Chapter IV: Analysis of Data

In this chapter the details of quantitative analysis of dependent and independent variable is discussed. The purpose of quantitative analysis is to provide exploratory evidence as how macroeconomic factors including unemployment influence new firm formation at a national level over a period of time. The data collected on a time series from 1994 to 2014 on various economic indicators is presented in Table 2 (Appendix). The selection of study period is based on availability and age of data which starts from 1994 which also matches with the positive changes in Indian economy.
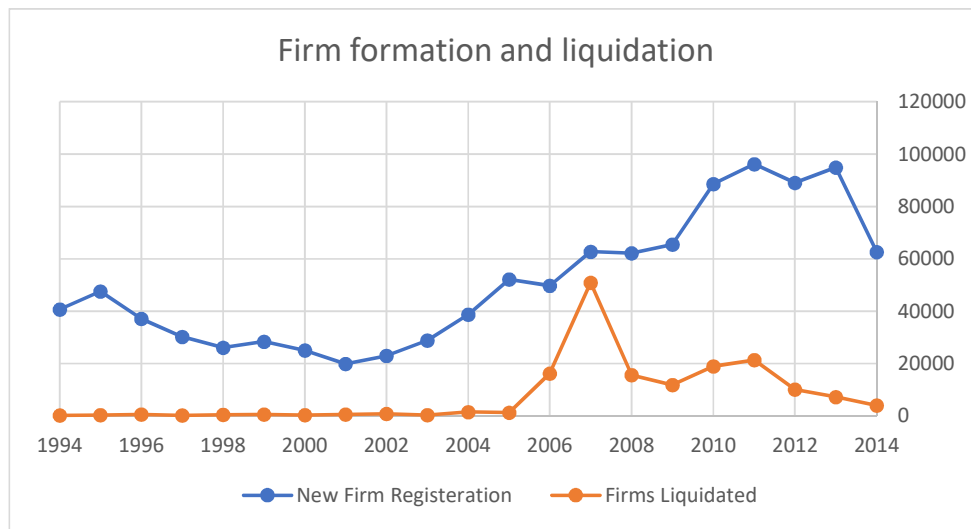

Figure 4: Number of Firms created and ceased

Firstly, the data showed a linear increase in the number of firms created over the years which correlates with increasing size of Indian economy (Figure 4). The data of number of firms created per year had firms of all possible sizes. MCA presents the data in various brackets of authorised capital ranging from less than ten lakhs to more than one crores. Basic analysis of size (authorised capital) of

firms showed that that 87% of new firms created had authorised capital value of less than 25 lakhs as shown in Figure 5 and Table 3. In other words, almost 90% of the new firms registered per year have authorised capital of less than or equal to 25 lakhs which would fall under the category of small firms and businesses. This relates to the theoretical framework of entrepreneurial activity discussed in Chapter II. Furthermore, studies on small businesses by other researchers and scholars can be compared and correlated for interpretation of results of this study.
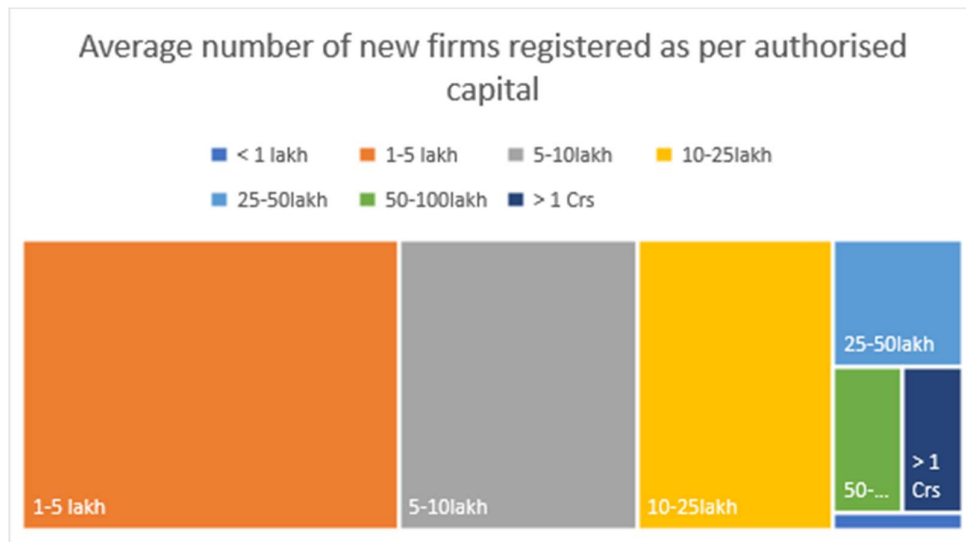

