'WHY DO SOME NATIONS SUCCEED AND OTHERS FAIL IN INTERNATIONAL COMPETITION?' FACTOR ANALYSIS AND CLUSTER ANALYSIS AT EUROPEAN LEVEL

Ion Popa, Simona Cătălina Ștefan
Faculty of Management, Management Department, The Bucharest University of Economic Studies, Bucharest, Romania
iipopa@yahoo.com
simonacatalina_2006@yahoo.com

Abstract: *As stated by Michael Porter* (1998: 57), *'this is perhaps the most frequently asked economic question of our times.' However, a widely accepted answer is still missing. The aim of this paper is not to provide the BIG answer for such a BIG question, but rather to provide a different perspective on the competitiveness at the national level. In this respect, we followed a two step procedure, called "tandem analysis". (OECD, 2008). First we employed a Factor Analysis in order to reveal the underlying factors of the initial dataset followed by a Cluster Analysis which aims classifying the 35 countries according to the main characteristics of competitiveness resulting from Factor Analysis. The findings revealed that clustering the 35 states after the first two factors: Smart Growth and Market Development, which recovers almost 76% of common variability of the twelve original variables, are highlighted four clusters as well as a series of useful information in order to analyze the characteristics of the four clusters and discussions on them.*

Keywords: Cluster Analysis, competitiveness, Factor Analysis

JEL classification: C38, O11, O47

Introduction and background

Competitiveness is a complex concept, long debated by economists worldwide. The literature considers competitiveness as one of the essential dimensions of business performance.

Competitiveness issues (definition, measurement, determinants, growth strategies) can be addressed by presenting elements of differentiation and interrelation elements at microeconomic level (firms, organizations) and mezoeconomic level (economic sectors or branches), macroeconomic (country level) or mega economic (a number of countries - such as the European Union). (Ioncică, Petrescu and Ioncică, 2008: 76)

'Why do some nations succeed and others fail in international competition?' As stated by Michael Porte (1998: 57), this is perhaps the most frequently asked economic question of our times. Unlike competitiveness at the firm / organization level, the competitiveness at the national level has still no universally accepted definition, despite the widely use (sometimes abuse) of this concept. While it is clearly defined what means a competitive company / organization, not the same can be said about the notion of a competitive nation / country. The first which attempted to explain this concept were the classical theories of international trade competitiveness of countries, respectively the absolute advantage theory developed by Adam Smith, David Ricardo's theory of comparative advantage and even the theory of endowment with production factors (Heckscher-Samuelson Ohlim). (Ioncică, Petrescu and Ioncică, 2008: 77) In the current international context, characterized by economic globalization, the growing trend of liberalization of international trade and

technological progress, there are modern theories of competitiveness, such as the Michael Porter's (1998) competitive advantage theory.
Porter (1990) (1998) has emphasized that national competitiveness has become one of the central preoccupations of the government and industry in every nation. 'In the modern global economy, prosperity is a nation's choice. Competitiveness is no longer limited to those nations with a favorable inheritance. Nations choose prosperity if they organize their policies, lows and institutions based on productivity.' (Porter, 1998: 105) In this context, Michael Porter and K. Schwab (2007) define national competitiveness as the 'set of institutions, policies and factors that determine the level of productivity of a country.'
As pointed out by Popa and Vlăsceanu (2014: 19), at global and European level there are concerns regarding the ways of stimulating competitiveness by raising the employment rate of the labor force, increasing the GDP and aiming at a decent standard of life for the population. They identified a number of factors related with competitiveness, such as economic development, infrastructure, development, legislative regulations, technological access, labor market flexibility, the quality of business environment, while Klaus Schwab (2014) recognize a set of factors likely to be important for competitiveness and growth: investment in physical capital, infrastructure, education and training, technological progress, macroeconomic stability, good governance, firm sophistication and market efficiency.
The complexity of this influences make the mission to capture the essence the competitiveness in boosting the living standards of the population very difficult (Popa and Vlăsceanu, 2014: 19), therefore, along time at the international level, considerable efforts were undertaken aiming at compiling synthetic indicators able to capture all possible influences of various sets of factors.

