Journal of Social and Economic Statistics

*Vol. 2, No.1, Summer 2013* *© 2012 Published by JSES.*

# *COUNTRY CLASSIFICATION BASED ON LABOUR MARKET-RELATED CHARACTERISTICS* *

**Emilia ŢIŢAN[a], Constantin MITRUŢ[a], Adrian OŢOIU[a] , Remus DUMITRESCU[b], Daniela MANEA[a]**

**Abstract**

*This paper attempts to do a classification of countries based on the labour-related values prevalent in their societies. Using data from the World Values Survey, wave 4, an attempt is made to create country clusters based on four classes of values relevant to labour market behaviour: 1) values important to the society as a whole, 2) societal perceptions, attitudes and behaviours 3) attitudes toward work, and 4) economic and political environment. After using principal component analysis to reduce dimensionality of the data on the last three dimensions, cluster analysis was done on a mix of variables and factors. Our findings are partially consistent to the existing classifications of countries. Some results appear to challenge some of the existing classifications, as we found some less expected associations that appear to be correct, given the fact that they occur for several classes of values.*

**Keywords**: cultural clusters, labour market characteristics, societal values, cluster analysis, principal components analysis

**JEL Classification**: A13, J01, Z10

**Author's/Authors' Affiliation**

[a] – Bucharest Academy of Economic Studies, Department of Statistics and Econometrics, otoiu.adrian@gmail.com

[b] – University of Bucharest, Department of Physical Education

*****An earlier version of this paper was presented at *The 6th International Conference on Applied Statistics*, November 2012, Bucharest.

## 1. Introduction

When carrying out analysis and making international comparisons, countries are often ranked according to different criteria, in order to make meaningful comparisons. This type of classification stems from the fact that different countries are quite diverse in many respects, that there are other countries which have common characteristics, and that comparisons are more meaningful when made between similar countries.

The most widely known country classifications are based on the relative wealth of the countries. Thus, World Bank makes its classification based on gross national income per capita. A much broader concept is used by the UNDP, which publishes the HDI (Human Development Index) with which it ranks countries on the basis of four major sub-indexes which correspond to a long and healthy life, knowledge and a decent standard of living.

While these classifications are widely used and they make sense at an overall level, their relevance could be limited if the scope of analysis and comparisons focuses only on a few characteristics. For example, some of the richest countries are quite different in terms of the extend of social protection, provision of free health care services, and labour force institutional environment. Thus, it is widely known that counties like France and Germany have higher levels of social protection than the US. Similarly, it is well known that labour force attachment for women is lower in southern Europe compared to Northern Europe.

Thus, several attempts were made throughout the history of economic thought to classify countries according to their labour force characteristics. In his book 'The Protestant Ethic and the Spirit of Capitalism' Max Weber (1920) has pointed out that hard work and frugality were two characteristics in countries with Protestant roots, where people tend to be more focused on effective working practices compared to people in many Catholic countries. In his book, 'Capitalism Against Capitalism', Michel Albert (1991) has classified developed countries based on their cultural and economic system characteristics into two main categories, and distinguishes paternalistic economic systems like Germany and Japan from Anglo-Saxon countries which have laissez-faire policies. In Europe, it is observed by Fargher, Kesting, Lange and Pacheco (2008) that former Communist countries share common cultural characteristics that are different from Western countries, while in some respects they are similar to them.

However, few quantitative approaches have been done in order to formalize and validate these findings. A paper by Gupta, Hanges and Dorfman (2002) has come up with a cluster classification based on the GLOBE database, and checked the likelihood that a country belongs to one of the following clusters (South Asia, Anglo, Arab, Germanic Europe, Latin Europe, Eastern Europe, Confucian Asia, Latin America, Sub-Saharan Africa and Nordic Europe), using discriminant analysis on societal cultural values and practices, and leadership variables.

Ronen and Shenkar (1985) have done a study based on several surveys carried out by other authors and have synthesized them into a comprehensive picture. The scientific approach consisted in performing clusterizations and amending the results with the results obtained by the other authors, fitting somehow the results on the existing findings.

However, none of these approaches has produced an universally accepted classification of countries, and, moreover, they often consist in a review of other studies where results are merely matched to previous findings to see if the aprioric information and conclusions can be validated by the data analysis. We consider that this approach contradicts the scientific nature of exploratory analysis itself. The quantitative methods used belong to the 'unsupervised learning' class (term used in data mining and machine learning), where it is assumed that there is little or no prior knowledge of the relationships in the data, and that the analysis is expected to find these relationships. Therefore, the approach would be first to discover the relationships and then see if they bring some new insights against current research and mainstream opinions, and, in the context of clustering of countries, unfold some unexpected characteristics that challenge the existing assumptions.

This paper attempts to come up with a classification of countries based on the values that relate to the labour market behaviour of their citizens. We believe that these traits are essential in defining labour market attachment and behaviour in general. Thus, clustering countries based on these characteristics may lead to an improved analysis and to new insights as to how different countries may show similarities or differences on certain values, attitudes and perceptions that are prevalent in their societies.

