CHAPTER 5

METHODOLOGY

Research methodology is arrived based on the research objective, scope and theory built. Further, previous literature gave the positive indications of doing an empirical study with a structured questionnaire based quantitative research. This chapter includes the process design, data collection steps, survey research procedure, sampling design, sample size adequacy, elite interview, instrument development and testing, reliability & validity tests and statistical tools of analysis with respective software.

5.1 RESEARCH PROCESS

The type of research work undertaken is basic in nature and the research design is descriptive one. The research process was executed in three phases as described below:

- Phase one included literature survey, defining the problem, scoping of study objectives, identifying constructs and the associated contributing factors and development of a theoretical framework. The extensive literature survey resulted in finalising the objectives of the study and developing the theoretical framework.
In phase two of the study, the researcher observed the latest advances in the chosen area of the study and developed insights through interactions with various experts in the field. This furthered the efforts in research design and instrument development.

In phase three of the study, data collection, analysis, interpretation and validation of results were carried out.

5.2 DATA COLLECTION

Both primary and secondary data were considered for the study. The details and sources considered are discussed here:

**Primary data:**

Survey research and interviewing the experts (elite interview) were the main sources of primary data collection. Survey research using a ‘structured questionnaire as a research instrument’ was used for collating information from the respondents identified for the study. As a new postulation on ‘BIs playing the role of innovation intermediaries’ was made, elite interview was used to validate the findings from the survey.

**Secondary data:**

The secondary data collection focused on information from previous literature, reports from OECD, infoDev, ‘Indian Science Parks and Business Incubators Association (ISBA), NBIA, NKC, NSTEDB, and UKBI. In addition, information from books on BI & OI, articles in periodicals and blogs of opinion leaders from the innovation and incubation domains were collated.
5.3 SURVEY RESEARCH

This research work is a basic research study carried out using population survey method. Survey research is one of the most important areas of measurement in ‘Social research’ applications. This encompassed measurement processes involving questioning of respondents. Two broader categories of survey research are questionnaire and interviews. Questionnaires are usually paper-pencil based instruments or web forms wherein the respondents fill details online. Interviews are conducted either in person or over phone or using internet.

5.4 AIM

The research design step was essential to determine the optimal research design, data collection method, and to explore the relationship with the variables through in-depth interview with focus groups.

As part of the research, various stakeholders in the incubation ecosystem i.e. incubation managers, incubatee start-ups, business accelerators, Government officials in charge of innovation and incubation programs, donor agencies connected to innovation and incubation programs, R & D and OI experts from industry, new product development team members, angel investors, SME and start-up consultants, service providers comprising IPR attorneys, legal experts from India and few other countries were interviewed by the researcher during participation in conferences, seminars, and other events. The interactions were primarily to understand the various factors that influence success of incubatee firms in an incubator. The learning out of these interviews and views of the various stakeholders coupled with that of prior literature information was the base for the researcher to develop a questionnaire. This provided an understanding of the issues related to incubatees’ success, incubation environment influencers’ impact on
performance of start-ups and role of an incubator as an innovation intermediary. Incubators having at least one year of operation were considered for the survey by the researcher. Incubator facilities, infrastructure support services, access to financial support, marketing, talent, mentoring, legal and networking were factored as elements contributing to the performance of the incubatee firms. The preliminary literature survey and the interview resulted in developing a detailed list of factors that influence incubatee performance. These too were helpful in finalising the parameters to evaluate the performance of the incubatee firms.

5.5 SAMPLING DESIGN & DATA COLLECTION

In India, business incubators are promoted by Government agencies namely ‘Department of Science and Technology (DST)’, ‘Department of Biotechnology’, ‘Ministry of Agriculture’, ‘Ministry of Communication and Information Technology’ and ‘Ministry of Micro Small and Medium Enterprises’. Barring DST, all other programs were started after 2009 and for want of in depth experience and understanding of the incubation systems, only the DST promoted incubators were considered for the study. The DST incubator programs are well structured and homogeneous in nature.

