Chapter 3

Logistic Regression Model of Occupational Choice
– A Pilot Study

3.1 Introduction

The Self Help Group (SHG) members selected in a sample have several occupational choices. In choosing their occupations, they make a binary choice between the Entrepreneurship /Self Employment & No Entrepreneurship/No Self Employment (The basic form of entrepreneurship is assumed to be Self Employment). Different factors, however, influence SHG member’s choice. To be able to predict and explain these decisions, the study employs discrete choice models, given that the outcome variable is not continuous, to identify the factors and provide information on the direction and magnitudes of their effects. To analyze this choice problem we have used the logistic regression model. In specifying the model, the outcome variable takes the value of 1 if the option, say the Self Employment/ Business (SE/B) is chosen and 0 if the other option, i.e., the No Self Employment/No Business (NSE/NB) is chosen. Predicting the probability that one option will be chosen depends, however, on a vector of explanatory variables $x_i$, and a vector of unknown parameters. A dichotomous logistic regression model was used to estimate the significance of the factors that determine the probability of SHG member’s choice of occupation. The model has several desirable characteristics. It ensures, unlike the linear probability model, that the estimated probabilities will lie between 0 and 1.

A logistic regression model allows us to establish a relationship between a binary outcome variable and a group of predictor variables. It models the logit-transformed probability as a linear relationship with the predictor variables. More formally, let $y$ be the binary outcome variable indicating failure/success with 0/1 and $p$ be the
probability of $y$ to be 1, $p = \text{prob}(y=1)$. Let $x_1, ..., x_k$ be a set of predictor variables. Then the logistic regression of $y$ on $x_1, ..., x_k$ estimates parameter values for $\beta_0, \beta_1, ..., \beta_k$ via maximum likelihood method of the following equation.

$$\text{logit}(p) = \log(p/(1- p)) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

In terms of probabilities, the equation above is translated into

$$p = \exp (\beta_0 + \beta_1 x_1 + ... + \beta_k x_k) / (1+\exp (\beta_0 + \beta_1 x_1 + ... + \beta_k x_k))$$

This Chapter begins with section 3.1 which introduces to the logistic regression framework and research methodology used to discuss the results of the pilot study undertaken. Section 3.2 elucidates the important explanatory variables of the occupational choice model related to the data collected for pilot study. Section 3.3 discusses the assumption of the logistic regression model. Section 3.4 deals with the analysis of the logistic regression model results in R. Section 3.5 presents the interpretation of log odds ratio for the model. This chapter ends with section 3.6 on conclusion and policy implications.

### 3.1.1 Sample selection for Pilot Study

Primary Survey was conducted to collect data for Pilot Study. Clients (SHG members) of Self Help Promoting Institutions (SHPIs) in Mumbai were interviewed during this Survey. The clients of the following SHPIs viz. Municipal Corporation of Greater Mumbai (MCGM), Parisar Vikas, Creative Handicrafts, Asha Kiran Community Centre were selected for field survey. The total number of SHG members (respondents) interviewed was 364 from 141 SHGs. Survey was conducted in the months of July and August 2014. Approximately 2% of the total SHGs reported by the field level coordinators of the various SHPIs were selected in a sample. In the first place, all SHG members were divided on the basis of their affiliation to the SHPIs. Then SHGs from the field areas of respective SHPIs were identified with the help of the field coordinators of SHPIs. Finally those SHG members were interviewed on the
basis of their availability at their residence/work place on the day of field visit. Thus Multi Stage Stratified Sampling method was used to select the sample.

3.1.2 Types of self help promoting institutions (SHPIs) covered

Following categories of self help promoting institutions and groups promoted by them have been considered for conducting a survey for pilot study.

1. Municipal Corporation of Greater Mumbai (MCGM) – MCGM promotes SHGs of women under two main schemes viz. Swarna Jayanti Shahari Rojgar Yojana (SJSRY) & Mahila Bal Kalyan (MBK). Approximately 1000 SHGs are promoted by MCGM in all 24 wards of MCGM.

2. Creative Handicrafts (CH) – Fair Trade Organisation (NGO) – Approximate number of SHGs promoted by CH in the slums of Andheri-Chakala area of Mumbai is 70.

3. Parisar Vikas- Stree Mukti Sanghatana- Approximately 200 SHGs were formed and nurtured under Parisar Vikas Programme in the slum areas of Govandi in Mumbai. The Parisar Vikas programme was launched in the year 1998 by the Stree Mukti Sanghatana with the cooperation of the Municipal Corporation of Greater Mumbai (MCGM). The programme aims to address the problems of waste management and of self-employed women engaged in the ‘menial’ tasks of collecting waste.