## 2. Data and Methodology

Data comes from the 4th wave of the World Values Survey, conducted in 2005. This survey uses a representative sample drawn from each participating nation to get data on religion, gender roles, work motivations, democracy, good governance, social capital, political participation, tolerance of other groups, environmental protection and subjective wellbeing. Included nations range from rich to poor, from authoritarian systems to democracies, from all major cultural zones, to a total of 57 countries (depending on availability of financial resources and logistic support at the time of conducting the survey).

Among all variables and classes of variables that may have a direct impact on labour market behavior, we believed that there are four classes of variables that would have a major impact on labour force behaviour:

- variables that address general attitudes prevalent in a society: attitudes toward work, leisure, religion, family, friends
- variables that show societal perceptions, attitudes and behaviours with potential impact to the world of work
- variables that explicitly show attitudes toward work
- variables that describe the economic and political environment of a country.

A description of the variables used can be found in Table 1.

A more thorough discussion of the variables is included in section three, and precedes analyses which are focused on each and every variable class mentioned above.

The following technique was used in analyzing the data and forming the clusters: first an initial clustering was run on all variables to get a basic initial result. Afterwards, an exploratory analysis was done on each variable class in order to see the relationships between them. After obtaining a clear picture of the variables and their structure, we proceeded in carrying out a separate clusterization for each main variable class, as shown in table 1 using K-means clustering. For the case of highly correlated variables, we ran principal components analysis and decided whether we can use factors instead of variables in the clusters. Decision was made based on the percentage of total variation explained by the factors obtained. As a rule of thumb we used factors wherever 80% of more of the variation was explained by the first two or three factors. In the end, clustering was done on all variable classes using the factors previously chosen, and taking into consideration the validity of the preliminary results obtained.

**Table 1. List of variables used for the analysis**

| **Societal values** | **Societal perceptions, attitudes and behaviours** | **Attitudes toward work** | **Economic and political environment** |
|---|---|---|---|
| - family important | - men make better business executives than women do | - first choice, if looking for a job | - having a democratic political system |
| - friends important | - one of main goals in life has been to make my parents proud | - second choice if looking for a job | - governments tax the rich and subsidize the poor. |
| - leisure time | - I seek to be myself rather than to follow others | - to develop talents you need to have a job | - religious authorities interpret the laws. |
| - politics important | - live up to what my friends expect | - humiliating to receive money without having to work for | - people choose their leaders in free elections. |
| - work important | - I decide my goals in life by myself | - people who don´t work turn lazy | - people receive state aid for unemployment. |
| - religion important | - satisfaction with the financial situation of household | - work is a duty towards society | - civil rights protect people's liberty against oppresion |
| - less importance placed on work | - most people can be trusted | - work should always come first | - the economy is prospering |
| - more emphasis on technology | - membership in organizations | | - criminals are severely punished. |
| - greater respect for authority | - woman as a single parent | | - people can change the laws in referendums. |
| - more emphasis on family life | - being a housewife fulfilling | | - women have the same rights as men. |
| | - men make better political leaders | | |
| | - university is more important for a boy | | |
| | **Desirable child qualities** | | |
| | independence, hard work, feeling of responsibility, imagination, tolerance and respect others, thrift saving, determination perseverance, | | |

| | | | |
|---|---|---|---|
| | religious | | |
| | **Trust** | | |
| | family, neighbourhood, people you know, people you meet the first time, people of another religion, people of another nationality | | |
| | **Justifiable** | | |
| | claiming government benefits, avoiding a fare on public transport, cheating on taxes, someone accepting a bribe | | |

A first run of the k-means clustering algorithm revealed the following cluster structure:

**Table 2. Initial clustering using all variables**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| Australia | Jordan | Andorra | Ethiopia | Bulgaria | Burkina | China | Brazil |
| Canada | Vietnam | Argentina | Indonesia | Moldova | Ghana | Georgia | Malaysia |
| Finland | | Chile | Morocco | Serbia | India | Romania | Mexico |
| Norway | | Cyprus | Turkey | Ukraine | Mali | | Thailand |
| Sweden | | Germany | | | South Africa | | |
| Uruguay | | Poland | | | Trinidad | | |
| USA | | S. Korea | | | Zambia | | |
| | | Slovenia | | | | | |
| | | Spain | | | | | |

A first inspection of the data reveals the shortcomings of this analysis. We could observe that, in some cases, similar countries were grouped together as in cluster 1, 5 and 7, while other clusters included countries that were completely different in terms of levels of development and cultural heritage. It was thus clear that a refined approach based on a more careful exploration of variables within variable classes was necessary in order to derive a more meaningful classification of countries.

## 3. Main Results and Interpretation

An analysis of the correlations among variables revealed that in three classes there are groups of variables exhibiting strong correlations, while this was not generally the case when we looked at variables that belong to different classes. Details of the observed correlations that were used for dimensionality reduction within variable classes will be given below as we detail results of the partial analyses performed on each of the four variable classes.

### *3.1. Societal Values Analysis*

The first group of variables analyzed referred to general values that are prevalent in a country, and reveal some of the characteristics of that society. The variables incorporate answers given on questions that asked respondents what they consider important (work, leisure, family, friends, work, politics, religion) and what are the desired changes in some of them (more emphasis in family life, less importance to work, more emphasis on technology, greater respect for authority).

