Worldwide clustering of the corruption perception

Michal Paulus, Ladislav Kristoufek

Institute of Economic Studies, Faculty of Social Sciences, Charles University, Opletalova 26, 110 00, Prague 1, Czech Republic

Abstract

We inspect a possible clustering structure of the corruption perception among 134 countries. Using the average linkage clustering, we uncover a well-defined hierarchy in the relationships among countries. Four main clusters are identified and they suggest that countries worldwide can be quite well separated according to their perception of corruption. Moreover, we find a strong connection between corruption levels and a stage of development inside the clusters. The ranking of countries according to their corruption perfectly copies the ranking according to the economic performance measured by the gross domestic product per capita of the member states. To the best of our knowledge, this study is the first one to present an application of hierarchical and clustering methods to the specific case of corruption.

Keywords: clustering, corruption perception

1. Introduction

Corruption has been understood as an important aspect forming societies as well as economic connections for a long time. Recently, it has also become more evident that grasping corruption quantitatively can yield important additions to models in social sciences and especially economics. Various studies examine corruption in fields of microeconomics, macroeconomics, institutional economics or international trade and its influence on economic performance, efficiency and decision making. Corruption is thus in the centre of interest of many empirical and theoretical works [1, 2, 3].

The most prominent topic regarding corruption is its impact on economic growth showing a negative influence of corruption on the growth [4, 5]. Another attractive domain of corruption effects is in the inequality studies, which show that corruption increases inequality [6]. The increased corruption also tends to boost public investments but these yield lower returns [7]. The higher corruption not only lowers the public investment returns but

---

also the level of private investments with a higher public-to-private investment ratio as a result [4, 8, 9].

In general, the empirical studies reveal negative effects of corruption on economic performance in terms of lower growth, private investments, higher inequality or wasting of talents. Except for the studies concerning the macroeconomic performance, corruption studies are becoming more popular in the international trade area as well. The theoretical effect of corruption on mutual trade is not straightforward as it can have both negative and positive effects. In the case of an adverse impact, corruption can be viewed as an additional tax because of bribes, a complicated enforcement of mutual agreements or an institutional safety [10]. The other perspective is completely opposite. Corruption helps to facilitate mutual business and to guarantee basic rights when the institutions are of a low quality.

When we abandon theoretical considerations and explore empirical literature, we cannot unanimously conclude that corruption has only a negative role in FDI (foreign direct investments) or trade, although it seems to be mostly the case [11, 12, 13]. Studies disrupting the prevailing negative effects of corruption are present as well. Egger & Winner [14] show that corruption can stimulate FDI in the short as well as the long term. They argue that a corrupt country attracts investments from other corrupt countries. In other words, a country attracts investments from countries with a similar institutional quality [15]. Brada *et al.* [16] go further and they develop a theoretical model showing that the most beneficial position in the FDI engagement is to be a "medium corrupt" country. Such a country has advanced technologies, good institutions but also knowledge of the corruption culture. Hence they can be engaged in an FDI mediation with corrupt as well as non-corrupt countries. If a country is highly corrupt, it has a problem attracting investments from low corrupt countries and vice versa.

Such an ambiguous role of corruption is found in the international trade analyses also. Horsewood & Voicu [10] report adverse effects of corruption on mutual trade. On the other hand, de Jong & Bogmans [17] find that corruption can stimulate imports under special circumstances. Some further results can be found in Refs. [18, 19, 20].

Analysis of corruption has been of interest in the interdisciplinary research as well. Shao *et al.* [21] demonstrate a power-law functional dependence between corruption level and economic factors, such as country wealth and foreign direct investments per capita. Podobnik *et al.* [22] report that an increase of the corruption perception index leads to an increase of the annual GDP per capita growth rate between 1999 and 2004. In addition, the authors introduce a new measure to quantify the relative corruption between countries based on their respective wealth as measure by GDP per capita. Podobnik *et al.* [23] further propose a production function (divided into the private and public sectors) where GDP depends on market capitalization, the public or private section workforce and competitiveness level to quantify the public sector efficiency. It is demonstrated that the less corrupt countries receive more investments that the more corrupt ones. Finally Podobnik *et al.* [24] argue that it is improbable that a country would rapidly improve its corruption rank.

Even though the effects of corruption on various economic indicators and measures

have been often investigated, an important message remains hidden beneath the results. The reported findings suggest that there are at least several groups of countries which are differentiated according to their corruption level and these groups tend to trade with each other more and they invest into similar countries as well. However, there is no study examining the actual clustering of countries according to their corruption or their corruption perception. Here, we aim to fill this gap in the literature by bridging an economic topic with tools of complex systems and interdisciplinary physics – clustering analysis. In the following section, we describe the utilized methodology in some detail. The next section describes the analyzed dataset followed by results presentation and interpretation. The last section concludes. We find that countries indeed cluster according to their corruption levels and we identify four well-divided clusters. Moreover, these clusters nicely correspond to a stage of development of the member countries. The level of corruption is thus found to be tightly connected to country's development.

