STRUCTURAL EQUATION MODELLING [SEM] REGARDING THE SUITABILITY OF KERALA AS AN IT INDUSTRIAL HUB

Dr. Jojo K. Joseph
Associate Professor and Head
Post Graduate Department of Commerce,
Marian College Kuttikkanam, Peermade Idukki Dist. Kerala 685 531.

ABSTRACT

Information Technology [IT] industry is one of the largest industry of the world. Due to its ability to generate wealth, power, knowledge and social welfare many countries around the world are looking for ways to promote the development of this industry. IT industry is presumed to have immense potential to change the economic landscape of Kerala. This article analyzes the fitness of Kerala as an IT industrial hub. For this purpose the article analyzes three factors viz. [i] the quality and supply of manpower in the Kerala [ii] the availability of infrastructure facilities and [iii] effectiveness of the various promotional efforts taken by the State though the different agencies functioning for the development of the industry. A Structural Equation Model [SEM] is constructed to indicate the extent of contribution of these factors for the development of the industry, interconnection and interdependence between the factors and their influence on the suitability of the state as IT industrial hub.

Key words: Kerala, IT Industry, Manpower, Infrastructure, Promotional Agencies, Structural Equation Modelling

Introduction

IT as an industry has immense potential to change the economic landscape of Kerala [Vasudevan, 2001]. IT industry which is less capital intensive and has high potential for employment, is ideally suited for Kerala due to its ability to generate opportunities and employment with very low pressure on land, environment, power and other resources. Kerala with its universal literacy, vast pool of skilled and English conversant
manpower, high quality of life indices coupled with the advancements in the field of communication systems, offers a fertile soil for the growth of this industry [Carayannis, 2002]. The advancement of this industry could help the State to a great extent in solving problems like unemployment of the educated youths and migration of the youngsters to places outside the State for want of opportunities here [State Planning Board, 2006]. Thus the State can ensure a better future by effectively exploiting the potential of this industry.

The Research Question

A fine beginning for the IT industry was made by the State in the early 1970s with the establishment of the Kerala State Electronics Development Corporation (KELTRON) as a public sector undertaking which pioneered the electronic industry in the State [http://www.Keltron.org]. A more specific acceleration was given by forming the Electronic Technology Parks [Technopark] in the beginning of 1990s for the development of electronics and software industries [http://www.technopark.org]. But the growth of the industry since then is not satisfactory when compared to the potential of the State. Other States that have entered the field simultaneously or much later have gone far ahead of Kerala. Even in 2005-06 which is almost a decade back the number of IT industrial units registered in Andhra Pradesh, Tamil Nadu and Karnataka were 3.2 times [http://apit.ap.gov.in, 2007] 3.7 times [http://www.elcot.com, 2007] and 4.5 times [http://www.bangaloreitbt.in, 2007] higher than that of Kerala. At the same time the exports from these States were 28.2 times, 31.9 times and 85 times greater than that of Kerala. It is very evident that despite the excellent opportunities and the pressing need for the development of the IT industry, its volume of growth in Kerala is not commendable. In such a paradoxical context this research article addresses the question regarding the role and contribution of three major factors such as the suitability and availability of manpower, infrastructure facilities and the performance of the promotional agencies in the state to make Kerala as an attractive IT industrial hub.

Objective of the Study

The review of literature showed that factors like Manpower Supply, Infrastructure Facilities, Investment Climate, Government Policies and Support, Promotional Efforts etc. are all significant in the development of ICT industry in any place [State Planning Board, 2006]. From among these different factors, the objective of this article is to make an analysis through Structural Equation Modelling regarding the suitability of Kerala as an IT industrial hub taking into consideration three important factors such as Manpower and Infrastructure available in Kerala as well as the contribution of the Promotional Agencies.

Major Factors of the Study

[i] Manpower Supply in Kerala for IT industry

IT industry is a labour intensive industry. Availability of suitable manpower is a major pre-requisite for the growth of this industry. Due to the great demand for suitable manpower with the right skills set, there is a natural tendency for this industry is to move towards places where there is sufficient supply of quality manpower. Kerala is a state blessed with abundant supply of educated and English literate human resource. This large talent pool is an attraction of the state. The state is in the forefront also in creating E-literate and digitally competent human resource. This scenario gives the state a cutting edge in the human resource requirement of the industry.
The human resource requirement of this industry is unique. The IT professionals are expected to possess good academic knowledge and technical skills in computer programmes and operations [Chun, H. 2001]. Further they are expected to have a handful of other traits which are collectively known as ‘soft skills’. Aspects like managerial ability, leadership skills, human relations skills, aptitude for team work, linguistic and communication skills especially in English, logical thinking and values for the exercise of the profession constitute the soft skills. The nature of work in this field call for lot of team work, keeping smooth and cordial human relationships, accepting challenges, managing changes, sharing of responsibilities, handling different and difficult situations, effective communication with subordinates, peers, superiors, and clients and that make soft skills an essential requirement of ICT professionals. The cost of the human resource available in a place and the rate of attrition/retention are also important aspects. The availability of such suitable human resource in plenty in an area makes it attractive for the industry (Brunner H.P., 1995).


