Erzsébet Németh – Bálint Tamás Vargha – Katalin Ágnes Pályi

The Scientific Reliability of International Corruption Rankings

SUMMARY: Owing to its hiding nature, corruption is difficult to measure. However, measurements lacking sufficient methodological grounds and the rankings so created pose just as much of economic risk as the corruption itself. The objective of the present study is to examine the methodological compliance of the calculation of the most well-known corruption perceptions index after having reviewed the specific literature and by using the method of document analysis. The most important criticisms formed against corruption measures in the international specific literature also highlight that in the most cases the independence of the organizations issuing the corruption measures, the transparency of data sources used and the applied methodology are not ensured. It follows from all of the above, and due to improper compilation of the respondent group and the inadequacies of the summary of the data sources - in the absence of methodological substantiation - the results are not suitable for comparing the countries, to draw scientific conclusions, and they do not even make a diagnosis related to the phenomenon examined which could serves as an efficient tool for making the appropriate policy decisions.

Due to the concealed nature of corruption its measurement is challenging. Yet, methodologically flawed quantification of corruption and subsequent ranking of nations implies equivalent economic risks as corruption itself. This study aims to investigate the scholarly and methodological propriety of the most popular corruption perception index, using document analysis as research method. Critics point out that in most cases the independence of the organizations issuing the measurements, the transparency of data used and the applied methodology is not ensured for. Consequently, and due to further methodological fallacies – inadequate composition of respondents and aggregation of data – results are unsuitable to enable a comparison of countries, or to draw scientific conclusions, nor does it provide a diagnosis that could be an effective tool to design policy interventions.

Not only corruption but the measurement of corruption, and as a result the ranking and possible stigmatization of countries (see Lambropoulou, 2012) also have substantial social and economic effects. The result achieved by any country according to the

Email address: enemeth@metropolitan.hu

corruption perception index by now is announced publicly together with the rate of the GDP growth or the amount of direct investment (Galtung, 2006). Its impact is illustrated well by that the willingness to invest may be set back by the poor result achieved in the corruption rankings (Warren and Laufer, 2009). In 2004, Kenya was denied aids due to the poor results achieved in the corruption surveys of the World Bank and Transparency International (Sampson, 2010). While the corruption indices undertake to measure perception, having distorted this, the media already talks about corruption: the portal 'Index' published a news item on 29th January 2019 with the title 'Steadily at the Bottom', which said that *'Together with Romania, Greece and Bulgaria, Hungary is still among the most corrupt countries of the European Union.'* International corruption rankings could also be tools for exerting influence; the publication of the results of corruption measurements and the media campaigns built thereon constitute the parts of a package, in which counsellors, representatives of civil and aid organizations or even officers responsible for loans provided by the International Monetary Fund may demand that the countries concerned took over the institutional and regulatory solutions as well (Sampson; 2010, Sampson 2015, p. 121).

The Corruption Perception Index of Transparency International is unarguably the most successful product concerning with the measurement of corruption. However, compared to its impact achieved through the public, the methodologies and the persons whose answers are used to form the index, as well as how reliably the index shows the corruption sensitivity of a given country are given little attention.

Through the identification of the credibility, measuring and validity risks of the corruption measurement systems, our aim is to answer the question whether the international corruption

rankings – including the annually published Corruption Perception Index, which has the greatest impact – are reliable, whether the methodology used by them allows scientific conclusions to be drawn, as well as whether the rankings are suitable for substantiating public policy interventions.

MEASURING CORRUPTION

The issue of the measurability of corruption became the focus of academic interest in the 1990s, simultaneously with the rise of combatting corruption. However, the efforts aimed at quantifying the rate of corruption encountered the problem that they were attempting to make the unmeasurable measurable. Due to the latency inherent in criminal statistics, the objective law enforcement data could not be used directly, while in questionnaire surveys the high rate of refusal to answer limited this attempt. As a result, corruption measuring approaches – also included in *Figure 1* – had developed, such as:

PERCEPTION-BASED MEASURES: this typically questionnaire-based method is used for mapping issues (social trust, corruption) where the objective administrative data are missing for understandable reasons;

EXPERIENCE-BASED MEASURES: aims at mapping direct experiences, it is based typically on personal interviews, therefore it is costly;

EXTERNAL (EXPERT) EVALUATIONS: comprehensive overview on national or industry level or the assessment of micro-level phenomena;

STATISTICS USING ADMINISTRATIVE DATA: national, government, sectoral, etc. statistics;

CONSTRUCTION OF COMPOSITE INDICES: through the summary of different types of data collection carried out by other organizations and quantified results, the purpose of which is to concentrate the information and to ensure

Figure 1

CERTAIN INTERNATIONAL MEASUREMENT SYSTEMS RELATED TO CORRUPTION

Source: own edited

comparability by creating one single index-number;

METHODS MEASURING CORRUPTION RISKS: a system forecasting the likelihood of the occurrence of corruption and/or the rate of protection against corruption.

Different limitations apply in case of the different approaches, see more detailed in *Trapnell* (2015). The proper selection of the data sources is of crucial importance in each and every case, and in respect of perception-based measures it is also of special importance that the questions included in the surveys are unambiguous and can be interpreted the same by everybody, while in course of the construction of composite indices, the careful selection of the procedure used to aggregate the indicators involved in the model requires attention. No matter which approach is used, if a phenomenon is measured with scientific tools, the fulfilment

of the criteria of scientific substantiation shall be borne in mind at all times:

• Is the impartiality, i.e. the independence and lack of bias of the measure ensured?;

• Had the methodology applied and the data sources used been presented in at least with the level of detail that allows the measure to be repeated by an independent researcher, thereby allowing the results to be verifiable?;

• Is the methodology of the research itself suitable for realizing the objectives of the research?, Does it use any statistical method through which the distorting effects of the data collection can be eliminated and the accidental differenced, simultaneous actions can be by-passed?;

• Are the appropriateness and the representativity of the sample respondents ensured?

