Chapter - 04

Methodology and Data Evaluation

4.1 Introduction

This chapter explains the methods used to operationalise the framework which was discussed in the previous chapter. The factors for this study considered are level of e-banking effectiveness (dependent variable - LEE-CP/BP) and independent variables such as customer attitude towards e-banking (ATE), scope of e-banking applications (SEA), level of e-banking service quality (ESQ), security of e-banking (SEB), profitability of e-banking (PEB), considering e-banking as a strategic tool (EST), level of customer relationships management (CRM) and operational efficiency of e-banking (EOE). In addition it gives an overview of sampling and data collection procedures. Further, it discusses how raw data are summarized to a meaningful set for the convenience of handling analyses and interpretation. In the study bivariate and multivariate analyses are used to find the relationships among independent variables and level of e-banking effectiveness (dependent variable).

4.2 Overview of Methodology

Data are collected from customers of seven banks and employees of seven banks through questionnaire survey, delivered by hand, normal post and e-mail. The sampling method is purposive sampling and it enables us to use judgement to select cases that will best enable us to answer our research question and to meet research objectives.

The method of data evaluation adopted depended on the number of variables involved. In this study there were eight independent variables that positively influence the dependent variable, “Perceived E-banking Effectiveness”.

Each independent variable comprised of one (1) to three (3) dimensions. The dependent variable or the level of e-banking effectiveness comprised of five (05) dimensions. This

means that aggregate score denoted the degree to which the key respondents (customers and bank employees) agreed or disagreed with independent variables and dependent variable.

The correlation analysis is used to quantify the relationship (if any) between the independent and dependent variables. The purpose of correlation analysis is to measure the strength of the relationship between two variables. The correlation coefficient cannot be greater than 1 or less than -1. As defined, correlation is a number between +1 and -1 that reflects the degree to which two variables have a linear relationship.

To find out about the relationship between and independent variable and a dependent variable a correlation analysis (bivariate analysis) is performed. The relationship between dependent variable and two or more independent variables (multivariate) can be studied through multiple regression analysis.

A correlation coefficient, (r) for the data values indicates how strongly two variables are related. A partial correlation coefficient, (r) is computed between each independent variable and the dependent variable at the significant level of 1 percent and these are compared to see which one enables the most accurate predictions of scores on the dependent variable, the perceived level of e-banking effectiveness.

Multiple regression is used as a descriptive tool in this research study. It is often used to develop a self – weighting estimating equation by which to predict values for a dependent variable from the values for several independent variables. Thus, we might try to predict level of e-banking effectiveness on the basis of customer attitude towards e-banking (ATE), scope of e-banking applications (SEA), level of e-banking service quality (ESQ), security of e-banking (SEB), profitability of e-banking (PEB), considering e-banking as a strategic tool (EST), level of customer relationships management (CRM) and operational efficiency of e-banking (EOE).

Regression coefficient (b) enables prediction of respondent's score on a dependent variable, whereas correlation coefficient (r) enables assessment of the accuracy of those predictions.

In this study, relative impact of a set of independent variables on the dependent variable has to be assessed. SPSS computed the model and the regression coefficient. Most statistical packages provide various methods for selecting variables for the equation. The equation can be built with all variables or specific combinations or we can select a method that sequentially adds or removes variables (forward selection, backward elimination, and stepwise selection). Forward selection starts with the constant and adds variables that result in the largest R^2 increases. Backward elimination beings with a model containing all independent variables and removes the variables than changes R^2 the least. Stepwise selection, the most popular method, combines forward and backward sequential approaches. The independent variable and that contributes the most to explaining the dependent variable is added first. Subsequent variables are included based on their incremental contribution over the first variable and whether they meet the criterion for entering the equation (e.g., a significance level of .01). Variables may be removed at each step if they meet the removal criterion, which is a larger significance level than for entry.

The stepwise solution is a variation of the forward solution. A serious limitation of the latter is that variables entered into the equation are retained despite the fact that they may have lost their usefulness in the light of the contributions made by variables entered at later stage. The stepwise procedure is designed in such a way that statistical tests are performed at each step to determine the contribution of each variable already in the equation if it were to enter last. It is thus possible to discard a variable that was initially a good explanatory variable. Stepwise solution of the multiple regression procedure is designed to have in some blending the advantages of both of the procedures: backward as well as forward.

4.3 The Sample Size, Selection Procedure and Data Collection

The source of data for this research is primary data. This study is conducted among the e-banking customers and banks employees who are mostly involved in the e-banking operations, using a survey questionnaire, which contained structured and semi-structured questions. Appendix 1 and appendix 2 contain question for the eight independent variables and one dependable variable. These questionnaires were delivered by hand, normal post and e-mailed.