DST’s ‘National Science and Technology Entrepreneurship Development Board (NSTEDB)’ division has been promoting business incubation in India since the year 1986. Fifty-two such incubators were on the list of supported institutions and were operational during July 2014. All of them were taken as part of the ‘Universe’ and 100% enumeration has been applied through population survey method.

Sampling unit: Incubators having at least a year of operation, working on different thrust areas were considered for the interview by the researcher. Respondents were managers of such incubator organisations.
The data collection was done by circulating the questionnaire to the incubator managers of these fifty-two NSTEDB promoted incubators. Prospective respondents were explained over email about the objectives and purpose of the study. The study was carried out during the period from July 2014 to September 2014. 39 responses were received and it had a response rate of \( \frac{39}{52} \times 100 = 75\% \). It is quite significant considering geographic spread of the incubators and their priorities. Completed responses were scrutinised and these 39 data responses were considered for the study.

For the sample results Kolmogorov-Smirnov test has been applied for testing normality and it is proved to be not normally distributed. Therefore, non-parametric modelling such as ‘Partial Least Square’ (PLS) is used instead of ‘Ordinary Least Square’ (OLS). The results are presented in ‘Table A 2.1’ in appendix 2.

5.6 SAMPLE SIZE ADEQUACY OF THE STUDY

Sample size of 39 yielded power of the test as 61.1%. This is the practical limitation of the study

<table>
<thead>
<tr>
<th>Table 5.1</th>
<th>GPower3.0 Software Power Analysis: Post hoc: Compute achieved power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input: Tails(s)</td>
<td>One</td>
</tr>
<tr>
<td>Effect size (</td>
<td>r</td>
</tr>
<tr>
<td>(\alpha) err probability</td>
<td>0.05 (5% level of significance)</td>
</tr>
<tr>
<td>Total sample size</td>
<td>39</td>
</tr>
<tr>
<td>Output: Non-centrality parameter (\delta)</td>
<td>1.963961</td>
</tr>
<tr>
<td>Critical (t)</td>
<td>1.687094</td>
</tr>
<tr>
<td>Distributed frequency</td>
<td>37</td>
</tr>
<tr>
<td>Power (1-(\beta) error probability)</td>
<td>0.611413</td>
</tr>
</tbody>
</table>
5.7 ELITE INTERVIEW

As all the three parts of this study were exploratory in nature, in the absence of pre validated constructs, it was essential to validate the findings from the survey research. Hence, a group of BI managers with over 10 years of domain experience i.e. Government officials in charge of running the BI programs, start-up founders / entrepreneurs connected with BI programs, mentors and international incubation experts (program manager for developmental organisations) were identified. Using judgement and convenience methods, thirteen of them were requested to participate in the elite interview. This elite interview was essential to overcome the challenges due lack of pre validated constructs, sample size adequacy and possible respondent bias.

5.8 RESEARCH INSTRUMENT

Questionnaire design involved selecting question content, wording, measurement scales, response format and sequence of questions. The researcher gained required insights through his practical experience in business incubation domain, participation in national and international conferences (related to start-ups, innovation, incubation and open innovation domains), interactions with stakeholders concerned and literature review. These steps were helpful in confirming the existing knowledge on the incubation process and to develop a questionnaire with clear awareness of the content to be included.

Negative worded questions were not used to avoid confusion to respondents in answering the questions.
The questionnaire in this study had four parts:

i. ‘General information’ about the incubator examined legal status, year of establishment, type of host institute, incubation space, average incubation period, number of companies under incubation, number of graduated companies, major thrust areas of the incubator, number of ventures supported under seed funding, and services provided by the incubator.

ii. ‘Defining success of incubatee firms’ probed about seven ‘success qualifiers’ of incubatee firms.

iii. ‘Incubation environment influencers contributing to the success of incubatee firms’ had the determinants of incubatee success from the incubation services and corresponding contributing factors. i.e., ‘Access to infrastructure, mentoring, funding, talent, market and legal/IPR/accounting support’.

iv. Examining the potential of incubators becoming innovation hubs had items examining the ability, transformation, reputation and success factors emerging from the backdrop of incubators’ role as innovation hubs.

A copy of the questionnaire is presented in ‘Appendix 1’.

