

# Applied Statistics

## Study on Reporters Without Borders' Freedom of the Press Index

---

Clotilde CORON, Anne-Elisabeth COSTAFROLAZ, Julia CUVILLIEZ, Ester DAURENJOU  
2010/2011

Nowadays freedom of the press is a notion of great importance. We can indeed find behind this idea many fundamental human rights by which freedom of speech is directly affected. Hence it is legitimate for Reporters Without Borders (RWB), a non-governmental organization, to construct an index in order to measure freedom of the press, knowing that other organizations before RWB already took an interest in it, such as Freedom House, an independent watch organization in the USA. However Freedom House's index above all aims at measuring evolutions in time of freedom of the press in the world, while RWB's index sets itself as an objective to permit international comparisons at one time. The questionnaire has indeed changed over time, what makes impossible any diachronic comparison. So, through its ranking's publication, RWB attempts to measure freedom of the press' way of things in the world, and above all to point a finger at the worst countries by denouncing violations of freedom of the press. The index reflects the degree of freedom enjoyed by journalists and media, and the efforts undertaken by the authorities to guarantee the respect of freedom of the press.

Since its setting-up in 2002, this index has gained in influence and countries as well as institutions (World Bank...) have begun to take a closer interest in the ranking's results. This new situation led RWB to reconsider its questionnaire and ranking by wanting to consolidate them and to justify their structure. This is why they appealed to an ENSAE group of applied statistics: they wanted us to bring them the statistical elements missing to their methodology.

In this general note, we are going to summarize our principal conclusions through a thematic approach, which seems the best way to highlight our results. First we will present RWB's methodology as regards data and the questionnaire, as well as problems which came up at this first stage. Then we will look into the statistical data processing of replies (and also of non-response). Next we will analyze how countries are divided up in the ranking: is it possible to distinguish between big "classes" of countries? Afterwards we are going to emphasize how it is necessary to set up intermediate indexes, as a complement to the global index that often gives rise to misinterpretations of the ranking. And lastly we will try to pertinently model the obtained score with exogenous variables.

### **Methodology : RWB, data, questionnaire**

Every year since 2002 RWB carries out a questionnaire sent to several respondents in most of the countries. This questionnaire is addressed to RWB's partner organizations (15 pro-freedom of speech organizations on the five continents), to its 140 correspondents network, journalists, research workers, lawyers or human rights activists. A scale devised by RWB then allows giving each country a score and a relative position. In 2010 178 countries were ranked: those who sent back completed questionnaires filled by various sources. Other countries are not in the ranking because of a lack of reliable data. In order to simplify and to minimize the respondents' selection bias, the questionnaire is distributed in seven languages: French, English, Spanish, Portuguese, Persian, Arabic and Russian. RWB often knows the respondents, but the latter's identities are not declared to the authorities as a safety precaution and to guarantee honesty in answers. Three elements already reveal a bias issue in the answers' reliability: the fact that respondents are not anonymous to RWB but a priori selected the limited number of respondents by country (4 on average) as well as translation and interpretation issues.

In order to collect the necessary data to carry out the ranking, Reporters Without Borders made a questionnaire that takes up the principal criteria evaluating freedom of the press in each country. We could notice seven more or less distinct themes: physical violence; state responsibility in events aiming at the

journalists' integrity; indirect and direct threats and access to information; censorship and self-censorship; control of media; judicial, business and administrative pressure and a part about Internet and new media. Questions evolved over the years, disappearing or reappearing year by year. For instance questions about media presence abroad were suppressed in 2005 when questions about Internet and new media were added, a theme which kept on gaining importance in the following years. Moreover the number of questions has decreased – 50 questions in 2002, 43 in 2010 – and some themes were brought together in order for the questionnaire to get clearer and less redundant. Therefore the questionnaire's content but also its structure evolved year by year to adapt to the world situation.

So each country receives a score – which is the mean of the scores obtained by the country respondents' questionnaires – and a relative position in the ranking: for example in 2010 Finland gets first place, on a par with five other countries and with a zero score, while North Korea and Eritrea come in last with respective scores of 104.75 and 105. Let's keep in mind that the higher is the score, the less the country respects freedom of the press. We quickly analyzed the results' evolution: at the ranking top we find year after year the Scandinavian countries, but the ranking is far from rigid since some countries clearly bettered their positions (for instance the Maldives), while others, quite unexpected, lost several ranks, such as France which is at the moment 44<sup>th</sup> in the ranking.

So a quick study on the ranking already underlined several issues. Insofar as RWB wanted us to formulate new suggestions for its methodology, we tried to put into words some solutions. We will express these issues later in our note, but already remember that subjectivity and cultural heterogeneity may lead to a bias in answers, and that there is not any unanimity about a press freedom ideal. Moreover the constant methodology modification raises the issue of time comparability of data – issue that Freedom House's ranking does not encounter! –, this is why we advised to set up a precise methodology and to stick to it, in order to be able to follow the evolution and tendencies over the years. Let us present now the statistical processing of answers.

### Statistical processing of answers

Answers to the questionnaire have, depending on the questions, either a dichotomous structure (yes/no) or a rating structure (on a scale from 1 to 5 for example). RWB's research workers allocated points to these answers ; for instance in the first question, they allocated 2 points to answer « yes » and zero points to answer « no » ; while in question 9 they allocated the following points :

Nr of jailed journalists	1	2 to 5	6 to 15	16 to 50	> 50
Points associated	2	3	5	7	9

So this conversion of qualitative answers into points is carried out thanks to a system of weighting: if the question is deemed to be important, it will have a wider amplitude than a question deemed less significant. For example, question 22 (presence of media censorship) has a 5 times greater weight than question 1 (presence of torture and mistreatment against journalists). Let's notice that we can qualify this weighting, according to RWB, as a "wild" one, insofar as they allocated points by intuition, without justifying them statistically. Therefore we worked on this issue which we think plays a key role. A data analysis permitted us to highlight the questions affecting the most the final score (without any weighting of course): thus themes such as presence of torture, state watch of journalists, access to public information... are valued by the data analysis, what could then justify a special weight on questions about these themes, compared to the others.

