI + MPI + MSI + MPHI}{4}$$

MEI= Mean of economic indicators

MPI= Mean of personal dimension

MSI= Mean of social dimension

MPHI= Mean of physical dimension

After obtaining composite index, we compare and contrast the composite index between different occupational categories, categories based on years of migration and among the states. To understand the migrants' satisfaction level, a

psychological scale, known as 'Satisfaction with Life Scale' (SWLS) of Diener et al (1985) is used in the study. This can help us to gauge the subjective concept of well-being of the migrants at the destination. As per SWLS model, a 5 item scale is designed to measure judgment of one's life satisfaction (not a measure of either positive or negative affect).Participants/migrants indicate how much they agree or disagree with each of the 5 items using a 7-point scale that ranges from 7 strongly agree to 1 strongly disagree. Of course, the scale does not assess satisfaction with life domains such as health or finances. Lastly, we try to see how the SWLS is affected, by using a regression model. How do the variables of occupation, income, time of migration and dimensional indices discussed above impact the SWLS is also done.

$$SWLS = f(Occupation, Number\ of\ years\ of\ migration, \ldots \ldots \ldots etc)$$

Another scale, the General Happiness Scale (GHS) is also known as subjective happiness scale is used in the study. In order to measure the subjective happiness, four items are given and each of them has 7 options. For each question the options are different. The scale (items with options is attached in the Appendix section of the thesis at the end). All the scores are summed up and are continuous in nature.