4. Asha Kiran Community Centre of St. Anthony Church, Vakola, Santacruz East. Approximately 45 SHGs were managed by the Asha Kiran field coordinators in the nearby slums of Vakola, Santacruz East in Mumbai.
3.2 Explanatory variables for the entrepreneurial occupational choice model

The logistic model tries to estimate occupational choice of SHG members in Mumbai. To explain this choice, following explanatory/predictor variables were considered for the present study. The predictor variables are continuous, and categorical. Discrete Continuous variables includes Age (in completed years), Earning Members (EarningM), No. of Family members (NFM), No. of Old members (OldM), No. of dependent members due to illness / disability (Depe), no. of children. The Categorical variables includes, Income- 5 categories, Self Employed Partner/Father/Mother (SEPFM), Ownership of House in Mumbai (House), Perception about the role of education and training in developing sense of initiative (PERCe.t), Risk Attitude (RiskAtti), Perception of one’s own skills to implement and manage new business activity (PERCa.ms), Perception about lack of financial support (PERCfinance), perception about administrative complexities (PERCadminpro), Perception of insufficient information (PERCinform), Perception of infrastructural constraints (PERCinfra). The nature of these explanatory variables is discussed briefly as follows.

3.2.1 Age

It was believed that decision to choose Self Employment as one’s occupation and age of a person would have positive correlation initially up to a certain age, reaching peak at that age, would start having negative correlation thereafter. Mean age of the sample was 40.27 years, Median age was 40 years. SHG members’ age measured in completed years was reported to be between 18 and 65 years. Standard Deviation for the age was 8.94years. The coefficient of variation for age was 22.20%.

Age measured in completed years was included as a discrete variable in the model. Given the average age over 40 years for the sample, it was expected that most of the SHG members would have no time constraint while making this choice. It was expected that these members would be relatively free form child care and other household responsibilities after the age of 40.
Figure 3.1 Box plot for Age distribution of the sample
Figure 3.2 Occupational choice by SHG members according to Age

3.2.2 Age square

Literature on entrepreneurship highlights that mathematical form of the relationship between age and occupational choice of an individual may not be linear. To capture this non linear nature of relationship Age square was included in the model as a variable.
3.2.3 Income

Monthly income per household reported during survey ranged between Rs. 2000 and Rs.50,000. Average Monthly Family Income reported was Rs.11,619.51 per household. Median Monthly Family income was Rs.10000. Standard Deviation of Monthly Family Income was Rs.6,455.26 for the sample. The coefficient of variation for Monthly Family Income was 55.56%. Income was categorised into 5 levels viz. 20000 & up = 1, 15000 & up = 2, 10000 & up = 3, 5000 & up = 4, less than 5000 = 5. Income as a variable was considered as a factor having 5 levels in the model.

Figure 3.3 Monthly Income of the SHG members in Mumbai
Figure 3.4 Pie Chart showing Income levels for the sample
3.2.4 Earning Members in a family (EarningM)

Minimum reported number of earning members in a family was zero and maximum reported number of earning members in a family was 7. The mean number of earning members was 1.701 and median number was 2. In absence of the accurate measurement of the economic status of the households in slums of Mumbai, it was believed that the number of earning members per household would be a proxy for the economic status of a family. It was expected that greater number of earning members in a household would have positive influence on the SHG member’s confidence of dealing with the potential risk of incurring a loss from a new business.
Figure 3.6 Occupational Choice according to Number of Earning Members in a family
3.2.5 Number of Family Members (NFM)

The average number of members in a family was 4.83. Median number of family members was 4. The smallest reported number of size of a family was 1 and the largest reported family size was 18. It was believed that greater number of members is associated with the greater family/other household responsibilities for women. Thus greater family size might have adverse impact on her decision to work/enter an occupation.

Figure 3.7 Sample distribution related to Number of Members per household
3.2.5 Number of Children below 8 years (Children)

The number of children below 8 years - this variable was reported to be between 0 and 8. Mean number of children below 8 years was 0.5192. Median number of children below the age of 8 was zero. Most of the women in the sample were over 40 years of age; therefore average number of young children was small for the sample. It was believed that this variable would have inverse relationship with the decision to choose self employment as one’s occupation. Greater number of younger children adversely affects the availability of women’s time for work.

![Line Graph showing number of children for the sample](image-url)

Figure 3.8 Line Graph showing number of children for the sample
3.2.6 Number of Old Members (OldM)

Minimum reported number of old members was zero and maximum number of old members in a family was 2. Mean number of old members was 0.2005 and the median value for the variable was zero. An old person in the households might put additional burden on women due to special attention and care required by the old person. Sometimes presence of an old person can help women in managing house & work responsibilities well. Thus effect of this variable on the decision to choose an occupation may not be definite.

Figure 3.9 Bar Chart showing number of old people per household for the sample
3.2.7 Number of dependent members due to illness/disability (Depe)

Mean number of dependent in a family was 0.06319 and median number of old members was zero. Minimum reported number was zero and maximum reported number was 1. This variable was considered to be inversely related to women’s choice of an occupation. Too many people in a family make more demands on her time. This reduces availability of time for work other than household chores.

![Number of dependent people in a household](image)

Figure 3. 10 Graph showing number of dependent people in a household
3.2.8. House owned by a household in Mumbai (House)

Most of the SHG members surveyed had owned house in Mumbai. The percentage of members owning house was 85.99. Only 14.01% of the members were living in rented house. Ownership of house in Mumbai was believed to be the indicator of relatively stable residential status of the member. Migratory character of the group has adverse impact on risk taking abilities of women.

