HOW TO IMPROVE THE INNOVATION LEVEL OF A COUNTRY? A BAYESIAN NET APPROACH

Esma Nur Cinicioglu[a*], Gündüz Ulusoy[b], Şule Önsel Ekici[c], Füsun Ülengin[d], Burç Ülengin[e]
[a]Istanbul University, School of Business, , Avcılar, Istanbul
[b]Sabancı University, Manufacturing Systems Engineering, Tuzla, Istanbul
[c]Doğuş University, Industrial Engineering Department, Acıbadem, Kadıköy, Istanbul
[d]Sabancı University, Faculty of Management, Tuzla, Istanbul
[e]Istanbul Technical University, Management Engineering Department, Maçka, Istanbul
esmanurc@istanbul.edu.tr, gunduz@sabanciuniv.edu, sonsel@dogus.edu.tr,
fulengin@sabanciuniv.edu, ulenginbur@itu.edu.tr

WORKING PAPER

SABANCI UNIVERSITY
last revision **August 25, 2014**

HOW TO IMPROVE THE INNOVATION LEVEL OF A COUNTRY? A BAYESIAN NET APPROACH

Esma Nur Cinicioglu[a*], Gündüz Ulusoy[b], Şule Önsel Ekici[c], Füsun Ülengin[d], Burç Ülengin[e]
[a]Istanbul University, School of Business, , Avcılar, Istanbul
[b]Sabancı University, Manufacturing Systems Engineering, Tuzla, Istanbul
[c]Doğuş University, Industrial Engineering Department, Acıbadem, Kadıköy, Istanbul
[d]Sabancı University, Faculty of Management, Tuzla, Istanbul
[e]Istanbul Technical University, Management Engineering Department, Maçka, Istanbul
esmanurc@istanbul.edu.tr, gunduz@sabanciuniv.edu, sonsel@dogus.edu.tr, fulengin@sabanciuniv.edu, ulenginbur@itu.edu.tr

ABSTRACT

This study aims to provide strategic guidelines to policy makers who are developing strategies to improve their country's innovativeness. In this paper, we claim that innovation cannot be related only to some factors inherent in the environment of a country, nor is it a single entity to be managed without any linkages to the rest of the actors comprising the competitiveness of a country. Hence, a comprehensive study on innovation should cover the interaction between competitiveness indicators and innovation. For this purpose, the innovation performance of 148 countries is analyzed using an integrated cluster analysis and a Bayesian network framework. These countries are first clustered based on the average values of their competitiveness indicators representing 12 pillars and several sub-pillars adopted from the Global Competitiveness Reports of World Economic Forum for the 2009-2012 period. As a result, five appropriate clusters emerge: Leaders, Followers, Runners Up, Developing Ones, and Laggers. A factor analysis is then conducted to reveal the main characteristics of each cluster in terms of competitiveness indicators. Subsequently, a Bayesian network is constructed and sensitivity analyses are performed to reveal important policies for each cluster.

Keywords: *Innovation, Competitiveness, Cluster analysis, Bayesian networks*, *Sensitivity analysis*

1. INTRODUCTION

Innovation is an important component of competitiveness. Fagerberg (2004) claims that innovative countries have, for example, higher productivity and income than the less-innovative ones. Innovation, moreover, plays a significant role in creating differences in performance and competitiveness not only among regions and countries but also firms (Günday *et al.*, 2011).

Innovation is the key to higher productivity and greater prosperity of nations (UK Innovation Report, 2003). High quality science and research and development (R&D) alone, however, are not sufficient for the realization of important economic and social objectives. New ideas and inventions, the initial steps leading to innovation, are obviously important but may possess little or no economic and social impact unless commercialized (Aarikka-Stenroos and Sandberg, 2012). Griffiths *et al.* (2009) examine the governmental, economic and technological factors that impact sustainable economic growth. For this purpose, they analyzed 34 nations using ordinary least squares, discriminant analysis, and mediated regression. The results show that while the environment and influence of government have positive impact on innovation ecology, sufficient resources for R&D, human capital and early seed funding are the key leading instruments for innovation.

An example of the indication of the recognition of the importance of innovation on the wealth of nations is the emphasis on innovation in the European Union (EU) Lisbon Strategy formulated by the Lisbon Council in March 2000 (Lisbon European Council, 2000). In order to assess the innovation level and as a benchmarking tool for the EU Member States and a few selected countries, The European Innovation Scoreboard (EIS) was initiated in 2000 at the request of the Lisbon European Council. Currently, the EIS provides a comparative assessment of the innovation performance of the 27 EU Member States (EU27) under the EU Lisbon Strategy based on 25 indicators (Innovation Union Scoreboard, 2011). The EIS 2012 includes innovation indicators and trend analyses for EU27 as well as for Croatia, Turkey, Iceland, Macedonia, Norway and Switzerland. Based on their innovation performance, EIS 2012 classifies the countries as *innovation leaders*, *innovation followers*, *moderate innovators and modest innovators*.

Mate-Sanchez-Val and Harris (2014) investigate the national differences of Spain and the UK to highlight how the structural characteristics influence the innovation levels of the companies. The UK is selected as an example of a leader group in terms of innovation according to EIS while Spain is in the follower group. The researchers used two samples comprising private manufacturing firms to explain the firm's innovation based on a two-step Heckman model. Based on the results of the model, they suggest that Spain lags behind the UK in terms of the capability to benefit from R&D and public support, an attribute which is important in promoting innovation. In the UK, on the other hand, international markets are more important for the improvement of innovation. Mate-Sanchez-Val and Harris argue that Spain should focus on promoting market relationships between co-located firms and a greater exposure to internationalization will benefit both countries. However, their study only includes a small subset of the factors that may shape innovation. Furthermore, they do not elaborate on how innovation will increase when those factors are improved.

Padilla-Perez and Gaudi (2014) analyze the stage of development in Central America in terms of science, technology and innovation. They use a questionnaire survey conducted with high-level government officials in each country. The results of the survey show that although Central American governments have initiated the development of public organizations and institutions to support science, technology and innovation, those initiatives are still few and face a number of barriers. Those barriers are: a high-level of political support, low tax revenue, long-term planning and implementation of science, technology and innovation policies, financial resource and institutional culture, domestic public debate on specifically needed policies, weak national commitment of these factors as a source for social and economic development, weak relation between private enterprises and universities, the degree of institutional development, coordination among public organizations, the education system that does not generate enough human resources in terms of quantity and quality and the financial system that is akin to support innovation.