The survey was conducted among a sample of 300 e-banking customers and 70 employees of banks that provide e-banking in Sri Lanka. For customers, these questionnaires were distributed based on level of e-banking activities of selected banks. These questionnaires were distributed among selected bank employees who are highly involved in e-banking operations in their banks such that each bank gets an equal number of questionnaires.

The sample of e-banking customers to be included in for data collection is tabulated below in the Table 4.1 based on level of e-banking activities of selected banks.

Table 4.1 Sample Size for E-banking Customers

Banks	Number of Respondents	Percentage (%)
Bank - 1	74	25
Bank - 2	61	20
Bank - 3	47	16
Bank - 4	31	10
Bank - 5	40	13
Bank - 6	26	9
Bank - 7	21	7
Total (Sample Size)	**300**	**100**

The sample of bank employees to be included in for data collection is tabulated below in the Table 4.2 as follows:

Table 4.2 Sample Size-Bank employees

Banks	Number of Respondents
Bank - 1	10
Bank - 2	10
Bank - 3	10
Bank - 4	10
Bank - 5	10
Bank - 6	10
Bank - 7	10
Total (Sample Size)	**70**

From the each selected bank, 10 key respondents were selected based on their involvement in e-banking operations for this research purpose.

4.4 Finalization of the Questionnaires

4.4.1 Questionnaire – Customer's Perspective

The questionnaire was prepared in the following manner. Table 4.3 shows it clearly.

Table 4.3 Questionnaire – Customer's Perspective

Part – I This is about e-banking customer's personal profile such as, gender, age, level of education, occupation and monthly income.
Part – II This is about customer's e-banking information such as, the number of banks they have accounts, e-banking services they are using, types of channel they use, etc.

Part – III

This is about customer's perspective on e-banking

No	Key Concept	What is measured?	Indicator	Related questions in the questionnaire
01	Customer attitude towards e-banking (ATE)	What kinds of attitudes customers have about e-banking?	Trust	01, 02, 03
			Customer acceptance	04, 05, 06, 07
			Promotional activities of e-banking by banks	08, 09, 10
02	Scope of e-banking applications (SEA)	Degree to which e-banking applications fulfill customers' banking requirements	Comprehensiveness of e-banking services/ Range of services	11, 12
03	Level of e-banking service quality (ESQ)	Degree of quality of services	Efficiency	13, 14
			Reliability	15, 16, 17
			Responsiveness	18, 19, 20
04	Security of e-banking (SEB)	Degree of security	Privacy	21, 22, 23,
			Integrity	24
			Authentication	25, 26, 27
05	Level of e-banking effectiveness – Customer's Perspective (LEE-CP)	Degree of e-banking effectiveness	Overall satisfaction, cost and time savings, usefulness, range of services and security and reliability of e-banking	28, 29, 30, 31, 32

4.4.2 Questionnaire – Bank's Perspective

The questionnaire was prepared in the following manner. Table 4.4 shows it clearly.

Table 4.4 Questionnaire – Bank's Perspective

Part – I This is about bank information such as, URL, year they launched web site and e-banking, main reasons for introducing e-banking, types of channel used by the customers, e-banking services they are providing, information about customers who use e-banking etc.				
Part – II This is about bank's perspective on e-banking				
No	**Key Concept**	**What is measures**	**Indicator**	**Related questions in the questionnaire**
01	Profitability of e-banking (PEB)	Degree of revenue growth through e-banking	Revenue Growth	01, 02, 03, 04, 05
02	Considering e-banking as a strategic tool (EST)	Degree to which considering e-banking as a strategic tool	Competitive Strategy	06, 07, 08,
			Distribution Channel Strategy	09, 10 11, 12, 13
03	Level of Customer relationship management (CRM)	Strength of customer relationship by usage of e-banking	Interaction	14, 15, 16,17
			Accessibility	18, 19, 20, 21
04	Operational efficiency of e-banking (EOE)	Degree of operational efficiency	Reduce Costs	22, 23
			Increase productivity and speed	24, 25, 26
05	Level of e-banking effectiveness – Bank's Perspective (LEE-BP)	Degree of e-banking effectiveness	Achieving overall objectives of bank (short-term and long-term), cost and time savings, usefulness, strategic importance and better customer relationships	27, 28, 29, 30, 31, 32, 33

4.5 Method of Data Evaluation

The method of data evaluation adopted depends on the number of variables involved. In this study there were 08 independent variables that positively influenced the dependent variable, 'level of e-banking effectiveness'. Therefore this study resulted in a large volume of raw data for each variable are summarized by means and standard deviation, which must be suitably reduced, so that those can be handled easily and can be used for further analysis.

