Chapter III

Data and Methodology

3.1. Introduction

This chapter deals with the data and research methodology used for the present study. The rigour of research outcomes and conclusions mainly depends on the quality of the research design and the overall evaluation methodology. The research methodology should be robust. Hence, the development of research methodology becomes both essential and relevant for conceptualising framework and implementing any programme for the rural poor. The main objective of this study is to find out the actual effect of microfinance programme on income, employment, poverty, inequality and financial inclusion of the rural poor. This demands a robust evaluating methodology in order to extract the actual impact of the programme. In order to ensure methodological rigour, an impact evaluation[8] must estimate the counterfactual, that is, what would have happened had the programme never taken place or had they not participated. Since assessing the impact of a microfinance programme is subjective to the design methods and data, therefore, this chapter is dedicated to design and develop research methodology for the purpose of this study. This would provide guides for formulating policies and interventions that will help financial inclusion of the poor and ultimately poverty alleviation prompting inclusive growth. First of all, we deal with the conceptual framework of impact evaluation problem and then based on its state research methodology.

3.2. The Conceptual Framework of Impact Evaluation

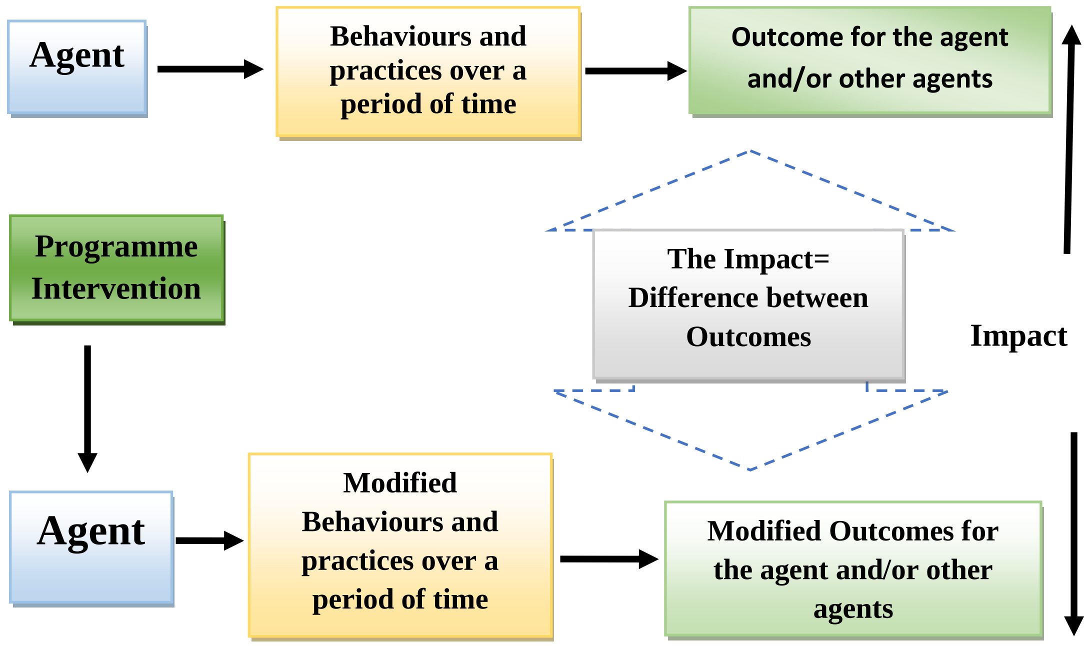
3.2.1. Formulation of the Evaluation Problem

The main objective of impact evaluation of a programme (intervention) is to assess the extent to which the programme has changed outcomes of agents, where agents may be individual, enterprise, household, community, and policymakers etc. In the microfinance studies, Hulme, (2000), provided a constructive model of impact chain for "impact assessment". The 'impact assessment' model is demonstrated in

[8] The terms 'impact evaluation', 'impact assessment' and 'programme evaluation' refer to the same thing.

Figure 3.1. The idea of impact assessment is to define the difference between the outcomes of "agents", which have experienced a policy intervention, against the outcomes that would have occurred without any intervention. Based on this model, the process of impact evaluation includes three steps: defining "agents" (assessment units), defining "outcomes" (assessment indicators) and assessing methods (Figure 3.1).

Assessing the impact of any intervention involves speculation of inference about the outcomes that would have been observed for participants had they not participated in the programme. The framework that guides in evaluation analysis to formalize this problem is the potential outcome approach, also known as Roy–Rubin model (Roy, 1951; Rubin, 1974). The approach is originally applied to evaluate the impact of programmes in labour markets but now-a-days enormously applied to other fields such as health care, education and rural credit markets. The model is built on three pillars of individuals, treatment and potential outcomes. The model suggests that there are two potential outcome exists for each individual in their status with treatment and without treatment.



Source: Adopted from Hulme, (2000)

Figure-3.1 Model of Impact Assessment

The model can be explained as follows:

Let us suppose that there is a microfinance programme assigned to target poor group of population. Then, in the case of binary treatment variable, the treatment indicator D equals 1 if she/he participates in the programme, and D equals 0 otherwise. Furthermore, let us suppose that Y denotes the observed value of the outcome of interest[9]. This variable can receive two potential values corresponding to the treatment variable, that is, $Y=Y^1$ if $D=1$, and $Y=Y^0$ if $D=0$. Then, the treatment effect[10] Δ for each individual is defined as:

$$\Delta = Y^{i1} - Y^{i0} \tag{3.1}$$

Where, Y^1 is the outcome of treated individual and Y^0 is the outcome of untreated[11]. Y^1 and Y^0 are mutually exclusive and so, only one of the potential outcomes is observable for each individual. The unobserved outcome is called counterfactual outcome and so, Δ is not observed for anyone. Hence, one has to concentrate on (population) treatment effects (Caliendo and Kopeinig, 2008).

To evaluate the individual effect that is due to observed outcome for each individual is given by:

$$Y = DY^1 + (1-D)Y^0 \tag{3.2}$$

Where, Y is the observed outcome and equal to Y^1 for participants and Y^0 for non-participants. The estimation of impact of a programme for an individual household is almost impossible because the counterfactual outcome is not exactly known (Heckman, et al, 1997). In fact, the impact of the programme can be estimated for a group of people by finding an adequate control group[12].

Two parameters are most commonly estimated in impact evaluation literature. These two parameters are the **Average Treatment Effect (ATE)** and the **Average Treatment Effect on the Treated (ATT)**[13].

---

[9] Y can be a vector of outcomes, but for simplicity a single outcome of interest is considered.

[10] The term ‘treatment effect’ refers to the effect that a person benefits from being treated or participated in the programme.

[11] For simplicity, the subscript ‘i’ is dropped from the formula.

[12] The terms ‘control group’, ‘comparison group’, ‘non-participants’ and ‘non-treated’ refer to a group of individuals who did not participate in the programme.

[13] There are other parameters such as local average treatment effect, marginal treatment effect, or even the effect of non-treatment on the untreated (ATU) which measures what impact the programme would have on the non-participants if they had participated in the programme, etc. that may be used in impact evaluation (Heckman, et al., 1997).

ATE is defined as the expected effect of treatment on a randomly drawn individual from the population. It measures the mean impact of the programme which is obtained by averaging the impact across all the individuals in the population. ATE is simply the difference of the expected outcomes after participation and non-participation as follows:

$$ATE = E(\Delta) = E(Y^1) - E(Y^0) \qquad (3.3)$$

Where, $E(\cdot)$ represents the average or expected value.

ATE is important when the treatment has universal applicability of random assignment. This parameter also includes effects on persons on whom the programme was not actually intended. Therefore, this parameter might not be of relevance to policy makers as it includes the effect on persons for whom the programme was never intended (Heckman, 1997). Actually, most microcredit programmes are targeted at a specific group of households who had been selected to receive microcredit.

The most prominent evaluation parameter that provide a direct measure of the desired impact of the programme on the target group is the so-called ATT (Average Treatment effect on the treated). ATT is defined as the average effect on treated who actually participate in the programme. This parameter measures the effect of the microcredit programme on those individuals who participated.