A major competitiveness study on a global scale has been conducted by World Economic Forum (WEF) for some time now. WEF has been publishing the Global Competitiveness Report since 1979. Over time they changed their measurement model and introduced the Global Competitiveness Index (GCI) in 2004 with the aim of combining the Growth Competitiveness Index and the Business Competitiveness Index that they have been publishing (Blanke and Lopez-Claros, 2004). After a period of testing, GCI replaced these two indices in 2009. WEF publishes GCI for more than hundred countries on the basis of over 100 criteria in its Global Competitiveness Reports each year. These rankings serve as benchmarks for policy-makers and other interested parties for judging the competitive success of their country within a global context. Based on GCI, WEF measures the competitiveness of countries, providing a source of data for all sides concerned, including public and private sectors, which they can employ to develop policies pertaining to factors leading to higher levels of competitiveness (WEF, 2012).

The results of the WEF report indicate that there may be other barriers related to other factors of GCI, which are not cited by Parilla-Perez and Gudi (2014) perhaps as a result of lack of a inside perspective of the respondents.

Furman *et al.* (2002) underline that the national innovative capacity depends on the strength of a nation's common infrastructure, the environment for innovation and the linkages between these two. They find that R&D manpower and spending, the extent of intellectual property protection, openness to international trade, the share of research performed by the collaboration of universities and the private sector, the degree of technological specialization and the country's knowledge stock are the basic reasons behind the variation across countries. In fact, those reasons constitute a small subset of the factors used to calculate the GCI of a country. The authors also emphasize that quantitative analyses of the impact of R&D intensity and productivity on innovation can be of particular interest to researchers in the innovation system literature. They provide a regression based analysis but do not provide different road maps to different clusters of countries at different levels of competitiveness.

Porter and Stern (2001) state that the external environment in which a company operates is at least as important as are the internal factors, which determine the capabilities within the companies for creating and commercializing innovation. Therefore, the nation in which the company is heavily operating influences the company's innovation capability. They give Israel example where the innovative output of Israeli firms is also due to Israel's favorable environment for innovation. Their research is based on the innovation level of the OECD as well as that of emerging countries in order to explain the innovative output differences based on national circumstances. The authors claim that the national innovative capacity, the overall human and financial resources, the cluster-specific environment and the quality of linkages of a country are the main drivers of innovation output of the companies operating within this country. In fact, once again, all these factors constitute a small part of the total list of indicators used in GCI. Additionally, the authors do not provide a quantitative analysis to analyze how an improvement in these indicators will result in a jump in the innovation level of the country.

As demonstrated by the work cited above, there exists indeed a wide extent of literature with the purpose of identification of the main characteristics of a country for moving its way to the higher ranks of innovation. However, most of these researches are either based on a specific country's analyses or lack the broad perspective for the environment within which innovation evolves by considering only a small subset of the innovation enhancing factors. In that instance, a broader perspective is needed to explore the main drivers for a country's success in innovation. In this paper, we claim that innovation cannot be related only to some factors inherent in the environment of a country, nor is it a single entity to be managed without any linkages to the rest of the actors comprising the competitiveness of a country. Hence, a comprehensive study on innovation should cover the interaction between competitiveness indicators and innovation.

To the best of our knowledge, there is no research considering all the GCI-related factors when analyzing the interaction between competitiveness and innovation. The underlying dynamics of this interaction depend to a large extent on the characteristics of the development cluster to which a country belongs and, thus, may differ from one cluster of countries to another. Our study contributes to the literature by filling this gap. This work aims to provide a strategic guideline for policy makers by showing how a country in a specific development cluster can move from that cluster to a higher level.

In Section 2, the GCI and the GCI Model developed by the WEF will be presented. The proposed methodology to be employed here will be explained in Section 3. The country groupings obtained through clustering are displayed in Section 4 followed by the development of the Bayesian networks (BN) in Section 5. The results of the sensitivity and scenario analyses for improving innovation in different clusters of countries are given in Section 6. The interpretation of these results is discussed in Section 7. The paper concludes in Section 8.

2. GLOBAL COMPETITIVENESS INDEX AND THE GCI MODEL

As explained above, GCI is a comprehensive tool, bringing together in a framework a large number of factors to assess the competitiveness of countries. That is why GCI in this paper will be employed for analyzing the interaction between the levels of competitiveness and innovation of countries.

GCI measures the microeconomic and macroeconomic foundations of national competitiveness defined as a weighted average of three sub-indices, each of which contains data that are critical to driving competitiveness in national economies. These sub-indices are designated as Basic Requirements, Efficiency Enhancers, and Innovation and Sophistication Factors. Each sub-index consists of a number of so-called competitiveness pillars, which add up to 12 pillars in total. The sub-indices and their corresponding pillars are introduced in Figure 1. According to WEF, the dominant sub-index of a country becomes the key for the classification of its national economy. This association among the sub-indices and classes of national economies is displayed in Figure 1.

Although the 12 pillars of competitiveness are analyzed separately, they are not independent and tend to reinforce each other (Sala-i Martin *et al.*, 2011). Strength in one area can have a positive impact on other areas or vice versa. For example, a strong innovation capacity (in pillar 12) can only be reached with a healthy, well-educated and trained workforce (in pillars 4 and 5).