[ii] Infrastructure Facilities

State of the art infrastructure facilities are highly essential for an attractive IT destination. Besides the basic infrastructure needed by industries in general, this industry requires IT specific infrastructure too. The various elements of infrastructure facilities required for the ICT industry was identified through a focused review of literature. The review made it clear that the infrastructure requirement of the IT industry globally is almost the same. For instance a report which compares and contrasts Ireland, an attractive destination for the information technology industry with Poland and Philippines, two emerging nations in this industry [Chun, 2001] observes that the core competencies that exist in the field of telecommunication, transportation network including road, rail and air, modern health care systems, digital literacy, presence of supporting facilities like industrial parks, quality education system, hospitality industry etc accounted for the progress of Ireland in IT industry. Similarly the articles analysing the impact of the infrastructure on the growth of the IT industry in Nigeria [Oyebisi, 2003] establish that the non availability of sufficient infrastructure like power, telephone services, updated technology, education and legal framework etc are militating against IT industrial development in Nigeria. Thus the review of the articles cited above along with several other ones reviewed indicates that sufficient availability of quality infrastructure facilities is an important pre requisite influencing the growth of IT industry.


[iii] Promotional Agencies for IT Industry

Strong promotional measures are essential for the growth of IT industry in any place. There are several agencies functioning in the State for the promotion of this industry. The important ones are the [i]
Technopark at Thiruvananthapuram [ii] Infopark at Kochi [iii] Cyber Park at Kozhikode and [iv] Other Agencies like the Department of Information Technology, the IT Mission and the Thiruvananthapuram unit of the Software Technology Parks of India (STPI-T).

Methodology of the Study

This study is based on both primary and secondary data. Primary data was collected from 112 top and middle level managerial personnel from the IT industrial units functioning in the State through a pretested interview schedule. Based on the collective responses of the sample respondents a Structural Equation Modelling [SEM] is evolved to understand the interconnection and interdependence between the factors under study, the extent of their contribution as well as influence on the suitability of the state as IT industrial hub.

The reliability score (Cronbach’s Alpha) of the responses collected in respect of the 12 variables of the factor manpower is 0.817. The reliability score of the responses collected for the factor infrastructure is 0.872. The reliability score of the data collected for the factor Promotional Agencies is 0.7329. These values of the alpha show that the different variables selected for evaluating the factors manpower, infrastructure and promotional agencies can give more than 81 per cent, 87 per cent and 73 per cent coverage of the aspects influencing those factors and are therefore reliable.

Structural Equation Modeling [SEM] of the Factors Influencing IT Industrial Development of the State

Structural Equation Modelling [SEM] is a family of statistical techniques which incorporates and integrates path analysis and factor analysis. SEM is a very general, very powerful multivariate analysis technique that includes specialized versions of a number of other analysis methods as special cases. It is a hybrid technique that encompasses aspects of confirmatory factor analysis, path analysis and regression [http://www.statsoft.com, 2010]. SEM grows out of and serves purposes similar to multiple regression, but in a more powerful way which takes into account the modelling interactions, nonlinearities, correlated independents, measurement error, correlated error terms, multiple latent independents each measured by multiple indicators, and one or more latent dependents also each with multiple indicators. SEM is viewed as a confirmatory rather than exploratory procedure. It tests whether the variables are interrelated through a set of linear relationship by examining the variances and co-variances of the variables. The IBM made Statistical Package for Social Sciences [SPSS] named AMOS (Analysis of MOment Structures) version 20 is used in constructing this model.

Key Concepts and Terms in SEM

1. Structural Equation Modelling Process: The SEM process centres around two steps [i] validating the measurement model and [ii] fitting the structural model. The former is accomplished primarily through confirmatory factor analysis while the latter is accomplished primarily through path analysis. It starts by specifying a model on the basis of theory. Each variable in the model is conceptualized as a latent one, measured by multiple indicators. Based on a large \( n>100 \) representative sample, factor analysis is used to establish that indicators seem to measure the corresponding latent variables, represented by factors. When the measurement model has been validated, the researcher proceeds to ‘fit a model’ which measures the extent to which the covariance predicted by the model correspond to the observed covariance in the data.