It is expressed as:

$$ATT = E(\Delta | D=1) = E(Y^1 - Y^0 | D=1)$$

or,
$$ATT = E(Y^1 | D=1) - E(Y^0 | D=1) \qquad (3.4)$$

The expected value of treatment effect 'Δ' is the difference between expected outcomes with and without treatment for those who actually participated in treatment. The second term '$E(Y^0|D=1)$' is the average outcome of treated individuals in the absence of the treatment, that is, had they not received the treatment. This term is never observed and is called as counterfactual outcome of the participants.

In the case of non-participant (untreated or control) group, the Average Treatment Effect on the Untreated (ATU) which measures what effect the microfinance programme would have had on the non-participants if they had participated in the programme, can be estimated as:

$$ATU= E\ (Y^1 - Y^0|D=0) \tag{3.5}$$

Where, ATU is the measures the effect of microfinance on the non-treated, i.e., on those who had not participated into the programme.

The problem is that all of these parameters depend on counterfactual outcomes which are not observable. Accordingly, the computation of ATT and ATE or ATU is not straightforward as equations as equations (3.3), (3.4) and (3.5) contain unobserved counterfactual level of the two groups. This is called a problem of missing data which is at the heart of impact evaluation problem.

The crux of the evaluation problem is the difficulty in determining the counterfactual $E(Y^0|\ D=1)$ which is never observed once they received the treatment (Baker, 2000; Hulme, 2000). But, we do observe a corresponding outcome $E(Y^0|D=0)$ of control individuals.

Thus, we can calculate:

$$\delta= E(Y^0|D=1)-E(Y^0|D=0) \tag{3.6}$$

Now, if $\delta=0$, the condition in equation (3.6) becomes,

$$E(Y^0|D=1)=E(Y^0|D=0) \tag{3.7}$$

If the condition expressed in equation (3.7) holds, the observed outcome $E(Y^0|D=0)$ can be substituted for the counterfactual outcome of treated individuals. That is, we can use mean outcome of non-participants as a proxy for the counterfactual outcome of participants or the treated.

Thus,

$$ATT= E\ (Y^1 \mid D =1) - E\ (Y^0 \mid D =0) \tag{3.8}$$

In randomised experiment, all the characteristics of the individuals are proportionally distributed between the treated and non-treated groups and thus the observable term $[E(Y^0|D=0)]$ can be substituted as a perfect counterfactual outcome for the unobservable term, $[E(Y^0|D=1)]$ to compute the ATT in (3.8). However, the identifying assumption in equation (3.7) is likely to hold in randomised experiment but not in non-randomised experiment. This is because the components which determines the participation decision also determines the outcome variable of interest simultaneously.

Consequently, estimating ATT by OLS method using the different in sub-population means of participants [E $(Y^1 \mid D = 1)$] and non-participants [E $(Y^0 \mid D = 0)$] may lead to a biased estimate as a result of selection bias (Rosenbaum and Rubin, 1983; Caliendo, 2006). The selection bias occurs because participants and non-participants are selected with different outcomes, even in the absence of the programme. The selection bias might arise from observable characteristics such as age or education differences. Furthermore, selection bias in a programme evaluation emerges when unobservable factors such as entrepreneurial ability or motivational factors might play a decisive role in determining participation decision. In brief, selection bias is the outcome of missing data problem.

In equation (3.8), adding and subtracting the term $E(Y^0 \mid D = 1)$, we have:

$$ATT = E(Y^1 \mid D = 1) - E(Y^0 \mid D = 1) + E(Y^0 \mid D = 1) - E(Y^0 \mid D = 0)$$

or, $$ATT = E(Y^1 \mid D = 1) - E(Y^0 \mid D = 1) + \delta$$

or, $$ATT = \delta + SB$$

or, $$SB = \delta - ATT \qquad (3.9)$$

The term, SB in equation (3.9) is the self-selection bias. The selection bias is the difference between the counterfactual for treated units and observed outcomes for untreated units which is to be eliminated or reduced to the minimum to find the true impact of the programme.

In randomised experiment, the control group is generated through random assignment and hence serves as a perfect counterfactual free from selection bias. The programme impact on the outcome indicators being evaluated can be measured by simply the difference between the means of the samples of the treatment group and the control group. But it is not possible to construct treatment and comparison groups through experimental design in non-random method. The selection bias in a non-experimental context is often sizable. Therefore, selection bias has to be eliminated in order to identify the true parameter ATT.

There are a number of methods available to deal with the selection bias which are discussed in the next point 3.2.2.

3.2.2. Methods of Impact Evaluation

In order to overcome this problem of selection bias, different methods have been developed so as to arrive at the true estimate of impact of a programme. A review of various methods to evaluate the impacts have been provided by Hulme, (2000), which include sample survey design, rapid appraisal, participant-observation, case studies, participatory learning and action, including a description of the key features of each method. Each of these methods have its own strengths and weaknesses and hence suggested to adopt pluralistic approaches instead of a single method to avoid the weaknesses of individual methods (Hulme, 2000). The most widely suggested and used econometric methodologies in literature are randomisation, matching, difference in difference (double difference) and instrumental variable (IV) methods (Imbens, 2004; Ravallion, 2005). All these methods try to address the problem of missing data in the sense that the participants' outcome of what had they not been participated is unobservable as long as there is counterfactual. These are discussed one by one in brief.

a) Randomisation

The idea of this method is to overcome the missing data problem by assigning the programme randomly to the agents. If treatment is random, non-participants should form the control group in the absence of the programme. In this case, there should be no difference on average between the two groups besides the fact that the treatment group had access to the programme. Because, in this setting, every member of the eligible population has equal chance of being participated or not participated. Thus, the difference of the outcome indicators between participant and control group can be attributed to the impact of the programme.

However, in observational studies, microcredit programmes are hardly randomised. So, the impact assessment may result in biased estimates as a result of the existence of confounding variable (Becker and Ichino, 2002). In microcredit programmes, randomisation means exclusion of some eligible households from the microfinance programme. Nevertheless, this approach is unable to ensure the control group to be completely unaffected by the presence of the microfinance programme. Further, such experiments are often expensive to implement.

b) Matching Method

The basic idea of the matching method is to find a control group that possesses similar traits in terms of observable characteristics as the treated group. The idea is to estimate the counterfactual outcome from the best matched eligible control group in terms of observable factors. Heckman, et al., (1997), showed that by selecting sufficient observable factors that any two individuals with the same values of these factors will display no systematic differences in their reactions to the programme intervention. The implied condition for the matching method to work pivots on finding the enough common support between the groups, which allows a consistent comparison. The region of common support can be constructed by using various matching techniques such as propensity score matching, nearest neighbour matching, Caliber and Radius matching, and Kernel matching. The most widely used type of matching is propensity score matching in which the comparison group is matched to the treatment group on the basis of propensity score P(X). The P(X) is defined as the predicted probability of participation given a set of observed characteristics of the individuals:

$$(Y^1, Y^0) \perp D | P(X)$$

Where, $\perp$ denotes independence.

This implies that observations with similar P(X) will have the same distribution of observable and unobservable irrespective of assignment to treatment. Thus, the treatment is virtually randomised and as a result treated and control group becomes on an average observationally identical (Becker and Ichino, 2002). The high dimensional matching problem of numerous observables is reduced to a 'one dimensional problem, given that P(X) is known [Heckman, et al., (1997); Cobb-Clark and Crossley (2003)].

c) Instrumental variables (IV)

The main idea of the Instrumental variable **(**IV) method to find a variable that determines participation of programme but does not the outcome. The IV affects the observed outcome only indirectly through the participation decision, hence causal effects can be identified through a variation in this instrumental variable (Imbens and Angrist, 1994; Rubin, et. al, 1996). Blundell and Dias, (2000), stated that a valid instrument has to satisfy the following three conditions:

i) the instrumental variable determines programme participation;
ii) the instrumental variable is uncorrelated with the programme outcomes; and
iii) the instrumental variable is not completely determined by the observable factors in the model.