The corruption measuring systems in the scientific literature

Despite the systems developed for measuring corruption being rather popular, the scientific community has warned to the reliability and validity risks thereof numerous times. The critical opinions formed in the scientific literature usually align with the criteria of scientific substantiation presented above, and they make negative observations related to the independence of the issuing organizations, the limited access to the methodology and the data sources used, as well as the inadequacies of the methodology.

The independence and objectivity of the issuing organizations

An important source of legitimacy for the organizations which construct the corruptions measures and rankings is that these organizations themselves and the measuring performed by them are free from bias, therefore they provide an objective picture of the corruption phenomenon. However, if the organization performing the evaluation follows an own public policy target system (See Századvég, 2016) or accepts financial support from a country which the organization evaluates, then the requirements of the independence and objectivity of the evaluation cannot prevail completely. With respect to one of the most popular corruption index-generating organization, Transparency International (TI), *Sampson* (2010; 2015) pointed out that the budget of TI is made up mostly by the donations of the ministries of foreign affairs of European states which at the same time are also the subjects of the corruptions survey carried out by TI annually. Furthermore, in addition to the government aids, the organizations of the TI receive donations provided by private entities as well, and the donors often include large corporations with

global investor interests (e.g. Siemens AG) or non-governmental organizations, foundations and civil organizations active in shaping the political agenda (e.g. Ford Foundation, USAID), which traditionally operate in the foreign policy interest of the United States (see Krige, 1999). While sustaining themselves from financial contribution of the countries concerned by the measurement system, the organizations engaged in measuring corruption also actively contribute to the politicization of the issue of corruption, and in some cases thereby undermining the success of the public policy measures and the approaches which are impartial in terms of party policy, too (Heywood and Rose, 2014).

The transparency of the methodology and the data sources

Numerous authors warn of the risks entailed by the lack of transparency of the methodology and the data sources. *'The lack of transparency in the production of measures of corruption is an ironic, and particularly embarrassing situation, (…). In particular, the lack of access to the data – which raises questions about the use of commercial data sources by (…) researchers – has meant that basic rules of scientific inquiry, such as the enablement of independent researchers to engage in replication exercises and proper validation tests, have not been followed.'* (Hawken and Munck, 2009, pp. 7-8)

Heywood and Rose (2014), as well as *Sampson* (2010) note that the organizations publishing the corruption indices usually refrain from having academic debate about the complexity and accuracy of the measure. *Hawken és Munck* (2009) argue that the organizations publishing the indices had not carried out the critical analysis of the reliability of the data sources before the development of the indices and as theoretical substantiation, as well as the posterior tests did not concentrate on measuring the reliability of the data used

either but merely on the correlations between the indicators. In addition, the co-authors expressed their concern because until the publication of their study, TI had not publish the data sources used for the Corruption Perceptions Index (CPI) and acquired on the market. In summation, multiple critiques raises concern about the problem that the methodology and the data sources used are not presented with the detail which would allow independent researchers to repeat and verify the research.

The composition of the sample respondents

Composite perception indices usually rely on indicators where the data providers are business operators, experts and – as the case may be – households and population groups. The construction of the sample respondents shall be suitable for providing authentic information about the phenomenon examined, in this case, about the corruption perception of the countries examined. It is debatable why experts and business operators are predominant among the respondents when the countries are the subjects of the corruption perception surveys, and whether it is even possible at all to create a true picture of the perception of the citizens if the research is carried out on a sample which does not represent the population concerned. In connection with the CPI *Galtung* (2006) also established that among the sample respondents not only the business operators and the experts are overrepresented but as a result the index shows the perspective of almost exclusively men with high standard of living, while it disregards the opinion of poor people and women, as well as the perception of the informal economic operators. Furthermore, in his recommendation for the optimal weighting of corruption indicators *Knack* (2006) argues that in course of the construction of composite indices, the survey conducted on samples

representing the national population shall be given a higher weight.

The scientific quality of the construction of composite indices

In most cases, the organization issuing the composite index does not participate in the data collection, therefore the scientific value added by it consists of the selection and summary of the data sources. According to *Sampson* (2010), although composite indices promise that through the joint use of multiple indicators they eliminate the accidental effects, however, they also obscure the concrete phenomenon measured by the data collection and the exact circumstances of the data collection. In connection with this limitation *Apaza* (2009) points out that composite indices are unable to provide a detailed picture of the nature of corruption, and consequently they cannot serve as an adequate tool for the fight against corruption. According to *Søreide* (2006), another limitation of composite indices is that they are unable to appropriately inform the reader about the scale of the phenomenon measured. *Heywood and Rose* (2014) pointed out that -as the case may be – the data sources used for the composite indices designated to measure the corruption perception have incompatible working definitions, and thus the summarization of the results arising from different conceptual starting points could lead to the blending of life phenomena and corruption transactions of different types. In addition to the above, the aggregated index may also conceal the measuring differences between the indicators selected for the purpose of the aggregation, since the indicators used are not always in accord. *Hawken és Munck* (2009. p. 5) highlighted that the average correlation of indices used for both the CPI and the CCI of the World Bank is 0.692, which means that approximately 50 percent of the variance remains unexplained. This strengthens the

assumption that the corruption phenomenon which is measured by the indicators eventually summarized is not the same. The authors also highlighted that the measurement differences between the indicators used are not random, and all these result in that the aggregation does not balance out the trend-like measurement differences but – to the contrary – they import it to the composite index. The problem is intensified further by that it is the availability of the data source that determines what weight the data source will have in the evaluation of the country concerned.

