CHAPTER 3
RESEARCH METHODOLOGY

3.1 Data Base

The study is mainly based on primary data collected through field survey. Some parts of the study are also assisted by secondary data pertaining to area, production and yield of the crops.

3.1.1 Primary Data

For the data collection, field survey was conducted during December, 2015 to June, 2016. The primary survey has been undertaken with the help of pre-tested schedule from the selected contract and non-contract farmers. Besides, a separate schedule was also constructed for the company officials to study procurement operations of the companies and discussions with the key persons of the agri-business companies involved in the contractual practice was also covered.

3.1.2 Secondary Data

Besides primary data collection, the study also uses secondary data from the various issues of Statistical Abstract of Punjab, Handbook on Horticulture Statistics, Indian Horticulture Database, Agricultural Statistics at a Glance, etc.

3.2 Sampling Design

The Punjab Agro Foodgrain Corporation was contacted for attaining a list of companies involved in contract farming of Punjab, but information provided by them was not updated. Thus the companies selected for the study were on the basis of personal contacts. The following companies were considered for the study (Table 3.1). Three companies, one MNC- PepsiCo and two local companies- Paras Spices Pvt. Ltd. and Rana Sugars Limited involved in processing of value added food products was selected for the study. The discussion with company officials included information about total number of farmers, selection criteria for area and farmer, market destination, type of contract, etc. (Appendix A).

Table 3.1
List of the Companies Studied

Companies	Crops
PepsiCo	Potato
Paras Spices Pvt. Ltd.	Chicory
Rana Sugars Limited	Sugarbeet

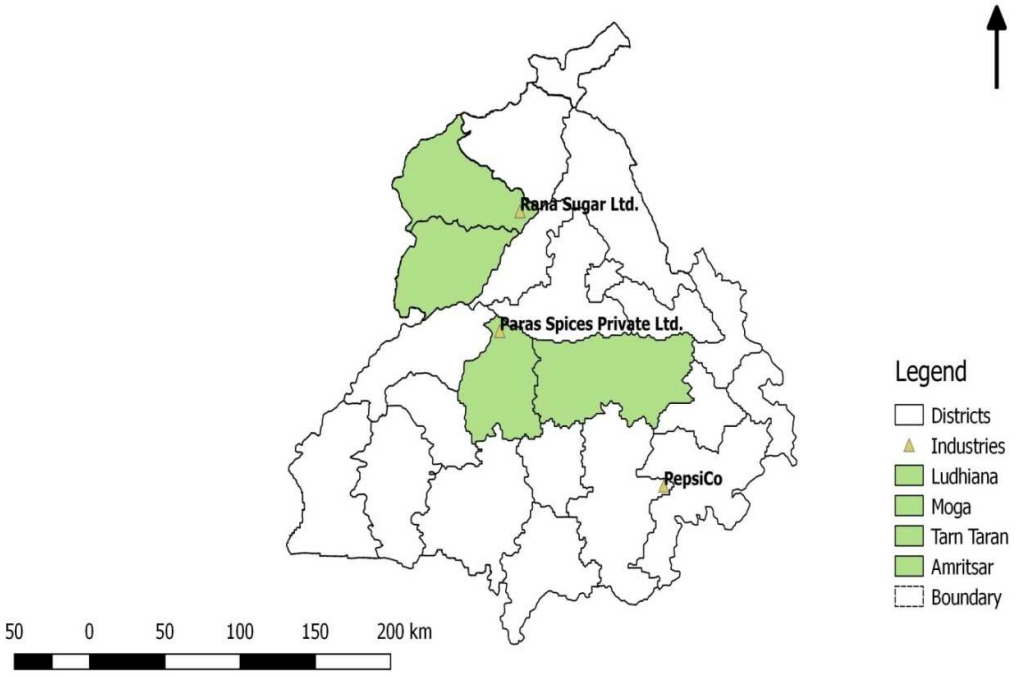
The selection of districts was determined on the basis of maximum area under contracted crops for all the three selected companies. Thus, in case of PepsiCo Ludhiana and Moga district, in case of Paras Spices Pvt. Ltd., Moga and in case of Rana Sugars Limited, Tarn Taran and Amritsar districts were selected. In case of Paras Spices Pvt. Ltd. only one district was selected as about more than 90 per cent of chicory farmers were located in the concerned district only.