However, in most cases it is hard to get instrumental variables capable of capturing the problem of endogeneity.

d) **Double Difference (DID)**

This approach can be used to recover the average effect of the microcredit programme on those individuals who entered the programme. The idea is to have a baseline survey before implementation of the programme and a follow up survey after implementation of the programme. The method is based on the critical assumptions that there is a common time effect across groups and no compositional changes within each group. Based on this assumption, 'Difference In Difference' (DID) allows one to estimate the coefficient of the average effect of the treatment on the treated, that is, ATT (Blundell and Dias, 2000). The main advantage of this method is that it allows for the selection bias of the programme based on some unobservable factors. However, the method has two disadvantages. The first is the requirement of the panel data that has to be collected before and after the programme implemented. The second is the time invariant assumption of the unobservable variables that are unchanged over time that affect the programme selection. This assumption might be violated under non-experimental data in which the households in both groups are systematically different and unbalanced in the pre-programme attributes that are possibly related to the outcome (Athey and Imbens, 2006).

The above discussed methods can be used in tandem or separately to measure the impact depending on the nature of the study and availability of data. In our study, both randomisation and DiD methods are out of question because of the fact that clients self-select and that the data availability is rather a cross-sectional data. Moreover, it is not possible to find a valid instrument variable to be used in minimising bias. Consequently, the econometric technique that will be applied to solve the evaluation problem is 'matching' which is more precisely referred to as ‘propensity score matching’ (PSM) method. We are next going to discuss the methodology associated with ‘propensity score matching’ method.

3.2.3. Propensity Score Matching (PSM) Method

The most widely used non-experimental method in drawing causal inferences in programme evaluation is the Propensity score matching (PSM) method. The treatment and comparison groups are usually selected after the intervention of the programme by using PSM method. PSM constructs a statistical comparison group based on a model of probability of participation in the microfinance programme using observed covariates. It assumes selection can be explained purely in terms of observable characteristics. It takes into account that any selection on unobservable is trivial and do not affect outcomes in the absence of treatment. The idea of PSM is to match participants and non-participants on their observable characteristics. The method contrasts the outcomes of participants with the outcomes of “comparable” non-participants wherein differences in the outcomes between the two groups are attributed to the programme (Heckman, et. al, 1998).

The advantage of matching over other regression methods is that it is less demanding with respect to the modelling assumptions. Regression models depend on the functional form of relationship such as linear, log-linear, etc. which may be inaccurate and which propensity score matching avoids. Specifically, matching does not require functional form assumptions for the outcome equation of linearity. It is a non-parametric model. Furthermore, there is no need for the assumption of constant additive treatment effects across individuals with matching method. Instead, the individual causal effects are unrestricted and individual effect heterogeneity in the population is also permitted. According to Dehejia and Wahba, (2002), it is invaluable for cross-sectional survey data because, given the nature of survey data, resurveying thousands of units at a later period might be problematic and costly.

Since conditioning on all relevant covariates is limited in the case of a high dimensional vector X, Rosenbaum and Rubin, (1983), suggested the use of so-called balancing scores, that is, functions of the relevant observed covariates X such that the conditional distribution of X given the balancing scores is independent of assignment into treatment. The balancing score is nothing but the propensity score, P(X) and matching procedures based on this score are known as propensity score matching (PSM). The propensity score P(X) provides a conditional probability of participation given a vector of pre-treatment observed characteristics X and its value ranges between zero and one. Symbolically,

$$0<P(X)<1$$

The value of P(X) can be computed as follows:

$$P(X)= Pr(D=1|X)=E(D/X)$$

Where, as defined above, D and X refer to participation dummy and a vector of observed control variables respectively.

The PSM method matches a participant from the treatment group with a non-participant from the control group with similar observable characteristics in order to infer the impact of a programme (Caliendo and Kopeinig, 2008). This method compares the outcomes of participants with those of matched non-participants of the programme, where matches are performed on the basis of propensity score P(X) estimated by using observed characteristics. However, there may lie some observations of participants and control groups that cannot be matched due to significant differences in their observable factors. They are called outliers. These outlier participants cannot be matched using their income and inclusion of these unmatched participants in evaluating the impact may produce misleading results. An important feature of the matching method is that, after the treated and control participants are matched, the unmatched participants in the matching process are discarded and not used in estimating the impact of the programme. The matching is performed within the overlapping or common support region only. So, the matching procedure performed by PSM method can significantly reduce bias in evaluating impact of programme (Heckman, et. al., 1996; Rosenbaum and Rubin, 1983; Setboonsarng and Parpiev, 2008).

Thus, the matching approach can be applied to estimate the impact of a microfinance programme. The PSM method first estimates the propensity score for each participant and non-participant in a microfinance programme on the basis of observed characteristics and then compares the mean outcome of the participants with that of the matched (similar in terms of propensity scores) non-participants. The main purpose of PSM is to select comparable non-participant households among all non-participants to generate a control group, and then compare the outcome of the treatment and matched control groups. The PSM depends on the crucial assumption that, among non-participants, those with the similar characteristics to participant borrowers should have the same outcomes as what the borrowers would have had without participation in the programme. This assumption is called "conditional independence assumption" (CIA) or "unconfoundedness" (Rosenbaum and Rubin, 1983). The implication of this assumption is that the treatment and control units with the similar propensity score have the same probability of assignment to the treatment as in randomised experiments (Dehejia and Wahba, 2002). The assumption is explained below:

<u>Assumption 1</u>: Unconfoundedness or Conditional Independence Assumption

The first assumption in propensity score matching is referred to as the conditional independent assumption (CIA), and is expressed as:

$$(Y^0, Y^1) \perp D | X \tag{3.10}$$

Where, $\perp$ denotes the statistical independence of (Y^0, Y^1) and D conditional on X, given a set of observable covariates, potential outcomes are independent of treatment assignment. Based on Assumption 1, the outcome distributions of participants and non-participant groups are defined as follows:

$$E(Y^0 \mid X, D = 1) = E(Y^0 \mid X, D = 0) \tag{3.11a}$$

and, $$E(Y^1 \mid X, D = 1) = E(Y^1 \mid X, D = 0) \tag{3.11b}$$

Equations (3.11a) and (3.11b) imply that the participant outcomes have the same distribution that non-participants would have experienced had they not participated in the programme. Like randomisation, matching balances the distributions of all covariates X in the treatment and comparison groups and makes virtually comparable to a randomised approach. Given the conditional independence assumption, the matching process is analogous to constructing an experimental dataset in that, conditional on observed characteristics, the selection process turned out to be random. Heckman et al., (1997), show that the missing counterfactual means can be constructed from the outcomes of non-participants and participants as follows:

$$E(Y^0 \mid X, D = 1) = E(Y^0 \mid X, D = 0) = E(Y^0 \mid X) \tag{3.12a}$$

and, $$E(Y^1 \mid X, D = 1) = E(Y^1 \mid X, D = 0) = E(Y^1 \mid X) \qquad (3.12b)$$

Thus, the counterfactual outcomes is deduced from the outcomes obtained from participants and non-participants of programme. In order that both sides of equations (3.12a) and (3.12b) are simultaneously defined for all covariates X, the Assumption 2 of common support or overlap condition should be satisfied.

Assumption 2: Common Support or Overlap

The second assumption is referred to as the overlap or Common Support assumption, written as:

$$0 < Pr(D=1 \mid X) < 1, \text{ for all values of } X \qquad (3.13)$$

This assumption implies that the support of X is equal in treatment and control groups, that is, S=Support (X| D=1)∩Support(X | D=0). It suggests that for each participant there is another non-participant with a similar X. The common support region ensures that any combination of characteristics observed in the participation group can also be observed among non-participants. The basic idea of this condition is to discard all participants, whose propensity score is smaller than the minimum and higher than the maximum in the non-participants. Therefore, participants who fall outside the common support region are to be discarded in the estimation of treatment effect. The literature encourage matching to be performed over the common support region only when there are regions where the support of *X* does not overlap for the treated and non-treated individuals (Lechner, 2000; Caliendo, 2006). Blundell, et al., (2005), suggest that interpreting the estimated effects has to be redefined as the mean treatment effect of those individuals falling within the common support only. The proposition implies that observations with the same propensity score have the same distribution of all vector of covariates, X.