The construction of rankings

The establishment of rankings of countries is a misleading portrayal of the rate of corruption from several points of view. The approaches of *Mungiu-Pippidi and Dadašov* (2016), Dadašov, *Hefeker and Lorz* (2017), as well as *Ferwerda* (2017) are consistent in that corruption occurs under specific sectoral, industry or even international circumstances, which do not necessarily align with country borders. *Sampson* (2010) describes this as the ranking of the countries concealing the stakeholders who are on the 'active corrupting party' side of the corruption transaction, which compared to the country evaluated could – as the case may be – be a foreign company or investor as well. *Malito* (2014) established that the methodology of the CPI and the CC does not allow for the comparison of countries and comparison of the temporal change in the performance of the countries. Having examined the development of the CPI for a longer period, *Heywood and Rose* (2014. pp. 7-9) established that the indices calculated annually and per country show significant consistency and lack of change, namely there is no country which would show trend-like movement. It follows from the above that the annual measurements are not suitable for the back measuring of public policy interventions.

All this questions the publication of the rankings on an annual basis, which could serve public relations purposes at the most. *Malito* (2014) and *Søreide* (2006) also pointed out that in the 'expert-based' perception survey repeated on an annual basis, the fact that experts learn about the preceding measurement results could inevitably lead to self-confirming feedback, moreover, the stigmatization could even become a self-fulfilling prophecy, see *Warren and Laufer* (2009).

The interpretation of the questionnaires

Malito (2014), *Johnston* (2004), *Laufer* (2006) and also *Søreide* (2006) pointed out that in the surveys which assess multiple countries with the same questionnaire, the expressions, concepts and categories used therein may be interpreted differently by the respondents subject to the different social contexts. While in some cultures the trips and representative business dinners may be claimed legally in the 'marketing' cost category, in other countries the same practices could be considered as prohibited influencing. An example for this is the system of medical gratuity. Whether the financial gratuity or gift given after the use of the medical service constitutes corruption abuse is subject to varying considerations in each country. However, the same phenomenon could be subject to different consideration even within the same country: for example, according to the FCPA regulation of the USA, the benefit provided to any foreign decision-maker is legal up to a certain threshold: it is considered as *'facilitation payment'* (Strauss, 2013), however, in the territory of the USA, such transaction conducted with a government official would be illegal.

The factors distorting the corruption perception

Measuring the perception of corruption is nothing more than summarizing the respondents' opinion on the extent certain

institutions, types of transactions or countries are corrupt according to the respondent. Although for a long time the perception indicators had been regarded as the best available tools for measuring corruption, simultaneously with the spread of the perception measures more and more researchers were interested in the question what relationship the corruption perception showed with the corruption experiences.

According to *Zaman and Rahim* (2009), no matter how complex our measurement method is, in course of measuring corruption perception we cannot be sure that the results in fact provide information about the corruption. *Treisman* (2007) thought along the same lines when he called the perception indices based on experts panels 'subjective indices' which reflect the conclusions of the interviewed.

Treisman (2007) and *Weber Abramo* (2008) both established that corruption perception does not show a strong correlation with the corruption experience. Meanwhile *Donchev and Ujhelyi* (2014) came to the conclusion that the level of corruption experience does not predict the expected value of the corruption perception, and the connection between the corruption perception and experience is not linear. At the same time, their study sheds a new light on the effects of democratic political or social institutions. It is shown that protestant nature, the level of economic development or the strength of the central government of a given society – without decreasing the corruption experiences – adjust the corruption perception downwards.

Further factors which distort corruption perception were identified by *Olken* (2009), who established based on his results that ethnic or religious heterogeneity increases corruption perception, at the same time – paradoxically – in communities of greater religious and ethnic heterogeneity, the indicator designated to measure actual corruption is lower. The author accounted this for the level of social trust and social control. In addition, *Gutmann, Padovano and Voigt* (2015) showed that corruption perception may be higher among respondents who are unemployed, have low income, or if their answers were raised during economic downturn, as well as corruption perception is lower if the answers were given by optimistic societies with high economic growth rates and low level of income inequality, while in countries hallmarked by strong political competition, the rate of corruption perception is typically higher. Accordingly, corruption perception is affected by numerous factors in addition to corruption. *Jahedi and Mendez* (2014) acknowledged that the subjective indicators are diverted by economic and cognitive factors, and those may show negative correlation with the objective data, however, according to them this is not sufficient reason for disregarding the data sources. In the opinion of *Roca, Orme and Brown* (2010), the rate of corruption perception is also related to the representation thereof in the media. This could be the reason why countries where the press operates under political control are able to show relatively good corruption perception results. Based on media content analysis and the comparison of the CPI figures, *Németh, Körmendi and Kiss* (2011) established that the media – while fulfilling its social duty by revealing and presenting the corruption cases – contributes to the increase of the rate of corruption perceived by the population.