Table 3.2
Distribution of Operational Landholdings in Selected Districts of Punjab during 2010-11

Categories	Ludhiana	Moga	Amritsar	Tarn Taran	Punjab
Marginal	8981 (12.70)	8946 (17.30)	8675 (12.41)	7682 (13.03)	164431 (15.62)
Small	11320 (16.01)	10486 (20.28)	16905 (24.18)	12869 (21.82)	195439 (18.57)
Semi-medium	20530 (29.04)	16238 (31.40)	26037 (37.24)	21142 (35.85)	324515 (30.83)
Medium	23093 (32.67)	13674 (26.44)	15924 (22.77)	15256 (25.87)	298451 (28.35)
Large	6770 (9.58)	2368 (4.58)	2383 (3.41)	2018 (3.42)	69718 (6.62)
All	70694 (100)	51712 (100)	69924 (100)	58967 (100)	1052554 (100)

Source: GoP, 2015

Map 3.1:
Location of the Companies along with their Farmers in Punjab



3.2.1 Data Collection

The schedule was prepared for the farmers (Appendix B). A complete list of the farmers was obtained from the company officials. A sample of 50 farmers in each company was taken through the stratified random sampling. The company-wisepopulation of contract farmers was divided into farmer category strata. From each stratum, sample was taken in such a way that proportion of farmers in each farmer category in the sample was similar to that in the contract farmer population. Therefore, a sample of 50 potato growing contract farmers in case of PepsiCo, 50 sugarbeet growing contract farmers in case of Rana Sugars Limited and 50 chicory growing contract farmers in case of Paras Spices Pvt. Ltd. was taken for the survey. Another set of the farmers in the vicinity of the contract farmers was also taken to make the comparison of both categories of farmers with respect to their socio-economic characteristics, factors determining their participation in contract farming, type of cropping pattern followed, etc. In case of PepsiCo, alternative non-contract farmers were potato growers. Since, domestic firms

introduced new crops, sugarbeet and chicory by replacing wheat, the alternative non-contract farmers chosen were traditional wheat growing farmers. Thus, a sample of 50 potato non-contract farmers in the vicinity of the PepsiCo and 50 wheat farmers each in the vicinity of Rana Sugars Limited and Paras Spices Pvt. Ltd. were taken based on the proportion of the farmers in each category in each location through the stratified random sampling technique. Thus, in nutshell, the study was carried out with 150 contracted and 150 non-contracted farmers constituting a sample of 300 farmers.

3.3 Description of the Selected Districts

3.3.1 Ludhiana: The total geographical area of the district is 3767 sq km, which is divided into the flood plains of the Sutlej and upland plain area. Out of total operational holdings, 16 per cent and 12 per cent belongs to small and marginal holdings, respectively. The Ludhiana district consists of seven sub-divisions i.e. Ludhiana (East), Ludhiana (West), Samrala, Khanna, Payal, Raikot and Jagraon. The district is divided into 12 development bocks i.e. Ludhiana (East), Ludhiana (West), Machhiwara, Samrala, Khanna, Doraha, Dehlon, Pakhowal, Raikot, Sudhar, Jagraon and Sidhwan Bet. The net sown area is 2.99 lakh hectares and total cropped area is 5.92 lakh hectares in 2013-14. The literacy level (82.20 per cent) is higher than the state's literacy rate (76.70 per cent).

3.3.2 Amritsar: It is bounded by river Beas in the south-eastern side and river Ravi on the north-west side. It comprises 9 development blocks viz. Ajnala, Chogawan, Harsha Chhina, Jandiala Guru, Majitha, Rayya, Tarsikka, Verka and Attari with geographical area of 2.64 lakh hectares, out of which 2.22 lakh hectares are cultivable.The net sown area is 2.20 lakh hectares and total cropped area is 4.18 lakh hectares in 2013-14. The literacy rate in the district is 70.68 per cent less than the state's literacy level (Table 3.3).

3.3.3 Moga: Administratively, the district has been divided into three sub divisions (Moga, Baghapuran and Nihal Singh Wala) and five development blocks i.e. Moga-I, Moga-II, Kot-Ise-Khan, Baghapuran and Nihal Singh Wala. The net sown area is 1.94 lakh hectares and total cropped area is 3.82 lakh hectares in 2013-14. The literacy rate in the district is 76.27 per cent.