Assumptions 1 and 2 together are called "strong ignorability" (Rosenbaum and Rubin, 1983) and therefore, ATE in (3.3) and ATT in (3.4) can be defined for all values of X. However, 'strong ignorability' or unconfoundedness conditions are overly strong and demands mean independence for estimating (3.3) and (3.4).

The final step in PSM is to assess the matching quality (Assumption 3). This is because the conditioning is not done on all covariates but on the propensity score, and so one has to check the ability of the matching procedure to balance the relevant covariates. The performance of the match can be judged by using the Pseudo-R^2 and the mean t-test. The basic idea of the Pseudo-R^2 is to re-estimate the propensity score on the matched participants and non-participants, and then compare the Pseudo-R^2's before and after matching. The Pseudo-R^2 indicates how well the regressors X explain the probability of participation in the programme. After matching there should be no systematic differences in the distribution of covariates and this is indicated by low value of Pseudo-R^2. Furthermore, the t-test can be used to check whether there are significant differences in covariate means of treated and comparison units. Before matching differences are expected to be significant, but after matching the covariates should be balanced in both groups and hence no significant differences should exist. Under the matching method, Heckman et al., (1997), have provided mean independence as an alternative assumption for estimating ATT:

<u>Assumption 3</u>: Mean Independence

$$E(Y^0 \mid X, D=1) = E(Y^0 \mid X, D=0) \tag{3.14a}$$

and,

$$E(Y^1 \mid X, D=1) = E(Y^1 \mid X, D=0) \tag{3.14b}$$

The mean impact of treatment on the treated (ATT) based on the above assumption can be written as:

$$^{\mathbf{ATT}}\Delta_{\mathrm{PSM}} = E_{P(X)}\,(Y^1 - Y^0 \mid X, D=1)$$

$$^{\mathbf{ATT}}\Delta_{\mathrm{PSM}} = E_{P(X)}\,(Y^1 \mid X, D=1) - E_{P(X)}\,(Y^0 \mid X, D=1 \mid D=1)$$

$$^{\mathbf{ATT}}\Delta_{\mathrm{PSM}} = E_{P(X)}\,(Y^1 \mid X, D=1) - E_{P(X)}\,(Y^0 \mid X, D=0 \mid D=1) \tag{3.15}$$

Where, the first term in equation (3.15) calculates the mean outcomes of participant group and the second term provides the calculation of the mean outcomes of the matched comparison group. The differences obtained by comparing the mean outcomes of the matches are the estimates of the effect of programme for these particular observations (Ravallion, 2001). The outer expectation is taken over the distribution of vector of covariates in the treated population (Caliendo, 2006).

According to Heckman, et al., (1997), and Dehejia and Wahba, (2002), the PSM method produces estimates with low bias if the datasets satisfy the following conditions:

(i) that the survey data for the treatment and control groups are collected using the similar questionnaire so that outcome measures are same;

(ii) that both treatment and control groups are drawn from the same area and locality; and

(iii) that the survey dataset contains a rich set of variables relevant to modelling participation and the outcomes.

The similarity of the treatment and control groups, in terms of observable characteristics, increases the likelihood of getting matches and hence reduces the bias. In addition, the PSM method allows controlling for potential bias such as non-placement and self-selection on observed characteristics in participation of programme. To control selection bias based on observable factors, a number of individual and household covariates need to be included. The variables that influence only participation but not outcome, the variables that influence the outcome but not treatment and the variables that influence neither treatment nor the outcome are not important to control for differences between the treatment and control group. So, in this study, the variables which influence both the treatment and outcome are used for matching and will be included in the Probit model to find propensity score. The next section 3.2.4 discusses the implementing strategies of PSM estimators.

3.2.4. Implementing strategies of Propensity Score Matching estimators

Theoretically, the households representing one matched pair are identical to each other except the access of loans from the microfinance programme. Therefore, matching is able to isolate the impact idiosyncratic factors that have on the outcome variables by reducing observed heterogeneity between the participants and non-participants.

The procedure of implementing PSM estimation consists of two steps. In the first step, either logit or probit model is used to estimate the propensity score P(X) or the probability of participation conditional on control variables X, and then stratifies individuals into blocks according to their similar scores. It can be noted that, there is no strong advantage to using logit vs. probit model as both are preferred to be a linear probability model. The estimation of the binary model using Probit model follows the Cumulative Distribution Function (CDF) of the standard normal distribution and the logit model uses the cumulative logistic function. These models provide predictions on the likelihood that individuals or households participate in the microfinance programme conditional on X and yield similar results. This procedure may include stepwise selection models with repeating steps until the treatment and control groups are achieved. In the second step, the estimated propensity scores have to be used together with various average treatment effect estimators mentioned below to obtain estimates of the Average Treatment Effect on the Treated (ATT).

However, the implementation of propensity score P(X) is hardly equal for two observations. So, matching has to be performed on the basis of closeness of propensity score rather than equality. But the problem arises is that of the basis on which to determine the closeness of the propensity score while matching; the closer the propensity score, the better the match. There are four matching algorithms available that are most widely used. They are namely, Stratification Matching, Nearest-Neighbour Matching, Radius Matching and Kernel Matching. We now discuss these matching procedures one by one below:

i. **Stratification Matching**

This Matching performs matching by dividing the range of variation of the propensity score into intervals to ensure that within each interval test the average propensity scores of treated and control households do not differ (Becker and Ichino, 2002). The weighted average of these interval impact estimate yields the Average Treatment on the Treated (ATT). However, this procedure discards observations when either treated or control units are absent and therefore, is not recommended for data in which the treated and control groups are unbalanced.

ii. Nearest-Neighbour Matching

Nearest-neighbour matching ensures that each treated unit is matched to the comparison unit with the closest propensity score. Hence, for each treated observation a nearest neighbour is sought from the control unit with or without replacement based on the value of propensity score. The method works well once the distribution of the propensity score of both the groups (treated and untreated) are similar (Backer and Ichino, 2002). With replacement, it is possible that the same control unit can be a nearest neighbour for more than one treated unit. After matching each treatment unit with the control unit, the difference in their outcome is calculated and obtaining the average of these differences for the entire sample gives the ATT. The advantage of this matching is that while treated observations that are not matched are discarded in stratification matching, nearest neighbor matching takes into account each treated unit by matching it with control unit possessing the closest propensity score irrespective of the extent of closeness. The main problem with this matching is that the difference in propensity scores for a participant and its closest nonparticipant neighbour may still be very high resulting bad matches.

iii. **Radius Matching**

Another matching procedure is Radius Matching. With Radius Matching, the average treatment effect is computed by averaging over the unit-level treatment effects of the treated where the control unit(s) within a pre-defined radius of propensity score(s) is/are matched to a treated unit. If there is more than one control unit within a radius, then the average outcome of those control units is used. This approach can avoid poor matches and can overcome the drawback of stratification matching, so, the quality of matching rises (Caliendo and Kopeinig, 2008). Given the dataset, the smaller the radius, the better the quality matching becomes since matched control units and the treated units have close scores. However, Radius Matching uses those treated units that have control matches within a radius, so if the radius is very small, many treated units are not matched and hence dropped. Therefore, the ATT by the radius matching estimator is no longer representative of the population of the treated units (Becker and Ichino, 2002; Caliendo and Kopeinig, 2008). Instead, other variants in matching estimators are applied to estimate impact of microfinance programme like Kernel matching.

iv. **Kernel Matching**

This matching method is used to match all treated with a weighted average of all controls with weights that are inversely proportional to the distance between the propensity score of treated and controls (Arun, et al., 2006). Becker and Ichino, (2002), showed that Kernel Matching provides a solution to the problem of discarding observations in Radius Matching because the Kernel Matching estimator possesses a smaller variance since information from all or nearly all control units is used. However, one drawback of this approach is bad matching because few or many far-distance control units may be used to match with one treated unit (Caliendo and Kopeinig, 2008).