5.9 RELIABILITY ANALYSIS

From this analysis, the applicability of the ‘Structured Model’ was ensured for the assumptions. A construct is reliable if it measures a dimension the same way, every time with minimum variance. Cronbach’s alpha is the one of the popular measures of reliability. Alpha ranges from 0 to 1. A pre validated construct is said to be reliable if alpha value is more than 0.7. A
newly proposed construct is said to be reliable if alpha value is more than 0.50. Construct reliability analysis was done using Cronbach’s alpha measure using SPSS software to determine the reliability of the various dimensions factored in for the study.

The Cronbach’s alpha measure for the constructs of the study is presented in Table 5.2.

Table 5.2 Reliability and 'Average Value Extracted (AVE)'

<table>
<thead>
<tr>
<th>Construct</th>
<th>Composite Reliability</th>
<th>AVE</th>
<th>Croanbach Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>0.689920</td>
<td>0.269265</td>
<td>0.624105</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.856856</td>
<td>0.602715</td>
<td>0.810187</td>
</tr>
<tr>
<td>Mentoring</td>
<td>0.689920</td>
<td>0.269265</td>
<td>0.624105</td>
</tr>
<tr>
<td>Funding</td>
<td>0.736382</td>
<td>0.389390</td>
<td>0.713527</td>
</tr>
<tr>
<td>Talent</td>
<td>0.636461</td>
<td>0.311295</td>
<td>0.774357</td>
</tr>
<tr>
<td>Market</td>
<td>0.817682</td>
<td>0.440142</td>
<td>0.762015</td>
</tr>
<tr>
<td>Legal</td>
<td>0.819820</td>
<td>0.537290</td>
<td>0.734386</td>
</tr>
<tr>
<td>OI Hub</td>
<td>0.858234</td>
<td>0.378211</td>
<td>0.804032</td>
</tr>
</tbody>
</table>

From Table 5.2, it is understood that the ‘Composite reliability coefficient’ and Croanbach alpha coefficients are greater than 0.60. Therefore, it shows high reliability of the constructs as well as the model.

5.10 VALIDITY

Since 39 responses were taken as sample for population size of 52, the Wald-Wolfowitz run test is applied to ensure the randomness of the sample (Wald & Wolfowitz 1940). The test result in table A2.1 (Appendix 2) shows that the sample is a random one. Hence, the random sample could be
used for future statistical analysis. The instrument was tested by experts in the field and was found to be valid by them (Friedman & Rafsky 1979). In addition, related literature references were considered while choosing the main constructs and their contributing factors.

5.11 STATISTICAL ANALYSIS TOOLS

The following statistical tools were used in the study:

5.11.1 Structural Equation Modelling (SEM)

Structural Equation Modelling (SEM) portrays one of the two ‘computer based statistical fit software packages’: PLS-PA and LISREL/AMOS software. These methods substitute regression equation approaches used in econometrics. SEM models are generally suitable for theory testing and model development as they permit both confirmation and exploratory models.

SEM is a collection of statistical techniques that allow a set of relations between one or more independent variables (IVs), either continuous or discrete, and one or more dependent variables (DVs), either continuous or discrete, to be examined. Both IVs and DVs can be either measured variables (directly observed), or latent variables (unobserved, not directly observed). SEM is also referred to as causal modelling, causal analysis, simultaneous equation modelling, analysis of covariance structures, path analysis, or ‘Confirmatory Factor Analysis (CFA)’. The latter two are actually special types of SEM (Ullman 2006).

5.11.2 Descriptive

Descriptive statistics is the discipline of quantitatively describing the main features of a collection of information / data in a more meaningful
way in order to arrive at certain patterns from the data or simply the quantitative description itself. Descriptive statistics are generally used to abridge a sample and hence are not used to learn about the sample’s population. So, the interpretation is that descriptive statistics, unlike inferential statistics, are not developed on the basis of probability theory. Measures of central tendency, variability or dispersion are some common measures used to describe a data set. Measures of central tendency include the mean, median and mode. The measures of variability include the standard deviation (or variance), the minimum and maximum values of the variables and skew. Descriptive statistics brings out abstracts about the sample and the interpretations that have been made. Such abstracts either could be quantitative or graphs. These summaries become the basis of the initial description of the data or they may be sufficient enough for a particular investigation (Mann 1995).