There were no significant correlations observed among these variables, except for the importance placed on work and the importance placed on religion, which confirmed once again Max Weber's (1920) findings.

Factor analysis carried out on these variables are shown in Appendix 1. While the first 5 factors explained 84% of the total variance, the scree plot was inconclusive in enabling us to decide how many factors to use.

Therefore, we chose to perform clustering on all variables for this class as such, for which we obtained the following clusters.

The analysis showed a high heterogeneity within the clusters, which almost always contained countries that are known to be dissimilar with respect to their societal values and tradition e.g. Brazil and Georgia, Poland and USA. In some cases, clusters groped countries with similar levels of economic development (e.g. cluster 2, cluster 9, cluster 10), while in some other a mix of rich and poor countries were grouped together (cluster 7 and 8). Another consideration in looking at the results with a certain degree of caution was the relatively unbalanced structure of the clusters, where some contained 1 to 2 countries, and others 9 and over.

**Table 3. Initial clustering using all variables**

| 1 | 2 | 3 | 4-6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|
| Brazil | Andorra | Chile | China | Argentina | Bulgaria | Burkina | Japan |
| Georgia | Australia | Colombia | Vietnam | Canada | Finland | Egypt | Norway |
| Ghana | Britain | Peru | Indonesia | Cyprus | Germany | Mali | S. Korea |
| Guatemala | France | Romania | Ethiopia | India | Russia | Morocco | Sweden |
| Iran | Hong Kong | | Thailand | Italy | Slovenia | Zambia | |
| Jordan | Netherlands | | | Moldova | Taiwan | | |
| Malaysia | N. Zealand | | | Poland | Ukraine | | |
| Mexico | Spain | | | Serbia | Uruguay | | |
| Rwanda | | | | USA | | | |
| S Africa | | | | | | | |
| Trinidad | | | | | | | |
| Turkey | | | | | | | |

### *3.2. Societal Perceptions, Attitudes and Behaviours Analysis*

The second group of values comprises a larger number of variables of a more heterogeneous nature. Core values considered desirable in shaping children's education as a preparation for a successful life ahead, the degree of trust that exist among people in a society represented by the degree of socialization, attitudes that are found to be justifiable by the members of a society, membership in organizations of different natures (e.g. work-related, leisure related, religious), perceptions of different roles to be played by different members of a society (men, women, lone parents), were chosen to classify countries into relevant clusters.

For this group there are strong correlations among certain groups of variables. Membership in different organizational types, where the strength of involvement was taken into consideration by indicating whether respondents were active members or passive members, showed a very strong degree of cross-correlation. Thus, activity or inactivity with respect to organizations tends to be very similar across different types of organizations, and define in a more general way the willingness of the members of a society to commit or not their free time towards a wide variety of goals that form the purpose of these societies.

**Table 4. Correlations among membership and activity in different types of organizations**

| | Church | Sport | Arts, Music, Recreation | Union | Political Parties | Environmental | Professional | Charitable/ Humanitarian | Consumer Organization |
|---|---|---|---|---|---|---|---|---|---|
| Church | 1 | 0.5 | 0.66 | 0.44 | 0.7 | 0.62 | 0.59 | 0.6 | 0.53 |
| Sport | 0.5 | 1 | 0.89 | 0.6 | 0.43 | 0.57 | 0.73 | 0.81 | 0.63 |
| Arts, Music, Recreation | 0.66 | 0.89 | 1 | 0.55 | 0.63 | 0.78 | 0.82 | 0.84 | 0.73 |
| Unions | 0.44 | 0.6 | 0.55 | 1 | 0.55 | 0.5 | 0.62 | 0.65 | 0.76 |
| Political Parties | 0.7 | 0.43 | 0.63 | 0.55 | 1 | 0.76 | 0.7 | 0.59 | 0.73 |
| Environmental | 0.62 | 0.57 | 0.78 | 0.5 | 0.76 | 1 | 0.83 | 0.77 | 0.82 |
| Professional | 0.59 | 0.73 | 0.82 | 0.62 | 0.7 | 0.83 | 1 | 0.91 | 0.77 |
| Charitable/ Humanitarian | 0.6 | 0.81 | 0.84 | 0.65 | 0.59 | 0.77 | 0.91 | 1 | 0.73 |
| Consumer Organization | 0.53 | 0.63 | 0.73 | 0.76 | 0.73 | 0.82 | 0.77 | 0.73 | 1 |

Another strong correlation was found on the beliefs that men make better political leaders, better executives and that going to university is more important for a boy.

Factor analysis confirmed these findings and enabled us to isolate factors that explain a large proportion of the variance in the data. For membership, results in Appendix 2 shows that the first two factors explain about 80% of the variance. Using this result in conjunction with the scree plot, we retained the first two factors for the subsequent clustering for this

variable class. For men roles in society, the first factor explained about 93% of the variance, and was retained for the final clustering performed for this variable class (Appendix 3).

Clusterization was then performed for the second variable class using the extracted factors (Table 5).