All these methods have their own limitation and strength. Therefore, one can employ either one or the other but their consideration in tandem give robust results. In summary, the use of propensity score matching to estimate the impact of the programme is expected to minimise biased results than those from OLS because matching (propensity score) compares participants only with similar non-participants based on observables. Nevertheless, the "similarity" of participants to non-participants is built on observed characteristics, so bias is likely to exist in the estimate when there are important unobservable. For example, if the selection bias based on unobservable counteracts that based on observables and affect both treatment and outcomes of interest. The assumption is easily violated if we are unable to control for variables of unobservable that affect both the treatment and outcomes (Bryson et al., 2002). This is a limitation of PSM that it fails to control for unobservable which may create a hidden bias. In the evaluation of job training programmes, Heckman, et al, (1997) has shown that the matching method applied to the control groups in the same labour markets using the same survey questionnaire would eliminate much of the selection bias associated with unobservable, though the remaining bias is still not altogether negligible. Notwithstanding, the derived impact depends on the variables used for matching and the quantity and quality of the available data and the procedure to eliminate any sample selection bias. Since our study focused only on rural poor households and the new entrants into the programme forms the subset of non-participant groups and they are also selected from the same locality close to the treatment groups, the same questionnaire is used for both, the disparity in unobservable between the participants and non-participants of the microfinance programme is expected not be large. Hence, the bias may be reduced and the reliability of the matching estimates is improved.

However, results may be subject to some limitations since the study is based on cross-sectional data and therefore caution will have to be taken while interpreting the data. The research methodology for the present study is based on the above analysis and is presented in the next point 3.3.

3.3. Research Methodology

3.3.1. Survey Design

The study is based on quasi-experimental design survey whereby comparison is made between two groups of respondents: the participants (treatment group) of the programme and non-participants (control group). The treatment group comprises of the members of Self-help group who have been benefitted from the microfinance services of Swarnajayanti Gram Swarozgar Yojana (SGSY) scheme and received bank loan (credit linked) through SHGs upto 2010. The respondents of the control group (non-participants) were selected from the newly formed groups who are eligible clients to reflect a comparable socio-economic group as similar as the treatment group. The new entrants are chosen from the same areas who have just entered the programme and did not receive any benefit of the programme. The use of new entrants as control group are suggested by the AIMS (Assessing the Impact of Microenterprise Services) Guidelines (Barnes and Sebastan, 2000). The use of new clients members not only ensures the eligibility of the control group (because they are eligible to get loans after six months of their active existence as per NABARD guidelines) but it is also believe to minimize some of the unobserved differences such as entrepreneurial ability, risk preference and motivation between the treated and controlled unit. Although the issue of why they not joined earlier to access the benefit of the programme may raise some questions (karnal, 2001), the fact that they have decided to give a go to it indicates a level of motivation and determination.

3.3.2. Outcome Indicators of Impact Assessment Applied

The main purpose of impact assessment of a programme (or any intervention) is to measure the extent to which the programme has changed the outcomes of the agents, where the agents can be defined as group of individuals, households, firms, cities, etc. The outcomes of interest are derived from the agents which are namely, income, employment generation, poverty indices, income inequality and financial inclusion indicators. The income of the respondents is worked out by taking into account income from income generating activities of self-help groups and other subsidiary occupation per year. Generation of employment refers to the days of employment generated per year through the various sources of activities, like agricultural crops, goat farming, piggery, group activity and self-employment activities. Indicators of financial inclusion taken into account are household's access to formal credit, savings, insurance and banking transaction services such as usage of ATM/cheque.

3.3.3. Analysis of Data

The data collected from the field are edited, analysed and interpreted carefully. Descriptive analysis including percentage and compare mean are used to present the data. The results and findings are presented with the help of statistical tables and diagrams. Statistical tools such as t-test, Chi-square test, correlation analysis, poverty measurement indices, Lorenz curve, Gini coefficient, Atkinson index, etc. are applied to find out the impact of microfinance programme. Propensity Score Matching (PSM) method was applied using the primary dataset to produce unbiased impact estimators of the evaluation of microfinance programme. Two matching algorithms, namely, Nearest-Neighbour and Kernel Matching methods were applied to estimate the treatment effect of microfinance programme. Computer software and statistical packages like Microsoft-excel, SPSS-22 and STATA-11 were used to apply various statistical techniques and to draw various graphs.

The methodology used in this study is divided as follows:

1. Descriptive analysis is used to answer research objective 1:
   Simple average, percentage and correlation are used to discuss the primary data collected from the SHGs.
2. Quantitative and diagrammatic analysis are used to examine the Research Objectives 2, 3 and 4:

The empirical analysis focused on two parameters of interest while estimating treatment effects of the microfinance programme. First, the impact of microfinance programme on outcomes who were actually treated- that is, the **Average Treatment Effect on the Treated (ATT)**. Second, what effect microfinance would have had on an individual drawn randomly from the population- that is, the **Average Treatment Effect (ATE).** The ATT and ATE will be identical when responses assume homogeneous of assignment to treatment among households. Both these two effect will differ, should the responses to assignment be allowed to vary across individuals. The policy makers concern is to determine whether microfinance had any impact on per-capita income and employment. The ATT provides answers to the question of the impact of microfinance. Another important concern is whether the expansion of microfinance programme in the selected area is worth considering. Therefore, ATE is required to go further and assess the opportunity of expanding microfinance in that area. For instance, if only individuals with the largest expected gains participate in microfinance, then ATE will be smaller value than ATT value. Thus, a generalization of the programme may generate a lower effect of microfinance than the one indicated by ATT. In this research, we employed ATE to assess the impact of microfinance on income, employment and financial inclusion (IFI). We employed ATT to evaluate the impact of microfinance programme on employment, income, poverty indices, income inequality and financial inclusion indicators. Since our research interest is to provide relevant policy implications for the targeted poor households that need microfinance, and there would never be an opportunity to estimate individual effects with confidence, we estimated the treatment effect of microfinance programme using ATT as proposed by Heckman et al., (1997).

The coefficient of ATE is calculated using equation (3.3) stated above which is as follows:

$$ATE = \Delta = E(Y|D=1) - E(Y|D=0) = E(Y^1) - E(Y^0)$$

The coefficient of ATT is obtained using propensity score matching (PSM) estimator equation (3.15) above which is:

$$^{ATT}\Delta_{PSM} = E_{P(X)}(Y^1 \mid X, D=1) - E_{P(X)}(Y^0 \mid X, D=0 \mid D=1)$$

Where, Y is the outcome of interest, D is programme participation dummy; D=1 if an individual is a participant of the programme; D=0 otherwise. X is a covariate of the observed factors including various socio-economic and demographic factors. It includes age of the respondents, education, marital status, female headed dummy,

household size, dependency ratio, family income are the control variables using which the propensity scores for treated and non-treated groups are estimated. The Propensity scores P(X), that is, the probabilities of participating in the microfinance programme based on the covariates are calculated using the probit model. The logit model can also be used to derive propensity score. The difference between the two lies in the distribution of the error terms. The logit model follows standard logistic distribution of errors whereas probit model follows the normal distribution of the error terms. The logit function is similar, but has thinner tails than the normal distribution and probit curve approaches more quickly than the logistic curve. But both the model yield similar results. As such, there is no strong reason for the preference of one over the other. Following much of the literature, probit model is preferred in our study. By formulation, Probit model is a binary choice model that can be estimated using the maximum likelihood estimation using cumulative distribution function (CDF) of the standard normal distribution. So, Probit regression model was estimated to compute the propensity score of participation of each individual belonging to both treatment and the control group. The participation status (D equals 1 if participant and D=0 otherwise) is treated as the dependent variable and control covariates are possible predictors for the model.
The Probit model is specified as:

$$P(X)=P(D=1|X)= \Phi(\beta_1 X_1 + \ldots\ldots\ldots\ldots +\beta_i X_i)= \Phi(X\beta) \qquad (3.16)$$

Where,

$0< \Phi(X\beta)<1$ for all values of X and Φ is the cumulative distribution function of the standard normal distribution. The parameter β are estimated by maximum likelihood estimator.