Figure 5: Distribution of number of new firms (average of ten years)

Secondly, the data has to be checked for independence or autocorrelation especially because it is a time-series data. Autocorrelation is generally found in time-series data which has to be checked before conducting any statistical analysis. In time series, it generally occurs due to sluggishness or inertia within the data. Since, most of the time series data is from WB site, it is assumed that there would not be any dependence among variables. The dependent variable for this study is number of new firms registered which is collected from DST and can be independent of other variables. Since the secondary data sources are reputed,

it is assumed that autocorrelation has not happened during data compilation by these sources.

It is known that autocorrelation of disturbances that arise most frequently in time-series data often leads to biased estimators in ordinary least squares regression. Data is , therefore, first checked for autocorrelation issues before statistical studies are conducted. There is a very popular test called the Durbin-Watson test that detects the presence of autocorrelation. To determine whether or not the disturbances in the data are autocorrelated, the Durbin-Watson test was performed. The results of the test was further examined with Durbin-Watson table (for n=21, F=1.161 for one regressor, k=1). The test indicated that autocorrelated disturbances are not present in the data set and normal descriptive multivariate analysis could be done.

Using IBM's SPSS version available at IIPA, correlation module is run on the number of new firm registration (dependent variable) and other factors (independent variables). The output correlation matrix are presented in Table 3 and detailed in Table 6 (Appendix). Since data is collected for R&D as percentage of GDP, another calculated variable for R&D expenditure in billion dollars is also introduced as an independent variable in the correlation test. For inflation variable, also, both inflation figure as GDP deflator as well as inflation based CPI are used.

Being a scale data, there was no need to go for additional correlation tests of Spearman or Kendall's Tau Correlation which are generally done for ordinal data. The Pearson coefficient and level of significance (two-tailed) are symbolically represented by r and p respectively. The Pearson correlation coefficient is further

interpreted as per the guidelines for significance of correlation coefficients (Figure 6).

The snapshot of descriptive test results is tabulated in Table 1 and 2 below and detailed results are in tables in appendix. The input data for correlation test is appended in the appendix.

Table 1: Mean and Standard Deviation

	Mean	Std. Deviation
New Firm Registration (number)	50911.48	24885.12
Firms Liquidated (number)	7776.33	12161.45
Inflation (%), GDP Deflator	6.11	2.29
Inflation (%) CPI	7.48	3.06
Unemployment (%)	3.95	0.30
Lending Rate %	12.14	1.90
Exp on R&D (%of GDP)	0.75	0.08
Exp on R& D (USD),Billion	7.49	5.15
GDP (Current USD),Billions	950.57	588.02
GDP/capita (Current USD)	807.89	431.16

Table 2: Correlations statistics

			1	2	3	4	5	6	7	8
1	New Firm	r	1							
	Registration	p								
2	Firm liquidated	r	.54*	1						
		p	.01							
3	Inflation (%),	r	.41	.05	1					
	GDP Deflator	p	.06	.83						
4	Unemployment	r	-.73**	-.41	-.38	1				
	(%)	p	.00	.06	.09					
5	Lending Rate	r	-.54*	-.22	.23	.35	1			
	(%)	p	.01	.34	.33	.12				
6	R&D Exp. (% of	r	.68**	.49*	-.18	-.30	-.76**	1		
	GDP)	p	.00	.02	.42	.19	.00			
7	R&D Exp.	r	.89**	.48*	.09	-.68**	-.70**	.86**	1	
	(USD),Billion	p	.00	.03	.71	.00	.00	.00		
8	GDP (Current	r	.90**	.49*	.01	-.70**	-.71**	.85**	.99**	1
	USD), Billion	p	.00	.02	.68	.00	.00	.00	.00	

*. Correlation is significant at the $p< 0.05$ level (2-tailed).
**. Correlation is significant at the $p< 0.01$ level (2-tailed).

The results of Pearson's correlation coefficients, r, is interpreted as per following table with level of significance, p, less than 0.01 and 0.05 (two-tailed).

