Chapter: 3 RESEARCH METHODOLOGY

The entire procedure of modeling approach, beginning with the two set of Models both for Internal Financial Information and Current Market prices as an independent variable is depicted below

3.1 Problem Statement - As discussed earlier in Chapter One
3.2 Research Objectives- As discussed earlier in Chapter One
3.3 Research Design : Exploratory research
3.4 Sampling Design : Purposive sample design (discussed later in this chapter)
3.5 Source of Data

Secondary sources were used to collect the historical information (longitudinal) of sample companies, the details are as under:

The financial statement information (acted as idiosyncratic information) growth rates were used on yearly basis from 2000- 2015 extracted from Capitaline database of BSE top 500 companies as on 21st January 2015.

3.5 Sampling Frame, Methodology and Sample Size

A “super-population” can be considered as the one which is “finite” here BSE 500 companies database was acquired for the same purpose. The rest of the details can be seen in the section of purposive sampling design.

Table 1a: Distribution of Sample companies in terms of market capitalization and related details (Figures in 10 million INR units)
See above details, In case of the market capitalization, we can witness that it seems a well-diversified portfolio, in Cement sector, Birla cement, Heidelberg represent a small-cap category, Ramco and Prism under Mid-cap and ACC and Shree cement under large-cap stocks.

For Textiles Bombay dyeing, Vardhaman and Raymond Industries are small-cap, SRF and Arvind mills comes under mid-cap and Grasim Industries is Large-cap stock

For Automobiles, only HMT comes under small-cap, Ashok Leyland and TVS motors under mid-cap and Hero Motocorp, M&M, Tata Motors and Maruti Suzuki comes under large-cap category

Finally, for Tata Sponge & Iron, Bhushan Steel and Welspun comes under small-cap, while Tata Steel and JSW steel comes under large-cap stocks category.

So overall, total 9 companies representing small-cap, 6 companies under mid-cap and 9 coming under large-cap categories.
Overall from market capitalization of 662136 crores (or 6621.36 Billion INR), 68.14% is contributed by Automobiles, 12.84% by Cement sector, 10.18% by Steel sector and 8.25% by Textiles sector.

And, together the following were the information used from the accounting books for regression and hypothesis testing purposes. The matrix representation is also provided along with the variable information where m stands for a number of observations, hence $x_{an}$ represents m=year, n as accounting variable. So, for the year 2000-01 growth rates where m=1, is presented below
3.6 Tools and Techniques for Data analysis (Modeling)

Exhibit 1 Empirical Model for Forecasting employee cost volatility depiction

(SEE next page)

Source: Self-generated

3.6.1 Empirical Model building process
1. Initially, after gathering the last 16 years (1999-00 to 2014-15) aggregate information of 24 sample companies (based on Market capitalization ending 21st January 2015), a PCA (principal component analysis) and KMO-Bartlett was conducted for factor decomposition and Single-Factor Modeling.

2. These 16 yearly values were later converted to the first differences (benefiting in terms of making the “time series” scale invariant) and to an extent help in removing the problem of high autocorrelation at the first level (if any) before safely adopting the PCA for selection of the right explanatory variable entering in the empirical Modelling framework.

3. The use of Single-Factor Modelling (with decomposed set of independent annual financial information using PCA and KMO-Bartlett) is statistically relevant, since, the first differences and autocorrelation in the dependent series may pose problem with the number of observations and use more independent variables (bivariate or multivariate) Model will probably increase the chances of “invalidating” the Model (since Standard error of regression will increase due to decrease in the degree of freedom).

4. A suitable OLS-cum-autoregressive Model was then implemented for ascertaining the shadow prices or more precisely aggregate human capital (employee costs) for all the 24 companies.

5. For measuring the “Model fit” approach, and also to retain the maximum data in the dependent and independent series for forecasting purposes, along with the usual linear regression tests (Box-Jenkins Model diagnostics) including Heteroskedasticity of residual, Normality of residual and Autocorrelation of dependent series, using the “differencing” and “appropriate lag selection” was administered. This resulted, in the selection of 4 companies, whose, data series comply with the Box-Jenkins Validation,
and still at least a year or more of historical growth rates were retained for the forecasting purposes.

6. Model validation for regression inputs (Model fit) for both the Single-Factor Models for internal financial information and Current market prices based forecasting of employee costs were conducted. For this purpose, Root means square error (RMSE) – scale variant and Mean Absolute Error (MAE)-scale invariant error measurement techniques were used one after the other.