5.11.3 Correlation

Any statistical relationship between two random variables or two sets of data is depicted by the term ‘dependence’ in the field of Statistics. A general /broad class of statistical relationships involving dependence is known as ‘correlation’. Departure of two or more random variables from independence could be construed as correlation, then technically it refers to any of several more specialised types of relationship between mean values. ‘ρ’ or ‘r’ are representations of several correlation coefficients depicting the degree of correlation.

5.11.4 Confirmatory Factor Analysis (CFA)

In social research, a special method of factor analysis known as ‘Confirmatory factor analysis (CFA)’ is widely used. It is used for testing consistency of measures of a construct with a researcher's understanding of
the nature of that construct. CFA’s purpose is to test if the data fit is a hypothesised measurement model. Theory and/or prior art is the basis for hypothesised model. CFA was first developed by Joreskog (1967). In confirmatory factor analysis, the researcher first develops a hypothesis about what factors he/she believes are underlying the measures he/she has used and may impose constraints on the model. By imposing these constraints, the researcher is forcing the model to be consistent with his/her theory.

For some applications, the requirement of "zero loadings" (for indicators not supposed to load on a certain factor) has been regarded as too strict. A newly developed analysis method, "exploratory structural equation modelling" (Asparouhov & Muthen 2009), specifies hypotheses about the relation between observed indicators and their supposed primary latent factors while allowing for estimation of loadings with other latent factors as well.

5.11.5 Mann-Whitney U Test

Mann-Whitney U test (also called the Mann-Whitney-Wilcoxon (MWW), Wilcoxon rank-sum test (WRS) is a non-parametric test of the null hypothesis. The purpose is to ascertain that two samples come from the same population against an alternative hypothesis, especially that a particular population tends to have larger values than the other.

It has greater efficiency than the t-test on non-normal distributions, such as a mixture of normal distributions, and it is nearly as efficient as the t-test on normal distributions.

The Wilcoxon rank-sum test is not the same as the Wilcoxon signed-rank test, although both are nonparametric and involve summation of ranks.
5.11.6  Path Analysis

Path analysis is known as an analysis of a set of interdependent variables together. This statistical technique is used to examine causal relationships between two or more variables. The aim of PLS path modelling is to analyse the interdependent relationships between factors and the unidimensionality of factors in one analysis algorithm. Path analysis is a subset of ‘Structural Equation Modelling (SEM)’, the multivariate procedure that, as defined by Ullman (Ullman 2006), “allows examination of a set of relationships between one or more independent variables, either continuous or discrete, and one or more dependent variables, either continuous or discrete.”

PLS path modelling is a combination of confirmatory factor analysis and path analysis.

A SEM was conducted on the model proposed. The purpose was to ascertain the validity of the constructs and the paths postulated in the model. SEM provides a convenient framework for statistical analysis that includes several multivariate procedures. This is used to investigate whether the model developed for research is valid for data. The process of pre deciding factors and testing them for unidimensionality is known as confirmatory factor analysis. PLS was considered ideal, if the conditions relating to sample size, interdependence, or normal distribution is not met and if predictions are more important than parameter estimation. A Structural Equation Model (SEM) with all constructs used in the study was analysed using ‘Visual PLS 2.0’ for identifying significant relations between variables of interest in the study.

Visual-PLS is a Graphic-User-Interface program for ‘Latent Variables Path Analysis’ with ‘Partial Least Squares Version 1.8’ running on the Windows environment. It helps the researchers to prepare, edit, and do the PLS analysis more easily. This software allows the researchers to focus more on their appropriateness of the research model and data set rather than on how
to get the program to run, which will dramatically decrease the time and efforts needed to complete a PLS analysis.

Apart from the above, SPSS software was used to determine association between certain nominal variables. Cross tabulation was performed along with the chi-squared test using SPSS software.

The incubator demographic classifications were carried out by taking frequencies and percentages. Extensive data analysis was carried out using arithmetic means as a comparison tool on the variables influencing the performance of the incubators.