3.3.4 Tarn Taran: The district is divided into 8 development blocks, viz., Bhikhiwind, Chohla Sahib, Gandiwind, Khadur Sahib, Naushehra Pannuan, Patti, Tarn Taran, Valtoha. The net sown area is 2.17 lakh hectares that is almost double cropped and some area is even put to 3 crops a year. 40 per cent area is under tubewell irrigation and 60 per cent under canal irrigation. Out of total 58967 operational holdings, about 34 per cent is under small and marginal holdings. The total geographical area of the district is 241449 hectare. About 4192 hectare area of the district is beyond the border wiring; out of which 3597 hectare area is cultivable.

Table 3.3
Profile of Selected Districts

Characteristics	Ludhiana	Moga	Amritsar	Tarn Taran	Punjab
Geographical area (sq. km)	3767	2242	2683	2414	50362
Total Population (in lakhs)	34.98	9.95	24.90	11.19	277.43
Literacy rate (%)	82.20	76.27	70.68	67.81	76.70
Gross cropped area (000 hectares)	592	382	418	401	7848
%age of gross irrigated area to gross cropped area	100	100	100	100	98.5
Net sown area (000 hectares)	299	194	220	218	4145
Cropping intensity	191	198	196	196	191

Source: GoP, 2015

3.4 Analytical Tools

The data collected through schedule was tabulated and analysed. The farmer category-wise analysis was done across contract and non-contract farmers. The data pertaining to socio-economic characteristics, perceptions of farmers was analysed using simple descriptive statistical tools and techniques like mean, percentage, etc. supplemented by qualitative observations. The various other analytical techniques used are outlined as under:

The cost concept ‘A1’, ‘A2’, ‘B’ and ‘C’ was adapted to enlighten the economics of crop cultivation. Cost ‘A1’ included all the direct expenses incurred on crop production in cash and interest on working capital, whereas cost ‘A2’ included cost ‘A1’ plus rent paid for the leased-in land. Cost ‘B’ included cost ‘A2’ plus interest on the value of fixed assets. Cost ‘C’ included cost ‘B’ plus imputed value of family labour. The costs of hired and family labour were estimated on the basis of average market rates prevalent for hiring labour in the locality. Interest on the working capital was charged @ 12 per cent per annum whereas cost of fixed capital was considered @ 10 per cent of the total fixed assets (Johl and Kapur, 2001).

3.4.1 Simpson Index of Diversification

Horizontal diversification is the increase in number of crops grown in order to either increase or stabilize their income. To assess the impact of cropping pattern Simpson index is used to measure the extent of diversification.

$$Simpson\ index\ (SI) = 1 - \sum_{i=1}^{n} p_i$$

Where, i= 1, 2,…..n

P_i = proportionate area of traditional (wheat-paddy) crop in the gross cropped area

SI is bound by 0 and 1 with 0 implying complete specialization and 1 implying complete diversification.

3.4.2 Treatment Effect Model

Treatment effect model is used to explain determinants of participation in contract farming and whether participation in contract farming scheme affects farm income. To explain these relationships, unobserved factors that may affect both likelihood of participation and farm income was also taken into account. The study employed treatment effects model. As OLS model does not take into consideration possible selection bias in contract participation. If contract farmers tend to be more skilled than non-contract farmers, then they would have higher income regardless of whether they participated in the contract farming scheme. In this case, the coefficient on the participation dummy variable will include the effect of these unobservable characteristics in addition to the effect of contracting, thus over-

estimating the effect of contracting. This indicates that there is a correlation between independent and dependent variables. The correlation leads to results having inconsistent and biased estimates of the coefficient of variables between participants and non-participants of contract farming in the income model. By using participating probit model, an inverse mills ratio was computed for each observation and included this as an independent variable in the income model. Inverse mills ratio is the ratio of the probability density function over the cumulative distribution function. This term corrects for possible selection bias and yield unbiased and consistent estimates in the income model. This analysis is implemented as maximum likelihood estimation as all the parameters in both models are estimated simultaneously, rather than as a two- step procedure (Puhani, 2000; Warning and Key, 2002; Sharma, 2008; Miyata *et al.*, 2009; Sambuo, 2014).