7. The result was that out of 24 companies while, 18 companies were having the superiority of Internal Financial information based Model against the 6 were CMP based forecasting Model shows better (Lower RMSE) figures.

8. For forecasting purpose, Bootstrapping (Fixed X resampling with replacement) was utilized for generating ten different resamples for each sample company. This was acted as stressed scenarios for robust shadow price movements.

9. The use of smaller resample size is permitted since the researcher wish to retain the original sample series biasedness (particularly in the non-normal pareto-type distribution), since, larger bootstrapped samples will make these biases appearing close to zero, severely impacting the overall effectiveness of Modeling predictability in the present case (Markovich,2008).

10. Each 10 resampled values were used for generating bootstrapped regression coefficients, the bootstrapped explained variation (R-squared) and Unexplained variation (Bootstrapped intercept). These bootstrapped regression coefficients (beta’s), explained variations (R-squared) and Intercepts (alpha’s) were later compared for inter-firm (intra-sectoral) research for the same purpose.Jorion and
Zhang (2009) discussed that declining intercept values closely relate with better or improved credit standards in terms of utilization of quality of accounting information is concerned. Therefore, any drop in intercept value also serves as an indication of better explanatory power.

11. Further, as mentioned in the flow chart above, an attempt to calculate the bootstrapped (Equal weighted moving average (henceforth EQMA) volatility of predicted employee costs was calculated for each resample since after ascertaining the correct regression Model in step 3, for Model predictive reliability for each resampled set of fitted series, a separate Autoregressive test was not conducted, instead bootstrapped sampled were taken with replacement.

12. From the series of bootstrapped shadow price volatility (a.k.a employee cost volatility) regression a value-at-risk was calculated.

13. To calculate Value-at-risk measures on annual shadow price conditional volatility EQMA (Equal weighted moving average) Model was considered. This was preferred; due to first the simplicity of calculation since the exact conditional dependency due to illiquid nature of human capital prices is not in terms of past time-frame. Secondly, since it is an illiquid category of asset, the chances of high auto-correlation are pre-anticipated in the stressed simulations (resampled price series). Application of EQMA over bootstrapped replacement series help in linear regression Model validation for predicting employee costs.

14. At the end, all the 10 value-at-risk were further adjusted (deflated) using the ratio of employee capital to the total capital (employee costs+total share capital) and thus

---

Adjusted VaR were used for the analysis purpose (considering Re. 1 as invested capital).

15. Model Validation at the output level was necessary, and therefore at stage 2, the pre-adjusted VaR values of forecasted employee costs were tested for different confidence intervals (90%, 95%, and 99% respectively) using point-backtest method and the results of “superior” stocks where the Model validated were compared with the Model validation at the initial Model fitment (Regression diagnostic) stage.

16. Ang (2014) Life-cycle approach in long run comprising labor income risks was considered at the “firm-level” unlike at the individual level in the book, and the Model outcomes are compared for “Original/Empirical Model” versus “Smoothed Bootstrapped Model” for robust analysis. The detail description is explained later.

17. At the end, the justification of firm-specific (granulated) and sector-specific risk buffer was suggested with the use of various empirical Model outcomes at various stages described above.

---

3.6.2. **Purposive sampling design**

The choice of sectors were based on the most debt-ridden Industries in India, accordingly the figures published by The Hindu (December, 2017) explained that the four dominant sectors are Basic metals (Steel), followed by Cement, Textiles and Automobiles. Lowest on the scale were Rubber and Chemicals. Thus these sectors predominated in the study.

The author chose four core sectors namely Textiles, Steel, Automobile and Cement for the study, as most evidentially it relates with the “Raw labour characteristics” since almost all firms are bulk manufacturing units and at aggregate level a homogeneous structure can be ascertained.³ The author limited his use of *only* internal financial information i.e. use of aggregate historical financial statements mainly because he wishes to assess the idiosyncratic relationships of employee costs and other pertinent “internal business-specific” variables impacting judgment for inter-industry comparisons. Another reason of published data is to provide 'comparable fundamentals' across inter-industry analysis.