Data and methods
The main data source was the Competitiveness Dataset (xls), the row data on which it was build The Global Competitiveness Report 2014-2015 issued by World Economic Forum (2014a). The report assesses the competitiveness landscape of 144 economies, providing insight into the drivers of their productivity and prosperity. (World Economic Forum, 2014b). The initial twelve variable of our analysis were: Institutions, Infrastructure, Macroeconomic environment, Health and primary education, Higher education and training, Goods market efficiency, Labor market efficiency, Financial market development, Technological readiness, Market size, Business sophistication and Innovation.
Our analysis covers all the 28 European Union member states, namely: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxemburg, Malta, Netherlands, Poland, Portugal, Romania, Slovak Republic, Slovenia, Spain, Sweden and United Kingdom plus Iceland, Macedonia, Montenegro, Serbia, Turkey, Norway and Switzerland. The analyses within this research rely on a dataset referring to the years 2013 and 2014 and were developed by employing the IBM SPSS Statistics 22.0 software package.
We followed a two step procedure called "tandem analysis". (OECD, 2008) First we will employ a Factor Analysis in order to reveal the underlying factors of the initial dataset followed by a Cluster Analysis on the object scores of the first two factors which aims classifying the 35 countries according to the main characteristics of competitiveness resulting from Factor Analysis.

1. Factor analysis is a multivariate analysis, which aims to explain the correlations manifested between a number of variables, called indicators or tests, using a smaller number of uncorrelated factors called common factors. (Ruxanda, 2009: 64) The model is given by (OECD, 2008):

$$
\begin{aligned}
X_1 &= a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + e_1 \qquad (1)\\
X_2 &= a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + e_2\\
&\dots\dots\dots\dots\\
X_Q &= a_{Q1}F_1 + a_{Q2}F_2 + \cdots + a_{Qm}F_m + e_Q
\end{aligned}
$$

where X_i *(i = 1,…,Q)* represents the original variables; a_{i1}, a_{i2}, ..., a_{im} are the factor loadings related to the variable X_i; F_1, F_2,...,F_m are *m* uncorrelated common factors; and e_i are the *Q* specific factors supposed independently and identically distributed with zero mean.

2. Subsequently we performed a Two Step Cluster Analysis, following a two-step procedure, as described by Bacher, Wenzig and Vogler (2004: 4), Chiu et al. (2001) and The SPSS TwoStep Cluster Component - Technical Report (2001: 3): Step 1: Pre-clustering of cases - a sequential approach is used to pre-cluster the cases. The aim is to compute a new data matrix with fewer cases for the next step. Step 2: Clustering of cases - a model based hierarchical technique is applied. Similar to agglomerative hierarchical techniques, the (pre") clusters are merged stepwise until all clusters are in one cluster. As we said above, the input variables in the Cluster Analysis were the scores of the two factors resulting from Factor Analysis.

Results and discussions

Descriptive statistics

In Table 1 there are displayed the descriptive statistics (mean, standard deviations and variance) for the initial twelve variables that will be used within the Factor Analysis procedures employed within the research, as well as the Romania's values related to the same variables. The scores given for Romania are all below the average of the 35 European countries analyzed, except for those for the criterion of Macroeconomic environment and Market size, where our country got a score above average.

From the above presented data, we conclude that Romania occupies a pretty bad position among European countries in terms of their competitiveness, but a much clearer picture will result from analyzes conducted throughout this paper.

Table 1: Descriptive Statistics (mean, standard deviation and variance)

		N	Mean	Std. Dev.	Variance	Romania
1	Institutions	35	4.487	0.849	0.721	3.557
2	Infrastructure	35	5.073	0.748	0.560	3.650
3	Macroeconomic environment	35	4.914	0.937	0.879	5.196
4	Health and primary education	35	6.264	0.331	0.110	5.507
5	Higher education and training	35	5.188	0.505	0.255	4.629
6	Goods market efficiency	35	4.686	0.407	0.166	4.180
7	Labor market efficiency	35	4.393	0.523	0.274	4.042
8	Financial market development	35	4.344	0.637	0.407	4.118
9	Technological readiness	35	5.267	0.709	0.503	4.485
10	Market size	35	4.163	1.005	1.011	4.437
11	Business sophistication	35	4.566	0.698	0.487	3.768
12	Innovation	35	4.085	0.870	0.758	3.283

Source: made by authors with SPSS Statistics 22.0

Factor analysis

Preliminary interpretation

In order to determine if our dataset is suitable for factor analysis, we first look to the *Correlation Matrix* to check if there is a patterned relationship among variables. Since the *Bartlett's Test of Sphericity* is significant ($p < .01$), the overall *Kaiser - Meyer - Olkin Measure (KMO) of Sampling Adequacy* (with a value of .945) and individual values for each variable (found in the diagonal of the *Anti - Image Correlation* Matrix) fall into the acceptable range (above .50), we can conclude that our dataset is suitable for factor analysis. The only one exception is *Market size* (.418) which we will keep under observation.