2. Methodology

The clustering analysis has now quite a long history in economic and financial applications mainly due a pioneering work of Mantegna [25] utilizing minimum spanning trees and hierarchical trees to uncover a hierarchical structure in the US stocks. The methodology has then been applied to various markets such as stocks and stock indices [26, 27, 28, 29, 30, 31, 32, 33, 34, 35], foreign exchange rates [36, 37, 38, 39], import/export networks [40], interest rates [41], and commodities [42, 43, 44]. Most of the studies focus on correlation analysis, using the correlation matrix as a starting point. The dissimilarity (distance) matrix is constructed using the correlations which then allows for the spanning and hierarchical trees construction. However, our analysis cannot use such an approach and we must apply an alternative one due to a specific nature of the corruption data and a different aim of the analysis.

We analyze a corruption clustering of various countries. For each country, we get a time series of the Freedom from Corruption (COR) index of the Heritage Foundation which ranges between 0 and 100 with higher values meaning lower corruption (more details are given in the next section). We are primarily interested in a common level of the index as well as in a co-movement between indices of different countries. Therefore, we cannot use a standard correlation procedure which demeans the series and thus the information about the corruption level is lost. Instead of a correlation coefficient, we utilize the simple Euclidean distance d_{xy} between time series (or in general vectors) $\{x_t\}$ and $\{y_t\}$ defined as

$$d_{xy} = \sqrt{\sum_{t=1}^{T} (x_t - y_t)^2} \tag{1}$$

where T is the time series length. Even though there are various possibilities for defining distances [45, 46], we opt for the most basic Euclidean one[1]. We thus have a distance

---

[1]We have applied other distance metrics as well and the results remain qualitatively practically intact.

between time series as a measure of dissimilarity between corruption levels in pairs of countries forming a distance matrix D, which is symmetric with zeros on the diagonal.

For the hierarchical structure construction, we utilize the average linkage clustering approach (also known as UPGMA). The algorithm starts with nodes in a network treated as a separate cluster each. The closest pair is identified based on the lowest distance in the distance matrix D (except for the diagonal elements). This closest pair forms a new cluster and it is substituted into the distance matrix instead of the original two elements. The new cluster is assigned the average of the distances of all elements of the newly formed cluster. This is repeated until we have a single cluster [47, 48, 49]. Such construction can be seen as a cautious one as it uses the average distance for the new cluster instead of minima or maxima of the alternative approaches[2].

The resulting hierarchical structure is standardly reported in the form of a dendrogram – a tree diagram (from Greek *dendron* for a tree and *gramma* for drawing). The diagram shows the most important connections as an output of the clustering analysis and a possible hierarchical structure in the analyzed dataset. The most standard way of drawing a dendrogram is in a form of roots of a tree which is mostly of a rectangular shape. For large datasets and mainly a large number of time series, alternative methods can be utilized for a better graphical representation[3].

3. Data

We examine the corruption clustering using the Freedom from Corruption (COR) index of the Heritage Foundation between years 1996 and 2014. Only the countries with statistics for each year within the analyzed period are included in the final dataset. This accounts for 134 countries. The Freedom from Corruption is based primarily on the Corruption Perception Index (CPI) constructed by the Transparency International. The CPI is a 0-10 range index with 0 meaning the most corrupt country and 10 the opposite. The COR is constructed by simply multiplying the CPI by 10. Hence the COR is on 0-100 scale. If the CPI index is unavailable for a country, the Heritage Foundation estimates the index using expert information from other reliable sources [4]. However, it needs to be noted that the number of countries not covered by the CPI is very low.

The CPI is not an index of an objective actual state of corruption but of the perception of corruption in the public sector of a country. The Transparency International uses several sources to derive the index[5]. The institutions ask country experts to assess the corruption

---

[2]In the Results section, we also provide two alternative clustering linkage methods – complete [47] and Ward's [50] algorithms.

[3]The analysis has been performed in RStudio 0.98.1028 via R 3.1.1 using the *ape* and *stats* packages.

[4]U.S. Department of Commerce, Economist Intelligence Unit, Office of the U.S. Trade Representative and official government publications of each country. Additional information can be found at http://www.heritage.org/index/freedom-from-corruption.