Figure 3.11 Bar plot showing occupational choice by SHG members on the basis of ownership of house
3.2.9 Professional status of the spouse/parents (SEPFM)

In the sample 75.55% of SHG members had no self employed spouse/parent. Only 24.45% of the total members had self employed spouse/parent. Those who had self employed family members were found to be engaged in following businesses. Auto rickshaw operator, taxi operator, tourist car operator, tempo operator, water tanker operator, flour mill operator, Kirana shop owner, hardware shop owner, general stores owner, garage owner, Security Agency owner, Screen Printing & Designing, selling Vada Pav, Chinese Bhel, Sandwiches, Sprouts, glass bangles, roasted corns on streets, selling garlic in exchange for plastic/scrap, selling milk packets, ironing of clothes, pandal decorator, civil work contractor, Painting Contractor, Lawyer, Tailor, Electrician, Carpenter, Welder, glass fabrication business, vegetable vendor, Florist, selling onions, potatoes, peanuts etc. on cart. Presence of a self employed family member was believed to be important factor in promoting entrepreneurial endeavour by a person.

Figure 3.12 Bar Plot showing occupational choice on the basis of professional status of spouse/parents
3.2.10 Risk Attitude (RiskAtti)

Risk attitude was believed to be the pre requisite for choosing self employment over non self employment. To capture the risk attitude of the respondent, answer to the following question was consider as a proxy variable. To what extent do you agree with the statement: *One should not start a business if there is a risk it might fail?* Their responses were codes as 1 if they answer ‘strongly agree’ or ‘agree’ and 0 otherwise. Those with code 1 were considered risk averse and those with code 0 were believed to be risk lovers. Thus dummy variable was used with 'strongly agree' or 'agree'=1, and ‘can’t comment’ or 'disagree' or 'strongly disagree'=0.

Figure 3.13 Diagram showing response by the SHG members to a question related to risk attitude
3.2.11 Perception about the role of education and entrepreneurship training by SHPI (PERCe.t)

Respondents were asked the following question to get a response and then it was taken as a proxy for their perception about the role of education and entrepreneurship training in developing entrepreneurial attitude. To what extent do you agree with the statement: *My school education & entrepreneurship training by SHPI helped me to develop my sense of initiative (entrepreneurial attitude)*? Response was coded for 'strongly agree' or 'agree'=1, and for ‘can’t comment’ or 'disagree' or 'strongly disagree'=0. It was taken as dummy variable.

![Graph showing response by the SHG members to the question related to entrepreneurial attitude](image.png)

Figure 3.14 Graph showing response by the SHG members to the question related to entrepreneurial attitude
3.2.12 Perception of one’s own skills to implement and manage new business activity (PERCa.ms)

To include the respondent’s perception about their own skills to implement and manage new business activity, response to the following question was considered as a proxy variable. To what extent do you agree with the statement: *I believe that I have the necessary ability & the appropriate skills to successfully implement and manage new business activity?* Dummy Variable with 'strongly agree' or 'agree'=1, and ‘Can’t Comment’ or 'disagree' or 'strongly disagree'=0 was used in the model.

![Graph showing response by the SHG members to a question related to perception about own skills to implement and manage new business activity](image)

Figure 3.15 Graph showing response by the SHG members to a question related to perception about own skills to implement and manage new business activity
3.2.13 Perception of lack of Financial support (PERCfinance)

Financial support acts as an important factor influencing the decision to start new business activity. To gauge the perceived availability/unavailability of financial support, following question was asked to the respondents. To what extent do you agree with the statement: *It is difficult to start one's own business due to a lack of available financial support?* Response to this question was taken as dummy variable with 'strongly agree' or 'agree'=1, and ‘Can’t Comment’ or 'disagree' or 'strongly disagree'=0.

![Graph showing response by the SHG members to a question related to perceived availability/unavailability of financial support](image)

*Figure 3.16 Graph showing response by the SHG members to a question related to perceived availability/unavailability of financial support*
3.2.14 Perception of administrative Complexities (PERCadminpro)

The complexity or the simplicity of the administrative procedures to set up a new business activity does affect the effort of an individual to start new business. The perceived difficulty level can adversely affect the plan to start new business. To capture the respondent’s perceived image about the administrative procedures following question was asked. To what extent do you agree with the statement: *It is difficult to start one's own business due to the complex administrative procedures?* The response to this question was treated as dummy variable with 'strongly agree' or 'agree'=1, and ‘Can’t Comment’ or 'disagree' or 'strongly disagree'=0.

![Graph showing response by the SHG members to a question related to perceived image about ease/complexity of administrative procedures](image)

Figure 3.17 Graph showing response by the SHG members to a question related to perceived image about ease/complexity of administrative procedures
3.2.15 Perception of insufficient Information (PERCinform)

The appropriate information related to how to start a new business & other relevant facts-the new business opportunities, profile of the prospective customer, market segment, degree of competition, sources of finance etc. would facilitate decision to start business. Access to information (related to business activity) can be judged by the confidence of the respondent about ease to obtain information on how to start a business. The response to the following question was taken as a proxy variable to include perception of insufficient information in the model.

To what extent do you agree with the statement: *It is difficult to obtain sufficient information on how to start a business?* Dummy variable with 'strongly agree' or 'agree'=1, and ‘Can’t Comment’ or 'disagree' or 'strongly disagree'=0 was included in the model.