THE RELIABILITY AND METHODOLOGICAL SUBSTANTIATION OF THE CORRUPTION PERCEPTION INDEX

Since 1995, Transparency International has been publishing annually the results of the Corruption Perception Index designated to measure the corruption perception, as well as the ranking of the countries evaluated

established based on such results. The composite index of the CPI is calculated by using 13 data sources, which include both directly available data sources and data sources which are available only commercially. In the following we provide a detailed presentation of the findings concerning the criteria of scientific substantiation, based on which we want to answer the question whether the CPI index of Transparency International is reliable, suitable for drawing scientific conclusions, and whether as a result its role is justified in substantiating policy decisions.

The independence of the organization releasing the index

Transparency International receives direct or indirect donations from multiple government agencies of countries graded by the measurement system.[1] The other main source of funding of the organization is made up of the donations received from market operators, large corporations and foundations. According to the data of the annual accounts[2], the largest donor of Transparency International has been the European Commission since 2015, however, numerous government bodies and foundations are also providing subsidies in excess of 1 million Euros per annum to the organization. The main donors are included in *Table 1*, in case of the donor organizations attributable to governments, with the last three years' CPI score and ranking of the countries.

In course of its CPI calculation Transparency International evaluates those donor countries as well which support its operation with significant amounts. Based on the comparison of the CPI figures of the largest donor countries it can be shown that these countries regularly achieve good results in the CPI evaluation and they are typically in the first ten places.

The methodology of the CPI index and the transparency of the data sources used

Despite the fact that Transparency International publishes the description of the 13 data sources used for the calculation of the CPI on its website,[3] as well as that in a summarized data table it also specifies the data sources it had used for calculating the CPI score of each country, Transparency International does not reveal the methodology and the content of the data sources with sufficient detail in numerous respects.

Certain data sources are inaccessible. The methodological description published by TI does not explain that the answers to exactly which questions are reflected in the WJP and VDEM indicators.[4] In addition, of the 13 data sources used to calculate the index, the accessibility of five data sources is significantly limited. Thus, the content and detailed methodology of the country risk evaluation carried out by EIU, the international risk assessment prepared by the PRS Group, and the risk assessment prepared by IHS Markit (Global Insight) are accessible against payment only. Similarly, also limited in accessibility are the base data used for the competitiveness ranking of the International Institute for Management Development, in which case the scores of each country are made publicly available, however, the yearbook containing the detailed data is available only against payment. The PRS Group and the EIU organizations do not publish information as to whether the basis of the risk analysis are factual data or expert perceptions. In case of multiple data sources neither the composition of the expert respondent group, nor the principles of the composition thereof is accessible (for more details see the parts titled *'Is the representativity of the sample respondents ensured? In the composition of the expert panels, are expertise and impartiality ensured?')*.

Table 1

STATEMENT ON THE DONOR ORGANIZATIONS SUPPORTING THE OPERATION OF TI

The largest donors in 2017	The amount of the donation in 2017 (EUR)	CPI score (and ranking) in case of donor countries		
Government bodies		**2018**	**2017**	**2016**
Department of Foreign Affairs and Development (DFARD) **Canada**	1.970.391	81 (9)	82 (8)	82 (9)
Federal Ministry of Economic Cooperation and Development (BMZ) **Germany**	1.499.457	80 (11)	81 (12)	81 (10)
Dutch Ministry of Foreign Affairs **Netherlands**	1.200.000	82 (8)	82 (8)	83 (8)
Development and Cooperation Agency (SDC) **Sweden**	1.040.957	85 (3)	84 (6)	88 (4)
Department of Foreign Affairs and Trade (DFAT) **Australia**	864.657	77 (13)	77 (13)	79 (13)
Department of State of the USA **USA**	763.074	71 (22)	75 (16)	74 (18)
Multilateral institution				
European Commission	4.580.722			
Foundations				
BHP Billiton Foundation	1.335.930			
Organizations of the Open Society Institute Foundation Open Society Institute Foundation to Promote Open Society Open Society Initiative for West Africa	705.389 213.225 17.808			
Corporations, market operators				
Siemens AG	856.835			

Source: 2017 financial statement of TI

It cannot be established how the data sources listed are rid of factors foreign to corruption. There are numerous data sources among the data sources used which evaluate factors which are unrelated to corruption or the perception of corruption. In the approximately 70-question questionnaire of the SGI index of Bertelsmann Stiftung, only one single question is about the perception of corruption.[5] Country Policy and Institutional Assessment of the African Development Bank measures 5 dimensions, of which however only one dimension (Structural Policies and Regulation), and within that only the evaluation of the transparency and accountability of the public sector can be attributed to corruption. Although the SGI Codebook refers to that in course of the calculation of the CPI only the relevant question is taken into consideration,[6] the methodology documents published by Transparency International do not explain how the data sources used in course of the scoring are rid of factors foreign to corruption.

With regard to the issues presented here, further information is provided by *Table 2,* which contains the main characteristics of the data sources used in course of the calculation of the 2018 CPI index.

The scientific substantiation of the index calculation methodology

It is easily understandable that if we want to compare the same phenomenon based on a different number of indicators which do not measure the same phenomenon, then it does not satisfy the criteria of scientific rigour. The CPI scores are developed from data originating from different data sources, after standardization and through simple averaging. In course of the calculation of the index, in respect of the evaluation of the countries there

are however significant differences depending on the number of the data sources the score of the index calculated for the country concerned is based on, as well as subject to the data sources used. Therefore, depending on the data sources available for the evaluation of the country, in respect of different countries the same data may be taken into consideration differently. This eventuality of the availability of the data sources used raises methodological risks from the aspect of comparability. The CPI was calculated by using 10 data sources in case of merely 5 percent – including Hungary – of the 180 countries in total which were evaluated in 2018. The drawback lies in that in case of 40 countries five or even less data sources had been available for the evaluation *(see Figure 2).* Therefore, the corruption perception characteristic typically for countries in Africa, Asian and the Pacific region is based on significantly less data.