The selection mechanism by the probit model can be motivated as a latent variable model.

Let us suppose that there is an auxiliary random variable, then

$$D^{*}= X\beta + \varepsilon$$

Where, $\varepsilon \sim N(0, 1)$. Then,

$$D = \begin{cases} 1 & \text{if } D^{*} > 0 \quad \text{that is,} \qquad -\varepsilon < X\beta \\ 0 & \text{otherwise} \end{cases}$$

Where,

$$Pr(D=1|X) = Pr(D^*>0=Pr(X\beta+\varepsilon>0)$$

$$= (Pr =\varepsilon>- X\beta) = (Pr =\varepsilon < X\beta) = \Phi(X\beta)$$

$$Pr(D=0|X)= 1- \Phi(X\beta)$$

and, $D^*=1, \text{ if } D^*= X\beta + \varepsilon >0$

D^* is a latent variable. In our case, D equals 1 if an individual has participated and accessed microfinance and 0 otherwise. X is a vector of individual and household characteristics included in the participation equation. Φ denotes the standard normal cumulative distribution function. The vector of characteristics X included in the model are as follows:

X_1 =Age of the respondent (in years);

X_2 = educational attainment of the respondent (in years);

X_3= Marital Status dummy; (1=married; 0= otherwise)

X_4 = Household size total number of people living in household;

X_5= Agricultural landholdings (in bigha);

X_6= gender of head of the household dummy (1=female; 0=male);

X_7 = Economic dependency ratio

X_8= Occupation of the respondent dummy (1=housewife; 0= otherwise);

X_9= Household Monthly income (in Rs.);

X_{10}= Distance to residence from bank (in kms.);

X_{11}= Age-squared

X_{12}= Education-squared

The covariates included in probit model to compute the propensity score are discussed in detail in the next chapter (Chapter-IV).

After obtaining propensity score, the quality of the matching procedure to balance the relevant covariates are checked by using two measures, namely, Pseudo-R^2 and the t-Test. The Pseudo-R^2's after matching is compared with the before matching. The low value of Pseudo-R^2 after matching indicates how well the regressors X explain the participation probability. It shows that after matching, there are no systematic differences in the distribution of covariates. Moreover, the t-tests was conducted to ensure that the mean propensity score is not different for the treated and control units in order to ensure that a good comparison group is constructed from the selected covariates. After the balancing property is fulfilled and common support is defined, they are used for matching purposes. Finally, Nearest-Neighbour-Matching and Kernel (with the default bandwidth of 0.06) Matching algorithms are used to perform the PSM for the impact assessment of microfinance programme on employment, income, poverty, inequality and financial inclusion.

3.3.4. Assessing the Impact on Poverty and Inequality

The ultimate target of the microfinance programme is to reduce poverty and inequality through financial inclusion, so, we looked at the effect of the programme on poverty and inequality indicators directly. We have used three Foster, Greer and Thorbecke (FGT) weighted poverty index for the quantitative poverty assessment (Foster, et. al., 1984). According to the United Nations, (2001), the most important purpose of a poverty measure is to enable poverty comparisons. Therefore, the choice of FGT is due to its decomposability of the overall population into sub-groups, which allows for comparison. In order to measure the inequality, we used three common measure of inequality, namely, Lorenz Curve, Gini coefficient and Atkinson index.

The FGT measure for the sub-group i^{th}, that is, Pα is calculated using formula (3.17) as follows:

$$P_\alpha = \frac{1}{N}\sum_{i=1}^{q}\left[\frac{Z - Y_i}{Z}\right]^{\alpha} \tag{3.17}$$

Where,

Z is the poverty line, N is the number of people in the sampled population, q is the number of poor people, and Yi is the per capita per month income for person i in ascending order for all households. The measures are defined for $\alpha \geq 0$, where α is a measure of the sensitivity of the index to poverty.

When α=0, we have the headcount index (P_0) which measures the incidence of poverty. The head count ratio measures the proportion of population under the poverty line.

When α=1, we have the poverty gap index (P_1) where P_1 measures depth of poverty. The poverty gap ratio the total amount that is needed to raise the poor from their present incomes to the poverty line as a proportion of the poverty line and averaged over the total population.

When α=2, we have the squared poverty gap index measuring P_2 the severity of poverty. The squared poverty gap takes inequality among the poor into account and means that the larger gaps count for more than the smaller gaps, and hence it captures the severity of poverty.

The Planning Commission (GOI, 2013) estimated the official poverty threshold for defining poverty line of ₹828 and ₹1008 per capita per month for rural and urban areas for the state of Assam for the year 2011-12. Since the area under study is basically rural in nature, and the field survey was conducted in 2013, we used the official poverty line of ₹828 per month per capita specified for the rural areas of Assam for the year 2011-12. Accordingly, the official poverty line, that is, ₹828 per capita per month income is used in our study.

The income inequality among the participants and non-participant households is estimated with the help of inequality measures such as Lorenz Curve and Gini coefficient and Atkinson index.

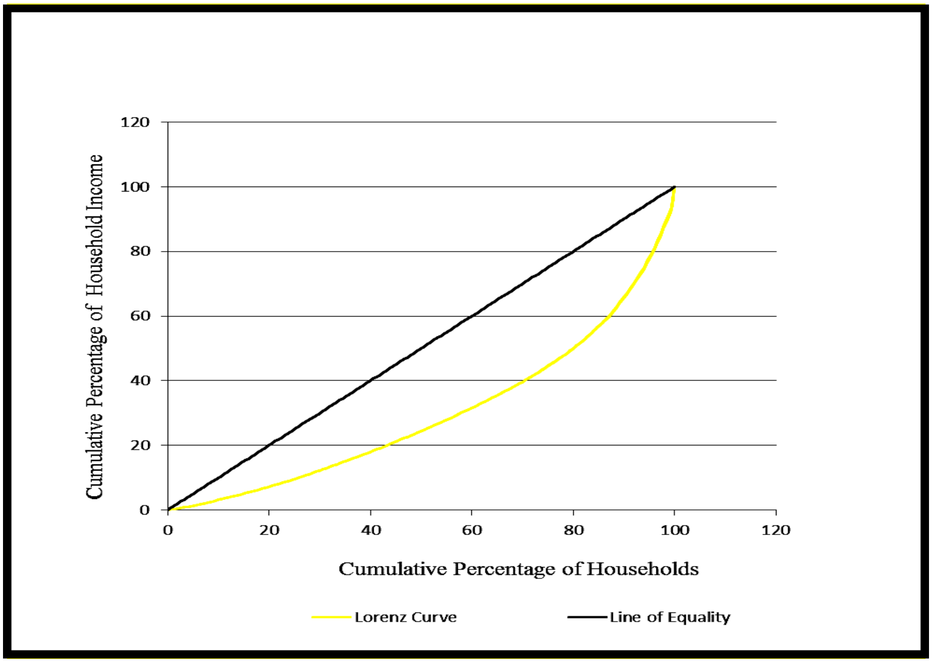


Figure-3.2: Lorenz Curve

Lorenz curve is a graphical representation of inequality of income as shown above in Figure-3.2. The cumulative percentage of population is plotted on the X-axis and the cumulative percentage of income on the Y-axis. The straight line represents the same income for every individual and is called the line of perfect equality. While the other curved line that shows the actual distribution of income is known as Lorenz curve as shown in Figure 3.2. The difference between the line of perfect equality and Lorenz curve shows the inequality in the distribution of income.

The Gini coefficient can be calculated from the individual incomes in the population as follows:

$$G = 1 - \frac{1}{2n(n-1)\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}|Y_i - Y_j| \tag{3.18}$$

Where, Y is an observed value, n is the number of values observed and μ is the mean income. The value of Gini coefficient ranges between 0 and 1. A low Gini coefficient indicates a more equal distribution, with 0 corresponding to perfect equality, while higher Gini coefficient indicate more unequal distribution, with 1 corresponding to perfect inequality.