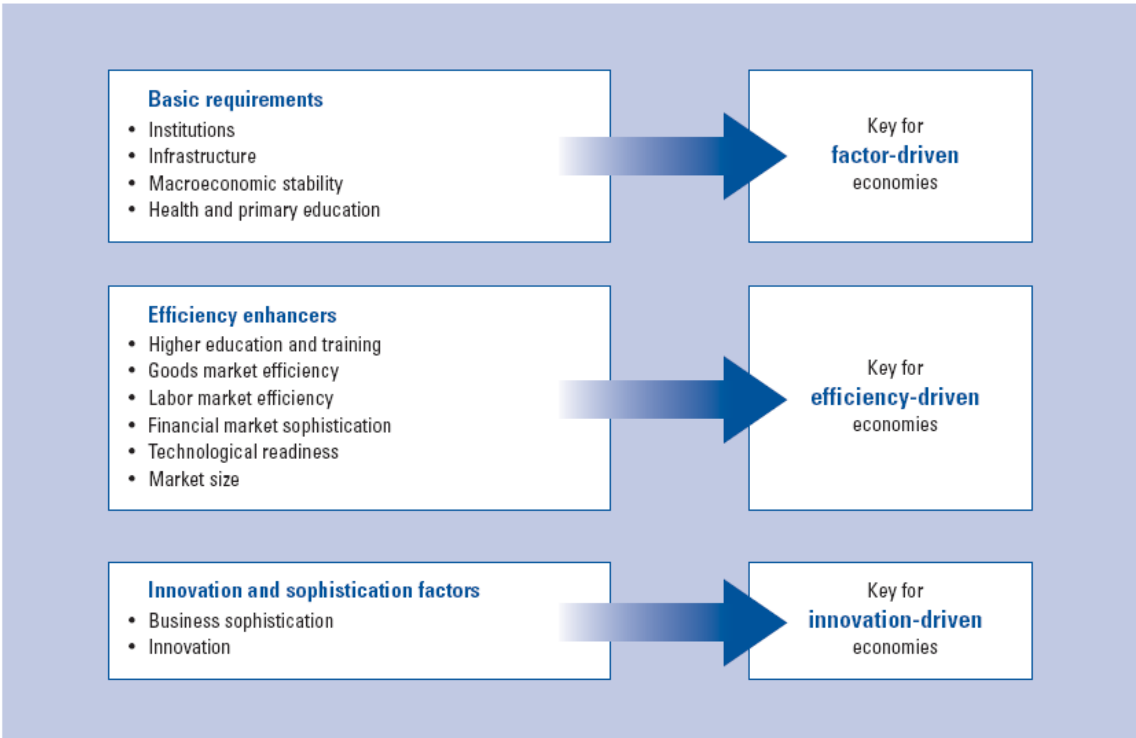

Figure1. The GCI model (WEF, 2012)

Table 1.The stages of development as employed by WEF (WEF, 2012)

	Stages of Development				
	Factor Driven Stage 1	Transition from Stage 1 to 2	Efficiency Driven Stage 2	Transition from Stage 2 to 3	Innovation Driven Stage 3
GDP Per Capita Thresholds (USD)	<2000	2000-2999	3000-8999	9000-17000	>17000

When composing GCI, the weights of the sub-indices are determined according to the stage of development of countries (Sala-i Martin *et al.*, 2008). There are three stages of development, surrounded by two transition stages, resulting in 5 clusters of countries. These stages of development are based on the Gross Domestic Product (GDP) per capita (Table 1) for a large number of countries, but for economies with a high dependency on mineral resources such as oil, the GDP per capita is not indicated as the sole criterion. In this study, however, since the grouping of countries as practiced by WEF does not provide a diversified innovation level for each cluster, a new classification is made based on cluster analysis employing the set of pillars and sub-pillars given in the Global Competitiveness Reports (WEF, 2009, 2010, 2011, and 2012). Subsequently, a solution based on the BN method for investigating the complex structure of the relationship between innovation and a set of competitiveness pillars is proposed.

3. THE PROPOSED METHODOLOGY

The proposed methodology is based on an interactive and visual decision support system that will be provided to the policy makers to guide them in their search for a means of improving the innovation level and assessing the impact of improvements in innovation on the competitiveness of their country in this fiercely competitive global environment. Competitiveness will be represented and investigated through the competitiveness indicators (in short: indicators). The indicators are taken here to correspond to various pillars and sub-pillars of GCI mentioned in Section 2. Innovation, for example appears as one of the twelve main pillars of GCI as published in the Global Competitiveness Reports of WEF. Innovation's importance is especially emphasized by the statement that the standards of living of countries in the long run can only be enhanced via improvements in innovation (WEF, 2012).

Figure 2 summarizes the framework of the proposed methodology. Initially, a cluster analysis is conducted in order to group 148 countries according to their similarities and dissimilarities based on the pillars and sub-pillars. Here, a hierarchical agglomerative technique is employed (Hair *et al.*, 2010).

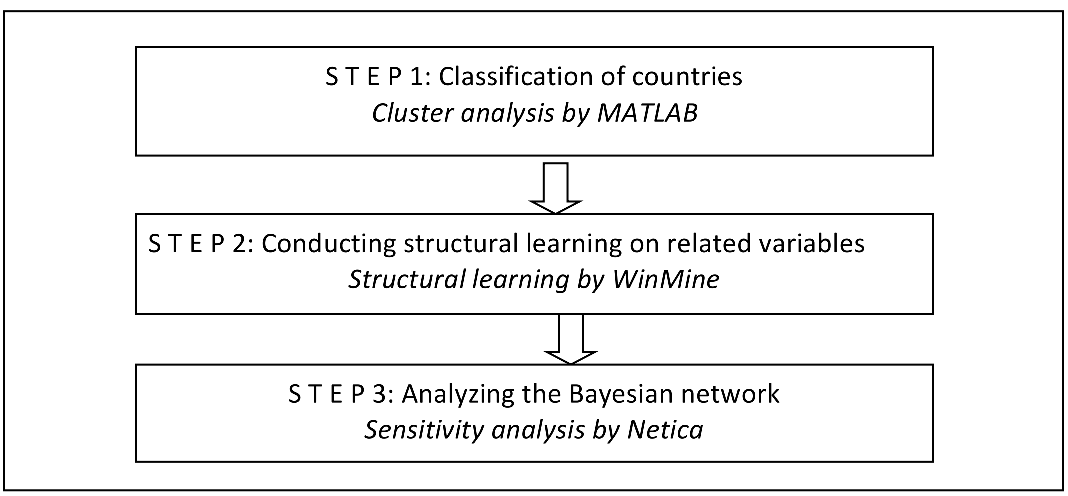

Figure 2. The framework of the proposed solution methodology

Subsequently, a BN is developed through structural learning using the tool WinMine (Heckerman *et al.*, 2000). As mentioned earlier, the pillars and sub-pillars of the GCI model introduced in Figure 1 will be employed to explore interrelations between the factors, which form a complex dependency structure for all the components involved. Hence, in this research, BNs will be used to discover the dependency structure between the pillars and sub-pillars and to evaluate their effect on innovation performance of countries.