![Graph showing response by the SHG members to a question related to insufficient information related to new business](image-url)
3.2.15 Perception of infrastructural constraints (PERCinfra)

Access to infrastructural facilities - space, electricity, water, utensils, furniture, fuel etc. - goes long way in influencing decision to set up business. It also affects the cost structure of the business and thereby affects profitability & sustainability of business. To understand perception of infrastructural constraints following question was asked to the respondents.

To what extent do you agree with the statement: *It is difficult to have access to basic infrastructural facilities such as access to place, markets, power, fuel, equipments etc. to start a business?* The response to this question was treated as dummy variable with 'strongly agree' or 'agree'=1, and ‘Can’t Comment’ or 'disagree' or 'strongly disagree'=0.

![Figure 3.19 Graph showing response by the SHG members to a question related to perception of infrastructural constraints](image-url)
### 3.3 Assumptions of logistic regression model

Assumptions of the logistic regression model are as follows.

1. Binary logistic regression requires the dependent variable (occupational choice decision by the SHG member) to be binary. Here it was assumed that the SHG members have two options while choosing their occupation viz. choose to be self employed/ in business activity (SE) or not self employed/ in no business activity (No SE). They were observed to be SE or No SE at the time of field survey. All other possible levels of involvement in business activity of the member is classified as No SE if the member reports as not active in SE at that point of time when the field survey was conducted.

2. The categories (groups) viz, SE & No SE are assumed to be mutually exclusive and exhaustive; SHG member can only be in one group and every SHG member must be a member of one of the groups.

3. It was assumed that each observation to be independent. The independent variables (predictors) were assumed to be independent from each other. This was to ensure that the model should have little or no multicolinearity.

4. The logistic regression assumes linearity of independent variables and log odds. Whilst it does not require the dependent and independent variables to be related linearly, it requires that the independent variables are linearly related to the log odds. Otherwise the test underestimates the strength of the relationship and rejects the relationship too easily, that is being not significant (not rejecting the null hypothesis) where it should be significant. To solve this problem, most independent variables were categorized i. e. transforming most independent variables into ordinal level and then included in the model. Continuous variables such as Age in completed years, Earning Members (EarningM), No. of Family members (NFM), No. of Old members (OldM), No. of dependent members due to illness (Depe), no. of children are considered as discrete variables resulting into count data.
5. It is assumed that the dependent variable (occupational choice of the SHG members) is determined by the independent variables. These independent variables mainly explain variation in the dependent variable.

### 3.4 Analysis of the Logistic Regression Model Results

An extract of the summary of the model I for the logistic regressions in R shows the following result (Table 3.1) and the specification of the Multiple Logistic Regression model in R language is as follows.

(a) 1. Specifications (Model 1):

```r
glm(formula = Ochoice ~ Age + Incomelevel + EarningM + NFM + OldM + Depe + Children + SEPFM + House + PERCa.ms + PERCe.t +RiskAtti + PERCfinance + PERCinfra + PERCinform + PERCadminpro, family = binomial(link = "logit"), data = mydata)
```

2. Logistic Regression Output in R

Deviance Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1.7687</td>
<td>-0.9103</td>
<td>-0.5091</td>
<td>1.0588</td>
<td>2.0379</td>
</tr>
</tbody>
</table>
Table 3.1 Fitted Logistic Regression Model of Occupational Choice by the SHG member on Age, Income Level, Earning Member (EarningM), Number of Family Members (NFM), Old Members (OldM), Dependent members (Depe), No. of children (Children), Self Employed Partner/Father/Mother (SEPFM) etc. from Primary Survey, n= 364

| Predictor         | Estimate | Std. Error | z value | Pr(>|z|)    |
|-------------------|----------|------------|---------|------------|
| (Intercept)       | -6.35385 | 1.21611    | -5.225  | 0.000000174*** |
| Age               | 0.04613  | 0.01499    | 3.076   | 0.002095 **  |
| Incomelevel       | 0.14703  | 0.13486    | 1.090   | 0.275618    |
| EarningM          | 0.42395  | 0.19321    | 2.194   | 0.028218 *  |
| NFM               | -0.18946 | 0.09889    | -1.916  | 0.055395 .  |
| OldM              | -0.42630 | 0.28922    | -1.474  | 0.140494    |
| Depe              | 0.29918  | 0.55562    | 0.538   | 0.590253    |
| Children          | 0.12304  | 0.17883    | 0.688   | 0.491452    |
| SEPFM             | 0.63558  | 0.29141    | 2.181   | 0.029182 *  |
| House             | 0.68122  | 0.37689    | 1.807   | 0.070684 .  |
| PERCa.ms          | 1.67689  | 0.56433    | 2.971   | 0.002964 ** |
| PERCe.t           | 1.05341  | 0.29723    | 3.544   | 0.000394 ***|
| RiskAtti          | -0.25905 | 0.29246    | -0.886  | 0.375753    |
| PERCfinance       | 0.50077  | 0.28806    | 1.738   | 0.082141 .  |
| PERCinfra         | 0.77652  | 0.35668    | 2.177   | 0.029475 *  |
| PERCinform        | 0.01592  | 0.27000    | 0.059   | 0.952981    |
| PERCadminpro      | -0.55518 | 0.27629    | -2.009  | 0.044495 *  |

