India's Ease of Living Index Report 2018 – A data driven approach

Deva Dutta Dubey

Associate Professor, RICS School of Built Environment, Amity University Maharashtra

Abstract : **Government of India, Ministry of Housing and Urban Affairs released the first ever Ease of Living Index Report 2018. The report ranked 111 cities in India on the basis of multiple facets of urban living. The underlying data for each of the attributes was also provided. The ranking was based on a-priori weightages to different dimensions of the analysis. This paper attempts to refine the ranking and applies the technique of Partial Least Squares Path Modeling, a technique which may be an alternative for such analyses. The results show that the regression coefficients estimated through the analysis show some variation compared to the weights assigned in the report. Results prima facie indicate that emphasis of different facets of urban living is different. It is also different for different city sizes when the cities were considered as large and small. In addition to the PLSPM Model, additional models have been prepared including neural network models. The neural network and Random Forest models appear acceptable based on accuracy of fit of the models, as represented by Pseudo R2. The report for 2019 is yet to be released.**

This paper is an academic paper having relevance for research in urban planning with its boundaries touching upon economics and land use. With passage of time and with evolution of the concept of ease of living and forming an index, we may witness different variables coming into play and having different loadings on the outcome of ranking of cities.

It could be useful for students of Real Estate Programs in the country and abroad as also for those pursuing Urban Planning and related academic disciplines. Familiarity with this concept and its variation will help them become better data science professionals as they seek gainful employment in various enterprises associated with urban planning and real estate development.

Keywords **– Ease of Living Index, Governance, Cities, Urbanisation, PLS-PM, Neural Networks**

I. INTRODUCTION

Government of India, Ministry of Housing and Urban Affairs (MOHUA) has released the Ease of Living Index (EOLI) Report 2018 in India. The report covers various facets of city and urban living and attempts to rank cities. The report attempts to ascertain and quantify the impact of chosen variables on the well-being of a region and its inhabitants. The index was first launched in 2017 and has been subject to refinement based on feedback from stakeholders. The measure looks at a broader view of factors encompassing several aspects of urban living.

It appears that the measure of EOLI, properly defined and measured, may be a reasonable indicator arising out of such an effort. An appropriate focus of the EOLI may be a city unit / urban agglomeration. This is particularly important considering that more and more people are shifting to cities in search of better economic and social opportunities and that poses a new set of challenges and aspirations for cities to cater to the requirements. The challenges appear in the domains of institutional, social and economic infrastructure not to mention the physical infrastructure. As per estimates of different agencies of the government, in case of India, the urban population is expected to reach high level by 2050.

Considering scarcity of resources and ever pressing need to address human needs and wants, there is an urgent need to manage the resources well for sustenance, well-being and eventual prosperity, as the administrative systems grapple with the constantly changing aspirations of the populace. So far, the EOLI 2018 has been released and the same is analysed in this paper together with some alternative data driven analyses.

II. REVIEW OF LITERATURE

'Ease of Living' is a term used in urban policy and programming but it has no standard definition (IPSOS, 2019). It is dependent on the perspective that one has. However, it is generally linked to many dimensions of city life including physical situation in a city, cultural underpinnings of the city, economic prospects of the city and safety and security prevailing in the city, as perceived by the resident.

Ministry of Housing and Urban Affairs, Government of India (MOHUA) has recently published a set of criteria which help in estimating ease of living and these are embodied in a set of standards which cover various aspects of urban living. The very first report has been published last year and it is likely to evolve with passage of time. The standards are envisaged as a bare minimum skeletal framework for evaluating urban living. It is an attempt to understand key aspects of urban living and will help in providing need-based intervention by the governments. Such interventions are expected to be aligned to Sustainable Development Goals (SDG) too. Each of the indicators have underpinning with SDG which have been referenced. The index will help identify areas where there may be gaps in the urban planning process and may help in filling those gaps. The overarching goal is to improve the quality of life in Indian cities. The methodology is presented in the relevant document referenced in the Methodology Section and appears to be a static exercise, looking to the supply side of the thematic situation, based on the overall infrastructure

spends incurred by the government / other service providers. Data is sourced from relevant departments and there is a provision for collecting high quality data from source departments / providers of services in the cities following a well defined and high quality collection process. There is also a provision for year on year comparison of the index to enable the government to find out annual incremental changes and help in documentation of the progress. The exercise is good, however, considering possibility of reach (i.e. services not reaching the targeted residents) due to many potential factors, it represents only part of the ease of living story.