**Table 5. Clusterization based on societal perceptions, attitudes and behaviours**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| Brazil | Norway | Chile | Argentina | Andorra | Egypt | France | Ethiopia | India |
| Mexico | Sweden | China | Bulgaria | Australia | Georgia | Germany | Ghana | Mali |
| Peru | | Cyprus | Burkina | Britain | Jordan | Spain | Indonesia | |
| | | Italy | Poland | Canada | Morocco | | Rwanda | |
| | | Malaysia | Romania | Finland | Turkey | | S Africa | |
| | | Moldova | Russia | Netherlands | | | Trinidad | |
| | | S. Korea | Serbia | Switzerland | | | Zambia | |
| | | Slovenia | Ukraine | US | | | | |
| | | Taiwan | Uruguay | | | | | |
| | | Thailand | | | | | | |
| | | Vietnam | | | | | | |

In this case, clusterization was more clear and yielded more intuitive results. Only few countries belonged to clusters where most countries were from another region of the world. Thus, all Anglo-Saxon countries except South Africa belonged to one cluster; that same cluster grouped countries that were similar in terms of culture and development (Finland the Netherlands and Switzerland). Norway and Sweden formed a distinct cluster, as did Brazil, Mexico and Peru, while most former Communist countries from Eastern Europe were gathered into one cluster into which two Latin-American countries found their way in. Perhaps this latter finding is a new insight that merits further attention. Surprising results were found for cluster 7 which grouped France, Germany and Spain. Cluster 3 grouped together several former communist countries and southern European countries along with Asian countries. Also, Indonesia and Trinidad clustered together with African countries.

In conclusion, clusterization performed on societal attitudes, perceptions and behaviours was far more consistent with existing research than the one based on societal values. While it is known that applying quantitative methods to cluster countries is fraught with inconsistencies (i.e. countries that should not normally be grouped with other countries), this may give some insights as to how reality has evolved past common knowledge and engrained beliefs, and make us reconsider how countries are similar in some respects.

### *3.3. Analysis of Attitudes Toward Work*

This is perhaps the most relevant group of variables with respect to the purpose of this paper, variables that address in a direct manner the way citizens of a given nation think and behave relative to work and labour markets. Two variables include a score for four major attributes of a job: money, feeling of being important, feeling safe, or an amenable work environment. In order to refine the importance of these major features of a job, original WWS data was reprocessed. Based on the answers obtained for the questions first choice and, respectively, second choice if looking for a job, we have computed scores for each of the four attributes where the first choice was given a rating of 1 and the second choice a rating of 0.75. When the first and second choice coincided, the aggregate score was revised down to 1.25 so as to account for a strong preference for a given attribute, by considering at the same time that respondents did not in fact answer correctly the 'second choice when looking for a job' question. Other variables included general attitudes towards work such as: work should always come first, it is humiliating to receive money without having to work for, etc.

Very strong correlations were found among these variables. Among key work attributes, income and safety had a high positive correlation, while being inversely correlated with the feeling of importance given by a job, and amenable work environment attributes. Perceptions that emphasized the importance of work in society were also highly correlated amongst themselves.

**Table 6. Correlations among key work attributes**

| | vimp | vinc | vlike | vsa |
|---|---|---|---|---|
| vimp | 1 | -0.76 | 0.72 | -0.74 |
| vinc | -0.76 | 1 | -0.68 | 0.4 |
| vlike | 0.72 | -0.68 | 1 | -0.68 |
| vsa | -0.74 | 0.4 | -0.68 | 1 |

**Table 7. Correlations among perception relating to work**

| | v51 | v52 | v53 | v54 |
|---|---|---|---|---|
| v51 | 1 | 0.78 | 0.64 | 0.76 |
| v52 | 0.78 | 1 | 0.72 | 0.75 |
| v53 | 0.64 | 0.72 | 1 | 0.83 |
| v54 | 0.76 | 0.75 | 0.83 | 1 |

| | |
|---|---|
| v51- | it´s humiliating to receive money without having to work for |
| v52- | people who don´t work turn lazy |
| v53- | work is a duty towards society |
| v54- | work should always come first |

Results obtained using factor analysis were in line with these findings, and enabled us to isolate factors that explained a large proportion of the variance in the data. For both groups of variables, the first two factors explained over 75% of the variance, and they were further retained for doing the clusterization.

Clusterization results were rather surprising (Table 8). Different types of countries were mixed within many of the clusters. To some extent, results have captured some similarities between some countries if we think about economic development or cultural background (clusters 5-6, cluster 8, cluster 1 to give a few examples). However, there is a lot of heterogeneity within many clusters that question the usefulness and validity of the analysis, and raise the question of the relevance of using variables revealing attitudes toward work for country clusterization.