Significant differences can be noticed within the group of European states as well. The Nations in Transit index of Freedom House and the Transformation Index of the Bertelsmann Stiftung are counted in the CPI score of merely 11 former socialist countries – including Hungary – of the 31 European countries. In case of the Transformation Index it is a risk that the figure of the index reflects the opinion of only 2 experts interviewed per country. The methodological audit prepared by the Joint Research Centre and published also by Transparency notes that in case of countries evaluated based on less than five data sources, the standard error is significantly higher, and they refer to that it would be advisable to specify the minimum number of data sources as at least five sources.[7]

Overall, due to the eventuality of the availability of the data sources, and thus the differences in the composition of the data sources used, with respect to certain countries the data used for the calculation of the CPI

SUMMARY OF THE DATA SOURCES USED FOR THE CPI

CPI 2018 data sources	Period	Data providers
AFDB African Development Bank Country Policy and Institutional Assessment	2016	Expert groups, their identity, number, selection criteria are not accessible
SGI Bertelsmann Stiftung Sustainable Governance Indicators	2018	91 experts in total (typically 2 per country, 3 to 4 persons per occasion)their identity is known, however, the criteria for their selection are not accessible + statistical, quantitative data
TI Bertelsmann Stiftung Transformation Index	2017-2018	246 experts in total (typically 2 per country) their identity is partially known, however, the criteria for their selection are not accessible
EIU Economist Intelligence Unit Country Risk Service	2018	Limited access
FH Freedom House Nations in Transit	2018	Expert groups, their identity, number, and the selection criteria are not accessible
GI IHS Markit Global Insight Business Conditions and Risk Indicators	2017	More than 100 country experts their identity and the selection criteria are not accessible
IMD World Competitiveness Center World Competitiveness Yearbook Executive Opinion Survey	2018	Survey done by interviewing approx. 6300 businessmen from 63 countries, their identity and the selection criteria are not accessible
PERC Political and Economic Risk Consultancy Asian Intelligence	2018	Survey done by interviewing 1802 businessmen (approx. 100 respondents/country), their identity and the selection criteria are not accessible
PRS Group International Country Risk Guide	2018	Limited access
WB World Bank Country Policy and Institutional Assessment	2017	The evaluation is performed by the employees of the World Bank, their identity, number, and the selection criteria are not accessible
WEF World Economic Forum Executive Opinion Survey	2018	Survey done by interviewing 12274 businessmen from 140 countries,due to the large number of respondents, their identity is not accessible, however, the selection criteria are accessible (random sample selected from groups according to sector and company size)
WJP World Justice Project Rule of Law Index Expert Survey	2017-2018	Survey done by interviewing 110 000 households and 3000 experts from 113 countries The number of respondents is known; the identity of the expert, the sampling criteria and the criteria for selecting the experts are partially accessible
VDEM Varieties of Democracy	2018	Approx. 2000 experts their identity and the selection criteria are not accessible

Source: own edited based on the data published by the organizations conducting the surveys

**THE DATA SOURCES BY NUMBER USED FOR THE EVALUATION
OF THE COUNTRIES EVALUATED BY TI**
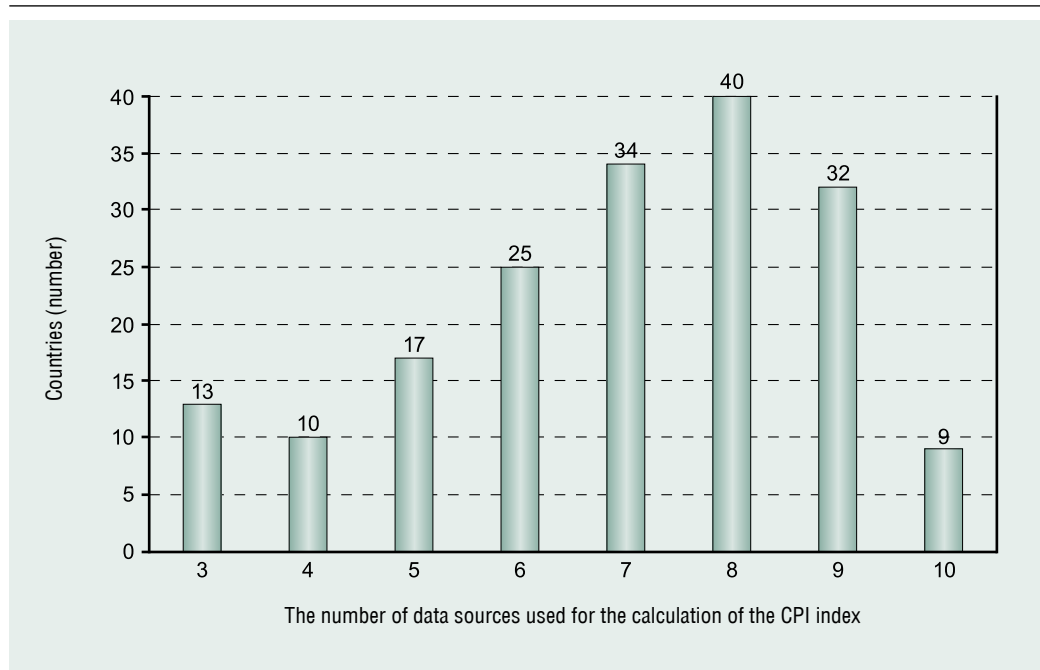
Source: own edited based on the data published by TI