Signif. codes: 0 ‘****’ 0.001 ‘***’ 0.01 ‘**’ 0.05 ‘.’ 0.1 ‘ ’ 1
Dispersion parameter for binomial family taken to be 1
Null deviance: 472.08  on 363  degrees of freedom
Residual deviance: 395.78  on 347  degrees of freedom
AIC: 429.78
Number of Fisher Scoring iterations: 5
1-pchisq(472.08,363) = 0.00009727919
1-pchisq(395.78,347) = 0.03621475

Source- Logistic Regression Model results in R for the primary survey data
Following predictors (explanatory variables) viz. Income Level of the SHG member, number of old people, number of dependent people (due to chronic disease/illness/old age), number of children below the age of 8 years, risk attitude of the SHG member and perception of the SHG member about the ease of access to information related to how to start new business activity, have emerged as statistically insignificant variables for this sample in the first specification of the Model. It was decided to drop those statistically insignificant predictors & obtain new estimates. This process was repeated till we found the model having better fit. Eventually some more variables were dropped viz. ownership of House, Perception of lack of Financial support (PERCfinance).

An extract of the summary of the model II for the logistic regression in R shows the following result (Table 3.2) and the specification of the Multiple Logistic Regression model in R language is as follows.

(b) 1. Specifications : (Model 2)

```
glm(formula = Ochoice ~ Age + Age.2 + EarningM + NFM + SEPFM + PERCa.ms + PERCe.t + PERCinfra + PERCadminpro, family = binomial(link = "logit"), data = mydata)
```

2. Logistic Regression Output in R

Deviance Residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.8643</td>
<td>-0.9104</td>
<td>-0.5123</td>
<td>1.0940</td>
<td>2.0726</td>
</tr>
</tbody>
</table>
Table 3.2 Fitted Logistic Regression Model of Occupational Choice by the SHG member(Ochoice) on Age, Square of Age (Age..2), EarningMember(EarningM), Number of Family Members (NFM), Self Employed Partner/Father/Mother (SEPFM) etc. from Primarty Survey, n= 364

| Predictor       | Estimate  | Std. Error | z value | Pr(>|z|) |
|-----------------|-----------|------------|---------|---------|
| (Intercept)     | -1.388961 | 2.120378   | -0.655  | 0.512433|
| Age             | -0.153664 | 0.099691   | -1.541  | 0.123219|
| Age.2           | 0.002358  | 0.001183   | 1.994   | 0.046116 *|
| EarningM        | 0.365657  | 0.165773   | 2.206   | 0.027400 *|
| NFM             | -0.164314 | 0.080216   | -2.048  | 0.040521 *|
| SEPFM           | 0.676138  | 0.288118   | 2.347   | 0.018939 *|
| PERCa.ms        | 1.929971  | 0.561289   | 3.438   | 0.000585 ***|
| PERCe.t         | 1.117255  | 0.286507   | 3.900   | 0.0000964***|
| PERCinfra       | 0.825065  | 0.353443   | 2.334   | 0.019577 *|
| PERCadminpro    | -0.484224 | 0.255557   | -1.895  | 0.058122 .|

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 1

(Dispersion parameter for binomial family taken to be 1)
Null deviance: 472.08 on 363 degrees of freedom
Residual deviance: 402.08 on 354 degrees of freedom
AIC: 422.08

Number of Fisher Scoring iterations: 5
1-pchisq(472.08,363) = 0.00009727912
1-pchisq(402.08,354) = 0.0395559
1-pchisq(472.08-402.08,363-354) = 0.00000000001522693