In the third stage, several sensitivity analyses are conducted in order to determine the competitiveness indicators that have the highest impact on the innovation level in each cluster of countries and how increasing the innovation level affects the competitiveness level. Plausible strategies to be adopted in each cluster of countries are analyzed.

The innovation level is represented here through the 12th pillar of the GCI model, Innovation. When we look into the individual indicators under Innovation, we observe that this pillar indeed contains the means and policies for the development of an environment conducive to innovation. It is not a complete list but none of the major means and policies appears to be left out: Capacity for innovation; Quality of scientific research institutions; Company spending on R&D; University-industry research collaboration; Government procurement of advanced technology products; Availability of scientists and engineers; Utility patents; and Intellectual property protection. Although it is questionable whether patents are the cause or result of innovation, all the sub-pillars indeed impact the level of innovation in a country. When we investigate the indicators under Business sophistication, on the other hand, we observe the close relationship of several of these indicators with various phases of innovation process and its success in the market such as, Extent of marketing; Production process sophistication; Value chain breadth or Willingness to delegate authority. The latter, for example has been cited as one of the important innovation drivers for manufacturing firms (Günday, *et al.*, 2008).

4. DETERMINATION OF THE COUNTRY GROUPINGS THROUGH CLUSTER ANALYSIS

Since the purpose of this research is to investigate the complex structure of the relationship between innovation and competitiveness represented here through a set of competitiveness indicators, the country grouping used in this study must show a diversified innovation level for each grouping. As mentioned above, rather than using the five country groupings obtained by WEF based on GDP we will determine the country groupings of the 148 countries included by performing a cluster analysis using the common set of pillars and sub-pillars given in the Global Competitiveness Reports of 2009, 2010, 2011, and 2012. Although the pillars remain the same, there appear slight differences in the sub pillars of the GCI model throughout the years. The set of pillars and sub-pillars used are listed in Table 2, where the first digit corresponds to the pillar number and the second digit to sub-pillar designation as referred to in the Global Competitiveness Reports[1].

Justify our choice of constructing the clusters anew, we first obtain a learned BN using the available WEF grouping. However, that analysis reveals that the WEF grouping of countries does not form an indicator for the Innovation variable; hence, it is not possible to use this grouping for our purpose here. The details of the Bayesian network created using the WEF grouping are reported in Table A1 and Figure A1 in the Appendix.

[1] The subpillars 9A Technological adaptation and 9B Information and communication technologies (ICT) has been first used in the Global Competitiveness Report 2010-2011. Since the data for the year of 2009 is not available for these subpillars, the 9th pillar is only considered with the main pillar Technological readiness. Other than that all subpillars are included to the analysis.

Table 2. The set of pillars and sub-pillars used in the study

1A Public institutions
1B Private institutions
2A Transport infrastructure
2 B Electricity and telephony infrastructure
3 Macroeconomic environment[2]
4A Health
4B Primary education
5A Quantity of education
5B Quality of education
5C On-the-job training
6A Competition
6B Quality of demand conditions
7A Labor market flexibility
7B Efficient use of talent
8A Financial market efficiency
8B Trustworthiness and confidence
9 Technological readiness
10A Domestic market size
10B Foreign market size
11 Business sophistication
12 Innovation

The list of countries analyzed by WEF differs slightly in the last four years.[3] The total number of different countries present in this study is 148, a figure that corresponds to 558 lines of the data set.

In the clustering process, unlike the WEF model, no weights are assigned to pillars and their sub-pillars. Each pillar or sub-pillar equally contributes to the overall result. As a result of the hierarchical clustering process conducted using SPSS (2013), the best clustering result is obtained using Ward's minimum variance, taking Euclidian distance as the distance measure, with the number of clusters equal to 5 as listed in Table 3.

Although we do not intend to distinguish among countries within the same group with respect to their competitiveness level, we still would like to be able to reach some conclusions concerning the differences in competitiveness levels between clusters. According to cluster analysis, Cluster 1 is composed of countries at the lowest level of competitiveness while those in Cluster 5 are composed of highly competitive countries.

[2] The 3rd pillar Macroeconomic environment does not have any subpillars.
[3] The complete list of the countries absent in the reports between 2009-2012 are given in Appendix in Table A2.

Cluster 2 includes the countries, which still have a low competitiveness level but which can be accepted at the transition level between the first and the third cluster. Finally the countries assigned to Cluster 3 are average performing countries, and those in Cluster 4 are in a transition stage between the third and fifth clusters. The header of Table 4 states the suggested name of each cluster to which it will be referred to from here on.

Table 3. Clusters of countries

Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
LAGGERS	**DEVELOPING ONES**	**RUNNERS UP**	**FOLLOWERS**	**LEADERS**
Angola	Albania	Algeria	Barbados	Australia
Bangladesh	Armenia	Argentina	Brunei Darussalam	Austria
Benin	Belize	Azerbaijan	Chile	Bahrain
Botswana	Bosnia and Herzegovina	Bolivia	Costa Rica	Belgium
Burkina Faso	Cambodia	Brazil	Cyprus	Canada
Burundi	Cape Verde	Bulgaria	Czech Republic	Denmark
Cameroon	Dominican Republic	China	Estonia	Finland
Chad	El Salvador	Colombia	Iceland	France
Côted'Ivoire	Georgia	Croatia	Jordan	Germany
Ethiopia	Guatemala	Ecuador	Kuwait	Hong Kong SAR
Gabon	Guyana	Egypt	Latvia	Ireland
Gambia, The	Honduras	Greece	Lithuania	Israel
Ghana	Jamaica	Hungary	Malta	Japan
Guinea	Kenya	India	Mauritius	Korea, Rep.
Haiti	Kyrgyz Republic	Indonesia	Montenegro	Luxembourg
Lesotho	Lebanon	Iran, Islamic Rep.	Panama	Malaysia
Liberia	Macedonia, FYR	Italy	Poland	Netherlands
Madagascar	Moldova	Kazakhstan	Portugal	New Zealand
Malawi	Mongolia	Libya	Puerto Rico	Norway
Mali	Morocco	Mexico	Seychelles	Oman
Mauritania	Nicaragua	Peru	Slovak Republic	Qatar
Mozambique	Paraguay	Philippines	Slovenia	Saudi Arabia
Namibia	Serbia	Romania	Spain	Singapore
Nepal	Sri Lanka	Russian Federation	Tunisia	Sweden
Nigeria	Suriname	South Africa	Uruguay	Switzerland
Pakistan	Tajikistan	Syria		Taiwan, China
Rwanda	Trinidad and Tobago	Thailand		United Arab Emirates
Senegal		Turkey		United Kingdom
Sierra Leone		Ukraine		United States
Swaziland		Venezuela		
Tanzania		Vietnam		
Timor-Leste				
Uganda				
Yemen				
Zambia				
Zimbabwe				