Factor extraction and rotation

In order to determine the most appropriate number of factors to be extracted, first, we used the Kaiser's criterion (which suggests retaining all the factors with eigenvalues greater than 1.00) in conjunction with the scree test. This three-factor solution was not considered appropriate since the 3rd factor was correlated with only one variable (as one can guess, Market size - the variable we kept under observation from the beginning). Tabacknick and Fidell (2007) recommend that for something to be labelled as a factor, it should have at least three variables, so for the final solution we retained only two factors. The final model computed with Extraction Method: Principal Axis Factoring and Rotation Method: Varimax with Kaiser Normalization, could be considered a good fit since, looking at the *Reproduced Correlation Matrix*, we found that the nonredundant residuals with absolute values greater than .05 are < 50%.

Having determined the number of retained factors, now we are concerned with interpretation. Table 2 presents the eleven variables significantly correlated with the two factors, arranged in descending order of loading. As recommended by Stevens (2002) due to the small size of the sample only the significant individual factor loadings (above 0.7) were retained. The high values of Cronbach's Alpha coefficient of reliability, presented on the last row of the Table 2, mean that there is evidence that the individual indicators measure the same underlying construct (0.7 is considered as an acceptable reliability threshold).

The fact that, In the Rotated Factor Matrix, factor loadings of all relevant variables are highly and positively related to only one factor each and we obtained only two factors could represent a big advantage for factor interpretation and spatial representation.

Table 2: Rotated Factor Matrix

	Factor 1	Factor 2	Communalities
Infrastructure	.937		.912
Business sophistication	.820		.932
Higher education and training	.794		.824
Innovation	.776		.935
Health and primary education	.773		.652
Technological readiness	.756		.833
Financial market development		.876	.793
Labor market efficiency		.782	.724
Macroeconomic environment		.736	.542
Institutions		.724	.920
Goods market efficiency		.701	.798

			Total
Eigenvalues	4.952	4.077	9.029
% of variance	41.264	33.973	75.237
Cronbach's Alpha	.951	.877	

Source: made by authors with SPSS Statistics 22.0

Naming the factors

The two factors resulting from factor analysis cumulate 75.237% of common variability of all the original twelve variables and will replace them in all subsequent analysis, therefore it is essential to have suggestive names and, on the other hand, to reflect as faithfully as possible the structure on which each of them was built. A meaningful interpretation of the factors is not straightforward but a big help is the fact that factor loadings of all relevant variables are highly and positively related to only one factor and the variables seem to cluster in a fairly suggestive way.

The first factor (the 2nd column of the Table 2) cumulates 41.264% of the common variability of all the twelve original variable and is strongly and positively correlated with six of them: Infrastructure (.937), Business sophistication (.820), Higher education and training (.794), Innovation (.776), Health and primary education (.773) and Technological readiness (.756), therefore the most appropriate name to define it would be *Smart Growth*.

The second factor (the 3rd column of the Table 2) cumulates 33.973% of the common variability of all the twelve original variable and is strongly and positively correlated with five of them: Financial market development (.876), Labor market efficiency (.782), Macroeconomic environment (.736), Institutions (.724) and Goods market efficiency (.701), therefore the most appropriate name to define it would be *Market Development*.