**Table 8. Clusterization based on perceptions and attitudes towards work**

| 1 | 2 | 3 | 4 | 5-6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|
| Brazil | Bulgaria | Argentina | Guatemala | Norway | Egypt | Andorra | Germany |
| Italy | Burkina | Chile | Jordan | Switzerland | Ethiopia | Australia | Moldova |
| Mexico | Cyprus | China | Mali | Sweden | Ghana | Canada | Serbia |
| Peru | Georgia | Indonesia | Morocco | | India | Finland | Spain |
| Slovenia | S Korea | Malaysia | Rwanda | | Romania | Japan | Ukraine |
| Trinidad | Thailand | Poland | Turkey | | | USA | Uruguay |
| | Zambia | S Africa | Vietnam | | | | |

### *3.4. Analysis of Economic and Political Environment*

The last class of variables refer to the economic and political environment of a country and are focused on the different features of a democratic society (free elections, rule of law, human rights, social assistance, etc.). While there are strong correlations among these variables, these are rather sparse, and we cannot see a regular pattern of strong cross-correlation among variables that are similar in nature. Given all this, we considered useful to perform a factor analysis in order to capture the main drivers behind these variables that could help us cluster the countries.

Principal component analysis revealed the existence of three main factors that explain about 77% of the total variance. Using the results of the scree plot, we decided to pick the first three factors and run a clusterization on them, as adding an additional 4th or fifth factor would not improve by much the explanatory power of these factors.

Results obtained were consistent to a certain extent, but generated strange results. Cluster one contains a wide variety of countries, while cluster 4 grouped Argentina and Uruguay along with many developed countries from all continents. Most former Communist countries

were grouped together with China and Taiwan. To sum up, we obtained once again puzzling results that did not allow for a meaningful classification.

**Table 9. Clusterization based on perceptions about the economic and political environment**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| Cyprus | Ethiopia | Ghana | Argentina | Burkina | Bulgaria | Egypt | Andorra |
| Indonesia | Malaysia | Jordan | Australia | India | China | Georgia | Brazil |
| Iran | S. Korea | Mali | Finland | Morocco | Moldova | | Britain |
| Poland | Serbia | | France | Rwanda | Peru | | Canada |
| Trinidad | Thailand | | Germany | S. Africa | Romania | | Chile |
| Turkey | | | Netherlands | Zambia | Russia | | Mexico |
| Vietnam | | | Norway | | Taiwan | | Slovenia |
| | | | Sweden | | Ukraine | | Spain |
| | | | Switzerland | | | | |

***3.5. Comprehensive analysis using selected variables for all variable classes***

Using all variables and the factors obtained for the last three variable classes, we have performed a final clustering to see if dimensionality reduction and previous findings will enable us to come up with a more uniform classification of countries that would be more reliable.

**Table 10. Clusterization done using variables and factors from all variable classes**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| Bulgaria | Norway | Brazil | Argentina | Ethiopia | Andorra | Malaysia | Egypt | Ghana |
| China | Sweden | Mexico | Burkina | India | Australia | Thailand | Georgia | Indonesia |
| Moldova | | Peru | Chile | Mali | Canada | | Jordan | S. Africa |
| Poland | | | Cyprus | | Finland | | Morocco | Trinidad |
| S Korea | | | Germany | | Switzerland | | Romania | Zambia |
| Serbia | | | Slovenia | | USA | | Turkey | |
| Ukraine | | | Spain | | | | Vietnam | |
| | | | Uruguay | | | | | |

Results have a higher consistency, with many similar countries clustered together and very few unexpected results. Most Communist and former Communist countries belong to cluster 1, where we also have China and an outlier, South Korea. Norway and Sweden form a distinct cluster, as do Brazil, Mexico and Peru and Malaysia and Thailand. Some other cluster group together more diverse countries. Thus most English-speaking countries form a cluster where Andorra, Finland and Switzerland are also included. India is grouped together with

Ethiopia and Mali. Most Southern African countries are grouped together in a cluster where Indonesia belongs too. Surprisingly, Romania Georgia and Vietnam are grouped together with Northern African countries and with Turkey. Cluster 4 has a rather strange composition, grouping a few Latin American countries together with Spain, Germany, Cyprus and Slovenia.

While results may not be entirely satisfactory, even when accounting for the fact that clusterization is not a perfect technique, fraught with shortcomings, as Ronen and Shenkar (1985) have pointed out, we should not rule out as unreliable some surprising results that keep repeating after several clustering approaches were performed on a narrower set of variables.

Very often Argentina and Uruguay cluster with other European or Anglo-Saxon countries rather than with other Latin American countries. The Netherlands and Finland tend to group with Anglo-Saxon countries. Indonesia and Trinidad often group with other Southern African countries. Germany has a bit of trouble grouping together with other countries, and it is often found in groups that include intuitively dissimilar countries with respect to cultural background and development levels. Romania shows surprising affinities with Georgia and other Northern African countries. India founds itself clustered very often with African countries. Yet another surprising results is that Asian countries never group in a homogeneous cluster of a larger size as do Anglo-Saxon and former Communist countries.

## 4. Conclusion

A quantitative-based clustering of the nations based on their values relevant to labour markets and labour force attachment does not always yield intuitive results with respect to our prior knowledge and expectations as to how countries should be grouped together. This analysis thus faces strong challenges with respect to the results obtained and their defendability, given not only previous research and literature in the field, but also common knowledge about different countries and their cultural and societal characteristics.