2. Indicators: They are observed variables such as items in the survey instrument. Models using a minimum of four indicators per latent variable are recommended (http://www.chass.ncsu.edu, 2010).
3. **Latent Variables:** They are the unobserved variables or factors which are measured by their respective indicators. Each variable in the model is conceptualized as a latent one, measured by multiple indicators. SEM represents relationships among observed and unobserved variables using path diagrams. SEM diagram commonly has certain standard elements: ovals or circles represent latent variables, while rectangles or squares represent measured variables. Residuals which are unobserved are also represented by ovals or circles. Single headed arrows are causal relations and double headed arrows are correlation between indicators or latents.

4. **Loading:** The latent variables in SEM are similar to factors in factor analysis and the indicator variables likewise have loading on their respective latent variables. As in factor analysis, the loading can be used to understand the meaning of the factors/latent variables.

5. **Metric:** In SEM each unobserved latent variable must be assigned explicitly a metric, which is a measurement range. This is normally done by constraining one of the paths from the latent variable to one of its indicator [reference] variables by assigning the value of 1.0 to this path. Given this constraint, the remaining paths can then be estimated.

6. **Error:** An error term refers to the measurement error factor associated with a given indicator.

7. **The Critical Ratio and Significance of Path Coefficients:** When the critical ratio \( \text{C. R.} \) is greater than 1.96 for a regression weight, that path is significant at the 0.05 level i.e. its estimated path parameter is significant.

8. **The Critical Ratio and the Significance of Factor Covariance:** The significance of estimated covariance among the latent variables are assessed in the same manner, i.e. if they have a C. R. greater than 1.96, they are significant.

9. **Goodness of Fit Tests:** These tests determine if the model being tested should be accepted or rejected. If the model is accepted, the researcher will go on to interpret the path coefficients in the model. Among several different goodness of fit measures, the choice of which is a matter of dispute among methodologists (James J. Jaccard, 1996) Jaccard and Wan recommend the use of at least three fit tests while Kline (Kline, 1998) recommends at least four tests.

### Measures of Model Fit used in this SEM

The goodness of fit measures used in this model are the following:

- **Model Chi-square:** The model chi-square which is also called as discrepancy is the most common fit test. The chi-square value should not be significant if there is a good model fit. That is, the P-value of the chi-square should be greater than 0.05 to accept the model. In the model suggested here the P-value is 0.102. Since the P-value is greater than 0.05 it is accepted that there is a good model fit.

- **CMIN/DF:** The value of CMIN/DF is an alternative to the model chi-square value to decide the model fit [Carmines, 1981]. Several writers have suggested that the use of this ratio is a measure of model fit. The rule of thumb is that a ratio less than 5 is indicative of acceptable fit between the hypothetical model and sample data. The CMIN/DF ratio in the model adopted here is 1.272 which is less than 5 which again establishes a good model fit.

- **Root Mean Square Error Approximation [RMSEA].** By convention there is a good model fit if RMSEA is less than or equal to 0.05 (Browne, 1993). The RMSEA for the suggested model is 0.049 establishing good model fit.

- **The Goodness of Fit Index [GFI]:** GFI will always be less than or equal to 1. A value of 1 indicates a perfect fit. Conventionally GFI should be equal to or greater than 0.90 to accept the model. Here the GFI is 0.928 and hence the model satisfies very well this fit measure.

- **Comparative Fit Index [CFI]:** It compares the model fit of the generated model with a null model which assumes that the latent variables in the model are uncorrelated, to gauge the per cent of lack of
fit. CFI varies from 0 to 1. CFI close to 1 indicates a very good fit. By convention CFI should be equal to or greater than 0.90 to accept the model, indicating that 90 per cent of the co-variation in the data can be reproduced by the given model. Here the CFI is 0.966 which strongly indicates the suitability of model fit.

(vi) Adjusted Goodness of Fit [AGFI]: The AGFI takes into account the degrees of freedom available for testing the model. AGFI equal to 1 indicates a perfect fit while a measure equal to or greater than 0.9 indicates a very good fit. In the present model the AGFI is 0.878 which is almost equal to the cut off and is therefore indicative of a good fit.

Grouping of the Variables

To make the data suitable for the requirement of SEM and to avoid redundancy in inferences the 35 variables representing the three factors are compressed into 12 indicators. This was done by clubbing 31 similar and very closely related variables to form 8 new indicators. The remaining 4 variables are taken as such to form the remaining individual indicators. The details of the variables which are clubbed together and otherwise are shown below in Table 1. The new name adopted for each of the indicator in the SEM figure is also shown in the table.