[5]E.g. African Development Bank Governance Ratings, Economist Intelligence Unit Country Risk Ratings, IMD World Competitiveness Yearbook, Political Risk Services International Country Risk Guide or World Justice Project Rule of Law Index.

level in the country in many aspects. The data is then gathered and the CPI is constructed to reflect the level of corruption in a country perceived by its experts.

Of course, the CPI has some natural weaknesses. The most important one is that the CPI is derived from the experts' judgement. The results can be obviously biased by the "elite" position of the respondent. The perception of corruption could be different if the "ordinary ones" were asked. Another problem lays in the comparison of the CPI scores between years. Temporal changes in ranking can be caused by changes in the country CPI sample or methodology. To overcome these shortcomings, several alternative corruption indices have been developed (e.g. the Global Corruption Barometer, the Bribe Payers Index or the WB's Worldwide Governance Indicators – Control of Corruption index). Even though all the criticism points to the relevant CPI shortcomings, the index is still the key measure of corruption level across countries. The discussion about CPI validity shows difficulties concerning the corruption measurement which can never be objective and must somehow rely on perceptions. The CPI still offers the largest database of corruption indicator with sufficient historical time series and it remains the most suitable resource for cross-country corruption studies.

4. Results

We apply the average linkage clustering technique based on the Euclidean distances to find a possible hierarchical structure between corruption levels of 134 studied countries. In Fig. 1, we present the final depiction of clustering between analyzed states[6]. Due to a large number of the examined countries, we opt for a spherical representation of the dendrogram. In the centre of the figure, the distance between clusters is the highest and the further we get from the centre of the circle, the closer the clusters or countries are.

We observe that the network of countries breaks down into four well-defined clusters[7]. To help reader with orientation in Fig. 1, we separate the clusters by additional dashed lines. Furthermore, we list the countries according to their cluster membership in Table 1 as well. A very interesting clustering structure emerges.

Cluster #1 is mainly formed of developed countries, i.e. the USA, Japan, the Western states of the EU, Hong Kong, Singapore, Australia and New Zealand, and it forms a solid piece of network structure with the average COR of 82.92 ± 1.82. Members of this cluster thus have very low corruption levels. Compared to the values of other clusters, which

---

[6]In the Appendix, we also present Figs. 2 and 3 which are based on the complete and Ward's linkage algorithms, respectively. Even though the details of the hierarchical structure are not identical, the interpretation of the results is qualitatively very close to what we present for the average linkage clustering in the main text. The results are thus quite robust to the linkage algorithm selection.

[7]The definition of clusters is quite arbitrary here and we label the main clusters based on a visual inspection. If desirable, the number of clusters could be reduced to two – Clusters #1 and #2 together, and Clusters #3 and #4 together. The main message of the results would remain very similar. Either way, the reported results, their structure and implications presented in this section show that such cluster selection is reasonable.

are reported in Table 1, the first cluster is very well separated from the others and a considerable distance to other clusters is evident.

The cluster is clearly separated from the others not only in terms of the corruption level but also in terms of the economic development. The average GDP per capita[8] in the group is 52,138 current USD[9]. There is no country from the African continent. Both most developed countries from the North and Central America (the USA and Canada) are members of the cluster as well as is the only one and also the most developed country from the Latin America (Chile), the most developed states from Oceania (New Zealand and Australia), except for Kuwait the most developed nations on the Asian continent (Singapore, Japan, Hong Kong and UAE) and 13 most developed countries of Europe. The lowest GDP per capita is less than 7 times lower than the highest one. In other words, the cluster of the least corrupt countries is composed of the most developed nations across all continents. Hence we see that low corruption is clearly connected with economic development.

Cluster #2 is more heterogeneous as it contains various EU states such as Malta, Cyprus, Spain, Portugal and Slovenia (i.e. somewhere between the most developed Western states and the post-communist East) as well as a mix of Taiwan, Israel, Barbados, Botswana and others for which the connection can be seen usually only for small groups (such as a mini-cluster of oil exporters Oman, Bahrain and Kuwait). The average COR of 59.21 ± 1.47 still suggests a decent level of corruption.

This cluster has a lower average GDP per capita of about half the GDP of the most developed cluster (23,521 USD). It consists of the fourth most developed African state (Botswana), the third and fifth most developed countries of the Central and North America, and the second most developed country of the Latin America. Also five (including the first one) from the eleven most developed states in Asia are members of the cluster. The ratio between the lowest and the highest GDP per capita is less than 8. The cluster hence contains states which can be labeled as the upper-middle class of economic development. They exhibit significantly lower economic performance in terms of per capita GDP but still much higher than the other states. These countries are the "runner-up" leaders on their continents and they are trying to catch up the leading group. Hence we can regard the cluster as a "successfully transitioning" or a "catching up" group. The cluster is not only the second richest one but it also has the second lowest corruption level.