Factor scores

Using Bartlett method we produced factor scores. Factor scores are variables describing how much an individual would score on a factor and are correlated only with them own factor. These scores will be used later for cluster analysis. Another exceptionally helpful use of factor scores is the graphical representation in two-dimensional space of the position of the 35 states relative to latent unobservable characteristics represented by the common factors. On the basis of such representations, one can make an assessment of a global nature and are created prerequisites for performing multi-criteria comparisons. In Figure 1 are shown the 35 states in relation to the two factorial axes: Smart growth and Market development. As one can see, in terms of Market Development the best position is held by Norway and the worst position by Italy, while in terms of Smart Growth, Netherlands holds the highest score, and the lowest score - Bulgaria. If we look to the overall picture of the distribution of the 35 states in two-dimensional space created by the coordinates resulting from Factor Analysis, we can catch a grouping tendency, but for a greater scientific accuracy, in the subsequent research we will employ a Cluster Analysis procedure.

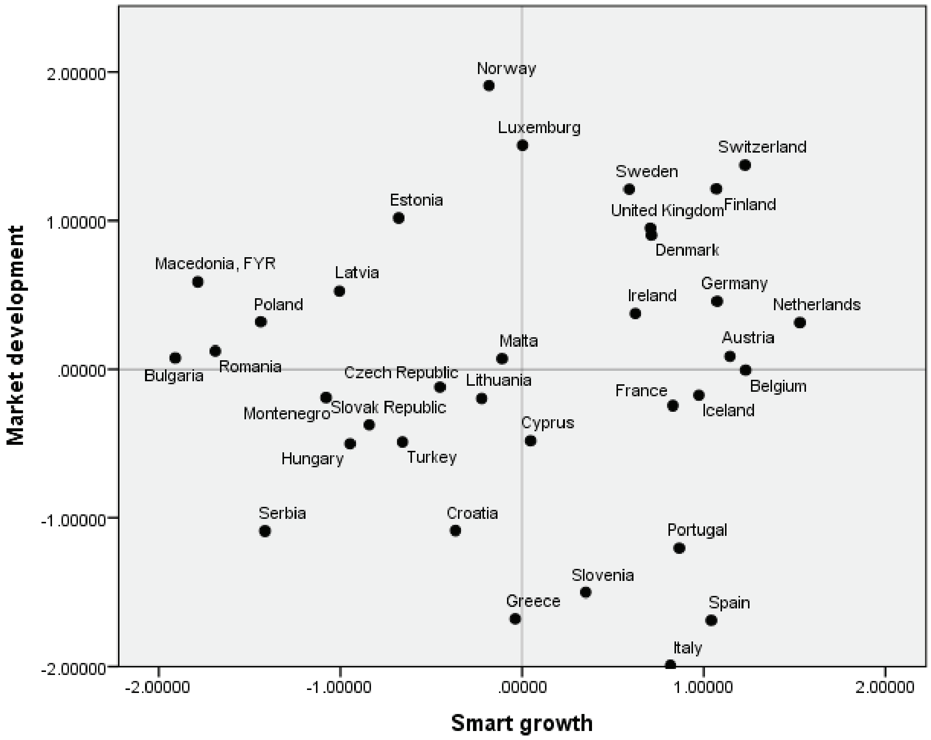

Figure 1: Scatter plot of observations
Source: made by authors with SPSS Statistics 22.0

Cluster analysis

For the purpose of identifying country groups that are homogeneous within themselves while also heterogeneous between each other based on the two new variables (named *Smart Growth* and *Market Development*), there was employed a Two Step Cluster Analysis, following a two-step procedure, as described above. Since the input variables were the scores resulted from the factor analysis, the assumption of independence and the normal distribution of the two variables were met. The evaluation variable was GDP per capita, theoretically known as related to competitiveness.This field will not be used to create the cluster model, but will give us further insight to the clusters created by the procedure.

The initial two cluster solution determined by SPSS was not considered appropriate for our analysis; therefore we set the number of cluster to four. The silhouette measure of cohesion and separation shows a good quality cluster solution (average silhouette = .6). The variable with the highest contribution to the determination of the final solution was Market Development (predictor importance = 1.00), while the contribution of Smart Growth was .76.

The cluster analysis revealed a series of useful information to analyze the characteristics of the four clusters and discussions on them. An overview of the four clusters based on the five variables depending on which they were formed, is presented below and is illustrated in Figure 2 and Table 3.

- Cluster 1 - include 7 countries (20% of total), namely: Switzerland, Finland, Sweden, United Kingdom, Denmark, Norway and Luxemburg, all in northern and western Europe. The most important predictor within the cluster is MD with a mean score of 1.295, followed by SG (mean score = .588). What particularizes the countries components of this cluster are the highest scores on both dimensions (factors) and the highest average value of GDP per capita, which is the evaluation variable (53,400 €). These data enables us to say that they would come on top in any ranking on the level of competitiveness.