are based on unequal requirements, not on the same methodology, along the lines of questions that are not the same and based on samples that are not the same. In addition, the methodology description published by Transparency International does not reflect that statistical methods used for the calculation of the CPI are capable of distinguishing the accidental simultaneous movements from proven correlations, or are able to provide explanation for the discrepancies between the scores of each country and each period, thereby ensuring the requirements of comparability and thus the ranking of the countries. This is not remedied even by the fact that after the compilation of the rankings, in 2018 TI calculated the significance of the change in scores of each country, starting from 2012.

The evaluability of the questions

Based on the questionnaires, in numerous cases the expert evaluate along the lines of questions the formation of which already does not allow for uniform and objective interpretation, for example: *'State capture by narrow vested interests'* (AFDB), *'To what extent does the government successfully contain corruption?'* (SGI), *'Are there general abuses of public resources?' 'Is there a tradition of a payment of bribes?'* (EIU), *'Does the public display a high intolerance for official corruption?'* (FH). The indicator of the NiT evaluates a factor the connection of which to corruption is not proven: *excessive state involvement in the economy*. In addition to all of the above, the questionnaires contain numerous issues which can be evaluated based

on the objective examination and analysis of the legislative and institutional environment or criminal statistics, and not based on the survey of experts' opinion: *'To what extent are public officeholders who abuse their positions prosecuted or penalized?'* (BF TI), *'Is there an independent judiciary with the power to try ministers/public officials for abuses?'* *'Is there an independent body auditing the management of public finances?'* (EIU), *'Are there significant limitations on the participation of government officials in economic life?' The legal protections of whistleblowers and journalists enjoy* (FH).

Is the representativity of the sample of respondents ensured?

The data sources used reflect the evaluation of expert groups in seven cases, where the data providers cannot be representative in respect of the population of the countries evaluated. The requirements of representativity are not ensured regarding even those data sources where the research had been conducted with interviewing an especially great number of respondents. In the framework of the World Justice Project Rule of Law Index 110 thousand households, 1000 persons per country were interviewed, at the same time, the data were collected in the 3 largest cities per country only, and as a result this index does not show the opinion of the population living in the countryside, in smaller settlements. The competitiveness ranking of the IMD, as well as the evaluation of Asian countries prepared by Political and Economic Risk Consultancy are based on the opinions of 100 businessmen per country on average, which does not ensure representativity either in respect of the population of the countries evaluated, or even the businessmen active in the country. The risk lying in the construction of the respondent groups is that such groups may be able to

ensure representativity in respect of the expert/businessmen population which in itself is difficult to define, however, where evaluations are conducted with the involvement of 2 to 3 experts per country, then not even this minimum requirement can be fulfilled.

Are expertise and impartiality ensured in the construction of the expert panels?

The majority of the indicators used for the calculation of the CPI is expert's evaluation. In course of the construction of such respondent groups it is a fundamental requirement that the special expertise and the impartiality of the respondents shall be ensured when selecting the experts. Overall, in case of the data sources examined, the accessibility of the identity of the experts is limited, while the criteria of their selection are not accessible at all. For example, in case of the V-Dem- and the Nations in Transit indices the information related to the identity and the exact number of the experts and to the selection criteria is inaccessible (about the V-Dem- see *McMann et al.,* 2016). In the framework of WJP's survey, a panel consisting of 3000 persons – lawyers proficient in civil law, labour law, healthcare and in the field of justice – had also been interviewed, however, the selection method of the respondents is inaccessible. Further risks lie in those data sources – for example the two indices of the Bertelsmann Stiftung – where the scores represent the opinion of merely 2 to 3 experts per country. All of this is worsened by that in case of Hungary it can be established that experts who are ideologically biased based on their participation in public life were asked to evaluate.[8] Having studied the country reports of the indicators used by Bertelsmann Stiftung, we established that in extreme cases the opinion of one single 'country expert' is behind the country reports. For the 2018

report about Hungary, the opinion of one single political science expert constituted the basis for the primary proposal for the scores of all sub-categories evaluated by the index.[9]

It poses a measuring risk that – based on the methodology descriptions available – it is not ensured that the expert evaluations are prepared by impartial persons who have special expertise. The expert evaluations may carry subjective elements, due to the small number of experts and the lack of transparency of the selection criteria.

GENERAL CONCLUSIONS

The measuring of corruption, and as result thereof the construction of international rankings may have severe consequences, may result in the public stigmatization of countries, as well as economic disadvantages. Therefore it is especially important to examine whether these international measurement systems provide an authentic picture, and whether they are scientifically reliable and valid.

Based on the specific literature, we can identify numerous risks affecting the scientific reliability and validity of the international indicators which measure corruption. The organizations which conduct the survey cannot be considered independent, they accept donations also from organizations which they rate. In a lot of cases the methodology used for constructing the index cannot be accessed to the extent that would allow the performance of independent researcher validation.

The CPI-based comparability of the countries is doubtful, since numbers and types of indicators and data sources different for each country are used in course of the development of the index. It is a risk affecting the validity of the ranking and the evaluation that the CPI does not use statistical methods which are able to distinguish the accidental simultaneous

movements from the scientifically affirmable correlations.