Cluster #3 is then the biggest one with 60 members and the average COR of 24.15 ± 0.78, i.e. the lowest value among clusters suggesting the highest level of corruption. The basket of countries is quite diverse here ranging from the new EU members (Bulgaria and Romania) to Russia and countries in between (Belarus, Moldova, the Ukraine) to China and India as well many African, Asian and a few Latin American countries. Several interesting implications arise from this structure. First, Bulgaria and Romania seem to

---

[8]Gross domestic product (GDP) for year 2012 is reported. GDP is a widely used measure of overall economic performance of a given country.

[9]Data source: World Development Indicators, WB, year 2012.

be institutionally very different from the rest of the EU as no other EU country falls into the most corrupt cluster. Even the newest EU member state (Croatia) is a member of the last cluster which possesses lower levels of corruption. Second, all BRIC countries (Brazil, Russia, India and China), which historically form a group of large, fast-growing economies, belong to the cluster. These countries are thus still well-behind the developed countries with respect to corruption environment and behavior. And third, various authoritarian regimes fall into this cluster as well – Azerbaijan, Iran, Laos, North Korea, Syria, Turkey, Venezuela and Vietnam. The corruption levels are thus tightly connected to the form of reign in a country. Note that most of the countries in the least corrupt cluster are standard, western-type democracies.

From the GDP per capita point of view, the third cluster is again quite diverse. The average GDP per capita is the lowest one with 3,888 USD. However, a range between the richest and the poorest country in the cluster is immense with a ratio of 53 – between Malawi (287 USD) and Russia (14,090 USD). Nonetheless, we can regard the cluster as a group of countries with the lowest economic performance. The cluster contains not only the poor countries or countries in civil war situations (e.g. Libya or Syria) but also transition states with a high future economic potential (e.g. Brazil or Turkey). It is likely the highest corruption can be a relevant obstacle for the development of these countries and hence the corruption-economic development correlation seems to be valid even in this case.

Cluster #4 is somewhere between the previous two clusters with the average COR of 41.28 ± 0.98. This is again quite a heterogeneous cluster (the ratio between the the richest and the poorest country is 45) with several interesting subgroups. First, the whole Visegrad group (the Czech Republic, Hungary, Poland and Slovakia), post-communist countries in the Central Europe, as well as the Baltic post-communist states (Latvia and Lithuania) are members of this cluster[10]. Second, the newest member state of the EU – Croatia – is also part of the cluster and it thus readily joins the post-communist countries of the region, overtaking Bulgaria and Romania. Third, some of developed countries such as Italy and Greece fall into this cluster as well which might give a notion about economic problems and its causes in these states. And fourth, practically the whole region of the North Africa (Algeria, Egypt, Morocco and Tunisia) is inside the cluster. From the economic perspective, the last cluster falls within the second and the third one with the average GDP of 9,751 USD. The connection between corruption and economic development is thus supported even by the last detected cluster.

5. Conclusions

We have inspected a possible clustering structure of the corruption perception among 134 countries. Using the average linkage clustering, we have uncovered a well-defined hierarchy in the relationships among countries. Four main clusters have been identified and they suggest that countries worldwide can be quite well separated according to their perception of corruption. Moreover, we have found a strong connection between corruption

---

[10]Estonia is not in the dataset due to incomplete time series.

levels and a stage of development inside the clusters. The ranking of countries according to their corruption perfectly copies the ranking according to the economic performance measured by the gross domestic product per capita of the member states. Even though our analysis does not (as it cannot) discuss potential causality between the two, the reported results remain promising. We believe that our analysis, which is the first one to the best of our knowledge, can be used as a starting point of various analyses in social sciences which struggle to find an appropriate mechanism of separating countries according to such a specific criterion as corruption. We are convinced that the appropriate clustering can be beneficial especially for studies concerning international trade where the corruption is gaining increasing attention and the studies lack rigorous differentiation of corrupt countries.

Acknowledgements

References

References

[1] S.J. Wei. How taxing is corruption on international investors. *The Review of Economics and Statistics*, 82:1–11, 2000.

[2] J. Svensson. Eight questions about corruption. *Journal of Economic Perspectives*, 19:19–42, 2005.

[3] A. Shleifer and R. Vishny. Corruption. *Quarterly Journal of Economics*, 109:599–617, 1993.

[4] P. Mauro. Corruption and growth. *The Quarterly Journal of Economics*, 110:681–712, 1995.

[5] P. Mauro. *Corruption and the global economy*, chapter The Effect of Corruption on Growth, Investment, and Government Expenditure, pages 83–107. Institute for International Economics, 1997.