Source- Logistic Regression Model results in R for the primary survey data
In logistic regression the coefficients derived from the model (e.g., \( \beta_j \), j = 0, 1, 2, ……9 indicate the change in the expected log odds relative to a one unit change in \( x_1 \), holding all other predictors constant. Therefore, the antilog of an estimated regression coefficient, \( \exp(\beta) \), produces an odds ratio. The test statistics to assess the significance of the regression parameters in logistic regression analysis are based on chi-square statistics. This is because the estimation technique, called maximum likelihood estimation, is used to estimate the regression parameters. The sign of the log-odds ratio indicates the direction of its relationship: + means a positive relationship between \( x \) (predictor) and the likelihood of a success, and - means a negative relationship. In order to get an intuitive sense of how much things are changing, we need to get the exponential of the log-odds ratio, which gives us the odds ratio itself.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate (Log Score)</th>
<th>Odds Ratio = ( \exp(\log \text{score}) )</th>
<th>Probability = ( \exp(\beta x)/1+\exp(\beta x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>( \beta_0 = -1.388961 )</td>
<td>( 0.249334228 )</td>
<td>( 0.199573679 )</td>
</tr>
<tr>
<td>Age (( x_1 ))</td>
<td>( \beta_1 = -0.153664 )</td>
<td>( 0.857560113 )</td>
<td>( 0.461659414 )</td>
</tr>
<tr>
<td>Age..2 (( x_2 ))</td>
<td>( \beta_2 = 0.002358 )</td>
<td>( 1.002360782 )</td>
<td>( 0.5005895 )</td>
</tr>
<tr>
<td>EarningM (( x_3 ))</td>
<td>( \beta_3 = 0.365657 )</td>
<td>( 1.441460737 )</td>
<td>( 0.590409141 )</td>
</tr>
<tr>
<td>NFM (( x_4 ))</td>
<td>( \beta_4 = -0.164314 )</td>
<td>( 0.848475559 )</td>
<td>( 0.459013675 )</td>
</tr>
<tr>
<td>SEPFM (( x_5 ))</td>
<td>( \beta_5 = 0.676138 )</td>
<td>( 1.966269318 )</td>
<td>( 0.662876195 )</td>
</tr>
<tr>
<td>PERCa.ms (( x_6 ))</td>
<td>( \beta_6 = 1.929971 )</td>
<td>( 6.889310449 )</td>
<td>( 0.87324621 )</td>
</tr>
<tr>
<td>PERCe.t (( x_7 ))</td>
<td>( \beta_7 = 1.117255 )</td>
<td>( 3.056452715 )</td>
<td>( 0.753479192 )</td>
</tr>
<tr>
<td>PERCinfra (( x_8 ))</td>
<td>( \beta_8 = 0.825065 )</td>
<td>( 2.282029092 )</td>
<td>( 0.69531044 )</td>
</tr>
<tr>
<td>PERCadminpro(( x_9 ))</td>
<td>( \beta_9 = -0.484224 )</td>
<td>( 0.616175163 )</td>
<td>( 0.381255187 )</td>
</tr>
</tbody>
</table>

Source – Primary Survey Data
A Receiver Operating Characteristic Curve (ROC) is a plot of sensitivity (the ability of the model to predict an event correctly) versus 1-specificity for the possible cut-off classification probability values $\pi_0$. ROC summarizes the predictive power for all possible $\pi_0$. The position of the ROC on the graph reflects the accuracy of the diagnostic test. It covers all possible thresholds (cut-off points). The ROC of random guessing lies on the diagonal line. The ROC of a perfect diagnostic technique is a point at the upper left corner of the graph, where the True Positive (TP) proportion is 1.0 and the False positive (FP) proportion is 0.

The Area Under the Curve (AUC) also referred to as index of accuracy (A), is an accepted traditional performance metric for a ROC curve. The higher the area under
the curve the better prediction power the model has. AUC = 0.743 can be interpreted to mean that a randomly selected individual from the positive group has a test value larger than that for a randomly chosen individual from the negative group 74.3 percent of the time.

### 3.5 Interpretation of log scores for self employment as one’s occupational choice

The multivariate logistic regression model results in R indicate that the SHG members choosing self employment as one’s occupation is a less likely event given all other predictors. Negative log odds of -1.39 indicate that the SHG members in Mumbai choosing self employment/business activity as one’s occupation is less likely/rare event. The odds in favour of choosing self employment as one’s occupation are only 0.25 vis-a-vis no self employment. Thus the chances of SHG members choosing self employment as one’s occupation are only 19.96% as compared to no self employment, given all other predictors. This result is statistically insignificant.

### 3.5.1 Perception about one’s administrative abilities, managerial skills & perceived role of formal education & entrepreneurial training: as determinants of Self Employment

Two very significant predictors which explains the choice of an occupation by the SHG members in Mumbai are Perception of one’s own skills to implement and manage new business activity (PERCa.ms) & Perception about the role of education and entrepreneurship training by SHPI(PERCe.t) (level of significance =0.001). Verheul, Uhlaner, & Thurik (2002) argues that the relationship between entrepreneurial activity and perception has been studied mainly from the viewpoint that perception influences entrepreneurial activity. In that context it has been argued that individuals make career choices based upon their perception of and the associated fit with a certain profession. In search of an individual characteristic that is distinctively entrepreneurial. Chen, Greene, & Crick (1998), proposed an
entrepreneurial self-efficacy construct (ESE) to predict the likelihood of an individual being an entrepreneur. ESE refers to the strength of a person’s belief that he or she is capable of successfully performing the various roles and tasks of entrepreneurship. It consists of five factors: marketing, innovation, management, risk-taking, and financial control (Chen, Greene, & Crick, 1998). The choice to engage in entrepreneurial activity is dependent upon whether individuals can identify with the characteristics and behaviours that are associated with entrepreneurship (Verheul, Uhlner, & Thurik, 2002). Verheul, Thurik, & Grilo (2006) mentions the fact that women feel that they lack the appropriate skills and knowledge for self-employment may reflect a lower entrepreneurial self-perception of women, where women are less optimistic and have less confidence in their own capabilities than men (Niederle and Vesterlund, 2006; Verheul, 2005).

Boz & Ergeneli (2013) cites the work of Boyd & Vozikis (1994) to highlight the importance of self-efficacy in entrepreneurship. Boz & Ergeneli (2013) states that in the context of self-employment, a person chooses to become self-employed because he or she feels confident in that area; this confidence leads to an interest in self-employment. The person’s confidence is shaped by his/her contextual factors such as ethnicity, age, gender, support system and past learning experiences. Perceived self-efficacy has been proposed as a central concept in entrepreneurship (Boyd & Vozikis, 1994).