The index reflects not only corruption but also the corruption perception poorly. The measures rely on non-representative samples. In case of the expert panels used as data providers, the principles of selection do not ensure lack of bias and the special expertise. The measuring risks are enhanced by the fact that the country ranking depicts the phenomenon of corruption as if only the passive corruption party, the state party was affected, and it disregards that the active corrupting party is a private operator, or – as the case may be – a foreigner compared to the country rated. This is one of the reasons why it is an important effort that the state shall establish relations with those operators of the private sector the lack of corruption of which the state had affirmed.

Despite the fact that CPI involves numerous indicators in the composite index, the CPI does not use tools which would allow the corruption risks and controls to be measured in an objective manner, and it does not examine the regulatory environment and institutional system responsible for reducing corruption either. In this matter there would be a chance to review and evaluate the laws, the system of sanctions, the special anti-corruption procedures and mechanisms, such as the guarantee of judicial independence, the regulation of conflict of interest cases, mechanisms managing whistleblowing. Another opportunity which goes beyond the horizon of the perception measures could be measuring the corruption risks, as well as the risk mitigating controls, based on the examination of the institutional factors and mechanisms.

The reliability of measuring corruption is essential for the prevention of corruption and for efficient action against corruption. However, the composite indices and the corruption rankings established based

thereon are unable to provide a detailed and accurate picture of the actual prevalence of corruption that would allow for a substantial professional debate about the opportunities of the actions against corruption. The fight against corruption is still a matter affecting the entirety of the society, which enjoys the support of the most important social, economic and state operators alike. Therefore development-facilitating measurement systems are necessary which can be used efficiently

in the reduction of corruption. It can be established unambiguously that compared to the corruption perception measure and the CPI-type indices, the assessment of the sectoral, institutional corruption factors which are below the national level, the identification of the specific corruption hazards, and pointing out the necessity of the establishment of the control reducing corruption are more adequate tools for developing the measures against corruption.[10]

Notes

[1] https://www.transparency.org/whoweare/account ability/who_supports_us/2

[2] https://www.transparency.org/whoweare/accountability/audited_financial_reports_with_independent_auditors_report/2

[3] According to the information published by Transparency International, the methodology of the CPI calculation consists of the following: (1) The data used for the calculation of the index are selected in accordance with the system of criteria specified by TI. (2) Afterwards, the source data are standardized. (3) The standardized data are aggregated through simple averaging. (4) Finally, the uncertainties and the standard error are quantified. (Source: Corruption Perceptions Index 2018 Full Source Description, https://www.transparency.org/cpi2018).

[4] Corruption Perceptions Index 2018 Full Source Description, https://www.transparency.org/cpi 2018

[5] see: http://www.sgi-network.org/docs/2018/basics/SGI2018_Codebook.pdf

[6] see: https://www.transparency.org/files/content/pages/2018_CPI_Methodology.zip

[7] https://www.transparency.org/files/content/pages/2018_CPI_2017_StatisticalAssessment.pdf, p. 16

[8] The 46 Central and Eastern Europe and South-East Europe experts include nine Hungarians: eight of them who work for Political Capital, as well as political scientist Attila Ágh.

[9] Dániel Hegedűs, political scientist, Humboldt University, Berlin

[10] This is the methodological approach which the State Audit Office of Hungary has been using for mapping the integrity controls and risks of the Hungarian public institutions since 2011, and of the state and local government-owned business associations since 2016, see Németh et al. (2017).

References

Apaza, C. R. (2009). Measuring governance and corruption through the worldwide governance indicators: Critiques, responses, and ongoing scholarly discussion. *PS: Political Science & Politics*, *42*(1), pp. 139-143

Dadašov, R., Hefeker, C., Lorz, O. (2017). Natural resource extraction, corruption, and expropriation. *Review of World Economics*, 153(4), pp. 809-832, https://doi.org/10.1007/s10290-017-0288-y

Donchev, D., Ujhelyi, G. (2014). What do Corruption Indices Measure? *Economics & Politics* 26(2), pp. 309-331, https://doi.org/10.1111/ecpo.12037

Ferwerda, J., Deleanu, I., & Unger, B. (2017). Corruption in public procurement: finding the right indicators. *European Journal on Criminal Policy and Research*, *23*(2), pp. 245-267, https://doi.org/10.1007/s10610-016-9312-3

Galtung, F. (2006). ,*Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices', in Sampford,* C., Shacklock, A., Connors, C., Galtung, F. (eds.). *Measuring Corruption,* Ashgate, Hampshire, UK and Vermont, US, pp. 101-130

Gutmann, J., Padovano, F., Voigt, S. (2015). Perception vs. Experience: Explaining Differences in Corruption Measures Using Microdata, SSRN: https://ssrn.com/abstract=2659349, http://dx.doi.org/10.2139/ssrn.2659349

Hawken A., Munck, G. (2009). Do You Know Your Data? Measurement Validity in Corruption Research, working paper, School of Public Policy, Pepperdine University (Malibu: California)

Heywood, P. M., Rose, J. (2014). ,Close but no Cigar': the measurement of corruption.