However, we consider that this stance may be flawed, on the grounds that critiques try to challenge the features of the statistical methods employed which yield counterintuitive results, as pointed out by Ronen and Shenkar (1985), while the same methods have found wide applicability in many scientific area and have yielded valid results. It is the purpose of these statistical techniques to find similarities and dissimilarities in the data based on the features of their underlying models, and structure the data accordingly. This is often referred in data mining as ‘unsupervised learning’ and has the purpose to structure data into similar classes based on the properties of the relatively large number of variables analyzed. In practice these methods often do have some degree of misclassification due to model

specification and data properties, but nevertheless yield useful results that led to important discoveries.

The two-step approach used in this paper has helped us to better understand that some societal attitudes and perceptions are highly correlated: membership in different organizations, men roles in society, desired attributes for a job and perceptions about work in general, and perceptions about the economical and political environment.

Partial clustering based on the four variable classes did yield mixed results. Only the clustering based on societal perceptions, attitudes and behaviours resulted in fairly homogeneous results, where most countries were clustered with others within their geographical regions and levels of development. Clustering based on attitudes and perceptions toward work has yielded fairly inconsistent results. Further to running a final clusterization using all uncorrelated variables and factors extracted for the variables for which high degrees of correlation was observed, we obtained clusters that were mostly homogeneous, with some surprising counterintuitive results.

We have thus observed that Anglo-Saxon, African, Latin American and Former Communist countries tend to group into homogeneous clusters. Asian countries tend to be included in different clusters, and only in a few cases they manage to form their own cluster. Finland, the Netherlands and Switzerland tend to cluster with Anglo-Saxon countries, while South Africa often clusters with other Southern African countries. Turkey, Romania and Georgia tend to have similarities with Northern African countries, while Argentina and Uruguay tend to enter in clusters dominated by European or Anglo-Saxon countries. Indonesia and Trinidad often cluster with Southern African countries, and India tends to enter in clusters with other African countries.

Norway and Sweden have strong cultural characteristics that put them into a distinctive cluster. Contrary to widely accepted beliefs, Germany does not cluster more often with other Northern countries than it does with other European countries.

## References

Albert, M. (1991). *Capitalisme contre capitalisme*, Editions du Seuil.

Gupta, V., Hanges, P., Dorfman, P. (2002). Cultural Clusters: methodology and findings. *Journal of World Business*, no. 37, pp. 11-15.

Hayek, F. (1986). *The Road to Serfdom*. Routledge and Kegan Paul.

Weber, M, (1920). *Etica protestanta si spiritul capitalismului*. ed. Humanitas, 2007, Original version: *Die protestantische Ethik und der Geist des Kapitalismus*. Gesammelte Aufsätze zur Religionssoziologie.

Ronen, S., Shenkar, O. (1985). Clustering Countries on Attitudinal Dimensions: A Review and Synthesis. *The Academy of Management Review*, **10** (3), pp. 435-454.

Fargher, S., Kesting, S., Lange, T., Pacheco G. (2008). Cultural heritage and job satisfaction in Eastern and Western Europe. *International Journal of Manpower*, **29** (7), pp.630-650.

The GLOBE survey [online] http://business.nmsu.edu/programs-centers/globe/, accessed on July 4, 2012

The World Bank [online] http://data.worldbank.org/about/country-classifications, accessed on July 5, 2012

World Values Survey [online] http://www.worldvaluessurvey.org/, accessed on July 5, 2012

## APPENDIX 1

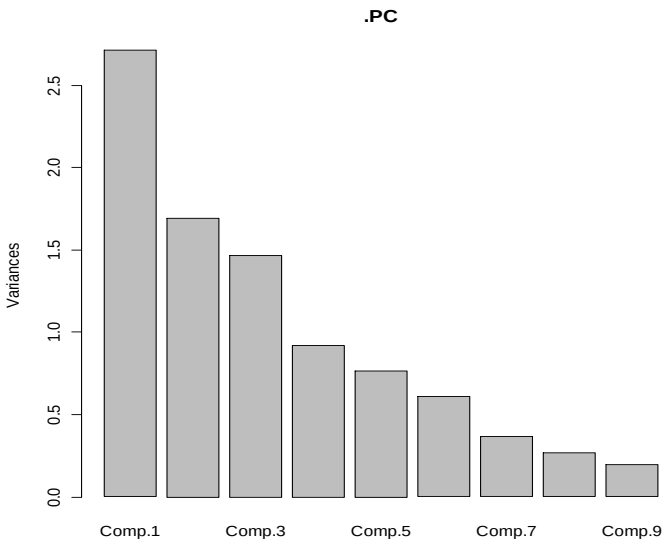
Factor analysis results for societal values

```
> .PC$sd^2  # component variances
   Comp.1    Comp.2    Comp.3    Comp.4    Comp.5    Comp.6    Comp.7
Comp.8
2.7120077 1.6932807 1.4684110 0.9223446 0.7653089 0.6095033 0.3660297
0.2680794
   Comp.9
0.1950347

> unclass(loadings(.PC))  # component loadings
        Comp.1     Comp.2        Comp.3      Comp.4      Comp.5      Comp.6
v4   0.2492426  0.3161775 -0.4768314195  0.24665878  0.40527280 -0.24494254
v5  -0.2921890  0.2348930 -0.5016442048 -0.31982510  0.04973023 -0.35172735
v6  -0.1585900 -0.1263259 -0.6100947207  0.03904647 -0.50341219  0.45384641
v7  -0.1949916  0.4322907  0.1140531690 -0.71525624 -0.01845069  0.13501319
v8   0.3954579  0.4162647 -0.0314105007  0.09811798 -0.14369554  0.50291781
v9   0.4292976  0.4498135  0.0159289481  0.07357262 -0.16588943 -0.14788923
v77  0.4414415 -0.2162650  0.0421015217 -0.45918131  0.26206095  0.22922099
v78  0.3924394 -0.1832563 -0.0004151541 -0.21226400 -0.62237592 -0.51049092
v79  0.3158615 -0.4313487 -0.3643461736 -0.22980793  0.27444702  0.06438254

> summary(.PC) # proportions of variance
Importance of components:
                          Comp.1    Comp.2    Comp.3    Comp.4     Comp.5
Standard deviation     1.6468174 1.3012612 1.2117801 0.9603877 0.87481937
Proportion of Variance 0.3013342 0.1881423 0.1631568 0.1024827 0.08503433
Cumulative Proportion  0.3013342 0.4894765 0.6526333 0.7551160 0.84015032
                           Comp.6     Comp.7    Comp.8     Comp.9
```

```
Standard deviation     0.78070694 0.60500390 0.5177638 0.44162731
Proportion of Variance 0.06772259 0.04066997 0.0297866 0.02167052
Cumulative Proportion  0.90787291 0.94854288 0.9783295 1.00000000
```