The spider diagram in Figure 3 depicts the 21 pillars and sub-pillars employed in the clustering of countries and the performance of the country groupings in these pillars. A general observation is that the clusters are embedded in each other with relatively few exceptions. The Laggers and the Developing Ones follow each other rather closely except in Health, Quantity of education, Primary education, and Electricity and telephony infrastructure in which the Developing Ones have a clear advantage. The Developing Ones and the Runners Up closely follow each other but the Runners Up display a better performance in Domestic as well as Foreign market size, and Quantity of education, i.e., establishing a broader base of education. They preserve their performance against the Followers in Domestic and Foreign market sizes and are also quite close to the Leaders. As a matter of fact, the Followers encompass the Runners Up except for these two competitiveness indicators. The higher performance of the Followers over the Runners Up is particularly visible in Public institutions, Private institutions, Electricity and telephony infrastructure, Transport infrastructure, Quality of education and Technological readiness. Leaders, on the other hand, encompass all others in all pillars and sub-pillars.

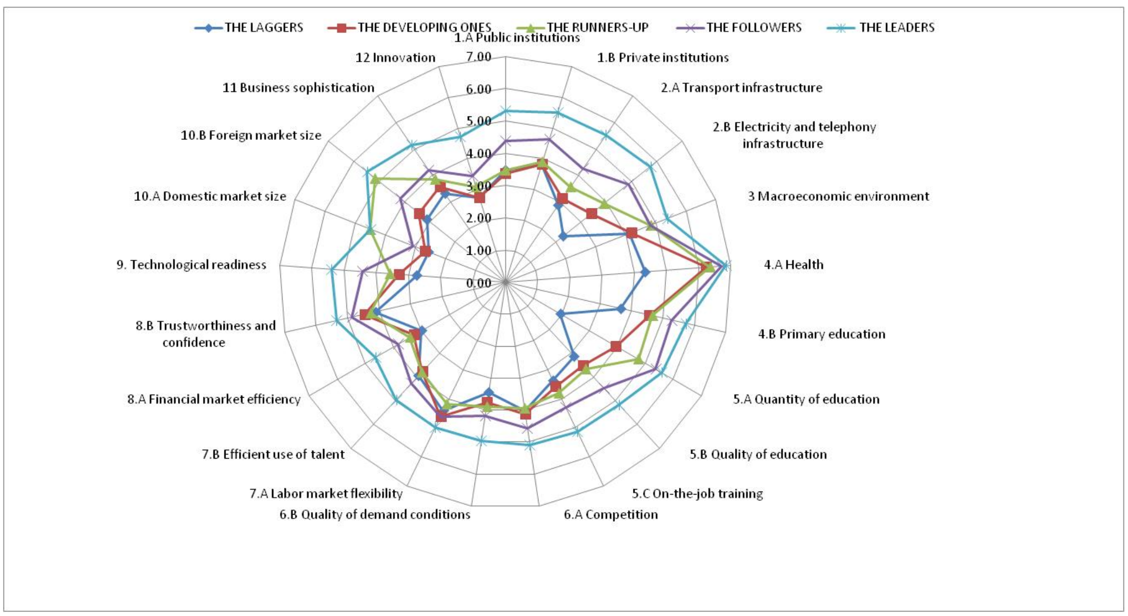

Figure 3. Country cluster performances in 21 pillars and sub-pillars of GCI

In addition to the visual analysis above, a variance analysis -ANOVA- is conducted followed by the Scheffe test used as a post-hoc test in order to decide whether significant differences exist concerning WEF competitiveness indicators among the clusters of countries. The ANOVA test revealed significant differences between the means of the clusters. As a result, it was possible to indicate similarity and difference among the clusters with respect to different indicators. Table 4 shows the results of the Scheffe tests conducted for each pillar and sub-pillar using different colors for different average values of the clusters. For a given pillar or sub-pillar, starting with the Laggers, each change in color corresponds to an increase or decrease for a cluster in the corresponding pillar or sub-pillar with 5% significance. Note that for some of the pillars and sub-pillars we have three leaps of their average values as we go from the Laggers to the Leaders and two leaps for the remaining ones except for the Electricity and telephony infrastructure for which we have four leaps. The pillars and sub-pillars with three leaps are Transport infrastructure, Quantity of education, Quality of education, On-the-job training, Quality of demand conditions, Financial market efficiency, Technological readiness, Business sophistication, and Innovation. The remaining ones are the pillars and sub-pillars with two leaps except for Labor market flexibility, Trustworthiness and confidence, Domestic market size and Foreign market size. These last competitiveness indicators do not steadily increase but rather fluctuate among clusters

Table 4.The results of the Scheffe tests

		1.A Public institutions	1.B Private institutions	2.A Transport infrastructure	2.B Electricity and telephony infrastructure	3 Macroeconomic environment	4.A Health	4.B Primary education	5.A Quantity of education	5.B Quality of education	5.C On-the-job training	6.A Competition	6.B Quality of demand conditions	7.A Labor market flexibility	7.B Efficient use of talent	8.A Financial market efficiency	8.B Trustworthiness and confidence	9. Technological readiness	10.A Domestic market size	10.B Foreign market size	11 Business sophistication	12 Innovation
Cluster 1	THE LAGGERS	3.48	3.81	2.89	2.28	4.11	4.33	3.65	1.97	3.13	3.38	4.02	3.44	4.42	3.94	3.00	4.12	2.75	2.55	3.11	3.33	2.74
Cluster 2	THE DEVELOPING ONES	3.37	3.83	3.14	3.40	4.19	6.23	4.57	3.95	3.53	3.58	4.12	3.75	4.61	3.77	3.26	4.46	3.30	2.66	3.42	3.59	2.74
Cluster 3	THE RUNNERS-UP	3.46	3.89	3.56	3.91	4.82	6.35	4.66	4.76	3.66	3.80	3.94	3.90	4.18	3.80	3.42	4.27	3.57	4.50	5.15	3.85	3.09
Cluster 4	THE FOLLOWERS	4.39	4.62	4.27	4.86	4.82	6.71	5.27	5.37	4.47	4.31	4.58	4.19	4.63	4.27	3.84	4.86	4.44	3.07	4.14	4.22	3.45
Cluster 5	THE LEADERS	5.31	5.51	5.51	5.75	5.38	6.84	5.73	5.59	5.17	5.13	5.11	4.97	4.99	4.96	4.64	5.35	5.39	4.49	5.47	5.14	4.73

