International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020
ISSN: 1475-7192

IDENTIFYING PARTNER COUNTRY'S IN GLOBAL INNOVATION OUTPUT INDEX INDICATORS: A MULTIVARIATE APPROACH

[1]Maria Danesa S. Rabia

ABSTRACT-- *One of the significant indicators of the country's development is having the presence of quality, and innovative outputs come from different sectors. This study examined the recent global innovation output index of 129 countries by employing factor analysis and non-hierarchical cluster analysis to construct factors and identify partners. Resulted to Eight common factors using the factor loadings and scree plot of 35 indicators. The factors differ substantially from the indicators used in previous data and also lead to different rankings of countries. As rankings are not that informative without further information, the distance between each country and the sample mean were considered and analyzed. Differences between countries are much more pronounced for the factors identified in the global innovation index than for individual country indicators. In the non-hierarchical method cluster analysis, the classification of the countries generated seven homogenous groups and was enhanced using multiple regression analysis to identify the predictors of the 2019 Index of Economic Freedom (IEF).*

Keywords--*Global Innovation Index, Global Economic Freedom, Factor Analysis, Multiple Regression Analysis*

I. INTRODUCTION

In a recent development, innovation is one of the drivers of economic progress and competitiveness. Understanding innovation definition has broadened research and development (R&D) awareness that not only restricts to laboratories, publish scientific papers but also could be more general, including social innovations and business model innovations (DOST, 2019). There is much information available on the Global Innovation Index (GII) as it determines the degree of research findings utilization generated and explored from the different fields of discipline. Innovation in emerging markets is seen as critical factor in inspiring people, especially for the next generation of entrepreneurs and innovators. More so, innovation is one of the significant indicators of the country's development, and each country has developed different outputs, but unfortunately implementation of different programs from cross-country differences still brought challenges.

The application of factor analysis on various global innovation output index indicators for 129 countries in the 2019 ranking is the answer to the challenge to endeavor cross-country differences and groupings and to find out whether global innovation output has more than one group. Factor analysis is an excellent instrument to identify what different indicators of a latent construct (global innovation) have in common and to separate common factors

[1] *PhD,**senillo_danesa@yahoo.mail**,09399189861 (text only).,Bohol Island State University, Calape, BoholPhilippines*

Received: 19 Feb 2020 | Revised: 28 Mar 2020 | Accepted: 25 Apr 2020

from specific factors. The outcome of factor analysis is to construct new global innovation factors. However, upon identifying different factors, another promising technique is the use of non-hierarchical cluster analysis to classify and identify partner countries. The purpose of cluster analysis is to find the arrangement of observations and clusters that maximize both within-group homogeneity and between-group (Borden, 2005; Everitt, Landau, Leese, and Stahl, 2011). The within-group homogeneity refers to the extent to which observations that are assigned to a given cluster share similar attributes on the variables included in the cluster analysis. Between-group heterogeneity refers to the extent to which clusters are dissimilar in the aggregate from other clusters concerning the analysis of the kind of variables. Cormack (1971) described that the aim of cluster analysis is to come up with "internal cohesion and external isolation" (p. 329) groups of observation. Thus, the optimal cluster solution places together into clusters those indicators that are most similar to the variables of interest and simultaneously create clusters that, on average, are most different from one another on the variables of interest.

II. OBJECTIVES OF THE STUDY

The purpose of this study is to identify the common factors of the country's Global Innovation Output Index indicators, to classify the partner country's that have commonalities and similarities and ascertain the predictors of the index for economic freedom using the global innovation output index indicators and test its significant difference.

III. METHOD

To come up a better measure of the global innovation output index and to determine whether the global innovation index has a multidimensional character, factor analysis (FA) was used. The first step in this study was to check whether the data used is suitable for factor analysis using the Kaiser-Meyer-Olkin measure of sampling adequacy testing, whether the partial correlation among variables is low. A test statistic above 0.6 indicates that the data is suitable (Kaiser, 1970). Another way of choosing is the test using Bartlett's test of sphericity, which checks the correlation matrix as an identity matrix to which the factor model is inappropriate (Lattin et al., 2003). The next step was to choose the number of factors to represent the output innovation index. The scree plot is one of the bases to look into the number of factors as well as with the use of the eigenvalues on the covariance matrix — initially, two ways of interpreting the graph. In the Kaiser's Rule, states that only factors with an eigenvalue exceeding unity should be kept (Kaiser and Dickman 1959). Another option is to check the number of dots found in the 'elbow' within the scree plot. After deciding on the number of factors, the naming of the factors that are the most difficult in interpreting the results of the study. In this case, the value generated from the factor loadings helped to locate on which dimension does the factor loads easily. Oblimin rotation was used as the rotation technique, which allows for correlation among the factors.