Beuningen and Schmeets (2013) created a social capital index for Netherlands. They combined a set of indicators into an index using partial least squares estimation. The index comprised two broad dimensions and each dimension having 3 levels. The index helps providing an overview of social capital in Netherlands in one integrated manner, and it can be monitored over time and it helps in comparing differing demographic groupings in the country. They obtained survey data from over seven thousand respondents and found the indicators to predict desired outcomes well. The model helps improve the understanding of the social capital concept.

III. RESEARCH METHODOLOGY

The ranking methodology followed in the Ease of Living Index comprises four principal pillars on which the index is based (MOHUA). The Ease of Living Index is expected to rank cities. These are termed as Institutional, Social, Economic and Physical. These represent the areas which the government is currently looking at from the point of view of improving the quality of life in Indian cities. The four pillars comprise a total of 15 themes which provide the data on those standards. The input values for the themes are in turn captured with the help of 78 indicators of which 56 are core indicators and 22 are classified as supporting indicators. For the purpose of computation of the index, weightages have been assigned to the pillars as follows – Institutional (25%), Social (25%), Economic (5%) and Physical (45%). It appears that the ease of living index is a weighted average score of the underlying

values of the indicator scores (NIUA.org). Exhaustive methodology for computation of the index is contained in the referenced document.

The research methodology proposed in this paper comprises application of the Partial Least Squares Path Modeling technique which appears to be eminently suitable for preparation of indices as the EOLI. In this methodology, constructs are created as Latent Variables (LV) and they derive their values from a clutch of Manifest Variables (MV). The overall philosophy revolves around obtaining LV scores from the MV scores. The PLSPM process involves iteration of the computed scores and their relationships and the algorithm proceeds back and forth till convergence is achieved. Once the LV scores are obtained, a regression is run between the estimated scores. The coefficients of the regression are included in the inner model while the loadings of the MVs on LVs are also numerically estimated and presented as the outer model. The basic model proposed to be applied is presented in Figure 1. The relationships between the MVs and corresponding LVs is traditionally called the outer model, while the relationships amongst the LVs is called the inner model. The inner model depicts how the constructs of the model are related with one another. The outer model depicts how the LVs are linked to MVs which are used as their 'images' captured through the variable scores. As can be seen from the model, the LVs Institutional, Social, Economic and Physical are the four pillars of estimation of the Overall Index which manifests itself through the Ease of Living Index. The variables used to measure or estimate each of the LVs are the categories which have been estimated in order to obtain the Ease of Living Index – they are therefore, the constituents of each of the Pillars. In order to substantiate the results, a boot strapping was performed and the model gave acceptable outputs.

Alternative analytical techniques have also been used to capture the findings from the survey in order to model the same. These are also discussed.
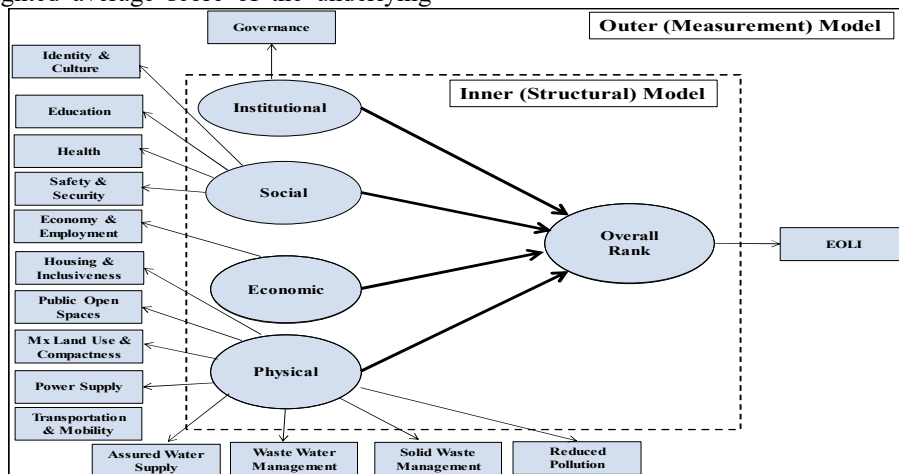

Figure 1: Layout of the Basic Model for PLSPM applied to Ease of Living Index

Source: Empirical Estimation

IV. DATA STRUCTURE AND SOURCES

Data comprises 23 variables showing different attributes of the cities. Overall, 111 cities have been rated. The individual scores are obtained from surveys and the raw scores have been obtained and used for computation of the index. The research is secondary in nature, and the principal data sources for this paper are sourced from Ministry of Urban Development Website and other related websites.