The equation for the study is:

$$Y_i = \alpha + \beta X_i + \mu C_i + \mu_i \qquad (1)$$

$$C_i^* = \gamma_1 + \gamma_2 Z_i + e_i \qquad (2)$$

$$C_i = 1 \; if \; C_i^* > 0, otherwise \; C_i = 0$$

Where Y_i is the gross revenue of the *i*th farmer, C_i is a dummy variable taking the value 1 if one participates in a contract with a contracting firm, and 0 if one does not participate in any contractual arrangement. *Xi* is a vector of the variables believed to affect the gross revenue and μ_i is a zero mean random variable; while *β*measures the impact of contracting on gross income. An OLS estimate of equation (1), is likely to be biased, because of the effects of unobservable factors. Thus, e_i (which contains within it the random unobservable factors) will be correlated with C_i. To correct for selectivity bias, equation (2) (probit) is estimated with a contract/independent producer as a binary dependent variable (C_i) and a set of explanatory variables *Zi*. Variables in *Zi* will overlap with variables in *Xi*. Identification requires that there should be at least one variable in *Zi* that is not in *Xi*. Then, predicted values (also known as the inverse Mills ratio) from equation (2) can be used as an instrument (of C_i) in equation (1) (Greene, 2003).

3.4.3 Data Envelopment Analysis

Technical Efficiency (TE) means that the transferring of physical inputs into outputs at the best level of performance i.e. without any wastage of resources to produce specific quantity of output (Charnes *et al.*, 1978). Coelli's (1996) 'A Data Envelopment Analysis (Computer) Program version 2.1' was used for the analysis of technical efficiency of farm. Data envelopment analysis (DEA) is a linear programming problem that provides a means of calculating apparent efficiency levels within a group of organizations. The efficiency of an organization is calculated relative to the group's observed best practice (Bhagavath, 1998). The purpose of DEA is to construct a non-parametric envelopment frontier over the data points such that all observed points lie on or below the production frontier.

Farmer is said to be technically efficient if he produces maximum feasible output from a given set of inputs or uses minimum amount of inputs to obtain a given level of output. So there are two measures of TE i.e input-oriented and output-oriented efficiency measure. The present study used input-oriented measure of efficiency.

Consider data are available on K inputs and M outputs for each of N farms. For the i^{th} farm in t^{th} time period, input and output data are represented by the column vectors x_i and y_i, respectively. The data for all N farms may be denoted by K×NT input matrix, X and M×NT output matrix, Y. The DEA model for TE under the assumption of constant returns to scale (CRS) is:

$$min_{\theta,\lambda}\ \theta$$

Subject to $$-y_i + Y\lambda \geq 0,$$

$$\theta x_i - X\lambda \geq 0,$$

$$\lambda \geq 0$$

Where ɵ is a scalar and λ is an N×1 vector of constants. This envelopment form involves fewer constraints than the multiplier form (K+ M < N+1), and hence is generally preferred. A measure of $ɵ_i$=1 indicates that the farm is completely technically efficient. However, the assumption of CRS is correct only as long as farms are operating at an optimal scale. In the case of agriculture, increased

amounts of inputs do not proportionally increase the amounts of output may cause the farm to operate at a non-optimal scale (Ahuja, 2007; Manjunatha *et al*., 2009). Using CRS DEA model will cause TE measures to be influenced by scale efficiencies. By adding convexity constraint, variable returns to scale (VRS) is instead assumed:

$$min_{\theta,\lambda}\, \theta,$$

Subject to $$-y_i + Y\lambda \geq 0,$$

$$\theta x_i - X\lambda \geq 0,$$

$$N1'\lambda = 1$$

$$\lambda \geq 0,$$

Where N1 is a N×1 vector of ones. This constraint makes the comparison of farms of similar size possible.

3.4.5 Garrett's Ranking Technique

The problems and benefits of involvement in contract scheme were prioritised using Henry Garrett Ranking Technique. As per this method, farmers have to assign the rank to major issues of dissatisfaction and benefits and the outcome of such ranking has been converted into score value with the help of the following formula:

$$Percent\ position = \frac{100\ (R_{ij} - 0.5)}{N_j}$$

Where

R_{ij}= Rank given for the i^{th} variable by j^{th} farmers

N_j= Number of variable ranked by j^{th} farmers

With the help of Garrett's Table, the percent position estimated was converted into scores (Appendix C). Then for each factor, the scores of each individual are added and then total value of scores and mean values of score was calculated. The factors having highest mean value was considered to be the most important issue among the farmers.