All indicators load the factor scores for the various dimensions (factors) were identified. These factors were used as indicators of the global innovation outputs of a country and analyze using linear regression, and exploration of data were done using pre-requisite k-means cluster analysis.

Data Preparation

The selection of indicators of global innovation output was based on two common rules. First, the data should be widely available for a large number of countries. Second, the data should be gathered in a consistent manner across countries. The data used were from the Department of Science and Technology Database Information System (DOST, 2019), from the Global Innovation Index (WIPO) and 2019 Index of Economic Freedom from Country Ranking.

Data on innovation output were grouped into two broad categories: the 35 national indicators of the global innovation output index.

IV. PRESENTATION, ANALYSIS AND INTERPRETATION OF RESULTS

Data Screening

The data on the global innovation output indicator were screened to ensure that the data were clean, useable, reliable, and valid for testing before further statistical analyses. Data screening was done to identify missing data, outliers, sampling adequacy, sphericity, and multicollinearity. For sampling adequacy and Bartlett's test of sphericity, Kaiser–Meyer–Olkin was used and indicated that the resulting data could be used for a factor analysis that resulted in 0.869, which is above the .6. It is very suitable for factor analysis. There is No missing data that was recorded. The data were subjected to check for multicollinearity and to detect outliers. Outliers distort the exact structure of the data and make the derived clusters unrepresentative of the population structure. Mahalanobis D2 (Hair et al. 1998) was used to test whether an observation had an outlier. Mahalanobis D2 estimates the standard deviation of the distances of the sample points from the center of mass. If the gap between the test point and the center of mass is more than one standard deviation, it is highly probable that the test point does not belong to the set and can be classified as an outlier. The Mahalanobis D2 measure indicated that less than 2% of the observations are outliers. However, the outliers did not affect our results, and these observations were therefore not deleted. Also, Multicollinearity can be a problem in cluster analysis because it distorts the weighting of variables in the different clusters. The resulting VIF is 3.089, which is within the tolerance label. Descriptive statistics were also explored to find out the outliers as one of the prerequisites for deciding the number of clusters for k-means cluster analysis.

Common factors from the global innovation output index indicators

Using factor analysis, this paper tried offers a better view of cross-country differences in the global innovation output indicators. The innovation output indicators are fundamental nowadays to determine the effectiveness of each country in working hard in attaining global's innovation. Factor analysis was used to examine the factor or group of indicators to come up better measures.

The individual ranking of a country's global innovation output indicators was analyzed. The extraction of the right number of signs out from various symbols of the global output innovation index was done using the scree plot (see Fig. 1). According to the Kaiser rule, eight factors could be identified. In this case, the multiple-factor model fits the data significantly better with the goodness-of-fit test statistic, which is 3892.284 with *df* 595 distributed and is highly significant at the five percent significance level, suggesting that the model is appropriate. There were Eight (8) factors were identified.

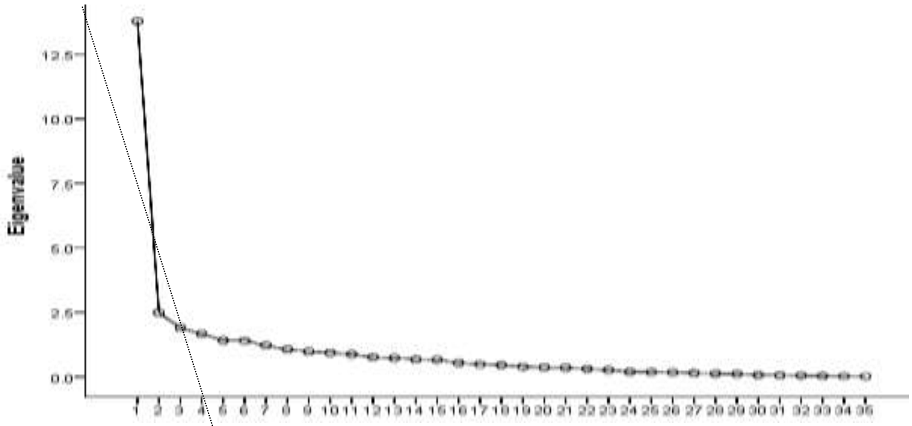

Figure 1: Scree plot of eigenvalues and the number of factors for indicators of the global innovation index among countries.

In addition, the factor loadings of the global innovation output data is very important, as it gives an idea on the structure of the matrix. The resulting matrix consists of a regression coefficient that multiplies common factors to predict variables. The explained variance of the first indicators, which is the knowledge output, is almost 40% of the aggregate that other variables in the dataset explain differences between the values of the variable of interest. The majority of the indicators have higher unexplained variance. This indicates that some of the amount that the other variables in the global innovation output data cannot be explained in terms of the changeable values of other variables in the data set.