Thus it was expected that those SHG members who perceived themselves confident of managing new business activity well, would likely to choose self employment as one’s occupation. The multivariate logistic regression model results confirmed this belief. The odds of choosing Self Employment as an occupation by the SHG members are 6.89 times higher among those members who perceive that they possess the necessary skills to implement and manage new business activity as compared to those who do not have similar perception.

Similarly the odds of choosing Self Employment as an occupation by the member of SHG in Mumbai are 3.06 times higher for those members who strongly perceive that their school education & entrepreneurship training (by SHPI) helped them in
developing a sense of initiative, as compared to others who do not have similar perception.

As stated by Block, Hoogerheide, & Thurik (2010), education is argued to be an important driver of the decision to start a business, higher the respondent’s level of education, the greater the likelihood that he/she starts a business. Per & Honig (2003) states that Human capital is not only the result of formal education, but includes experience and practical learning that takes place on the job, as well as non-formal education, such as specific training courses that are not a part of traditional formal educational structures. Thus it was believed that those who perceive formal education & practical training related to business activity as the important contributors in the development of sense of initiative would choose self employment as their occupation.

3.5.2 Perception about infrastructural constraints (PERCinfra): as a determinant of self employment

The odds of choosing self employment are 2.28 times higher for those SHG members who feel that it is difficult to have access to basic infrastructural facilities such as access to place, markets, power, fuel, equipments etc. to start a business, as compared to those who believe otherwise. This may be because those who are actively involved in business are experiencing these problems on a day to day basis. These economically active members perceive this as a major problem in starting a new business or expanding existing business. This result is statistically significant at 0.05 level.

3.5.3 Self Employed Partner/Parent: as a determinant of self employment

Gurley-Calvez, Harper, & Biehl (2009) cited the findings of several studies which have shown that parental or spousal experience with self-employment has a strong effect on entry into self-employment. Boz & Ergeneli (2013) argues that individuals
whose parents are entrepreneurs are likely to perceive that entrepreneurship is a feasible endeavour, thus triggering the development of their entrepreneurial intentions.

The log odds of choosing Self Employment as an occupation among the SHG members are 0.68 times higher in members who have Self Employed Partner/Father/Mother as compared to those who do not have Self Employed Partner/Father/Mother as a member of their family. If we take the antilog of the regression coefficient, \( \exp (0.676138) = 1.97 \), we get the crude or unadjusted odds ratio. The odds of choosing Self Employment as an occupation by the SHG member are 1.97 times higher among those persons who have Self Employed Partner/Father/Mother as compared to those who do not have Self Employed Partner/Father/Mother. The association between having a Self Employed family member and incident of choosing Self Employment as one’s occupation is statistically significant (0.05 significance level).

3.5.4 Number of Earning Members in a family (EarningM): as a determinant of self employment

The odds of choosing self employment as one’s occupation are 1.44 times higher for the members who have relatively more earning members as compared to those SHG members who have relatively lesser earning members in a household (significant at 0.05 level).

3.5.5 Number of Family Members: as a determinant of Self Employment

Nasir (2005), states that the occupational choice of men is not being substantially affected by family responsibility as marriage and presence of children in the family are not having any statistically significant impact on the occupational choice. However for females, marriage and children in the family play important role in the selection of occupation. For men, the family related factors may have an important role for taking part in the labour market activities rather than in the choice of a
particular occupation. Whereas for women, these factors may not have much relevance at the time of entry in the labour market but become important when choosing a particular occupation. All the respondents interviewed for this (pilot) study were women. Thus it was expected that the higher the number of family members in a household lesser would be the chances of SHG member (woman) choosing self employment as her occupation. Logistic regression model results confirm this belief.

Negative log odds of -0.16 indicated that the odds of choosing self employment are inversely related to the number of family members in a household (NFM). Higher is the number of family members in a household, less likely are the odds in favour of the SHG member choosing self employment as one’s occupation. It indicates inverse relationship between the log odds of self employment & number of family members (statistically significant at 0.05 level). Greater size of the family puts greater burden on women due to additional responsibilities of household chores. More family members will make more demands on her time; and thus will act as a constraint for choosing self employment as one’s occupation.

3.5.6 Perception of administrative Complexities (PERCadminpro): as a determinant of Self Employment

Verheul, Stel, & Thurik (2004) in a study based on Global Entrepreneurship Monitor\textsuperscript{47} data for 2002, mentions that business licensing may pose more problems for female entrepreneurs as compared to male entrepreneurs. This is because women usually have less previous experience with starting up a business, they may have less knowledge of government legislation and how to comply with it, posing particular problem or even discouraging them to start a business (Verheul, Stel, & Thurik, 2004).