- Cluster 2 - holds 7 countries (20% of total), namely: Netherlands, Germany, Belgium, Austria, Ireland, Iceland and France. The most important predictor within the cluster is SG with a mean score of 1.057, followed by MD (mean score = .115). From geographical point of view, this group is also composed mainly of countries in Western Europe. It seems that the overall level of competitiveness is comparable to that of the first cluster, but unlike it, the most important contribution and the highest score comes from the first factor - Smart growth. GDP per capita values are slightly lower than for cluster 1 but significantly higher than clusters 3 and 4.
- Cluster 3 - with 15 countries (42.9%) is by far the largest and the most heterogeneous (the highest values of the standard deviation of means for both factors) of the four clusters. It includes mainly countries in Central and Eastern Europe: Estonia, Malta, Cyprus, Lithuania, Czech Republic, Latvia, Slovak Republic, Turkey, Poland, Montenegro, Macedonia, Hungary, Romania, Bulgaria and Serbia. The most important predictor within the cluster is SG with a mean score of -.945, followed by MD (mean score = - .945). As one can see in Figure 2 and Table 3, cluster 3 has the lowest average scores for both dimensions of the four clusters, which makes likely location of these countries for the last places in the ranking of competitiveness. Moreover, the evaluation variable (GDP per capita) also recorded the lowest average value of the four clusters (10,621€).
- Finally, Cluster 4 holds the remaining 6 countries (17.1%) countries located mainly in southern Europe: Portugal, Spain, Italy, Slovenia, Croatia and Greece. The most important predictor within the cluster is MD with a mean score of -1.525, followed by SG (mean score = .444). What characterizes this cluster is that, despite quite high values for Smart Growth (similar to those of cluster 1), Market development presents unexpectedly low values, which may have negative repercussions on the overall level of competitiveness. Such opinion is strengthened by the fact that the average GDP per capita is only a third of the cluster 1.

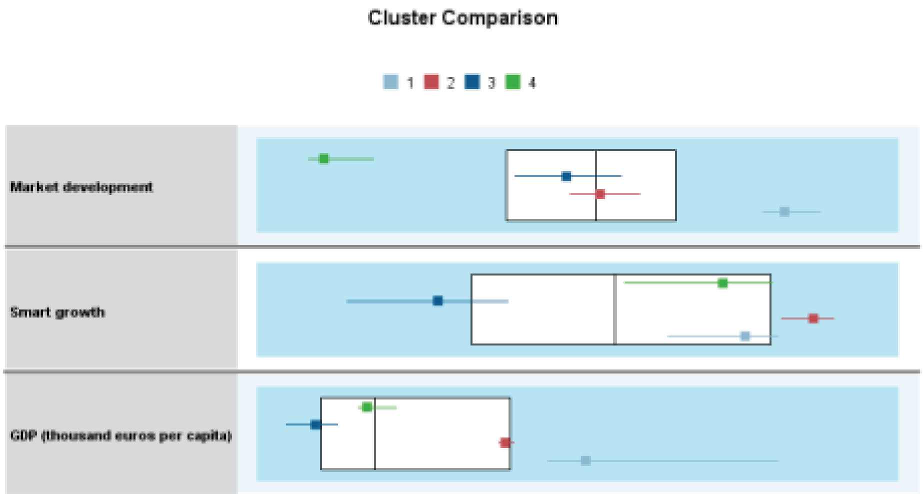

Figure 2: **Cluster comparison**

Source: made by authors with SPSS Statistics 22.0

Table 3: Cluster Cancroids

		Cluster				
		1	2	3	4	Romania
Cluster size		7 20.0%	7 20.0%	15 42.9%	6 17.1%	
Smart growth	Mean	.588	1.057	-.945	.444	-1.689
	Std. Dev.	.517	.290	.614	.560	
Market development	Mean	1.295	.115	-.048	-1.525	.123
	Std. Dev.	.344	.275	.527	.336	
GDP per capita (thousand euro)	Mean	53.400	34.514	10.621	18.050	7.100
	Std. Dev.	20.501	1.886	4.701	5.378	

Source: made by authors with SPSS Statistics 22.0

We assessed the predictive validity of the final cluster solution following the procedure recommended by Hair et al. (2010) investigating whether statistically significant differences exist across the clusters. Therefore, was performed a One-Way ANOVA analysis using GDP per capita as dependent variable and cluster membership as independent variable. GDP per capita was chosen due to its known theoretical relationship with the clustering variables, but not included in the cluster solution. As one can observe in Table 4, the results ($F = 31.758$, $p < .05$) proves the predictive capability of cluster solution for others key outcomes, which provides evidence of criterion validity.