[6] S. Gupta, H. Davoodi, and R. Alonso-Terme. Does corruption affect income inequality and poverty? *Economics of Governance*, 3:23–45, 2002.

[7] M.E. Haque and R. Kneller. Public investment and growth: The role of corruption. *CGBCR Discussion Papers*, 98:1–21, 2008.

[8] H. Davoodi and V. Tanzi. Corruption, public investment, and growth. *IMF Working Papers*, 97:139, 1997.

[9] S. Tarhan. Public investment and corruption in an endogenous growth models. *MPRA Paper*, 21319:1–35, 2008.

[10] N. Horsewood and A.M. Voicu. Does corruption hinder trade for the new eu members? *Economics: The Open-Access, Open-Assessment E-Journal*, 6(47):1–28, 2012.

[11] P. Egger and W. Hannes. How corruption influences foreign direct investment: A panel data study. *Economic Development and Cultural Change*, 54:459–486, 2006.

[12] M. Busse and C. Hefeker. Political risk, institutions and foreign direct investment. *European Journal of Political Economy*, 23:397–415, 2007.

[13] M. Habib and L. Zurawicki. Country-level investments and the effect of corruption some empirical evidence. *International Business Review*, 10:687–700, 2001.

[14] P. Egger and H. Winner. Evidence on corruption as an incentive for foreign direct investment. *European Journal of Political Economy*, 21:932–952, 2005.

[15] A. Cuervo-Cazura. Who cares about corruption? *Journal of International Business Studies*, 37:807–822, 2006.

[16] J.C. Brada, Z. Drabek, and F. Perez. The effect of home-country and host-country corruption on foreign direct investment. *Review of Development Economics*, 16:640–663, 2012.

[17] E. De Jong and C. Bogmans. Does corruption discourage international trade? *European Journal of Political Economy*, 27:385–398, 2011.

[18] P. Dutt and D. Traca. Corruption and bilateral trade flows: Extortion or evasion? *The Review of Economics and Statistics*, 92:843–860, 2010.

[19] R.K. Goel and I. Korhonen. Exports and cross-national corruption: A disaggregated examination. *Economic Systems*, 35:109–124, 2011.

[20] S. Thede and N.-A. Gustafson. The multifaceted impact of corruption on international trade. *The World Economy*, 35:651–666, 2012.

[21] J. Shao, P.Ch. Ivanov, B. Podobnik, and H.E. Stanley. Quantitative relations between corruption and economics factors. *European Physical Journal B*, 56:157–166, 2007.

[22] B. Podobnik, J. Shao, D. Njavro, P.Ch. Ivanov, and H.E. Stanley. Influence of corruption on economic growth rate and foreign investments. *European Physical Journal B*, 63:547, 2008.

[23] B. Podobnik, B.E. Baaquie, S. Bishop, D. Njavro, and B. Li. Growth versus government management improvement during economic downturn. *Scientific Reports*, 3:1612, 2013.

[24] B. Podobnik, D. Horvatic, D.Y. Kenett, and H.E. Stanley. The competitiveness versus the wealth of a country. *Scientific Reports*, 2:678, 2012.

[25] R.N. Mantegna. Hierarchical structure in financial markets. *European Physical Journal B*, 11:193–197, 1999.

[26] J.-P. Onnela, A. Chakraborti, K. Kaski, and J. Kertesz. Dynamic asset trees and portfolio analysis. *European Physical Journal B*, 30:285–288, 2002.

[27] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68:056110, 2003.

[28] J.-P. Onnela, K. Kaski, and J. Kertesz. Clustering and information in correlation based financial networks. *European Physical Journal B*, 38:353–362, 2004.

[29] G. Bonanno, G. Caldarelli, F. Lillo, S. Micciche, N. Vandewalle, and R. Mantegna. Networks of equities in financial markets. *European Physical Journal B*, 38:363–371, 2004.

[30] M. Tumminello, T. Di Matteo, T. Aste, and R. Mantegna. Correlation based networks of equity returns sampled at different time horizons. *European Physical Journal B*, 55:209–217, 2007.

[31] J.M. Brida and W.A. Risso. Multidimensional minimal spanning tree: The dow jones case. *Physica A*, 387:5205–5210, 2008.

[32] C. Eom, G. Oh, W.-S. Jung, H. Jeong, and S. Kim. Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series. *Physica A*, 388:900–906, 2009.

[33] T. Di Matteo, F. Pozzi, and T. Aste. The use of dynamical networks to detect the hierarchical organization of financial market sectors. *European Physical Journal B*, 73:3–11, 2010.