V. ANALYSIS AND FINDINGS

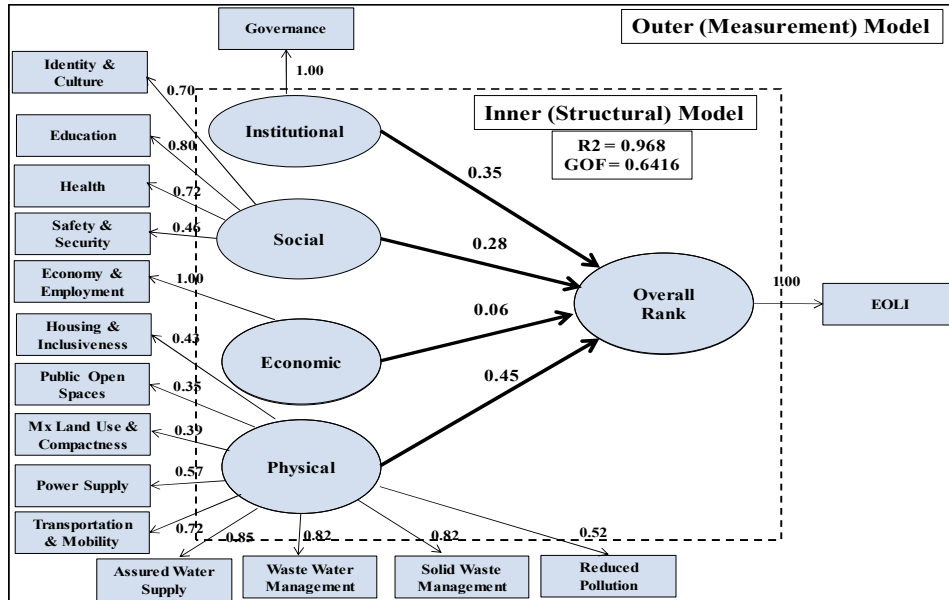A schematic layout of the model showing both inner model and outer model is presented below.

Figure 2: Model for PLSPM applied to Ease of Living Index with estimated Coefficients
Source: Empirical Estimation

Assessment of Model Fit

In terms of the specification for judging model validity (Sanchez 2013) the following model outputs are assessed.

Table 1: Table for assessing validity of model

Parameter	Model Standard for Acceptance	Model Outputs	Remarks
Inner Model - p values	< 0.5 is significant, else insignificant	All p values were less than 0.5	May be considered acceptable
Quality of Structural Model R Square	R2 < 0.3 is a poor fit, R2 < 0.6 is moderate fit and R2 > 0.6 is a good fit	R2 = 0.968	May be considered acceptable
Goodness of Fit Index	The higher the better, GoF > 0.6 is good	0.6416	May be considered acceptable
DG Rho	The higher the better, should be > 0.70	> 0.7	May be considered acceptable
Eigenvalues	1st eigenvalue should be >> 1, 2nd eigenvalue should be < 1	All values meeting criteria, except for Physical	May be considered acceptable
Loadings	Should be significant and values should be higher the better, should be > 0.7	All loadings are found to be significant and of acceptable value, though some are small	May be considered acceptable.
Cross Loadings	Should be highest for variables loading on its own LV	Highest for relevant LV	May be considered acceptable

Source: Empirical Estimation

Further to the above analysis, the model underlying data was bifurcated into large cities and small cities. The model was run again on the divided datasets and the findings are summarized in Table 2.

Table 2: Comparison of assumed Pillar Coefficients and Estimated Coefficients for each type of city

Category	Institutional	Social	Economic	Physical
Pillar Coefficients Assumed	0.25	0.25	0.05	0.45
All Cities	0.35 ⇧	0.28	0.06	0.45
Large Cities	0.36 ⇧	0.20	0.09 ⇧	0.48 ⇧
Small Cities	0.34 ⇧	035 ⇧	0.03	0.43

Source: Empirical Estimation

It can be seen from ***Table 2*** that the initially assumed values for each of the pillars was different than what is borne out by the data. A small difference was observed in the original estimates and the all cities estimates for the pillar of Economic. The pillar for physical was estimated to be the same at 0.45 while the pillar of Institutional was originally underestimated by about 0.10 and the social pillar was underestimated by 0.03.

