Chapter 4: Research Design

CHAPTER 4: RESEARCH DESIGN

A research design is a structure that helps to find out the answer(s) to research question(s). The research design enumerates step by the step work plan, data collection procedures, and statistical analysis plans. The research design is aimed at achieving objective of the study. The primary objective of the study is to find out the relationship between the internal risks of the firm on the value of the company.

4.1 Data

The collection of data and its management is one of the most important aspects in research design. The generalization of observation is possible only when appropriate sample is collected.

4.1.1 Sample:

The sample in this study consists of 497 companies listed in the BSE-500 stock exchange index. In the study, 11 industries considered as per NIC 2004 (Table A). The study covers Sixteen years of company data from 2001 to 2017. The data are secondary in nature and collected from audited annual reports that were available at the Capitaline database. As part of data management, the sample is processed through the following ways: First, the firms that have abnormal data (visually inconsistent) are omitted. Secondly, a few observations that included high negative (over -1%) and positive values (over +1%) for one of those variables been omitted. Lastly, some missing value in the data field been filled up by the average value of the two corresponding years.

Table A: Broad Group of Company as Per NIC 2004

Section A: Agriculture, Hunting, And Forestry
Section B: Fishing

Section C: Mining And Quarrying
Section D: Manufacturing
Section E: Electricity, Gas And Water Supply
Section F: Construction
Section G: Wholesale And Retail Trade; Repair Of Motor Vehicles, Motorcycles
And Personal And Household Goods
Section H: Hotels And Restaurants
Section I : Transport , Storage And Communications
Section J: Financial Intermediation
Section K: Real Estate, Renting And Business Activities
Section L: Public Administration And Defense; Compulsory Social Security
Section M: Education
Section N: Health And Social Work
Section O: Other Community, Social, And Personal Service Activities
Section P: Undifferentiated Production Activities Of Private Households And
Activities Of Private Households As Employers
Section Q : Extra Territorial Organizations And Bodies

4.1.2 Variables Discussion: Based on the objective of the study a discussion on two types of variables are required (1) Business valuation variables and (2) Risk variables

Business valuation variables:

According to different previous work such as Fakhruddin and Hadianto (2001), Damodaran (2005), Tasker (1998), Beatty, et al. (1999), Baker and Ruback (1999), Kim and Ritter (1999), Kaplan and Ruback, (1989) it was found that price multiples can able to evaluate a firm value in better way. Therefore, following price multiple taken as dependent variables in the research work.

PE ratio, PBV, PCEPS, EVEBIDTA, Market-Cap-Sales.

Independent Variable:

According to Chung (1989), Opler and Titman (1994), McConnell *et al.* (1995), Fama and French (1998), Aggarwal *et al.* (2008), Mollik (2008) etc. risk variables are one of the determinants of firm value. The following variables are taken as independent variables to find out relation between risk and value in Indian context.

DE Ratio=Debt-equity ratio, OR= Operating risk, FR=Financial risk, Industry classification (dummy variables). X = Total period was bifurcated in two-part (before 2008-09 and after 2008-09).