---
³ Beker, V. A., & Albusu, E. (2010). Raw labor: homogeneous or heterogeneous?. *Wageindicator.org*
Given that initial sample characteristics in terms of “labour costs” were kept identical, the empirical model will differentiate heterogeneity in terms of two distinctive dimensions, firstly, various stages of model outcomes, specifically at the model validation stages and secondly, in terms of inter-sectoral and inter-firm level differences. Hence, initially the researcher wish to retain “generic homogenous characteristics” corresponding to the labour costs so that the empirical outcomes can validate differentiated-wage-behaviour under the “dimensions” as stated before in this paragraph. No deliberate stratification of purposive sample was made for classifying companies on “size” basis hence out of the final number of sample companies selected, the size-effect study was done later for comparative purposes.

With regard to second dimension of study, generically as Galor and Moav (2004) supported the notion that why the physical capital accumulation is different than human capital accumulation, and therefore the concept of equality is more pertinent and important, once the economy shift gears towards human capital accumulation.

A model-based approach was used for the selection of sample and therefore was conducted on the basis of purposive basis, since, out of the 4 sectors selected, the companies which were fitting the “criteria” were considered for further statistical treatment.

The criteria for purposive selection of companies was to ensure that only those financial variables and companies in the core sectors selected were part of the model where all the variables with respective years available and no “missing data” occupies in the initial database. This lead to almost a identical basis for conducting research (in terms of time-period and the financial variables selected)

The financial statement information (acted as idiosyncratic information) growth rates were used on yearly basis from 2000- 2015 extracted from Capitaline database of BSE top 500 companies as on 21st January 2015.
The purpose of choosing 2000 to 2015 was mainly to cover the time-period somewhat equally between the so called financial crisis which struck the world during 2007-08. Since the empirical model proposed is also covering variedly the concept of human-capital pricing and its relationship with “mispricing” to some extent.

There reason for Hence, a “super-population” can be considered as the BSE 500 companies (which is coming from a non-random finite population which may be considered as all the listed companies in the BSE at the respective time when the researcher thought of choosing the database), while in this study under Model based approach, then this leads to selection of core companies from the four sectors defining “purposive sampling” out of super population is situated between the set of 43 companies (comprising 4 distinct sectors-the total companies in the respective sectors were 12 companies in Cement, 11 in Textiles, 10 in Automobiles, and 10 in the steel sector). Further, after filtering to the panel-fit basis only 24 companies with 6 cement companies, 6 textiles, 7 automobiles and 5 steel sectors companies used for the thesis purposes.

The reason of taking a standard and authentic database was desired and therefore BSE 500 supplied through Capitaline for a span of 15 years (annually) was used for the study.

And, together the following were the information used from the accounting books for regression and hypothesis testing purposes. The matrix representation is also provided along with the variable information where m stands for a number of observations, hence $x_{mn}$ represents m=year, n as accounting variable. So, for the year 2000-01 growth rates where m=1, is presented below

### 3.6.3 Important variable relationships

Income statement idiosyncratic information:

Sales Turnover (ST), Other Income (OI), Raw Materials (RM), Employee Cost (EC), Power & Fuel Cost (P&F), Other Manufacturing Expenses (OME), Selling
and Administration Expenses (S&A). Miscellaneous Expenses(MISC), Reported Net Profit(RNP)
Balance sheet idiosyncratic information:
Total Shareholders’ Funds (TSF), Total Current Assets (TCA), Total Current Liabilities (TCL), Net Current Assets (NCA), Revenue expenses in forex (REFx).
Hence, out of 14 financial variables, 13 were considered as independent, and EC was considered as a Dependent variable in the present analysis.

3.6.4. Underline functional equations and their description:
Functional equations are as under:

3.6.4.1 PCA equations:

a) Firstly, a correlation matrix is proposed which will

\[ S_x^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2 \]
\[ S_y^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2 \]  

And

\[ S_{xy}^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (y_i - \bar{y}) (x_i - \bar{x}) \]  

Then covariance will be

Hence correlation will be:

\[ r = \frac{S_{xy}}{S_x S_y} \]  

b) Now if we have \( M \) observations of \( N \) demeaned variables \( x_1, x_2, x_3, x_4, \ldots, x_{14} \)
where \( x_{ij} \) is the \( j^{\text{th}} \) observation of the \( i^{\text{th}} \) variable then the covariance matrix \( C \) is an \( N \times N \) matrix in which

\[ C_{ij} = \frac{1}{M} \sum_{k=1}^{M} x_{ik} x_{jk} \]

thus the only difference between the

\[ S_{xy}^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (y_i - \bar{y}) (x_i - \bar{x}) \]  
is that each variable \( x \) and \( y \) has been replaced with its matrix for example in the present example the \( X \) or \( Y \) matrices with 14 observations and 14 variables may look like:
\[ X = \begin{bmatrix} x_{11} & x_{21} & x_{31} & \cdots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1m} & x_{2m} & x_{3m} & \cdots & x_{nm} \end{bmatrix} \]

Here \( M \) and \( N \).  