In case of large cities, there was higher difference in a particular direction (i.e. higher or lower than the original weights), while in case of smaller cities, the divergence was of comparable magnitude but of the opposite direction.

We can infer the following from *Table 2*.

1. Analysis suggests that institutional pillar was considered more important than what was originally assumed in the index. The difference between the assumed value and the estimated coefficients is more for large cities than for small cities. This may be indicating that across cities, a different set of governance parameters / scale of operations are in play depending on the size of the city.

2. The analysis suggests that social categories were considered more important in case of estimation of all cities information, however, for larger cities, the estimate was lower than the assumed value but in case of smaller cities, it was much higher. This may possibly be indicating that depending on the size of the city, different parameters / forces are at play in the case of parameters included in the social pillar.

3. The analysis also gave difference in economic category between larger cities and smaller cities. This may be indicating that in case of larger cities, since cost of living is high, there is a high awareness about economic issues and other forces that determine the outputs. Whereas in smaller cities, economic forces could be different and there may be different underlying market mechanisms.

4. The analysis suggests that at the level of physical categories, there appear to be no substantial differences than what was originally envisaged. Though there exists a slight difference between larger cities and smaller cities. Perhaps it reflects the volume of requirement of physical category variables for large cities due to their size.

In addition to the PLS-PM model, additional models were also generated using Rattle and Neural Net add-in over Microsoft Excel and they are discussed briefly hereunder.

Table 3: Comparison of different models obtained from the data available

Parameter	Models Using R / Rattle				Models Using Excel Add-in
	Linear Model	Decision Tree Model	Random Forest Model	Neural Network	Neural Net – Excel Add-in
Model No. >	(2)	(3)	(4)	(5)	(6)
No. of cases	111	111	111	111	111
Variables / Particulars	All variables had significant coefficients (p-values), except Housing and Inclusiveness (0.1861) Power Supply (0.0609) Waste Water Management (0.2851)		No of trees : 500 No of variables in each split : 3 % Var explained : 88.48 MS Residuals : 0.12 Variable Importance (% IncMSE): Largest : Assured Water Supply (15.97) Smallest : Public Open Spaces (0.46)	Model using nnet Using a 15:10:1 network with 186 weights	Training Set : 78 cases, % bad predictions 20.51 (30% Tolerance) RMSE : 0.1451 MAE : 0.1128 SD of Abs Error : 0.09 Testing : 33 cases, % Bad Predictions 54.55% (30% tolerance) RMSE : 0.27 MAE : 0.20 SD of Average Error : 0.18
Adj R2	0.9646				0.9791 (Training, Linear Predictor)
Pseudo R2	0.9657	0.6417	0.9384	0.9403	RMSE Training Linear : 0.1451, NN : 0.025 RMSE Testing : Linear : 0.27 NN : 0.31

Remarks	3 variables having insignificant coefficients				All variables used in NN Model : Largest Coeff : Governance : 0.3547, Smallest Coeff : Waste Water Management : 0.0136
Comments	Though Adj R2 is high, it is not appropriate to use the model as 3 variables had unacceptable p-values.	Though Pseudo R2 is good, it is still not substantially better than the PLSPM model.	Good and acceptable model.	Good and acceptable model.	Well trained model but having somewhat unacceptable testing of the model as shown by high percentage of poor predictions.