[34] C. Eom, O. Kwon, W.-S. Jung, and S. Kim. The effect of a market factor on information flow between stocks using the minimal spanning tree. *Physica A*, 389:1643–1652, 2010.

[35] A. Nobi, S.E. Maeng, G.G. Ha, and J.W. Lee. Effects of global financial crisis on network structure in a local stock market. *Physica A*, 407:135–143, 2014.

[36] T. Mizuno, H. Takayasu, and M. Takayasu. Correlation networks among currencies. *Physica A*, 364:336–342, 2006.

[37] M.J. Naylor, L.C. Rose, and B.J. Moyle. Topology of foreign exchange markets using hierarchical structure methods. *Physica A*, 382:199–208, 2007.

[38] W. Jang, J. Lee, and W. Chang. Currency crises and the evolution of foreign exchange market: Evidence from minimum spanning tree. *Physica A*, 390:707–718, 2011.

[39] M. Keskin, B. Deviren, and Y. Kocakaplan. Topology of the correlation networks among major currencies using hierarchical structure methods. *Physica A*, 390:719–730, 2011.

[40] E. Kantar, B. Deviren, and M. Keskin. Hierarchical structure of Turkey's foreign trade. *Physica A*, 390(20):3454–3476, 2011.

[41] B. Tabak, T. Serra, and D. Cajueiro. The expectation hypothesis of interest rates and network theory: The case of Brazil. *Physica A*, 388:1137–1149, 2009.

[42] B. Tabak, T. Serra, and D. Cajueiro. Topological properties of commodities networks. *European Physical Journal B*, 74:243–249, 2010.

[43] L. Kristoufek, K. Janda, and D. Zilberman. Correlations between biofuels and related commodities before and during the food crisis: A taxonomy perspective. *Energy Economics*, 34:1380–1391, 2012.

[44] L. Kristoufek, K. Janda, and D. Zilberman. Regime-dependent topological properties of biofuels networks. *European Physical Journal B*, 86:40, 2013.

[45] K.V. Mardia, J.T. Kent, and J.M. Bibby. *Multivariate Analysis*. Academic Press, 1979.

[46] I. Borg and P. Groenen. *Modern multidimensional scaling. Theory and applications.* Springer, 1997.

[47] T. Sorensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. *Biologiske Skrifter*, 5:1–34, 1948.

[48] B.S. Everitt, S. Landau, and M. Leese. *Cluster analysis.* London: Arnold, 2001.

[49] D. Defays. An efficient algorithm for a complete link method. *The Computer Journal*, 20:364–366, 1977.

[50] F. Murtagh and P. Legendre. Wards hierarchical agglomerative clustering method: Which algorithms implement wards criterion? *Journal of Classification*, 31:274–295, 2014.

Table 1: **Countries in clusters.** Clusters from Fig. 1 are listed here and the average values with standard errors are reported as well. The countries are listed alphabetically within the clusters.

Cluster #1 – average 82.92 ± 1.82				
Australia	Austria	Belgium	Canada	Chile
Denmark	Finland	France	Germany	Hong Kong
Ireland	Japan	Luxembourg	Netherlands	New Zealand
Norway	Singapore	Sweden	Switzerland	United Arab Emirates
United Kingdom	USA			

Cluster #2 – average 59.21 ± 1.47				
Bahamas	Bahrain	Barbados	Botswana	Cyprus
Estonia	Israel	Kuwait	Malta	Oman
Portugal	Slovenia	Spain	Taiwan	Uruguay

Cluster #3 – average 24.15 ± 0.78				
Albania	Azerbaijan	Bangladesh	Belarus	Bolivia
Brazil	Bulgaria	Burkina Faso	Burma	Cameroon
China	Colombia	Cuba	Dominican Republic	Ecuador
Ethiopia	Fiji	Georgia	Guinea	Guyana
Haiti	Honduras	India	Indonesia	Iran
Ivory Coast	Kenya	Laos	Lebanon	Lesotho
Libya	Madagascar	Malawi	Mali	Mauritania
Moldova	Mozambique	Nepal	Nicaragua	Niger
Nigeria	North Korea	Pakistan	Paraguay	Peru
Philippines	Republic of Kongo	Romania	Russia	Suriname
Swaziland	Syria	Tanzania	Turkey	Uganda
Ukraine	Venezuela	Vietnam	Yemen	Zimbabwe