Journal of Public Policy, 34(3), pp. 507-529, http://dx.doi.org/10.1017/S0143814X14000099

Jahedi, S., F. Méndez (2014). On the advantages and disadvantages of Subjective Measures. *Journal of Economic Behavior & Organization,* 98(C), pp. 97-114, http://dx.doi.org/10.1016/j.jebo.2013.12.016

Johnston, M. (2004). Comparing corruption: Participation, Institutions, and Development, in Heffernan, W. C., Kleinig J. (eds.). *Private and Public Corruption*, Rowman & Littlefield, pp. 275-322

Knack, S. (2006). Measuring corruption in Eastern Europe and Central Asia: a critique of the cross-country indicators. The World Bank. SSRN: https://ssrn.com/abstract=923275

Krige, J. (1999). The Ford Foundation, European physics and the cold war. *Historical Studies in the Physical and Biological Sciences*, 29(2), pp. 333-361, http://dx.doi.org/10.2307/27757813

Lambropoulou, E. (2012). Myths and Realities about Corruption in Public Administration and its Discourse in Greece, *Amsterdam Law Forum*, 4(3). pp. 77-96

Laufer, W. S. (2006). The Importance of Cynicism and Humility: Anti-Corruption Partnerships with the Private Sector. *Development Outreach,* 8(2), pp. 18-21

Malito, D. (2014). Measuring corruption indicators and indices. Robert Schuman Centre for Advanced Studies Research Paper, 13, http://dx.doi.org/10.2139/ssrn.2393335

Mungiu-Pippidi, A., Dadašov, R. (2016). Measuring control of corruption by a new index of

public integrity. *European Journal on Criminal Policy and Research*, 22(3), pp. 415-438

McMann, K. M., Pemstein, D., Seim, B., Teorell, J., Lindberg, S. I. (2016). Strategies of validation: Assessing the varieties of democracy corruption data. V-Dem Working Paper, 23

Németh E., Körmendi G., Kiss B. (2011). Korrupció és nyilvánosság: a média hatása a korrupcióra és annak társadalmi megítélésére. *(Corruption and Publicity: The Media's Impact on Corruption and its Social Judgement) Pénzügyi Szemle/ Public Finance Quarterly* pp. 56, 57-65

Németh, E., Martus, B. S., Vargha, B. T., Gergely, Sz., Vasváriné, J. (2017). *Elemzés a közszféra integritási helyzetéről 2017, A hazai integritás helyzete a köz- és a magánszektor érintkezési felületén*, *(Analysis of the Integrity Situation of the Public Sector 2017, The Status of Hungarian Integrity in the Contact Area of the Public and the Private Sectors)* https://asz.hu/storage/files/files/Publikaciok/Elemzesek_tanulmanyok/2017/int_koz2017.pdf?ctid=1126

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, 115(2), pp. 200-249, http://dx.doi.org/10.1086/517935

Olken, B. A. (2009). Corruption Perceptions vs. corruption reality. *Journal of Public Economics* 93 (7-8). pp. 950-964, http://www.sciencedirect.com/science/article/pii/S0047-2727(09)00031-0

Roca, T., Orme, W., Brown, J. (2010). Fear and Loathing of the Corruption Perception Index: Does Transparency International Penalize Press Freedom? DOI: 10.2139/ssrn.1694211.

Sampson, S. (2010). Diagnostics: Indicators and transparency in the anti-corruption industry. In Jansen, S., Schröter, E., Stehr, N. (Eds.). *Transparenz:*

Multidisziplinäre Durchsichten durch Phänomene und Theorien des Undurchsichtigen (pp. 97-111), VS Verlag für Sozialwissenschaften

Sampson, S. (2015). The anticorruption package. Ephemera: Theory and Politics in Organization, 15(2), pp. 115-123

Shore, C. (2005). ‚Culture and Corruption in the EU: Reflections on Fraud, Nepotism and Cronyism in the European Commission', in Haller, D., Shore, C. (eds.) *Corruption: Anthropological Perspectives*, Pluto Press.

Strauss, E. N. (2013). Easing out the FCPA facilitation payment exception. BUL Rev., 93, 235

Søreide, T. (2006). Is it wrong to rank? A critical assessment of corruption indices. Chr. Michelsen Institute

Trapnell, S. E. (2015). User's Guide to Measuring Corruption and Anti-Corruption. *United Nations Development Programme,* https://www.undp.org/content/undp/en/home/librarypage/democratic-governance/anti-corruption/user-s-guide---measuring-corruption-and-anticorruption.html

Treisman D. (2007). What Have We Learned About the Causes of Corruption from Ten Years of Cross-National Empirical Research? *Annual Review of Political Science,* Vol 10, pp. 211-244, https://doi.org/10.1146/annurev.polisci.10.081205.095418

Warren, D. E., Laufer, W. S. (2009). Are corruption indices a self-fulfilling prophecy? A social labeling perspective of corruption. *Journal of Business Ethics*, 88(4), pp. 841-849, https://doi.org/10.1007/s10551-009-0316-5

Weber Abramo, C. (2008). How Much Do Perceptions of Corruption Really Tell Us? *Economics*:

The Open Access, Open-Assessment E-Journal 2, http://dx.doi.org/10.5018/economics-ejournal.ja.2008-3

Zaman, A., Faiz-Ur-Rahim (2009). Corruption: measuring the unmeasurable. *Humanomics*, 25(2), pp. 117-126, https://doi.org/10.1108/08288660910964184

Századvég (2016). Nemzetközi indexek módszertani korlátai és kritikája. (The Methodological Limitations and Critique of International Indices) https://szazadveg.hu/uploads/media/57e3d8e300 c7d/szazadveg-elemzes-a-nemzetkozi-indexek nek-komoly-hianyossagai-es-korlatai-tanulmany. pdf