Due to the high levels of correlations among the 21 variables, instead of individually using each variable, a dimension reduction is realized based on factor analysis. To ensure suitability for conducting factor analysis; this study employed the Kaiser– Mayer–Olkin (KMO) test and Bartlett's test of sphericity. The result of the KMO test was 0.88 and Bartlett's test of sphericity was high at 494.75 (associated with a probability value of 0.000). Both tests indicated the suitability of the variables for factor analysis. Principal component analysis was carried out to extract the factors. Eigenvalues (>1) were taken as criterion for the extraction of the principal components required for explaining the sources of variance in the data. The four principal components explain 84.6% of the total variance of the 21 variables. The principal component solutions were orthogonally rotated using Varimax rotation in order to add the advantage of improving the interpretability of the resulting factors. The factor analysis results are given in Figure 4; the factor groups, which constitute the basis of these clusters, in Table 5.The factors are designated as Efficiency Related, Basics, Market Size, Macroeconomic Environment and Labor Flexibility.

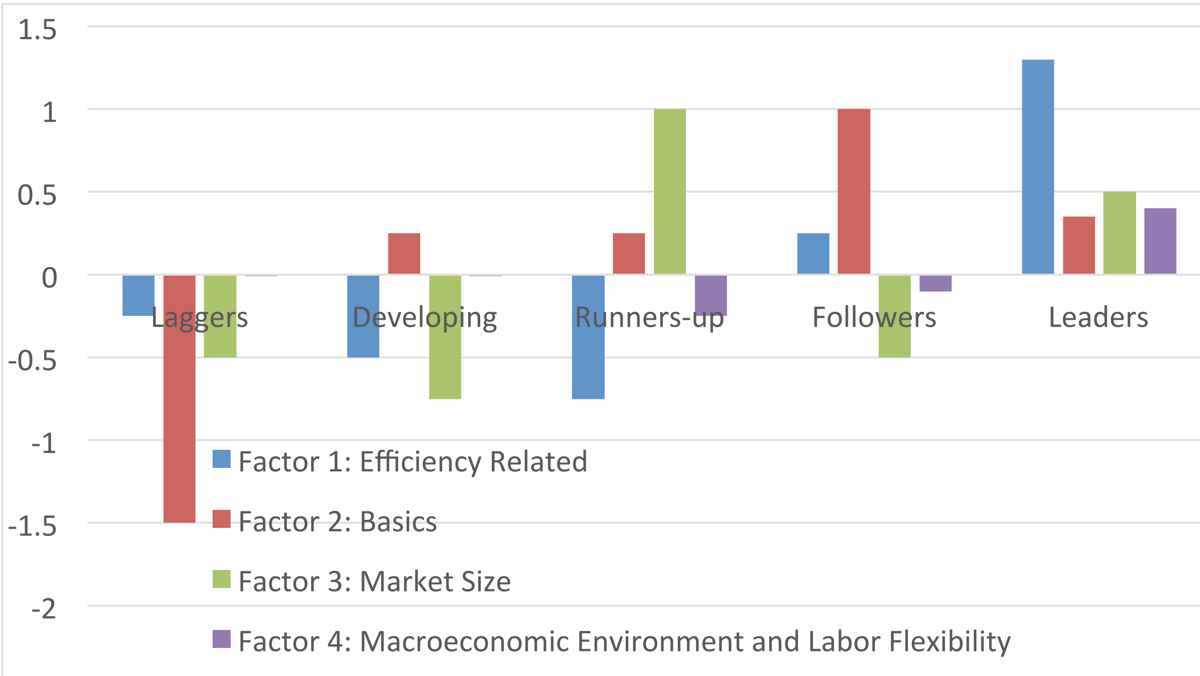

Figure 4. Results of the Ward method

When Figure 4 and Table 5 are analyzed together in detail, it can be seen that Laggers display relatively low performance on all the factors particularly in Basics, which is related with the basic requirements for a country. On the other hand, Developing Ones have only a positive performance in Basics. Runners Up have positive performance in both Basics and Market Size. Market size includes both domestic and foreign market sizes. The demographic characteristics of the countries in the Runners Up cluster are the determining parameters for this factor, which positively impacts the level of competitiveness of a country. The Followers' positive performance in the first and second factors indicates that countries in this cluster have accomplished the necessary improvements in the basics as well as most of the Efficiency Enhancers of the GCI model (Figure 1). Finally, Leaders have positive performance in all factors.

Note that the Leaders include, among others, a number of EU countries, which are designated as the *innovation leaders* by Innovation Union Scoreboard (2011) as was mentioned in Section 1. For example, Switzerland, Denmark, Finland, Germany, and Sweden, which are above the EU27 average in terms of innovation performance, are in this highest level of competitiveness group – the Leaders. Similarly, Austria, Belgium, France, Luxemburg, the Netherlands, Slovenia, and the UK that have an innovation performance level close to that of EU27 average are also members of the Leaders. They were referred to as *innovation followers* in the Innovation Union Scoreboard study. However, the Czech Republic, Malta, Poland, Portugal, Slovak Republic, and Spain that have an innovation performance below that of the EU27 average and are accepted as *moderate innovators* according to the Innovation Union Scoreboard are in the Followers cluster in terms of competitiveness. Greece, Hungary, and Italy, which are *moderate innovators* as well, are included in the Runners Up cluster here. These results show that there is indeed a significant relation between the competitiveness level and the innovation performance of a country. So far we have investigated the clusters and related data employing visual and basic statistical tools. In the following, we will make use BN analysis for that purpose.