4.2 Research Methods:

Descriptive Statistics & Statistical Testing: First of all correlation matrix of six dependent variables and five independent variables were calculated to find out a multi-co-linearity problem among the independent variables. Multicollinearity is a relationship among the independent variables in which there is an exact or perfect correlation between the predictor variables. When they present a perfect or exact association among the predictor variables, no reliable coefficients estimation of their individual is possible. It will bring in incorrect conclusions about the association between the dependent variable and independent variables, (Gujarati, 2004). The OLS estimated regression coefficients had affected by the multicollinearity problem. Some problems of multicollinearity are, high variance of coefficients which decrease the precision of estimation, wrong sign in estimated coefficients, it can inflate the estimated variance of predicted values, the parameter estimates and their standard errors, (Montgomery, 2001). Therefore, multicollinearity is a significant problem when we formulate some predictive models. Therefore, it is a very important task for us to find out a method that solves the multicollinearity problem. A number of different techniques had developed to solve the multicollinearity problem. The principal components method is one of the simplest methods among the different techniques (Naes and Indahl, 1998) to solve the said problem. The principal component analysis helps to solve the multicolinearity problem by identifying the most significant factors from the many independent variables. After solving the said problem Descriptive, analysis of the variables had carried out. To understand the basic features of the data sate, descriptive statistics analysis carried out. Descriptive Statistics help to present data in a manageable form by quantitative description. The large amounts of data can be expressible in a sensible simplified way with the help of descriptive statistics. They provide some of the measures and simple summaries about the data set. It helps in quantitative analysis of data sate. Its simple pictorial presentation helps to conduct graphic analysis of the data set. Descriptive statistics are different from inferential statistics. With descriptive statistics, we are simply describing what the data shows or what the data nature is. The inferential statistics helps to reach conclusions that extend beyond the present data set. For example, we use inferential statistics to estimate population characters depending on the sample data. We use descriptive statistics simply to describe what is going on in our data set. In descriptive statistics, a simpler summary of data set had done from lots of data set that helps to take quick understanding of the data feature. Every time we try to describe a large set of observations with a single indicator so that the risk of handling the original data or losing important detail had minimized. We often use financial time series data for firms' parameters measure. When the data set are time series data, one common character is the non-Stationarity of the data set. Generally, the assumption of Stationarity is not valid for real applications. Financial models with non-stationary time series data produce spurious and unreliable results, which lead to poor forecasting and understanding. So stationary is required to bring a good result. The non-stationary time series data become stationary if we transform them with the appropriate method. The differencing process of transformation helps to make data series stationary when the non-stationary data series is a random walk with or without a drift. On the other side, if the time series non-stationary data series exhibits a deterministic pattern trend, the stationary can be possible by de-trending the data series. When the non-stationary series combines the above two problem simultaneously, differencing and de-trending would apply, as de-trending could remove the deterministic trend and differencing could remove the trend in the variance. When time-series data become stationary, its statistical properties remain unchanged overtime period. We would not be able to use methods designed for stationary data series on non-stationary data series as it leads to completely misleading results. The question is how can we know whether a time series data is stationary or not? The stationarity of the data series can be checked through the unit root test. When data sate is panel data, the panel unit root test help to identify the said problem. The Levin, Lin and Chu; Im, Pesaran and Shin W-stat; ADF-Fisher Chi-square, and PP-Fisher Chi-square panel unit root test help to identify stationary of data series. After the Stationarity of the data series is confirmed, the ordinary least square (OLS) method were applied to find the explanatory power of independent variables on dependent variables. However, coefficients estimated from a simple OLS regression method bring misleading result if the data series contain autocorrelation problem. We can find out the autocorrelation or serial-correlation problem. The Durbin–Watson statistic (or Durbin's H statistic) helps to identify the autocorrelation problem in simple OLS regression results. However, the problem with it is that it is only applicable for testing the first-order autoregressive model (e.g. AR (1)) and for non-stochastic regression error. The serial correlation LM test of Breusch–Godfrey (BG) helps to solve the said problem. It finds out the autocorrelation problem by testing error terms in the regression model. The test is more general than the Durbin–Watson statistic test and it has no such restrictions but it is more statistically powerful than Durbin's H statistic. Whatever problem may be there, these two methods help the researcher to find out the autocorrelation problem in the model. According to the rule, DW value from 1.5 to 2.5, indicate no autocorrelation or serial-correlation in the data series. Values outside this range could be cause for concern. Field (2009) suggests that values under one or more than three are a definite cause for concern. The normality of the data series is also one important factor for applying the ordinary least squared method in any investigation. The normality of the data series can be tested by the Heteroskedasticity test of the regression residual under panel ordinary least square and robust model with the help of the Jerque-Bera statistic and Panel Cross-section Heteroskedasticity LR test. The outliner of regression residual also tested by using the Leverage plot and the Hat matrix ratio as Outliner helps to understand the pattern of Heteroskedasticity. The Robust-Least Square regression model brings good results when data are heteroskedastic and have more than two outliners. The long run or short-run cointegration some time creates problems to find out the impact of independent on dependent variables. It is because some time-dependent and independent variables automatically correlated with each other in a long time or short time without any valid causality. The cointegration test helps to detect the said aspect. Therefore, Pedroni and Johansen panel Cointegration test had used to find out the said problem. Pedroni residual cointegration test and Johansen panel Cointegration test are applicable when the variables had co-integrated in the same order. As there was no unit root in the level in any series, so they had integrated at the same level. If we have a long run or short run co-integration, but we run OLS, then the model will misinterpret the coefficient and the results may be biased. Therefore, to find out the applicability of OLS model the co-integration test had also conducted. The VEC model is applicable when the Johansen cointegration test result said that there is at least one co-integrated equation. The VEC model had used to find out the long run or short run cointegration between the dependent and independent variables. The Wald test used in the VEC model is to find out the short-run causality between the dependent and independent variables. Then the above hypothesis had tested first using linear regression models where unbalance panel data (comprised of the cross-section and time-series data) is used. In the current work, some assumption of OLS model does not match in OLS results so robust least square method and panel EGLS model had used sequentially to estimate the effect of the independent variable on the dependent vectors. However, after rigorous study of the result, it had found that instead of a linear regression model, polynomial model would bring a more consistent outcome as the Wald test on regression coefficient of categorical independent variables showed significant difference among the categorical variables. Lastly, after confirming different types of relationships categorically, the nonlinear model (polynomial) is applied. Hausman test had conducted to find out appropriate model applicability. To test the significance of the regression model, R^2 and adj.R^2 statistics are used and to test the applicability of the said regression model in the population, F or Rn-sq statistics had carried on. The probability percentage had used to confirm the significance of the results. Lastly, the Wald tests had done on different parts of the work to test whether the coefficients of the regression are equal or not.