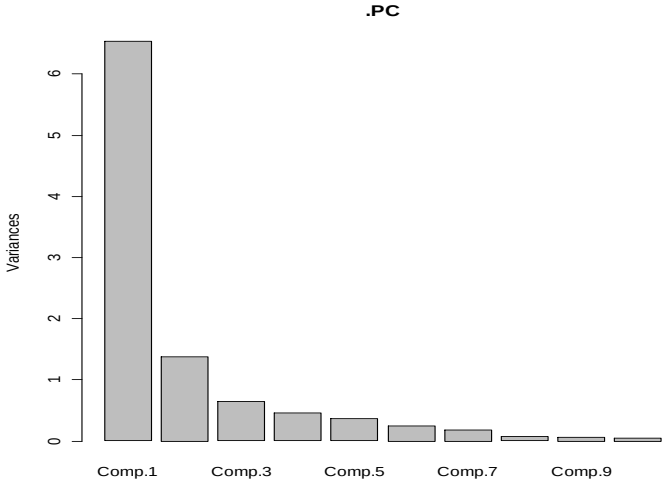

## APPENDIX 2

Factor analysis results for membership and activity in different organization types

```
> .PC$sd^2  # component variances
    Comp.1     Comp.2     Comp.3     Comp.4     Comp.5     Comp.6
Comp.7
6.53276792 1.37975340 0.64424649 0.45925092 0.36112608 0.24830519
0.18672128
    Comp.8     Comp.9    Comp.10
0.07405268 0.06454112 0.04923491

> summary(.PC) # proportions of variance
Importance of components:
                          Comp.1    Comp.2     Comp.3     Comp.4     Comp.5
Standard deviation     2.5559280 1.1746290 0.80264967 0.67768055 0.60093766
Proportion of Variance 0.6532768 0.1379753 0.06442465 0.04592509 0.03611261
Cumulative Proportion  0.6532768 0.7912521 0.85567678 0.90160187 0.93771448
                           Comp.6     Comp.7      Comp.8      Comp.9
Standard deviation     0.49830231 0.43211258 0.272126221 0.254049446
Proportion of Variance 0.02483052 0.01867213 0.007405268 0.006454112
Cumulative Proportion  0.96254500 0.98121713 0.988622397 0.995076509
                           Comp.10
Standard deviation     0.221889414
```

```
Proportion of Variance 0.004923491
Cumulative Proportion  1.000000000
```

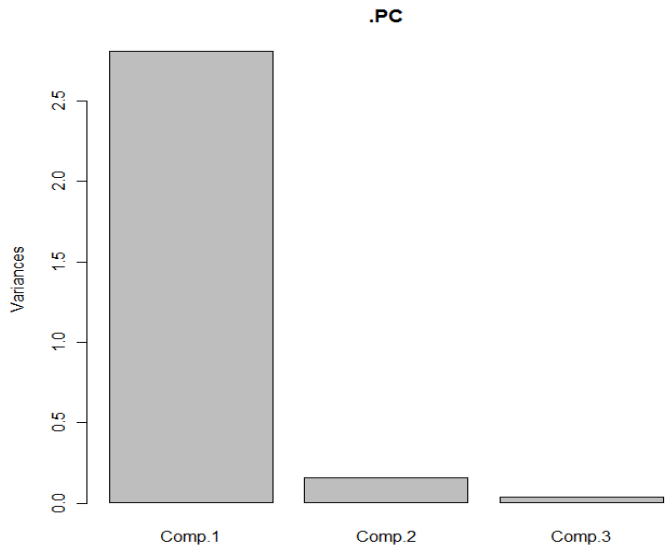



## APPENDIX 3

Men roles in society

```
> .PC$sd^2  # component variances
    Comp.1     Comp.2     Comp.3
2.80617733 0.15619634 0.03762633

> summary(.PC) # proportions of variance
Importance of components:
                          Comp.1     Comp.2     Comp.3
Standard deviation     1.6751649 0.39521683 0.19397508
Proportion of Variance 0.9353924 0.05206545 0.01254211
Cumulative Proportion  0.9353924 0.98745789 1.00000000
```