<table>
<thead>
<tr>
<th>Latent Variable</th>
<th>Variables Clubbed Together</th>
<th>Name of the Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manpower</td>
<td>Cost competitiveness + Retention rate</td>
<td>Cost and Retention</td>
</tr>
<tr>
<td></td>
<td>Domain expertise + Work experience</td>
<td>Experience</td>
</tr>
<tr>
<td></td>
<td>Managerial ability + Leadership + HR skills + Team work + Linguistic and communication skills + Logical thinking + Values for the profession</td>
<td>Soft Skills</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>Power + Water + Land</td>
<td>General Infrastructure</td>
</tr>
<tr>
<td></td>
<td>Telecommunications + Bandwidth + Facilities in IT parks + Space availability in parks + Incubation facilities</td>
<td>IT Infrastructure</td>
</tr>
<tr>
<td></td>
<td>IT education and training + P C and internet penetration + Peripherals + Hospitals + Schools + Hotels + Entertainment</td>
<td>Social Infrastructure</td>
</tr>
<tr>
<td></td>
<td>Roads and railways + International airline connectivity + Domestic airline connectivity + Urban transportation</td>
<td>Transport Infrastructure</td>
</tr>
<tr>
<td>Promotional Agencies</td>
<td>Department of IT + STPI-T + IT Mission</td>
<td>Other Agencies</td>
</tr>
</tbody>
</table>

Source: Compiled by the researcher
Similarly Table 2 given below shows the 4 variables which are as such considered as indicators, without being clubbed with any other item, in the SEM diagram.

**TABLE 2**

<table>
<thead>
<tr>
<th>Latent Variable</th>
<th>Variables Considered as Indicators and Included as such in the SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manpower</td>
<td>Academic Knowledge</td>
</tr>
<tr>
<td>Promotional Agencies</td>
<td>Technopark</td>
</tr>
<tr>
<td></td>
<td>Infopark</td>
</tr>
<tr>
<td></td>
<td>Cyber Park</td>
</tr>
</tbody>
</table>

Source: Compiled by the researcher

Using the 12 indicators shown in the above Tables 1 and 2, the following SEM is adopted to illustrate the interconnections and interdependence between the latent variables, the indicators and the extent of their contribution as well as influence on the suitability of Kerala as an IT industrial hub.

**Figure 1**

SEM Regarding the Suitability of Kerala as an IT Industrial
Inferences Emerging From the SEM

The diagram models the role of the three major factors on the suitability of Kerala as an IT industrial hub. The important measures of SEM from which interpretations can be drawn are the estimates of [i] Variance of the latent variables [ii] Co-variance between the latent variables and [iii] Regression Weights between the indicators.

1. Variance of the Latent Variables

The estimate measure of variance shows the relative contributions of each of the latent variables in the system. The following table shows the variance of each of the three factors which affects the IT industry of the State.

<table>
<thead>
<tr>
<th>Latent Variable</th>
<th>Estimate</th>
<th>S E</th>
<th>C R</th>
<th>P-Value</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>2.624</td>
<td>0.726</td>
<td>3.613</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td>Manpower</td>
<td>2.268</td>
<td>1.015</td>
<td>2.235</td>
<td>0.025</td>
<td>Significant</td>
</tr>
<tr>
<td>Promotional Agencies</td>
<td>0.145</td>
<td>0.044</td>
<td>3.281</td>
<td>0.001</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Source: Survey data

Those latent variables whose critical ratio [C.R.] is greater than 1.96 or P-value is less than 0.05 play significant role in the system. The C.R. and P-values for infrastructure, manpower and promotional agencies satisfy these norms indicating that they have the potential to significantly influence the IT industrial development of the State. This shows that in Kerala the industry can be fostered by further improving the supply and quality of infrastructure and manpower as well as by increasing the effectiveness of promotional measures.

2. Co-variance Between the Latent Variables

The co-variance between two latent variables are significant when the C.R. is greater than 1.96 and P-value is less than 0.05. Positive co-variance between two variables indicates that change in one variable will induce movement in the other in the same direction. Negative co-variance implies that change in one variable in one direction will lead to movement in the other variable in the opposite direction. Significant and positive co-variance between all the latent variables are desired for the optimum growth of the industry in the State. The following table gives the details of the co-variance observed in the State.