There were eight (8) factors that resulted during the factor analysis. Among these names were mind-based output, which has a group of five (5) indicators with .4 to .9 score, export-based output, ICT-based output, product-based, person-based output, ICT-publication based, property-based output, and quality standard based output. This means that values are within the range of at least .3 above. Data were explained in factor analysis that in every factor analysis, there is a matching of the number of factors as there are presence of variables. Every factor get an ample amount of the total variance in the observed variables, and the factors are always enumerated in order to show more variation, as explained. Any factor with an eigenvalue ≥.3, where the eigenvalues measure how much of the variance in the observed variables a factor explains. This also explains more variance than a single observed variable.

Classifying partner countries that have commonalities and similarities

The cluster analysis technique known as *k-means* is an iterative algorithm that attempts to generate the most appropriate fit of observations to clusters, given the number of clusters (k) selected by the researcher prior to the execution of cluster analysis (Everitt, Landau, Leese, and Stahl, 2011; Hunt and Jorgensen 2011). In other words, the researcher selected the number of clusters in advance, which is the k in "k-means" (Rapkin and Luke, 1993).

From the original data of 35 indicators, 8 factors identified using factor analysis were classified to identify the partner countries. There are Seven (7) classified clusters from the most prominent rank of countries in terms of innovation output. These are the countries of Czech Repupblic, China, Estonia, Ireland, Israel, Latvia, Malta, Netherlands, Rep Korea, Singapore, Sweden, Switzerland, the United Kingdom, the United States of America, Denmark, and Vietnam, which belong to cluster 1. Switzerland tops the ranking as the major innovative country.

Received: 19 Feb 2020 | Revised: 28 Mar 2020 | Accepted: 25 Apr 2020

Countries with commonalities and similarities in terms of the global innovation output index are arranged accordingly. Cluster 2 includes countries from Algeria, Bangladesh, Burkina, Guinea, Madagascar, Nepal, Niger, Pakistan, Rwanda, Togo, Uganda, Zambia, Bolivia, Malawi, Nigeria, cluster 3 Bahrain, Botswana, Brunei, Cambodia, Cameron, Cote D' Ivoire, Dominican Republic, Ecuador, Ethiopia, Kazakhstan, Morocco, Nicaruguay, Peru, Tajikistan, United Tanzania and Mozambique. However, the Philippines belongs to cluster 4, with partner countries include Armenia, Canada, Chile, Georgia, Greece, Honduras, Iran, Japan, Kuwait, Lebanon, Mexico, Mongolia, Montenegro, North Macedonia, Panama, Poland, Rep Moldova, Romania, Russian, Serbia, South Africa, Thailand, Tunisia, Turkey, Ukraine, and Uruguay. According to DOST (2019) in the data on High Tech Exports (% of the total net exports), the Philippines performed well this year with a score of 32.7 and ranking 1st out 129 countries, and among the ASEAN countries, Philippines ranked 1st together with Singapore, Vietnam, and Malaysia. Cluster 5 included Australia, Belgium, Bulgaria, Costa Rica, Cyprus, Croatia, Germany, Hongkong, Hungary, Iceland, Italy, Lithuania, Luxemburg, Malaysia, New Zealand, Norway, Portugal, Slovakia, Slovenia, Spain, Finland, France, and cluster 6 also includes countries from Argentina, Austria, Bosnia, Brazil, Colombia, Egypt, India, Indonesia, Jordan, Kenya, Mauritius, Trinidad, United Arab Emirates, while Cluster 7 Albannia, Azerbaijan, Belarus, Benin, Barundi, El Salvador, Ghana, Greece, Jamaica, Kyrgyzstan, Mali, Namibia, Oman, Paraguay, Qatar, Saudi Arabia, Senegal, Srilanka, Yemen, Zimbabwe.

Predictors of the Index of Economic Freedom (IEF)

The study used regression analysis to identify the variables of global innovation output index that solely predicts the country's ranking of the index of economic freedom. With the aid of available software, the generated output is presented in the equation below:

Regression equation for knowledge output:

2019 Index of Economic Freedom= -.406 + -.276 knowledge and technology output + -.257 net flows

Regression equation for creative output:

2019 Index of Economic Freedom= -.579 + -.628 Generic top-level domain + .184 industrial design

The knowledge output has 18 variables, creative outputs 17 a total of 35 variables for global innovation output. These were used to predict the economic freedom index of the ranking of 129 countries. Having a good economic performance of the countries is assumed as a major contributor to the number and kinds of products to be innovated. The global innovation output index was used to identify as the variables that can predict. The equation states that when all variables in the knowledge output are held constant, for every innovation output indicator, the global innovation index will decrease by about .257.