Table 5. Factors and related variables

Factors	Variables	Factor Loadings
Factor 1	1B Private institutions	0.841
	1A Public institutions	0.809
	8B Trustworthiness and confidence	0.778
	5C On-the-job training	0.777
	12 Innovation	0.770
	8A Financial market efficiency	0.757
	5B Quality of education	0.753
	11 Business sophistication	0.747
	2A Transport infrastructure	0.738
	6B Quality of demand conditions	0.733
	6A Competition	0.730
	9 Technological readiness	0.695
	7B Efficient use of talent	0.692
Factor 2	4A Health	0.881
	5A Quantity of education	0.836
	2B Electricity and telephony infrastructure	0.694
	4B Primary education	0.689
Factor 3	10A Domestic market size	0.918
	10B Foreign market size	0.870
Factor 4	7A Labor market flexibility	0.809
	3. Macroeconomic environment	0.557

5. DEVELOPMENT OF THE BAYESIAN NETWORK

For executing the BN analysis, the data released by WEF for the 2009-2012 period will be employed as the data set. The 21 different variables used in this study along with their minimum and maximum observed values within each year are provided in Table A3 in the Appendix.

The data set used to learn the BN constitute the data for the 21 variables considered for all 148 countries participating in the WEF study over the four years indicated. Additionally, one more variable indicating the cluster is added to the network. To learn the corresponding BN, the 1 to 7 scaled GCI data is first transformed into a form where each variable is discretized to have seven states, each having equal width of range. The WinMine software is used to learn the structure of the BN from data. Using WinMine we first divide the data into a 70/30 percent train and test split. All 21 variables are used both to predict and be predicted. As the model is built, the granularity of the learned network is determined by using a factor called Kappa. The values for Kappa range between 0 and 1.0. As Kappa value approaches 1.0, a much denser network is obtained. In this analysis, we want to evaluate the effect of all the different factors on countries' innovation performance and also see the dependencies between them. For that reason, in order to see the complete dependency structure between the variables, the Kappa value is taken as 1.0.

In our analysis, we evaluate the accuracy of the learned BN model using the log-score. Log score is a quantitative criterion, which evaluates how the provided model learned from the training data performs on the test set. The formula for the log score is given in Equation 2, where *n* is the number of variables and *N* is the number of cases in the test set.

$$Score(x_1, \dots, x_N) = \sum_{i=1}^{N} log_2 p(x_i|model))/nN \quad (1)$$

In our case, the provided model resulted in a log score of – 0.9886, meaning on average, the log probability that each variable assigns to the given value in the test case, given the value of other variables, is 50.40%. By using WinMine, the lift over the marginal value can be found as well, which is the difference between the log scores of the provided model and marginal model. A positive difference indicates that the provided model outperforms the marginal model on the test set. The lift over marginal value obtained by our model is 0.7009, meaning that the predictive capacity of the provided model is about 19.39% better than the marginal model.

As the next step of the research, the same BN is created using Netica software (2012). This way we can enter evidence for variables and observe the consequential change in the posterior probabilities. Additionally, by using Netica software, sensitivity analysis may also be performed, which allows the decision maker to identify the variables affecting the target variable most. The BN created using Netica software and the marginal probabilities of the variables in the network can be seen in Figure 5.

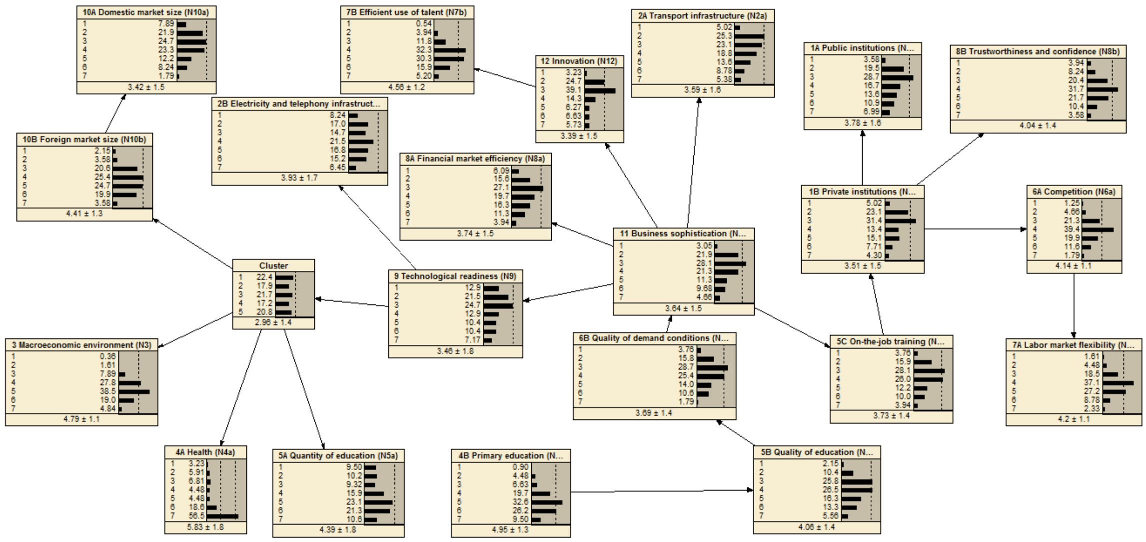

Figure 5.The BN created using Netica software

Notice that the network in Figure 5 is learned from the complete data set, including the data of all countries for the four years considered. This network is beneficial in giving us a complete picture of the interdependencies between the considered pillars. On the other hand, in order to provide strategic guidelines for improving the innovation performance of the countries, it is necessary for all country groups under consideration to analyze the interdependencies inherent in each of the clusters. For that reason, the data set is partitioned into five new data sets according the cluster sets identified previously. For each of the new data sets, the corresponding BN is learned. As a result, six different BNs are created for our analysis, one for the whole data set and one for each of the five country clusters. The corresponding logscores and lift over marginals of all BNs learned are reported in Table 7. As can be seen from the table, there is an improvement obtained for all of the BNs created indicating that these cluster-based BN models can be used efficiently in analyzing the interdependencies between pillars and sub-pillars.