<table>
<thead>
<tr>
<th>Latent Variables</th>
<th>Estimate</th>
<th>S E</th>
<th>C R</th>
<th>P- Value</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manpower ↔ Infrastructure</td>
<td>0.931</td>
<td>0.449</td>
<td>2.072</td>
<td>0.038</td>
<td>Significant</td>
</tr>
<tr>
<td>Promotional Agencies ↔ Infrastructure</td>
<td>0.299</td>
<td>0.100</td>
<td>2.982</td>
<td>0.003</td>
<td>Significant</td>
</tr>
<tr>
<td>Manpower ↔ Promotional Agencies</td>
<td>0.275</td>
<td>0.115</td>
<td>2.397</td>
<td>0.017</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Source: Survey data
As in the case of variance, here also, the three factors viz. manpower, infrastructure and promotional agencies have significant co-variance. In the model adopted here significant positive co-variance are found between [i] manpower and infrastructure [ii] promotional agencies and infrastructure and [iii] manpower and promotional agencies. The indication in the first case is that improvement in infrastructure facilities is a necessity for creating, attracting and retaining suitable manpower in the State. The relationship in the second case shows that when the promotional agencies increase their effectiveness in performance, it can result in the increased and improved supply of various infrastructure facilities and this can automatically make the job of the promotional agencies easy and effective. Similarly, improvement in manpower will make the promotional efforts easy and effective and the agencies could improve the manpower scenario if they are effective in their performance.

3. Regression Weights

These are the path coefficients that measure the effect sizes of each path in the model. When the C.R. is greater than 1.96 for a regression weight, that path is significant at 0.05 level. In the case of each latent variable one path is taken as the standard whose regression weight is assumed to be 1. Based on that, other paths are compared for interpretation. That is, if the contribution of the reference indicator is one in the latent variable, how much more or less is the contribution of the other indicators. Table 5 gives the output of the regression weights applicable to the model adopted here.

**TABLE 5**
Regression Weights of the Indicators

<table>
<thead>
<tr>
<th>Latent Variable</th>
<th>Indicators</th>
<th>Estimate</th>
<th>S E</th>
<th>C R</th>
<th>P-value</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manpower</strong></td>
<td>Cost and Retention</td>
<td>0.244</td>
<td>0.355</td>
<td>0.688</td>
<td>0.491</td>
<td>Not significant</td>
</tr>
<tr>
<td></td>
<td>Academic Knowledge</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>2.684</td>
<td>0.564</td>
<td>4.758</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>Soft Skills</td>
<td>4.122</td>
<td>0.967</td>
<td>4.262</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td><strong>Infrastructure</strong></td>
<td>Basic Infrastructure</td>
<td>0.756</td>
<td>0.126</td>
<td>5.976</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>IT Infrastructure</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Social Infrastructure</td>
<td>1.238</td>
<td>0.200</td>
<td>6.194</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>Transport Infrastructure</td>
<td>1.304</td>
<td>0.215</td>
<td>6.073</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td><strong>Promotional Agencies</strong></td>
<td>Technopark</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Infopark</td>
<td>1.188</td>
<td>0.218</td>
<td>5.459</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>Cyber Park</td>
<td>1.432</td>
<td>0.281</td>
<td>5.092</td>
<td>0.000</td>
<td>Significant</td>
</tr>
<tr>
<td></td>
<td>Other Agencies</td>
<td>2.856</td>
<td>0.514</td>
<td>5.553</td>
<td>0.000</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Source: Survey data

For the latent variable manpower indicator ‘academic knowledge’ is taken as the reference item. Of the remaining three paths two are significant. Cost competitiveness and retention rate, though they are positive aspects of the State for this industry, are playing only insignificant role in the latent variable when
compared to its other indicators. Experience is 2.684 times important than academic knowledge. Soft skills with the estimate value of 4.122 indicate that it is more than four times important than the reference item.

In the case of infrastructure facilities, IT specific infrastructure is taken as the standard. The remaining three paths are significant. The influence of basic infrastructure items like land, water and power are less in the variable compared to the reference item. But social and transportation infrastructures have higher importance. This indicates the need to improve those facilities grouped under these two heads in order to foster the IT industrial development of Kerala.

Among the promotional agencies Technopark is taken as the standard for comparison. All the remaining three paths are significant. The regression weights for Cyber Park, Infopark and Other Agencies which are higher than that of Technopark indicate their greater potential to influence the IT industrial development of the State.

**Conclusion**
The structural Equation Modelling adopted here clearly suggest that the manpower and infrastructure facilities available in the state are suitable for making the state an attractive IT industrial hub. Similarly the efforts being undertaken by the IT promotional agencies in Kerala are effective and in the proper direction. At the same time there exists considerable scope for improvement in each of these areas.

**References:**