Hence, the symmetric Covariance matrix equation will be \[ C = \frac{1}{M} X^T X \] 

In principal component a matrix \( V \) of eigenvectors which diagonalizes the \( C \) matrix like \[ V^{-1} CV = D \] 

The columns of \( V \) are orthogonal vectors of unit length and they define principal components - i.e. combination of data in directions leading to zero covariance, the diagonal elements of \( D \) are the variances of the each of the corresponding principal components.

c) For principal components sorting of elements of \( D \) is conducted, and this is similarly applied to \( V \), thus the fraction of the variance explained by each vector will be: 

\[ f_i = \frac{D_{ij}}{\sum_{k=1}^{M} D_{kj}} \]

### 3.6.4.2 Creation of Regression equations:

For a simple Uni-variate OLS regression look like:

\[ y_{EC_i} = \beta_1 + \beta_2 x_i + \epsilon_i \] 

\( y_{EC_i} \) = decision variable (Employee cost)
\( \beta_1 = \text{Const (drift)} \)

\( \beta_2 = \text{Regression parameter for } x. \)

\( x \) = the explanatory variable from PCA decomposition

\( \varepsilon_r = \text{stochastic error term} \)

**3.6.4.3 Bootstrapping with fixed matrix x resampling**

Midi and Imon (2009) introduced fixed x-Resampling method which is used for 15 explanatory variable observations in the series.

Method: in the fixed resampling the bootstrap replication is conducted when matrix X is fixed. We test the fitted values \( \hat{Y}_i \) for the Model, by the bootstrap responses. The steps summarized as

Step 1: Fit a Model to the original sample like to get the \( \hat{\beta} \) and the fitted values as, \( \hat{Y}_i = f(x_i; \hat{\beta}) \)

Step 2: Get the residuals \( \varepsilon_i = y_i - \hat{y}_i \)

Step 3: draw \( \varepsilon_i^* \) from \( \varepsilon_i \) and attach to \( \hat{Y}_i \) get a fixed x bootstrap values \( Y_i^* \). Where

\[
Y_i^* = f(x_i; \hat{\beta}) + \varepsilon_i^*
\]  \hspace{1cm} (9)

Step 4: regress the bootstrapped values \( Y_i^* \) on the fixed X to obtain \( \hat{\beta}^* \).

Step 5: repeat step 3 and step 4 for \( \hat{\beta} \) times to get \( \hat{\beta}^* \) ......... \( \hat{\beta}^{*b} \)

**3.6.4.4 EQMA Model for Conditional Variance of fitted time-series data**

Using Equal weighted moving average for Conditional variances (to recollect -here fitted shadow price series volatility for bootstrapped samples)

\[
\sigma^2_{\varepsilon n} = \omega \sigma^2_{\varepsilon (n-k)} + \lambda \sigma^2_{\varepsilon (n-k)} + (1 - \lambda) \mu^2_{\varepsilon (n-k)}
\]  \hspace{1cm} (10)
\( \sigma^2_n \) = The n period variance of the index series

\( \omega V \) = Long term weight*long term volatility

\( \lambda \sigma^2_{n-(n-k)} \) = The \( \lambda \) (decay rate) multiplied with the squared lagged variance

\( \mu^1_{(n-k)} \) = the lagged component, in the present paper the yearly lag is considered

\( (1-\lambda)\mu^1_{(n-k)} \) = here this is decay representing with the Growth rate, this denotes whether there is rapid decay or slow decay, in the case of rapid decay the mean reversion is fast.

For the present purpose, the decay rate was kept as 0.10 or 10% y.o.y. for 10 years simulated time period.