Table 7. Log scores and lift over marginals values of BNs created

BN	Logscore	Lift over marginal	Improvement obtained
The Laggers	− 0.807413	0.487418	0.163822
The Developing Ones	− 0.769154	0.409844	0.145103
The Runners Up	− 0.788488	0.515584	0.173969
The Followers	− 0.66602	0.487645	0.630243
The Leaders	− 0.659827	0.527949	0.193975
All Countries	− 0.988644	0.70088	0.193922

Detailed analyses of the BNs created from different clusters and from the whole data set are provided in Section 6. The analyses are in terms of sensitivity analysis and evidence observation conducted for the specific interest. Varying results of the clusters and their posterior outcomes are interpreted to constitute a strategic guideline for how to improve the innovation performance of a country.

6. SENSITIVITY AND SCENARIO ANALYSES FOR INNOVATION ENHANCEMENT

6.1 On How to Reach an Improved Stage of Innovation

After creating the BN structure from the data set, a sensitivity analysis is performed to determine the competitiveness indicators (variables that have the greatest effect on the Innovation variable. The results of the sensitivity analysis are in terms of the variance reduction (VR) rates corresponding to indicators. VR is the expected reduction in the variance of the output variable (Q) due to the value of an input variable (F). When information is supplied about the state of an input node, the output node distribution may shrink towards more probable values, reducing its variance (Nash et al. 2013). In other words, variance reduction is the difference between the variance of the output node (var(Q)) and the variance of the output node given the input node (Var(Q|F)).The variable with the greatest variance reduction rate is expected to be the one to most change the beliefs of the observed variable, hence, it has the highest explanatory power over the output variable. The variable with the greatest VR rate is expected to be the one that most alters the beliefs of the target variable. Table 8 lists the results of the sensitivity analysis performed on the five different cluster BNs for the Innovation variable. In each column, the indicators are listed in decreasing order of the percent VR (VR%). According to Table 8, it can be seen that Business sophistication and On-the-job-training are the two variables that appear in the top 5 indicator affecting the Innovation level of all country groups.

Table 8's sensitivity analysis offers information to which competitiveness indicators inherent in the environment most affect the Innovation level of the Cluster under consideration. The next step necessary is to analyze how to improve the Innovation level of the Cluster under consideration in order to scale up to the subsequent Cluster's Innovation level. With that purpose, to determine competitiveness indicators, where improvement is necessary for scaling up, the top five competitiveness indicators in the subsequent Cluster are considered. The logic of the proposed approach here is to identify the competitiveness indicators, which most affect the Innovation level of the subsequent Cluster and then to investigate the impact of an improvement obtained for these competitiveness indicators on the Innovation level of the current Cluster. For the analysis, referring to Table 4, only those top competitiveness indicators are considered where the competitiveness indicator mean values of the current and the subsequent Cluster under consideration significantly differ. That is, the Laggers have to benchmark the Developing Ones to improve their Innovation level. Table 8 gives the ranking of the indicators according to their explanatory power on Innovation level for each cluster. Among the 5 most important indicators of the Developing Ones, only On-the-job-training has a similar mean with that of the Laggers and so, it is not taken into account for this analysis. The remaining 4 indicators are considered as benchmark indicators and an analysis has been conducted by improving their level one at a time. In Table 9, the first column indicates the cluster under consideration. Each of the first four clusters is considered in this analysis with the fifth cluster ignored since it corresponds to the best one. The last column gives the seven states of Innovation and lists the probabilities for the highest three states. The first row of each cluster indicates the original states of the competitiveness indicators for that Cluster and the corresponding probabilities of Innovation. Then in the subsequent lines within the same Cluster, one of the identified competitiveness indicators is improved each time to the next higher state. Notice that the cases where the state improves to the higher level are framed in bold. The idea is to observe the improvement in Innovation through evidence observation. Notice the first case where we consider Cluster 1 (The Laggers). According to the original values, this Cluster on the Innovation variable has its highest probability for state 2, 50.2 %. When we improve the state of the variable Trustworthiness and confidence from its original state 4 to the higher state 5, then Innovation has its highest probability for state 3 with 44.7 %. In fact, as is also underlined by Fawcett *et al.* (2012), great innovator enterprises rely extensively on other members of their supply chain for the highest advantage from innovation. Similarly, Murphy (2002) analyses the social dimensions of innovation and reveals that trust is especially important for innovation as it improves the quality of information exchanges and encourages the development of strong intercommunity ties. Therefore, trust is at the heart of a collaborative capability.

Meanwhile it is also noteworthy that the evidence observation for the variable Trustworthiness and confidence does not only improve the stage of Innovation: a simultaneous improvement is also seen in Quality of demand conditions and Technological readiness. The last two mentioned cases are indicated with an upward arrow on the right hand side in the Table 9. It is already mentioned that demanding customers are one of the main drivers of innovation; for cultural reasons, customers may be more demanding in some countries than in others. For example, intense product competition and demanding customers encourage rising R&D spending and the development of new products and processes. However, this Quality of demand conditions in turn, depends upon the trust worthy and transparent financial markets with appropriate regulations to protect customers (Dobson and Safarian, 2008).On the other hand, as is also stated by Vize *et al*. (2013), the industry trust is of crucial importance for the firms, which are new to technology environment and likely to rely on an institutional trust mechanism to reduce their perceived risk and increase their technological readiness.

As indicated in Table 9, the same type of analysis is valid also for the three indicators Quality of demand conditions, Financial market efficiency, and Business sophistication. When, for example, the state of Business Sophistication is improved to state 3 from its original value of 2, we observe that the probability of Innovation occurring in state 3 increases to 72% from 36.8%, which is higher than the corresponding probability of 51.8% for the Innovation variable being in state 3 of the subsequent cluster, namely Cluster 2 (The Developing Ones). In fact, the quality of a country's overall networks and the quality of individual firms' operations and strategies are especially important for countries at an advanced stage of development. Note that Business sophistication includes local supplier quantity, supplier quality and state of cluster development among others. Parallel to our finding, Lai *et al.* (2013) explore the relationships among firms in a cluster on innovation performance and reveal that participation in industrial clustering is important for sustainable corporate development and that geographical proximity influences firms' innovation performance.