Source: Empirical Estimation

Simultaneous comparison of the 4 models obtained from Rattle is presented in *Figure 3*.
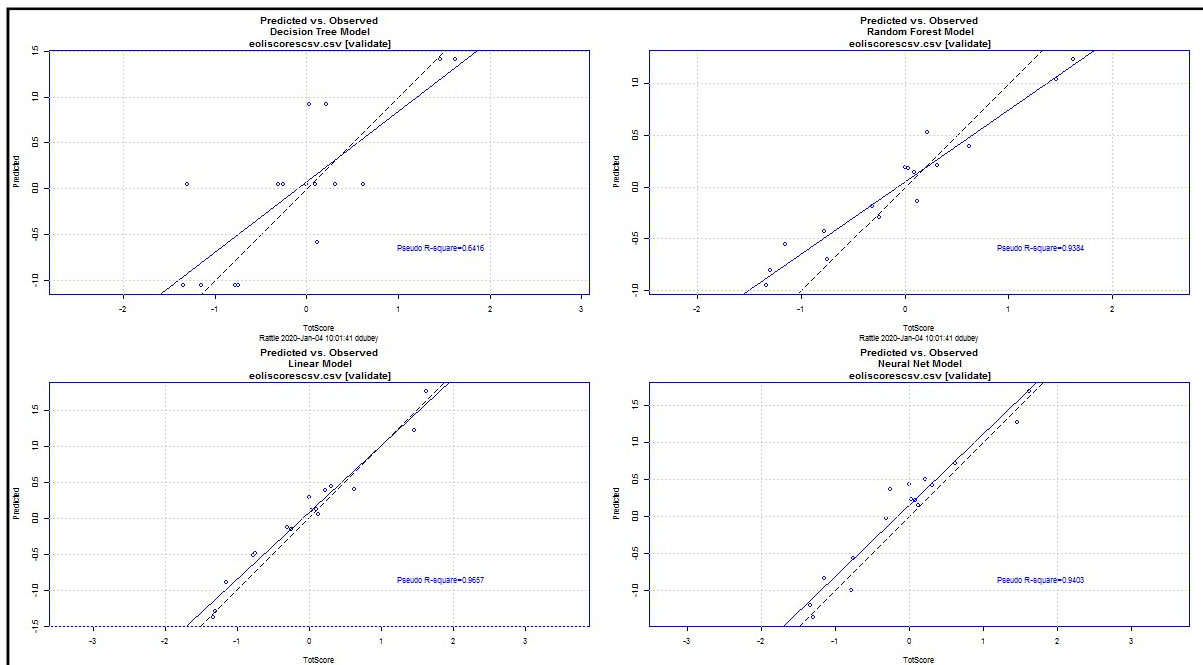

Figure 3: Evaluation of Models obtained from Rattle with estimated Pseudo R2

Source: Empirical Estimation

VI. CONCLUSIONS

In all six models have been prepared using the underlying data. These are the PLSPM model and five other models as contained in Table 3. The PLSPM appears to be acceptable in terms of its context and computational outcomes, subject to data availability. Neural network model comprising Neural Net using 15-10-1 neural layers appears acceptable with a high pseudo R2 followed by Random Forest Model which also indicates variable importance. The Decision Tree model and the Linear Model appear to have poor capability to represent the data as reflected in their lower values for pseudo R2. The neural network models described above appear to have a better inclusion of different dimensions of urban planning as it is based on actual data obtained through the data extraction process. This was also the objective of the exercise. With further modifications of the process, including identification of more variables, it is expected that we may obtain better models.

VII. SUGGESTIONS

The methodology for computation of Ease of Living Index may be modified so as to reassign increased weightages of scores for the different pillars. The method used assigns weightages to different pillars which add up to 1.00. Using a partial least squares path modeling method, coefficients for different pillars are computed as 0.35, 0.28, 0.06 and 0.45. These may be compared to fixed a-priori assigned weights of 0.25, 0.25, 0.05 and 0.45. Such an approach may be able to better represent the data obtained from during the exercise.

REFERENCES

[1] Sanchez G (2013), PLS Path Modeling with R, Trowchez Editions, Berkeley

[2] IPSOS, 2019, https://www.ipsos.com/en-in/ease-living-index-2018 , site accessed on March 4, 2019 at 1626 hrs.

[3]   MOHUA, 2019, https://easeofliving.niua.org/home/about_liveability site accessed on March 4, 2019 at 1645 hrs.
[4]   MOHUA,2019, https://easeofliving.niua.org/assets/upload/pdfs/ease-of-living-methodology.pdf site accessed on March 4, 2019 at 1655 hrs.
[5]   Jacqueline Beuningen & Hans Schmeets, 2013. "Developing a Social Capital Index for the Netherlands" Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement, Springer, vol. 113(3), pages 859-886, September.
[6]   https://competitiveness.in/ease-of-living-assessment-framework/
[7]   https://smartnet.niua.org/eol19/ease-of-living.php