3.6.4.5. Individual VaR 11th year equations:

\[
\text{VaR}_{\text{annual}} = -\left( \mu + Z\sigma \sqrt{n} \right)P
\]

Point backtest for a check of Normality condition:

For 99% C.I. = \( np - 2.58 \times \sqrt{np - np^2} np + 2.58 \times \sqrt{np - np^2} \),

For the 11th year the ranges as per point backtest rule are as follows:

1. For 90% confidence level lower limit is -0.41 while the maximum limit of breaches is 3.41 (rounded off to values without decimals)

2. For 95% confidence level lower limit is -0.90 while the maximum limit of breaches is 2.40 (rounded off to values without decimals)

3. For 99% confidence level lower limit is -0.84 while the maximum limit of breaches is 1.14. (rounded off to values without decimals)

3.6.5 Life cycle approach Modelling equations

(Specific mathematical interpretations)

3.6.5.1. Bootstrapped conditional volatility comparison
(Smoothed EWMA volatility\(^4\)) Model (Others) & Model (CMP)

\[
\tilde{\sigma}^2_{\text{Resample}_{x_{\text{model(others)}}}} = \text{Average smoothed bootstrapped Model (others) volatility}
\]

\[
\tilde{\sigma}^2_{\text{Resample}_{x_{\text{model(CMP)}}}} = \text{Average smoothed bootstrapped Model (CMP) volatility}
\]

\[
\sigma^2_{\text{Original}_{x_{\text{model(others)}}}} = \text{Average original empirical Model (others) volatility}
\]

\[
\sigma^2_{\text{Original}_{x_{\text{model(CMP)}}}} = \text{Average original empirical Model (others) volatility}
\]

\[
i_{S_{\text{Resample}}} = \text{Stock-like capital under smoothed bootstrapped case}
\]

\[
i_{B_{\text{Resample}}} = \text{Bond-like capital under smoothed bootstrapped case}
\]

\[
i_{S_{\text{original}}} = \text{Stock-like capital under original empirical Model case}
\]

\[
i_{B_{\text{original}}} = \text{Bond-like capital under original empirical Model case}
\]

**Scenario 1:**

\[
\left\{ \tilde{\sigma}^2_{\text{Resample}_{x_{\text{model(others)}}}} = \left(\sum_{n=1}^{t} \tilde{\sigma}^2_{x_{\text{other}}(n)} \right) \div n \right\} < \left\{ \tilde{\sigma}^2_{\text{Resample}_{x_{\text{model(CMP)}}}} = \left(\sum_{n=1}^{t} \tilde{\sigma}^2_{x_{\text{CMP}}}(n) \right) \div n \right\} = i_{S_{\text{Resample}}}
\]

(13)

**Scenario 2:**

\[
\left\{ \tilde{\sigma}^2_{\text{Resample}_{x_{\text{model(others)}}}} = \left(\sum_{n=1}^{t} \tilde{\sigma}^2_{x_{\text{other}}(n)} \right) \div n \right\} > \left\{ \tilde{\sigma}^2_{\text{Resample}_{x_{\text{model(CMP)}}}} = \left(\sum_{n=1}^{t} \tilde{\sigma}^2_{x_{\text{CMP}}}(n) \right) \div n \right\} = i_{B_{\text{Resample}}}
\]

(14)

**3.6.5.2 Original (empirical) conditional volatility comparison (EWMA volatility) between Model (Others) & Model (CMP)**

Scenario 1:

\footnote{Volatility = Square root of variance}
\[ \{ \sigma^2_{\text{Original \ X_{\text{model(others)}}}} \} < \{ \sigma^2_{\text{Original \ X_{\text{model(CMP)}}}} \} = I^S_{\text{Original}} \]  \hfill (15) 

Scenario 2:
\[ \{ \sigma^2_{\text{Original \ X_{\text{model(others)}}}} \} > \{ \sigma^2_{\text{Original \ X_{\text{model(CMP)}}}} \} = I^B_{\text{Original}} \]  \hfill (16)

3.6.5.3. Credit-rating migration Model

\[ \begin{align*} 
[I^B_{\text{Original}} ={}& \Rightarrow I^S_{\text{Resample}}] = \text{UPGRADE} \\
[I^B_{\text{Original}} ={}& \Rightarrow I^B_{\text{Resample}}] = \text{NOCHANGE} \\
[I^S_{\text{Original}} ={}& \Rightarrow I^S_{\text{Resample}}] = \text{NOCHANGE} \\
[I^S_{\text{Original}} ={}& \Rightarrow I^B_{\text{Resample}}] = \text{DOWNGRADE} 
\end{align*} \]