| Cluster #4 – average 41.28 ± 0.98				
Algeria	Argentina	Armenia	Belize	Benin
Cape Verde	Costa Rica	Croatia	Czech Republic	Egypt
El Salvador	Gabon	Ghana	Greece	Guatemala
Hungary	Italy	Jamaica	Jordan	Latvia
Lithuania	Malaysia	Mexico	Mongolia	Morocco
Panama	Poland	Saudi Arabia	Senegal	Slovakia
South Africa	South Korea	Sri Lanka	Thailand	Trinidad and Tobago
Tunisia	Zambia			
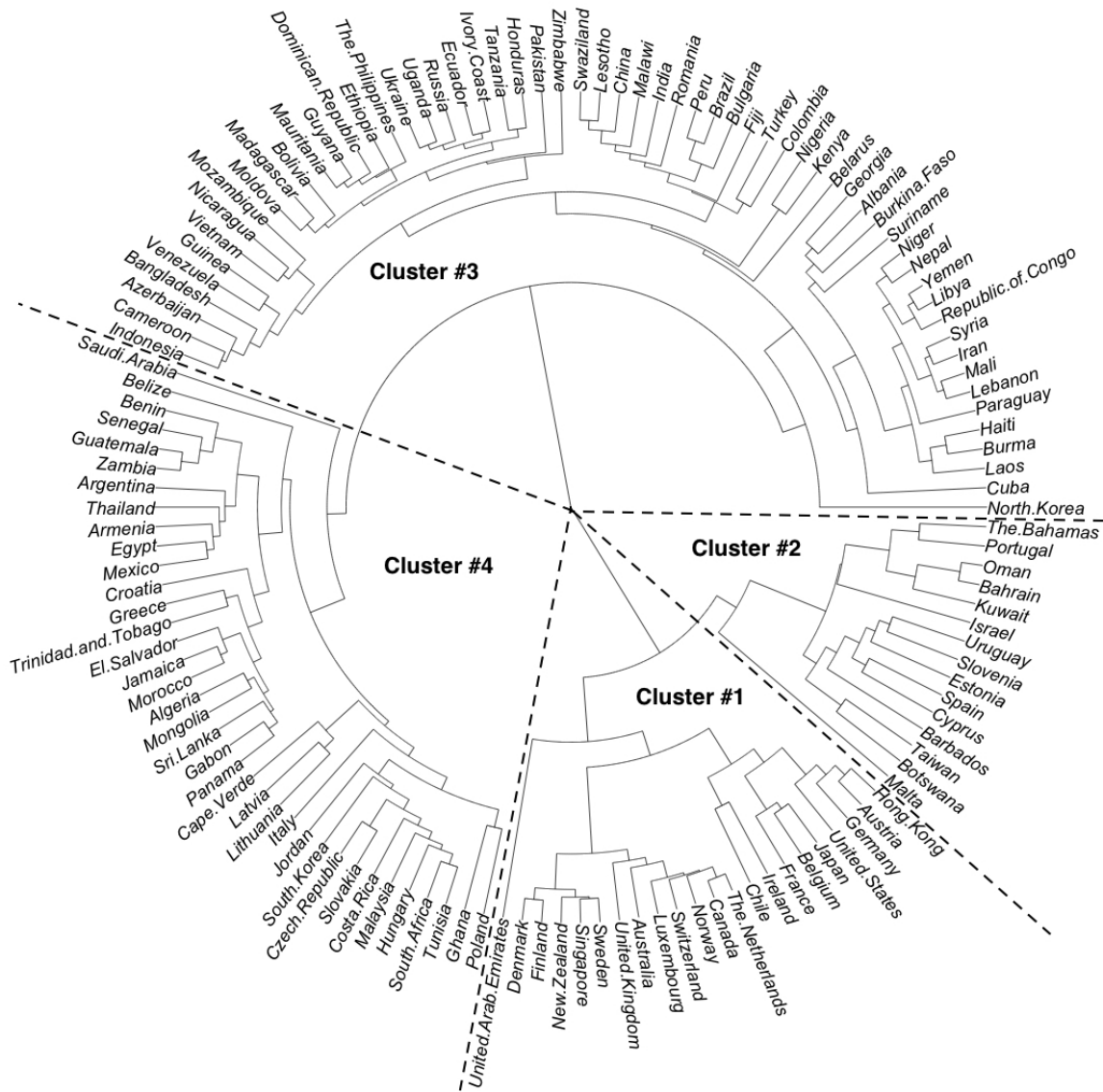
Figure 1: *Dendrogram of corruption perception.* The dendrogram is based on the average linkage clustering procedure using the Euclidean distance as a measure of dissimilarity. The spherical depiction gives a base of the tree in the centre and roots are growing centrifugally. The identified clusters are based on a visual inspection and for a better orientation, these are separated by dashed lines.