\textsuperscript{47} The Global Entrepreneurship Monitor (GEM) project is an annual assessment of the entrepreneurial activity, aspirations and attitudes of individuals across a wide range of countries. The GEM is a global study conducted by a consortium of universities. Started in 1999, it aims to analyze the level of entrepreneurship occurring in a wide basket of countries.
Negative log odds of -0.48 indicates that the odds of choosing self employment are less likely for those SHG members who consider the administrative procedure to start business activity in Mumbai very complex/difficult, as compare to those who perceive this procedure not very complex. Those who have actually tried to start business activity in their life at any time in the past have already experienced the hardships/complexities of the procedure, which at present is not perceived by them as difficult or complex. Most of the members who were involved in some business activity find this procedure difficult, at that point of time but once it is complete, they do not perceive it as that difficult. This negative association between the member’s perception about administrative problems as complex & choice of self employment is statistically significant at 0.1 level.

3.5.7 Age as a predictor to determine choice of an occupation

In the earlier specifications, the estimation results suggested that as age increases, the likelihood to choose self Employment as an occupation also increases. As the log odds were positive. This result was statistically significant at 0.01 level. Regarding this choice, however, it was observed that older members are less likely to take risk to start a new business activity. Both individual level and regional level studies have found an inverse U-shaped relationship between the age and start-up activity (Monsen, Mahagaonkar, & Dienes, 2012). Bönte, Falck, & Heblich (2009) in their empirical analysis found an inverse U-shaped relationship between age structure and start-up activity in a region at the group aged 40 and older. They have also found that entrepreneurs are more likely to start a business locally – in their own social environment. Regionally bound peer effects increase the likelihood of recognising an opportunity or by facilitating access to resources once such an opportunity has been seized. Rees & Shah (1986) considered age, as an important determinant of employment status of an individual. Its effect is non-linear, rising first and declining towards the end of life.

To capture this inverted U-shaped relationship, when a quadratic term for age in the estimations is included, the effects of the age are insignificant. However, Age^2 is statistically significant at 0.05 level. Odds ratio for age^2 is 1:1, i.e. there is 50%
chance that the SHG member would choose self employment as her occupation with unit increase in Age square.

3.6 Conclusion & Policy Implications

Thus, the perception of one’s own skills to implement and manage new business activity (PERCa.ms) & the perception about the role of education and entrepreneurship training (PERCe.t), significantly affect the occupational choice of the SHG members in Mumbai.

By citing work of Harper (1998) & Baumol (1990), Langowitz & Minniti (2007) states that the nature of political and economic institutions influences individuals’ perceptions. Institutional arrangements affect the quantity and type of entrepreneurial efforts. This is particularly important for women since perceptions of themselves and their environment play a greater role in their behaviour. Thus, to alter the way in which individuals think about themselves and their role in society may require very localized and specific interventions (Langowitz & Minniti, 2007). Langowitz & Minniti (2007) suggest that although risk tolerance plays some role in gender differences, the main difference is in the way in which men and women perceive themselves and their environment. Noticeably, perceptions and risk tolerance are both characteristics of the individual. They cannot be easily changed by exogenous actions such as, e.g., government intervention. Langowitz & Minniti’s (2007) empirical findings coax them to suggest that while policy can alter an individual’s incentives, the cultural factors that mould perceptions and risk profiles depend on the specific history of the place. They are path dependent and, as a result, only change slowly over time (Langowitz & Minniti, 2007).

Warnecke (2013), mentions that it is crucial to be sensitive to local cultures and to find ways to work within the context of local culture to support social change. Only then will initiatives (and the priorities underlying the initiatives) gain enough traction and popular support to become sustainable over time. Government support to promote SHGs and development of entrepreneurial spirit in Mumbai needs to be
viewed in the context of the results of this study. Warnecke (2014) presents an exploratory framework for considering five key characteristics of female entrepreneurship programmes which includes provision of money, goods, or services; skills training; business development services; professional networking services; and/or professional mentoring. Warnecke (2013) highlighting the role of the institutions in the entrepreneurship development & growth states that by shaping the likelihood that individuals will gravitate towards opportunity or necessity entrepreneurship, institutions shape the linkage between entrepreneurial activity, the fulfilment of human capabilities, and the quantity, quality, and distribution of economic growth. What needs to be emphasized, however, is that institutions influence these social relations. Institutions make certain forms of interaction harder or easier, but not to the same degree for everyone (Warnecke, 2013).

Results of this study suggest that public policy to generate self employment for women in city like Mumbai should focus on the institutional arrangements which would influence perception about entrepreneurship & help women in developing entrepreneurial attitude in future. The study also highlights that economically active SHG members strongly believes in the importance of formal education & hands on training programmes to equip women to take up challenges of new business activity. SHG members would have optimistic attitude towards self employment if complemented with relatively simpler administrative procedures to start new business. Local governing body has to take initiative to address the major infrastructural constraint faced by SHG members in Mumbai i.e. unavailability of space & other amenities to start new business activity. Thus government through appropriate policy intervention can utilise hitherto untapped manpower. Self-employment would be a ‘sponge’ providing subsistence for those who could not find better alternatives, and it can also absorb unproductive manpower (Pietrobelli, Rabellotti, & Aquilina, 2004).

Milla´n, Congregado, & Roma´n (2012) write that if the objective is the promotion of long-term successful self-employment, the prescription should not only be to facilitate entry by means of subsidies or guaranties, but also to support the acquisition of the necessary entrepreneurial human capital and to facilitate growth aspirations.