Table 4: Assessing cluster solution criterion validity

Variable	Cluster member	Cluster mean (thousand euro)	F	Sig.
GDP per capita	1	53.400	31.758	.000
	2	34.514		
	3	10.621		
	4	18.050		

Source: made by authors with SPSS Statistics 22.0

4. Conclusions and further research

The findings of the present research reflect that the pattern in the initial dataset revealed by the Factor Analysis indicated that there were two factors or dimensions of country level competitiveness in the selected data set. Furthermore, the first competitiveness dimension revealed in the data (Smart Growth) corresponded to the components of others composite indices of competitiveness and to the objectives set by the Strategy Euro 2020 (a smart, inclusive and sustainable Europe). The subsequent Cluster Analysis carried out with the scores of the two factors resulting from Factor Analysis as impute variables, highlighted four clusters as well as a series of useful information in order to analyze the characteristics of the four clusters and discussions on them.

This finding may lead us to future directions of research that should investigate possible causal relationships between the objectives of Euro 2020 and the competitiveness level of EU member countries. Another possible research direction could follow the evolution in time of the four cluster solution developed within this paper.

References

Bacher, J., Wenzig, K. and Vogler, M. (2004) 'SPSS TwoStep Cluster - A First Evaluation', *Arbeits- und Diskussionspapiere*, vol. 2.

Chiu, T., Fang, D., Chen, J., Wang, Y. and Jeris, C. (2001) 'A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment', Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, 263-268.

Hair, J.F., Black, W.C., Bobin, B.J. and Anderson, N.E. (2010) *Multivariate data analysis*, 7th edition, Pearson Prentice Hall.

Ioncică, M., Petrescu, E.C. and Ioncică, D. (2008) 'Macro and microeconomic approaches to competitiveness in the services sector', *Revista de Marketing Online*, vol. 2, no. 1, pp. 76-82.

OECD (2008) *Handbook on Constructing Composite Indicators. Methodology and user guide*, OECD Publishing.

Popa, I. and Vlăsceanu, C. (2014) 'Competitiveness and Outsourcing in the New Economy', *Revista Economică*, vol. 66, no. 1, pp. 16-27.

Porter, M.E. (1990) 'The Competitive Advantage of Nations', *Harvard Business Review*, Marth - April, pp. 71-91.

Porter, M.E. (1998) *The Competitive Advantage of Nations - with a new introduction*, New York: The Free Press.

Ruxanda, G. (2009) *Analiza multidimensională a datelor - suport de curs*, București: Academia de Studii Economice. Școala Doctorală.

Schwab, K. (2014) *The Global Competitiveness Report 2014–2015*, Geneva: World Economic Forum.

Schwab, K. and Porter, M.E. (2007) *The Global Competitiveness Report 2007-2008*, Geneva: World Economic Forum.

Stevens, J.P. (2002) *Applied multivariate statistics for the social sciences*, 4th edition, Hillstade, NS: Erlbaum.

Tabacknick, B.G. and Fidell, L.S. (2007) *Using multivariate statistics (5th ed)*, Boston.

The SPSS TwoStep Cluster Component - Technical Report (2001), [Online], Available: http://www.spss.ch/upload/1122644952_The%20SPSS%20TwoStep%20Cluster%20Component.pdf [3 January 2015].

World Economic Forum (2014a) *The Global Competitiveness Report 2014–2015. Competitiveness dataset*, [Online], Available: http://reports.weforum.org/global-competitiveness-report-2014-2015/ [30 Marth 2015].

World Economic Forum (2014b) *The Global Competitiveness Report 2014–2015*, [Online], Available: http://reports.weforum.org/global-competitiveness-report-2014-2015/ [18 April 2015].