4.2.2 Empirical Model:

Variables Discussion: There are two categories of independent variables in our panel data regression model. One is the natural independent variables and the other is the dummy independent variable. The overall debt-equity ratio (DE), financial leverage, operating leverage, industry category (dummy) and period (dummy) are taken in the first equation as the independent variables to assess the effect of firm-specific risks, industry nature and structural break on the firm value (FV). The effect of structural break accounted for in dummy variables 2008. The year 2009 is considered as structural break point as an economic, recession occurred at that time. The dummy for the structural break, for the unusual values, put one otherwise zero; this will give us a series of a single variable. I put it into the independent side of the empirical model. In econometrics and statistics, a structural break is an unexpected change over time in the parameters of regression models, which can lead to huge forecasting errors and unreliability of the model in general. If the dummy variables of structural break would be significant then OLS brings biased results. In the second, third and fourth models, categorical DE, FR, and OR are taken as dummy variables. These dummy variables represent categorical DE, FR and, OR of the firm. In the second, third and fourth models, it is assumed that the Debt-Equity ratio, FR(financial risk) and OR(operating risk) which represents patterns of capital structure configuration and company cost structure have a categorically different effect on the value of a firm (null hypothesis). Therefore, I have classified debt-equity ratio in eight types and treated them as dummy variables. As the independent variables are taken in, eighth categories in the second model so seven dummy variables taken to explain the effect of the dummy variables on the value of a firm. In second model eight dummy variables of debt-equity are taken which are DE zero here to D1DE, DE greater than 0 and less than equal 0.25 here to D2DE, DE greater than 0.25 and less than equal

0.50 here to D3DE, DE greater than 0.50 and less than equal 0.75 here to D4DE, DE greater than 0.75 and less than equal 1 here to D5DE, DE greater than 1 and less than equal 1.50 here to D6DE and DE greater than 1.5 and less than equal 2 here to D7DE and rest D8DE taken as categorical dummy variables. In third and fourth model, four dummy variables are taken which are, FR/OR greater than 1 and less than equal 1.50 here to D1FR, D1OR, FR/OR greater than 1.5 and less than equal two here to D2FR, D2OR and FR/OR greater than 2 here to D3FR, D3OR and rest D4FR, D4OR. The dependent variables are valuation ratios such as price to earnings ratio (PE), price to book value (PBV), price to cash EPS (PCEPS), economic value added by EBIDTA (EVEBIDTA) and market capitalization by sales (Market-Cap-Sales). The modulus values of risk factors had taken in the present work.