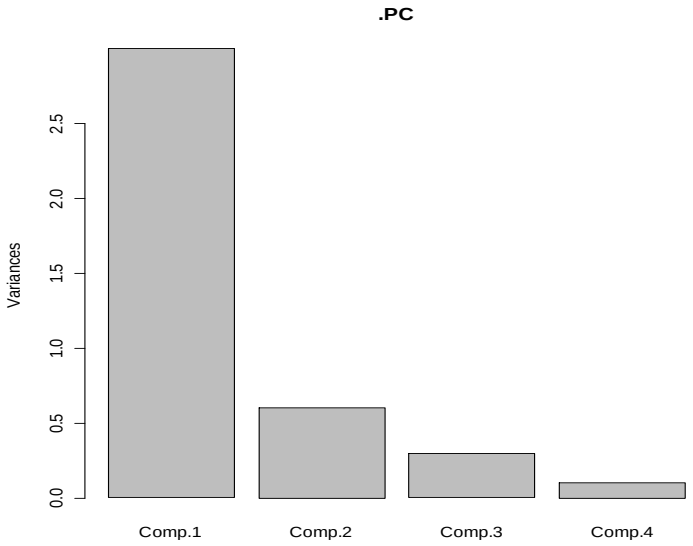



## APPENDIX 4

Attitudes and values towards work

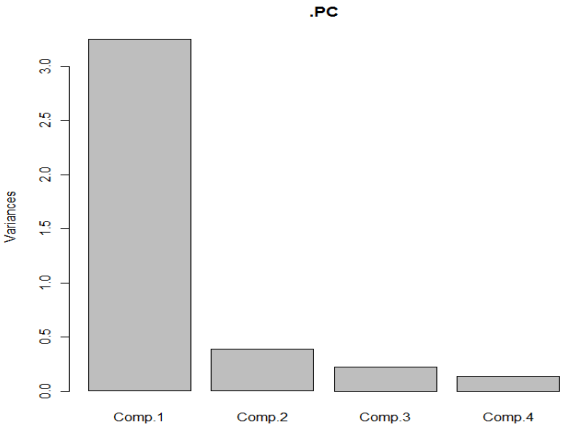
```
> .PC$sd^2  # component variances
   Comp.1    Comp.2    Comp.3    Comp.4
2.9988551 0.6024648 0.2963317 0.1023484

> summary(.PC) # proportions of variance
Importance of components:
                          Comp.1    Comp.2     Comp.3     Comp.4
Standard deviation     1.7317203 0.7761861 0.54436353 0.31991943
Proportion of Variance 0.7497138 0.1506162 0.07408291 0.02558711
Cumulative Proportion  0.7497138 0.9003300 0.97441289 1.00000000
```




```
> .PC$sd^2  # component variances
   Comp.1    Comp.2    Comp.3    Comp.4
3.2472431 0.3866417 0.2262620 0.1398532

> summary(.PC) # proportions of variance
Importance of components:
                          Comp.1     Comp.2    Comp.3     Comp.4
Standard deviation     1.8020109 0.62180519 0.4756700 0.37396947
Proportion of Variance 0.8118108 0.09666042 0.0565655 0.03496329
Cumulative Proportion  0.8118108 0.90847121 0.9650367 1.00000000
```



## APPENDIX 5

Economic and political environment variables

```
> .PC$sd^2  # component variances
   Comp.1    Comp.2    Comp.3    Comp.4    Comp.5    Comp.6    Comp.7
Comp.8
3.5141254 3.0686038 1.1401066 0.6261641 0.5545909 0.4358528 0.2458074
0.2046120
   Comp.9   Comp.10
0.1354699 0.0746671

> summary(.PC) # proportions of variance
Importance of components:
                          Comp.1    Comp.2    Comp.3     Comp.4     Comp.5
Standard deviation     1.8746001 1.7517431 1.0677578 0.79130532 0.74470863
Proportion of Variance 0.3514125 0.3068604 0.1140107 0.06261641 0.05545909
Cumulative Proportion  0.3514125 0.6582729 0.7722836 0.83489999 0.89035909
                           Comp.6     Comp.7    Comp.8     Comp.9
Comp.10
Standard deviation     0.66019150 0.49578962 0.4523406 0.36806230
0.27325282
Proportion of Variance 0.04358528 0.02458074 0.0204612 0.01354699
0.00746671
Cumulative Proportion  0.93394437 0.95852510 0.9789863 0.99253329
1.00000000
```