The Atkinson index of inequality is calculated as follows:

$$A_\varepsilon = 1 - \frac{1}{\mu}\left[\sum_{i=1}^{n} y_i^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} \tag{3.19}$$

Where: Y_i is the income of i^{th} individual and μ is mean income. ε is the inequality aversion parameter. The value of the Atkinson Index varies between 0 and 1. A lower Atkinson value represents more equal income distribution with 0 corresponding to perfect equality and a larger value indicates more inequality with 1 representing perfect inequality. In addition, the sensitivity parameter (ε) ranges from 0 to infinity. As the sensitivity index approaches 0, the Atkinson Index becomes more sensitive to changes in the income position of the higher income groups. As the sensitivity index approaches higher values, the Atkinson Index becomes more sensitive to changes at the lowest income groups.

The impact of the programme on poverty index for participants is expressed as follows:

$$\Delta_p = P(Y^1) - P(Y^0) \quad (3.20)$$

Where, Δ_p denotes the treatment effect on poverty indices. The first term of the right hand side of equation (3.20) is the measure of poverty level of participant households (treatment group) and the second term is counterfactual measure of poverty, that is, poverty indices of the non-participants households (control group). Regarding inequality, we measured the impact of the microfinance programme on reduction of inequality of the whole sampled population and then compared it with the non-participant control group. The impact on an inequality index is expressed as:

$$\Delta_p = I(Y^1) - I(Y^0) \quad (3.21)$$

Where, Δ_I refers to the treatment effect on inequality indices (Gini and Atkinson), $I(Y^1)$ denotes the inequality index of the participant households and $I(Y^0)$ is the inequality of the control group households.

3.3.5. Assessing the Impact on Financial Inclusion

Financial inclusion is a multidimensional concept and is defined by various authors in different ways (Mor and Ananth, 2007; World Bank 2005; Kamath, 2007; Sharma, 2008; GOI, 2008; European Commission, 2008; and Prathap, 2011). Financial inclusion may be interpreted as poor and low income households' access and usage of basic financial services which include savings, credit and insurance available from formal institutions in a manner that is reasonably convenient and flexible in terms of access and design and reliable in the sense that savings are safe and that insurance claim are paid with certainty. This definition is a modified one from the earlier definitions given by (Mor and Ananth, 2007; World Bank 2005; Kamath, 2007; Sharma, 2008; GOI, 2008; European Commission, 2008; and Prathap, 2011). The definition of financial inclusion has stressed that financial inclusion does not essentially focus on providing credit and offering facilities for savings alone, but also includes the whole gamut of financial services including, insurance, savings and money transmission mechanisms mode suited to the income pattern of the poor (Committee on Financial Inclusion, 2008). Thus, financial inclusion includes broadly households' permanent access to formal financial services- credit, savings, insurance, money transfers which are also the products and services of microfinance. This study has developed a Financial Inclusion Index (IFI) to measure financial inclusion as a composite measure that takes into account access to transaction services (uses of ATM/Debit Card/cheque), savings, credit and insurance. The calculation of the financial inclusion index is based on the mathematical concept of weighted average index numbers. The variables and weights are selected based on extensive literature available on the subject and the index has been constructed by using the indicators and corresponding weights as presented in table 3. The financial services selected for preparing Financial Inclusion Index (IFI) are discussed below:

a. Transaction Services

Financial inclusion basically promotes efficient payment mechanism and strengthens the resource base of the economy (Chakrabarty, 2009). The individuals increasingly need money transmission mechanism including services like debit cards, direct debit, automatic transfers, etc. require for storing money, saving and accessing money safely and for making payment to third parties (Kumar, 2002). The Financial Access Survey, (2010), has also probed the usage of other banking services like ATM in addition to deposit or credit (Kendall, et al., 2010). This study takes into consideration usage of transactions services like usage of cheque and usage of ATM/Debit card with the formal financial system as an indicator to measure financial inclusion.

b. Savings

Savings has been the trust area in financial inclusion programme (Government of India, 2008) and the idea of financial inclusion promotes thrifts and develops the culture of savings among the poor and low income group (Chakrabarty, 2009). Savings with a formal financial system determines the basic access to other financial services. Savings can be of various forms, namely, savings bank account, recurring deposits, and fixed deposits with any of the formal financial system including Commercial banks, Cooperative banks, Regional Rural Banks or SHGs. The study looks into household access to operational Savings Bank accounts of banks/post office savings, recurring deposit and fixed deposit accounts. Though savings with SHGs can be accounted as group savings compared to other individual based accounts, it was considered as an indicator for measuring the level of financial inclusion.

c. Formal Credit

Credit has been accorded prime importance in the concept of Financial Inclusion (The Committee on Financial Inclusion of India, 2008). Credit access and indebtedness of a family with formal service providers determines the level of well-being achieved, because access to credit is widely regarded as a financial service (Schilling, 2003). Chant and Link Associates, (2004), observes leverage through credit as a standard and critical financial strategy for a small business and lack of access to credit may place such a business at a distinct competitive advantage (Prathap, 2011). In this study we consider credit accessed by individual household from formal and semi-formal sources for measuring the level of financial inclusion. The formal source comprises of commercial banks, cooperative banks, Regional Rural Banks, while microfinance has been treated as semi-formal finance (Basu, 2006). The access and usage of credit by the household both from formal sources and semi-formal sources including SHG-bank linkage in three preceding years of the survey period have been used to develop the financial inclusion index.

d. Insurance

Insurance is considered a very important financial product because it provides coverage to the accidents or emergencies arising in a society affecting human lives, assets or livelihoods. Research suggests that those consumers who are least well placed to stand the risks are often those without insurance cover (Whyley and McCormack, 1997). Therefore, inclusion in terms of insurance is considered as one element in the financial inclusion index.

These variables are selected on the basis of extensive literature available on the subject and the appropriate weights are assigned (Rangappa, et al., 2009; Anjugam, 2011; Prathap, 2011) by using the judgement method. Furthermore, banking officials, knowledgeable villagers, farmers and researchers were consulted and based on their suggestions, some of the weightages of these variables have been modified in the present study according to the area/field condition. Finally, the following weightage distributions on selected variables demonstrated in table 3.1 were adopted to construct the financial inclusion index (IFI). The selected variables were put to response in the survey and by aggregating responses in each variable, the index was calculated. Hence, index was estimated with the help of the mathematical concept of weighted average index numbers.

The values assigned to each variable are given as either 1 or 0. The value '1' implies household having association with the formal/semi-formal source of finance and value '0' implies having no association with the specified source of finance. Table-3.1 shows the details of financial services selected and their corresponding weightages for developing the IFI.

Table 3.1: Construction of Financial Inclusion Index (IFI)

Indicators	**Sources of Finance**		**Weight**	**Total**
Formal Credit	a	From formal agencies directly and/or through SHG during Survey (during 2012)	30	50
	b	From formal agencies directly and/or through SHG during Survey (during 2011)	10	
	c	From formal agencies directly and/or through SHG during Survey (during 2010)	10	
Savings	d	Operating SB Account in Bank/Post office/Co-operative Banks	10	25
	e	Fixed Deposit or Recurring Deposit Account with Institutional Agencies	10	
	f	Savings in SHG	5	
Insurance	g	Any source/type of insurance	15	15
Transaction Services	h	Usages of ATM/Debit Card/Cheque	10	10
Total			**100**	**100**

As credit is the prime importance, the highest weight is given to borrowing from formal sources (50), followed by saving in formal agencies (25), insurance (15) and other banking services (10) according to the importance of the variables. Though relatively higher weightage (30) has been given to current borrowing (2012) from formal sources/SHG-bank linkage programme, additional weights have been assigned to borrowing from formal/SHG-bank linkage programme during 2010 and 2011 (10 for each). Since rural households often require credit for various purposes, small and frequent borrowing from the internal sources of the SHGs may outweigh participant households of SHGs. Therefore, borrowing from the SHGs which are institutional sources were considered for giving weights. Since saving is compulsory for the SHG members relatively lower weight (5) has been assigned to saving in SHG to avoid the possible overweight to households with SHGs. Usages of ATM/Debit Card or Cheque has been assigned a weight equal to 10.