4.5.1 Univariate Analysis

This analysis is performed on individual variables. Descriptive statistics is used to analyze each variable. Frequency distribution is used to count how many informants gave particular answers to each question on e-banking information.

'Percentages' is used to determine the fraction of informants who gave a particular response. The percentage figure gives a very good picture of the distribution of the variable. These frequency plots in presented as tables and pie charts.

Independent and dependent variables are measured using a 5-point scale. Regarding to the independent variables, the scale '1' corresponds to 'strongly disagree' and scale '5' related to 'strongly agree'. The 'mean score' can be considered as '3'. This is the margin that determines whether the respondent agrees or not.

Regarding the dependent variables, the scale '1' corresponds to 'very low' and scale '5' related to 'very high'. The 'mean score' can be considered as '3'. This is the margin that determines whether the respondent agrees or not.

4.5.2 Bivariate Analysis

The importance of bivariate analysis was to see whether two variables were associated. Two variables were said to be associated or related when the distribution of values on one

variable differed for different values of the other. This analysis answers the following questions

1. Did there exist association or correlation between the two variables? If yes, to what degree?
2. Was there a relationship between an independent variable and a dependent variable?

In a nutshell, the first was answered by the use of correlation technique and the second question by the technique of regression.

Pearson's Correlation (r) (Cooper – 2003): This provides a measure of index for the strength and direction of any linear relationship between two variables. The value of 'r' lies between -1 and +1. Positive value of 'r' indicates positive correlation, whereas negative values of 'r' indicate negative correlation. The value of 'r' nearer to +1 or -1 indicates high degree of correlation between two variables. That means when the points are more concentrated, the value of r is higher. Correlation coefficients between + or - .81 and 1.00 are considered very strong, between – or + .61 and 0.80 strong, between – or + .41 and .60 moderate and less than this as weak.

4.5.3 Multivariate Analysis

The relationship between the dependent variable and two or more independent variables (Multivariate) can be studied through multiple regression equation whereas correlation of multivariate population can be studied through multiple correlation techniques.

Multiple regression works on the principle that the more one knows about a person, the more accurately he can guess other attributes of that person. Here a whole set of information would be used to predict the dependent variable and it addresses the combined effect of many independent variables on dependent variable.

4.6 Hypothesis Testing

Hypothesis testing is the strategy for deciding whether a sample data offer such a support for a hypothesis that generalization can be made. Thus, hypothesis testing enables making probability statement about population parameters(s). In this process, it is standard to begin by assuming that in the population any given two variables are not associated (i.e. $r_p = 0$) and that correlation (.r_p) between perceived level of e-banking effectiveness and attitude towards e-banking is zero. This is a null hypothesis (H_0: r_p=0). If in the sample for example, observed association between the two variables is 0.3, then the discrepancy between the assumption of no association and the sample observation of association can be interpreted both as:

1. There is a sampling error (i.e. sample is unrepresentative despite the purposive sampling technique adopted)
2. Assumption of no association in the population is incorrect.

As we are interested in the population rather than the sample per se, and accept the first interpretation that the sample result is due to sampling error, then we maintain the assumption of no association in the population.

If on the other hand, accepting the second interpretation that the initial assumption of no association in the population is incorrect, then it leads to interpreting sample results as reflecting association that actually exists in the population. To work which interpretation is correct, a test of significance is used.

Symbolically we can express hypotheses as:

Null hypotheses H_0: $r_p = 0$

Alternative hypothesis H_a: $r_p \neq 0$

i.e. $r_p > 0$

$r_p < 0$

It is noted that H_a is of the type "whether greater or smaller". As there are one rejection regions, one tail distribution curve and therefore a one tailed test is used. Considering the purpose, nature of this study and the sample size, the significant level of 1% is adopted. Then it implies that H_0 will be rejected when the sampling result (i.e. observed evidence) has a less than 0.005 probability of occurring if H_0 is true. In other words, the 1 percent level of significance means the willingness to take as much as a 1 percent risk of rejecting the null hypothesis when it happens to be true. The null hypothesis could be tested as shown in section 4.6.1 below.

4.6.1 Correlation Analysis

Null hypothesis is defined as follows.

H_0: The two given variables are not associated ($r = 0$)

where r = correlation coefficient

The alternative hypothesis is defined as follows.

H_a: The two given variables are associated ($r \neq 0$)

If $r \neq 0$ where $p<0.05$, Reject H_0, Accept H_a.