Figure 6: Pearson's Coefficient for strength of association

	Coefficient, r	
Strength of Association	Positive	Negative
Small	.1 to .3	-0.1 to -0.3
Medium	.3 to .5	-0.3 to -0.5
Large	.5 to 1.0	-0.5 to -1.0

Based on above guidelines for interpreting results, highly correlated and moderately correlated variables are indicated with double asterisk and single asterisk respectively in Table 3 and 6. (Detailed table with all variables is at Appendix).

Above coefficients are studied to find the impact of macroeconomic and technology factors on new firm formation. These coefficients are graphically presented in Figure 6 to show its strength of association.

The results show that unemployment is significantly related negatively to new firm formation, $r=-0.73, p<0.01$. It is found, statistically, that unemployment has strong negative influence on new firm formation. However, R&D expenditure is positively related to firm formation and act as potential driver as is cost of capital determined by lending rate. R&D expenditure spurs growth in technology which in turn provides more entrepreneurial opportunities. It is noted that even failure of firms are also significantly related to R&D expenditure, $r=0.49, p<0.05$, which

means that new technology and innovations force old firms to give way to new one providing improved goods and services in the economy.

Negative significant relationship of firm formation to lending rate indicate expensive capital is a deterrent to new businesses. Another logical result is strong link between expansion of economy (GDP increase) and number of new firms, *r=0.90, p<0.01*.

Another interesting finding which came to notice from above result is the relationship of inflation to firm births. Inflation based on GDP deflator has not any significant association, *r=0.41, p>0.05* but inflation, based on CPI, has shown significant association, *r=0.51, p<0.05* (Table 6, Appendix). From result, it seems that general inflation in prices of consumer prices have significant association to firm formation. This needs to be further investigated in future study.
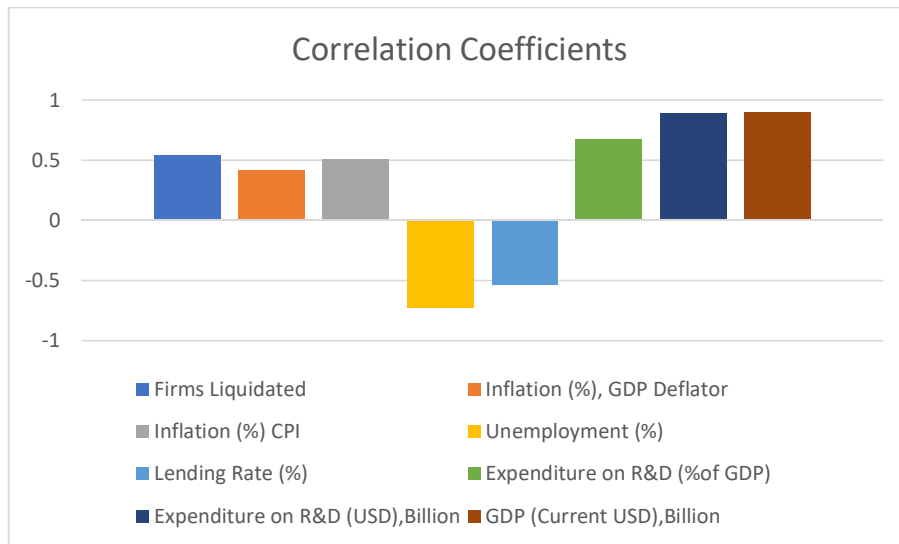

Figure 7:Pearson's Correlation Coefficients

Thus, the results of this study revealed a very interesting finding regarding the role of unemployment in new firm formation in India. The negative relationship indicates that labour availability due to unemployment is not antecedent to entrepreneurial activity. This result is quite contrary to results of similar studies done in US which show a positive relationship with unemployment. Another significant finding is the positive relationship of R&D spending on firm formation. It substantiates the view that technological opportunities increase with spending on R&D which spur firm formation.

Lending rate or cost of capital is negatively correlated to firm formation which indicate that accessibility of easy and cheap capital would lead to new firm creations. Higher lending rate makes capital scarce which affects form formation.

Finally, GDP is very strongly associated with firm formation, r =0.90, p<0.01, which explains that in an expanding economy, opportunities for entrepreneurs increase resulting in increase in new businesses.

Statistically. result show that the above macroeconomic aggregates influence the intensity of entrepreneurship. Policy makers through macroeconomic policies should consider to work on unemployment or labour infrastructure along with other economic and technological factors to encourage entrepreneurship in India.