A linear regression model analyzed the data primarily and then after confirming the nature of the relationship the nonlinear model had applied. The measurements for the variables had displayed as under. The primary empirical models are as follows:

- **Linear Model to Describe the Effect of industry nature, structural break, DE, OR and FR on Firm Values (H_1)**

Valuation-atio $=C_1+ß_1*DE+ ß_2 *OR + ß_3 *FR + ß_j D_i +X+e_1$ --------------------- (1)

DE= Debt-Equity Ratio, OR== Operating Leverage, FR=Financial Leverage, ß= coefficient value, D_i = Dummy Veriables, X= Dummy Veriable for structural brake (assume 2009), e = error term, C_1 = intercept.

Valuation-ratio=

C_2+$ß_1$*D1DE+$ß_2$*D2DE+$ß_3$*D3DE+$ß_4$*D4DE+$ß_5$*D5DE+$ß_6$*D6DE+$ß_7$*D7DE+

$ß_8$*D8DE + e_2 --------------------------------- (2)

DE= Debt-Equity ratio, ß= Coefficient Value, D1DE, D2DE, D3DE …. D8DE = categorical dummy veriables of debt-equity ratio , e_2 = error term. C_2 = intercept.

Valuation ratio=C_3+$ß_1$ *D1FR+$ß_2$ *D2FR+$ß_3$ *D3FR +e_3 ----------------------------------- (3)

ß= coefficient value, D1FR, D2FR, D3FR = categorical dummy veriables of Financial Risk , e_3 = error term. C_3 = intercept.

Valuation ratio=C_4+$ß_1$ *D1OR+$ß_2$ *D2OR+$ß_3$ *D3OR +e_4 ----------------------------------- (4)

ß= coefficient value, D1OR, D2OR, D3OR = categorical dummy veriables of Operating Risk, e_4 = error term, C_4 = intercept.

- **Nonlinear Model to Describe the Effect of DE, OR, FR, Effect of industry nature, and structural brake on Firm Values(H_1)**

  $Y= C+ \beta_1 DE + \beta_2 DE^2 + \beta_3 DE^3 + \beta_4 DE^4 + \beta_5 OR + \beta_6 OR^2 + \beta_7 OR^3 + \beta_8 OR^4 + \beta_9 FR + \beta_{10} FR^2 + \beta_{11} FR^3 + \beta_i D_j + \beta_{22} *X + \beta_{23} \text{Lag}(Y) + e$ where i =12 to 21 and j= 2 to 11. --------------------- (5)

DE= Debt-Equity Ratio, OR= Operating Risk, FR=Financial Risk, ß= coefficient, D_j= dummy veriables for industry category, e = error term, C = intercept, X= dummy veriable for structural brake (assume 2009).

4.2.3 Hypothesis:

To examine impact of financial risk on value of the firm

H_{10}: Financial risk has no impact on the value of the firm.

H_{11}: Financial risk has impact on the value of the firm.

To examine impact of operating risk on value of the firm

H_{20}: Operating risk has no impact on the value of the firm.

H_{21}: Operating risk has impact on the value of the firm.

To examine impact of debt-equity on value of the firm

H_{30}: Debt-Equity has no impact on the value of the firm.

H_{31}: DE has impact on the value of the firm.

4.3 Assumption of the Study:

(i) Data have no inflationary effect.

(ii) Time value of money (time value of money) not considered here.

(iii) There have no big structural change over the period.

(iv) Capitaline data Source is true and fair.