The index varies between 0 and 100. Value of '100' implies total financial inclusion and value of '0' implies complete financial exclusion. Value '1-29' implies low financial inclusion, value '30-60' implies medium financial inclusion and value '61-99' implies high financial inclusion.

The impact on financial inclusion Index is expressed as:

$$\Delta_{IFI} = IFI(Y^1) - IFI(Y^0) \qquad (3.22)$$

Where, Δ_{IFI} denotes the treatment effect on financial inclusion index. The term $IFI(Y^1)$ denotes the level of financial inclusion of the participant households and $IFI(Y^0)$ is the financial inclusion level of the control group households.

The t-test and Chi-square test were applied to test the significance of various results obtained from the analysis of surveyed data.

The t-test is applied to measure the mean difference between the participants of the microfinance programme and the non-participants in terms of quantitative variables. The null hypothesis (H_0) formulated is that both the samples come from the same normal population ($H_0: \mu_1 = \mu_2$) and there is no significant difference in their mean values. The alternate hypothesis (H_1) is that there is significant difference in the mean values of two samples ($H_1: \mu_1 \neq \mu_2$). To carry out the test, t-value was calculated using the following formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S} \cdot \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$S = \sqrt{\frac{\Sigma(X_1 - \bar{X}_1) + \Sigma(X_2 - \bar{X}_2)}{n_1 + n_2 - 2}} \quad (3.23)$$

Where,

$\bar{X}_1$= Mean value of the first sample

$\bar{X}_2$= Mean value of the second sample

n_1 = Size of first sample

n_2 = Size of second sample

S = Combined standard deviation of two samples

The degree of freedom is equal to n_1+n_2-2.

The estimated value of 't' is compared with the table value for degrees of freedom at a certain level of significance for acceptance and rejection of null hypothesis. If the estimated value of 't' is greater than the table value for n_1+n_2-2 degrees of freedom H_0 is rejected and H_1 is accepted. Conversely, if calculated t-value is smaller than the table value for n_1+n_2-2, H_0 is accepted and H_1 is rejected. The hypotheses are used throughout the analysis for testing the mean values between the groups.

Secondly, the Chi-square test as a non-parametric test was to test the relationship (independent or not) between the categorical variables and treatment assignment. This test shows how likely a categorical variable is independent of the distribution of the two groups. In order to test whether or not categorical variables are associated, the null hypothesis (H_0) formulated that there is no association between the attributes and treatment assignment. The value of Chi-square is calculated as follows:

$$\chi^2 = \sum \frac{(O - E^2)}{E} \quad (3.24)$$

Where,

O = Observed frequencies

E = Expected frequencies.

The expected frequency for any cell can be calculated as follows:

$$E = \frac{RT \times CT}{n}$$

Where,

RT = The row total for the row containing the cell.

CT = The column total for the column containing the cell.

n = Total number of observations.

The estimated value of χ^2 is compared with the table value for degrees of freedom; defined as (Number of columns-1) × (Number of rows-1) at certain specified level of significance. If the calculated value is greater than the Table value at a certain level of significant, the null hypothesis is rejected and the association between the attributes is considered significant. On the other hand, if the estimated value of χ^2 is less than the table value at a certain level of significance, the null hypothesis that attributes are independent is accepted. This hypothesis is used to test the association between categorical variables and treatment assignment throughout the discussions.

The estimated results for the microfinance programme impact estimators are presented and discussed in the next chapter, that is, in Chapter-IV.

3.3.6. Data

The study is mainly based on primary data. The primary source of data is being supported by the secondary sources for the comprehensive analysis of the problem under investigation. Primary data have been derived through field survey using interview method and secondary data have been accessed from reliable secondary sources.

Primary data aimed at capturing all the required information was collected with the help of a schedule especially prepared for the purpose of this study. The questionnaire was pre-tested in select area and then necessary correction and modifications were done in order to adapt all the needs of the study.

Furthermore, we have collected the required secondary data with regard to microfinance and SHG-Bank Linkage programme from the authentic sources like

NABARD (National Bank for Agriculture and Rural Development) publications, Status Reports of Microfinance in India published by Microfinance India, NSSO (National Sample Survey Organisation), DRDA (District Rural Development Agency) of respective districts, Block offices, Journals, Books, etc. to draw the profile of the study area as perspective of the study.

3.3.7. Sampling Procedure and Sample Design

The sampling frame for the study followed multi-stage purposive random sample selection method. The detailed sample selection process is discussed as follows:

In the first stage, two districts namely, Baksa and Udalguri districts were selected purposively. In the second stage, two development blocks from each of the district, namely, Jalah and Baska from Baksa district and Bhergaon and Udalguri from Udalguri district were selected to conduct the field survey. In the third stage, 60 SHGs (taking 15 SHGs from each block) are selected from the selected blocks spread over 35 villages in the study area[14]. In the last stage, three members from each SHG are randomly selected using random number table. Thus, a total of 150 member respondents comprises the participants of the programme due to their absentee of 30 members at the time of survey. Further, 180 respondents are also selected with the similar method and technique from the new entrants (non-participants) of the programme for interview from the same areas to form the control comparison group. Thus, the total sample size comprises of 330 households for the present study. The following table-3.2 shows the sample frame for the present study:

Table 3.2: Sample Frame for the study

Sl. No.	District Selected	Block Selected	No. of Surveyed		
			SHGs	Participants	Non-participants
1	Baksa	Jalah	15	40	45
		Baska	15	39	45
2	Udalguri	Bhergaon	15	34	45
		Udalguri	15	37	45
Total	**2**	**4**	**60**	**150**	**180**

[14] The Name of Villages where SHGs are located are shown in Appendix I.

3.3.8. Data Collection Approach

Data collection approach included survey of households, survey of SHGs and face to face interviews. Questionnaire was administered mainly to group member, but other family members were allowed to provide relevant information which could not be adequately supplied by the respondents.

The schedule for data collection included questions on general information intended to identify the respondents' demographic background such as age, marital status, religion, community, education, occupation, household income, operational bank A/C, usages of ATM/Debit Card or Cheque, household's number of loans obtained from formal or informal source, availability of basic amenities, and land holding. Some of the questions relating to information on respondents' sources of income, insurance cover, nature and days of employment, savings and expenditure pattern, ownership of assets, financial vulnerability, etc. were elicited**.**

The questionnaire also covered information about SHGs profile and activities, such as group size and structure, details about meetings, trainings, sources of SHG information, group maturity, saving per month, total amount of group saving, number of bank loans received, total amount of bank loan received, rate of interest, purpose of loan, etc. were collected. Moreover, an attempt was made to identify the problems relating to group activities faced by the SHGs participants in various process of group functioning.

3.3.9. Selection of Study Area

The study area covers two backward districts viz. Baksa and Udalguri districts of Assam. The districts are most backward and remote areas of Assam. The districts were selected purposively because of the large number of SHGs operating in the selected areas under the Swarnajayanti Gram Swarozgar Yojana (SGSY) scheme which is a comprehensively microfinance programme for financial inclusion and poverty alleviation through SHG-bank linkage model. Thus, the area is chosen in order to capture comprehensively the impact of the group based SGSY programme on financial inclusion (access to microfinance services) and poverty alleviation of the participant households of the programme. The study considered those members of SHGs who availed revolving fund/microloans from the microfinance programme of SGSY scheme at least two years back from the date of survey for sampling and interview. The field survey was conducted from May 2013 to October 2013.

Conclusion

In this chapter research design and methodology has been presented. Various conceptual and methodological problems in impact evaluation of the microfinance programme are discussed to develop research methodology for the present study. The details of research methodology developed, uses of questionnaire, sample design and data and data sources are also discussed.