D. Migration ratio (in times)

\[ \text{migrationratio}_{\text{smoothed resampled}} = \frac{\sigma^2_{\text{Resample \ X_{\text{model(CMP)}}}}}{\sigma^2_{\text{Resample \ X_{\text{model(others)}}}}} \]  \hfill (17) 

\[ \text{migrationratio}_{\text{original}} = \frac{\sigma^2_{\text{Resample \ X_{\text{model(CMP)}}}}}{\sigma^2_{\text{Resample \ X_{\text{model(others)}}}}} \]  \hfill (18)
3.7 Operational Definitions:

1. Migration ratio: Here migration ratio (smoothed resampled) is determined by taking average of 10 resampled Equal weighted moving average (EQMA) based human capital pricing variances through Model (CMP) divided by averaged of 10 resampled equal weighted moving average based human capital variances Model (Others). While, Migration ratio (original) is determined with the original (pre-bootstrapped) ratio of EQMA-based variance of Model (CMP) divided by EQMA-based of Model (others).

2. UPGRADE: when the smoothed (resampled) stock-like risk capital stands equal to or lowers than the bond-like original empirical model risk capital in the migration phase (from original to bootstrapped).

3. NO CHANGE: when the situation remain same in characteristics (irrespective of value) i.e. migrating from original empirical risk capital to smoothed bootstrapped risk capital the bond-like and stock-like characteristics remain same.

4. DOWNGRADE: It is the reverse of UPGRADE, when the stock-like original empirical risk capital remains same or more in comparison to smoothed resampled bond-like risk capital.

5. Stock-like risk capital under smoothed bootstrap case: It is situation where the smoothed bootstrapped model (others) volatility is equal to or lesser than the smoothed bootstrapped model (CMP) volatility.

6. Bond-like risk capital under smoothed bootstrap case: it is reverse of stock-like risk capital under smoothed bootstrap case.Stock-like risk capital under original empirical model: It is situation under original empirical model (others) variance having equal to or lesser than the original empirical model (CMP) variance.Bond-like risk capital under original empirical model: It is reverse of stock-like risk capital under empirical model case.

7. Average original empirical Model (others) volatility: Yearly Average of (average 10 fixed- X resampled) original bootstrapped model i.e. EQMA (equal weighted moving average) volatility derived from Model (others) prices.

8. Average original empirical Model (others) volatility: Yearly Average of (average 10 fixed- X resampled) original bootstrapped model i.e. EQMA (equal weighted moving average) volatility derived from Model (CMP) prices.

9. Average original empirical Model (others) volatility: Yearly Average of original empirical model i.e. EQMA (equal weighted moving average) volatility derived from Model (others) prices.
10. Average original empirical Model (CMP) volatility: Yearly Average of original empirical model i.e. EQMA (equal weighted moving average) volatility derived from Model (CMP) prices

11. Model (others): A linear single-factor model selected considering firm-level human capital (employee costs) growth rates as dependent variable and most suitable financial variable/factor growth rate as independent variable

12. Model (CMP): A linear single-factor model selected considering firm-level human capital (employee costs) growth rates as dependent variable and firm’s closing (annual) stock price as independent variable

13. Unadjusted Value-at-risk: On per INR basis, unadjusted stands for both Model (others) and Model (CMP) EQMA-volatility annual value-at-risk exposure quantified without adjusting the portion of human-capital/total capital where total capital is the sum of equity and human capital only (excluding debt).

14. Adjusted Value-at-risk: On per INR basis, adjusted stands for both Model (others) and Model (CMP) EQMA-volatility annual value-at-risk exposure quantified with adjusting the portion of human-capital/total capital where total capital is the sum of equity and human capital only (excluding debt).
### 3.8 Statistical tools used and its description:

<table>
<thead>
<tr>
<th>Sno</th>
<th>Statistical tool</th>
<th>Purpose</th>
<th>Usage</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>KMO-Bartlett test</td>
<td>Selecting of the single factor</td>
<td>Target KMO&gt;0.05</td>
<td>SPSS 25</td>
</tr>
<tr>
<td>2</td>
<td>PCA</td>
<td>Selection of the single factor</td>
<td>Eigen value &gt;1</td>
<td>Gretl</td>
</tr>
<tr>
<td>3</td>
<td>Human capital pricing model - OLS based</td>
<td>Regression diagnostics</td>
<td>Intercept, slope, SE of regression, RMSE, MAE</td>
<td>Gretl and MS Excel</td>
</tr>
<tr>
<td>4</td>
<td>Fixed X (independent) resampling</td>
<td>Predictive analytics using one year bootstrapping volatility</td>
<td>Small resamples created</td>
<td>MS Excel and Gretl</td>
</tr>
<tr>
<td>5</td>
<td>EQMA (Equal weighted moving average volatility)</td>
<td>For human capital pricing volatility</td>
<td>Used a lambda value and kept it constant. JP Morgan risk metrics model</td>
<td>MS Excel</td>
</tr>
<tr>
<td>6</td>
<td>Value-at-risk</td>
<td>For determining the EQMA volatility based risk capital exposures</td>
<td>VaR calculated on historical simulation basis</td>
<td>MS Excel</td>
</tr>
<tr>
<td>7</td>
<td>Point-back-test</td>
<td>Backtesting VaR for model validation purposes</td>
<td>Considered at 90%, 95% and 99% C.L.</td>
<td>MS Excel</td>
</tr>
</tbody>
</table>
3.9 Limitations of the study and needed modelling arrangements

To recapitulate the present study is based on the very objective to develop and compare the performance of robust illiquid pricing model based volatility and justifying the volatility scenarios in terms of its application for credit rating migration framework. Hence, the attention of the researcher was more the empirical modelling and internal consistency and validation as against its validation in terms of its application and acceptance to the entire pool of companies.

1. Hence, the first limitation is towards why only 24 companies, because although as mentioned earlier the selection of sectors were kept in such a way that labour (employee) emerge as ‘homogeneous’ and therefore large-scale fundamental sectors like Cement, Steel, Textiles and Automobiles were taken into consideration.

2. The present model is using limited sample size and hence can only be considered under as a micro-economic model. The basic premise is that any micro-level model can be scaled up to provide macro-economic justifications.

3. The sampling method used was purposive and not random, hence, the results of the present study is limited to the use of the companies taken under the sample.

4. The debt capital is completely ignored in the present work:

Firstly, because debt is non-marketable and here a “market-price” of equity was used in one of the illiquid pricing (human capital) pricing framework.

Equity defines principal stake (as owners) and Human equity (the proportional amount of human equity over sum total of financial equity capital and human capital) as Agents (employees).


Jiang and Zha () work put the emphasis that even Fama-Macbeth two-pass empirical pricing model used close to 25 companies as sample size in the portfolio setup. The paper also highlighted that the non-traded macroeconomic factors have low correlation with
latent factors due to they being idiosyncratic in nature. So a small size portfolio can be used to generate empirical justifications.

Pietersz and Pelsser (), presented their outcomes in terms of operational loss events where there is a need a bank-specific, customised classification system. hence, each bank should build their own operational risk factors and its risk capital arrangements. The paper further explained that as per Basel Committee on 2006, Paragraph 674, the claim is that under infrequent yet potentially severe losses on external data (which is called a public data (or pooled industry data) seems relevant. Further, the paper highlighted that for measuring the high operational loss severity a low frequency data (say a 1 year period) is suggested. Hence, needless to refer again, a one-year period was chosen as frequency for analysing the illiquid asset pricing volatility.

Risk capital is mainly for infrequent large losses, and hence for operational risk capital both the idiosyncratic factors and its frequency matters. One of the factor which is equally important is human capital and its linkage with firm specific information. Again, to quote “……a bank must be able to demonstrate that its approach captures potentially severe “tail” loss events. [Basel Committee on Banking supervision 2006, Paragraph 667].

Banks due to a diversified pool of credit portfolios (investment in firms) reduce their operational risk capital exposure and also the severity convergence among one bank to another remains low. In the present study however, we studied only one bank’s hypothetical portfolio which is well diversified to avoid large severe losses in low frequencies.

An important point worth mentioning is that our present data used yearly firm level information for 16 years. This is seemingly lower than 30 observations and as such does not aim for normal distribution criteria as usually find in other literature. But, it is worth debating whether Gaussian arrangement due to large degree of freedom truly provide the tailed event (extreme values) while considering the operational risk assessments at the banking portfolio level.