Similarly, if none of the countries will perform on the output generation, this means that each country should continue to enhance the quality of innovation output with regard to its major unique characteristics. So much For these, it would be of great help for the countries to continually strengthen the call for everybody to work on various indicators and align with the available resources of the nations. Meanwhile, in the creative output, two (2) indicators also predict the index for economic freedom. The equation states that when all variables are held constant, for every innovation output indicator, the index will increase to approximately.184.

The test of significance

Table 1 represents the test of significance of the variables of the global innovation output index as a predictor of the index of economic freedom. The 2019 ranking of countries of economic freedom was used as the dependent variable. Then, the 35 variables of the global innovation index were also divided into knowledge output and creative output and were used to identify what particular indicator that predicts the global innovation index. The results showed that using the knowledge output variables as the predictor, and there are three (3) variables that can predict the economic freedom, which is the knowledge and technology output, foreign investment, and net flows. In the creative outputs, only two variables predict the generic top-level domains and industrial designs by origin. The data shows a significant difference among the countries ranking of variables. This means that the country's performance on innovation greatly differs from each other.

Table 1: Analysis of variance for the test of the predictor of the index for economic freedom.

Indicator		Sum of Squares	Df	Mean Square	F	Sig.
	Regression	2336.376	1	2336.376	24.818	.000[b]
	Residual	11861.794	126	94.141		
Knowledge Output	Total	14198.170	127			
	Regression	3038.868	2	1519.434	17.020	.000[c]
	Residual	11159.302	125	89.274		
	Total	14198.170	127			
	Regression	4765.129	1	4765.129	63.649	.000[b]
	Residual	9433.041	126	74.865		
Creative Output	Total	14198.170	127			
	Regression	5212.634	2	2606.317	36.257	.000[c]
	Residual	8985.535	125	71.884		
	Total	14198.170	127			

**SIGNIFICANT AT THE 5% LEVEL OF SIGNIFICANCE*

V. CONCLUSIONS

The global innovation output index indicators for 129 countries of the 2019 data have a multidimensional character identified using factor analysis (FA). There were eight factors identified. Cluster analysis classified seven (7) clusters of partner countries, ranging from countries with very good innovation out not too good, cluster 1 identified as a country's generated more innovation output that includes Switzerland, which ranks first from total global innovation indicators, while the Philippines belongs to cluster 4. Both knowledge and creative outputs variables yield two (2) variables that predict the index for economic freedom. Differences among the country's innovation output were confirmed using the test for significance. Although there were countries that do not perform well but still the majority of them differ significantly, so each of the country has a different response to its innovation output. Furthermore, this strengthened our conclusion that both measures captured different factors of global output innovation and that cross-country comparisons should not only be based on one variable, but other variables also showed significant differences.

Received: 19 Feb 2020 | Revised: 28 Mar 2020 | Accepted: 25 Apr 2020

VI. RECOMMENDATION

Further analysis of the data should be done to minimize the problem encountered in analyzing the cross-country output innovation differences, especially for developing countries. Some variables for these countries show large and unrealistic swings and gaps. Also, the data dispersion within a country cannot be addressed in this study because the focus is on country-level data. However, extracting more factors did not give us more insights and worsened the interpretation of the results.

REFERENCES

1. Borden, V.M.H. "Identifying and Analyzing Group Differences." In M. A. Coughlin (ed.), Intermediate/Advanced Statistics in Institutional Research (pp. 132–168). Tallahassee, Fla.: Association for Institutional Research, 2005.
2. Cormack, R. M. "A Review of Classification." Journal of the Royal Statistical Society, Series A, 1971, 134, 321–367.
3. DOST (2019). Global Innovation Index. Database Information Center. From dost.gov.ph/knowledge_resources/database.html. Accessed on November 27, 2019
4. Everitt, B., Landau, S., Leese, M., and Stahl, D. Cluster Analysis (5th ed.). Hoboken, N.J.: Wiley, 2011.
5. Global Innovation Index. World Intellectual Property Organizations (WIPO). GII Countries ranking. From https://www.wipo.int/global_innovation_index/en/2019/. Accessed on November, 2019
6. Harris, C., & Kaiser, H. (1964). Oblique factor analytic solutions by orthogonal transformation. Psychometrika, 29, 347–362. doi:10.1007/BF02289601.
7. Hair, J., Anderson, R., Tatham, R., & Black, W. (1998). Multivariate data analysis. Upper Saddle River, New Jersey: Prentice Hall.
8. Kaiser, H. (1970). A second generation little jiffy. Pshychometrika, 35, 401–415. doi:10.1007/BF02291817. Kaiser, H., & Dickman, K. (1959). Analytic determination of common factors. The American Psychologist, 14, 425.
9. Kaiser, H., & Dickman, K. (1959). Analytic determination of common factors. The American Psychologist, 14, 425.
10. Lattin, J., Carrol, D., & Green, P. (2003). Analyzing multivariate data. Belmont, CA: Duxbury Press.