13

Appendix

Figure 2: *Alternative dendrogram of corruption perception – complete linkage.* The dendrogram is based on the complete linkage clustering procedure using the Euclidean distance as a measure of dissimilarity. The spherical depiction gives a base of the tree in the centre and roots are growing centrifugally. The identified clusters are based on a visual inspection and for a better orientation, these are separated by dashed lines. Structure of the clusters is very similar to the one based on the average linkage clustering in Fig. 1.
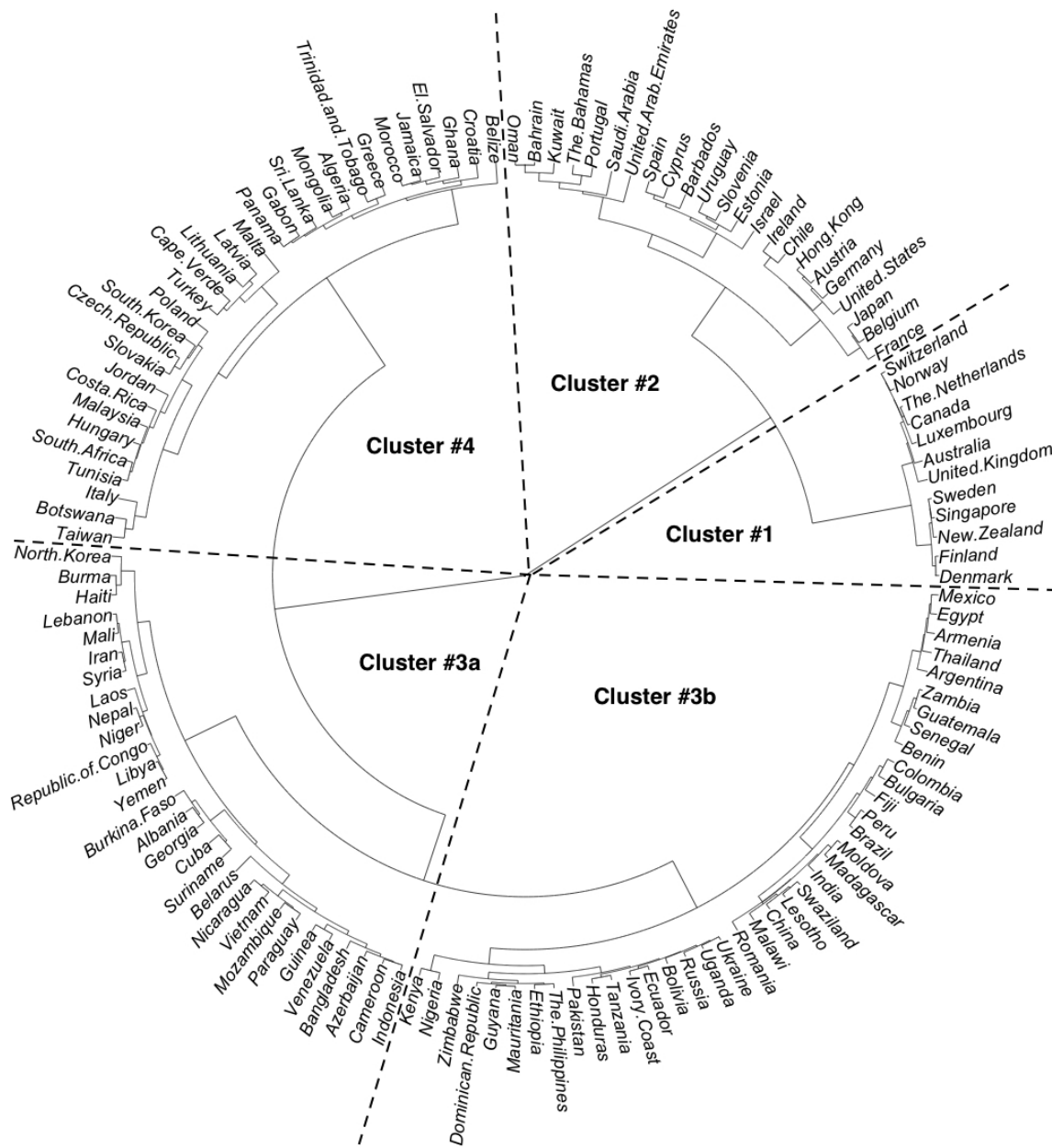
Figure 3: *Alternative dendrogram of corruption perception – Ward's linkage.* The dendrogram is based on the Ward's [50] linkage clustering procedure using the Euclidean distance as a measure of dissimilarity. The spherical depiction gives a base of the tree in the centre and roots are growing centrifugally. The identified clusters are based on a visual inspection and for a better orientation, these are separated by dashed lines. Structure of the clusters is more complex than for the average and complete linkage methods, such as splitting of Cluster #3 into two clusters. Nevertheless, the most important features of the network remain intact.