We also studied the impact of the randomness of the respondents' number by country. The fact that there are many respondents for each country makes indeed answers more objective since it permits the avoidance of a subjectivity bias. But there are sometimes really blatant gaps between the questionnaires of a same country, in which case RWB's research workers have to come to a decision. Moreover some states send back only one, what implies a subjectivity risk that can bother the processing of answers. We could unfortunately do nothing about this, for lack of extra data. However, for countries with many respondents, we noticed an increasing relationship between the answers' standard deviations and the final score of each state: the more variable the questionnaires within a country are (i.e. the more subjective the answers are), the worse the country's rank is!

Another important issue for the statistical processing is the strong presence of non-response in the data: they make indeed more difficult the results' analysis, since it brings in a bias, and since in the data analysis one of the axes is built on non-response (and extreme values of answers as well). We thought interesting to construct a model in order to explain non-response, having in mind that the higher the score of the country is (i.e. the more it

violates freedom of the press), a priori the more frequent non-response is. Some respondents may indeed censor themselves, preferring not to answer rather than to write an answer revealing a serious violation of freedom of the press in their country, and/or because a state violating this freedom is a fortiori a state where access to information is limited, and so where respondents do not necessarily have the information needed. We used a latent model with exogenous variables such as a freedom of speech index, respect of workers' rights, respect of physical integrity rights... which could explain the non-response phenomenon. The obtained model highlighted the respect of workers' rights and of physical integrity rights: the more the countries respect these rights, the less non-response there is. Now, after having analyzed the statistical processing of answers (and of non-response), we are going to show our study on the countries' distribution.

### Analysis of the repartition of the countries

In order to study the distribution of the countries in the ranking, and to see whether we could construct a few number of groups of countries, delimited by the results in the questionnaire (and also to go further than a simple opposition between "good" and "bad" countries), we did a data analysis. This analysis has underlined three groups: good countries (the countries which do not violate freedom of the press and have a score close to zero), bad countries with respect to physical violations (torture, mistreatment...), and bad states with respect to moral violations (censorship by the government, limited media ...). This result corroborates what we had already noticed when we first saw the questionnaire and the ranking: this ranking does not properly differentiate the top countries (i.-e. that have a score close to zero), and differentiates better countries that violate the most freedom of the press. We had then the idea of bringing in more "positive" questions (means to help journalists...), in order to differentiate better countries that are respectful of freedom of the press. For example, it would be interesting to ask questions about active measures (creation of funds to help journalists...), and not only about passive measures (to permit the creation of independent medias...).

We also carried out a classification, with the same goal: to reveal groups of countries. We obtained five classes, constituted by: the "good countries" (Germany, the United States, Sweden...), countries with a great censorship (Iran, Saudi Arabia...), countries with a strong state intervention against journalists (China, Libya...), countries with non-response and badly ranked (Congo, North-Korea, Niger...) and lastly countries with violations of the journalists' physical integrity (Cameroun, Mexico, Afghanistan...).

This classification made us think that the questionnaire, in place of creating one single index, often misunderstood since the only ranking aspect is underlined, could lead to the creation of many partial indexes: one about the government role in the economical and judicial situation of the press, another about the violation of the physical integrity of the journalists, one other about censorship... We examined this possibility in our work, showing that we could create intermediary indexes, based on the classification.

### Necessity to construct partial indexes

The creation of partial indexes could permit to avoid any focusing on the ranking, what does not seem right to us, as regards RWB's goal by creating this index. Nowadays indeed the different countries indeed wait for the ranking's release to know their scores and ranks, without thinking that the denomination « press violation » covers different meanings. For example, a country which has a bad ranking because it censors media does not have to take the same measures than a country badly ranked because it does not fight enough against the physical violence against towards journalists. On the one hand the ranking compares what is not necessarily comparable (physical violations, censorship...), and on the other hand, does not give information about what and how to improve.

The classification gave us ideas for the partial indexes' creation, which could correspond in fact to subdivisions of the questionnaire, and that one could regroup to land on the initial (or final !) ranking, if it is really necessary to have such a global ranking. We created three indexes: "violence and attacks" (VA), which measures physical violence against journalists; "access to information and censorship" (AIC), which measures the censorship of information "economic concentration and legislation" (ECL), which underlines the role of the government, taking into account administrative pressures for example.

These partial indexes are rich with respect to a statistical point a view, because they allowed us to use different tests. For example, we measured the independence between the ranks of countries in those different indexes. There was a lot at stake: does a country which violates a lot freedom of the press with acts of censorship,

violate it also necessarily with physical violence against journalists? Studying the correlations between the ranks got in these indexes, we could see that, even if the correlations are positive, they are not so elevated than we thought. For instance the correlation coefficient between ECL and VA is quite low (0,36).

Besides, the creation of partial indexes allowed us to underline “special” countries, which are well ranked in an index, and badly ranked in another index. It is the case for example of Afghanistan, which is the 173<sup>rd</sup> in the index of physical violence, and the 25<sup>th</sup> in the index of the economic concentration and legislation; we observe the opposite situation with Djibouti, which is the 10<sup>th</sup> in the index of physical violence, and the 157<sup>th</sup> in the index of the economic concentration and legislation. Partial indexes are a precious source of information for these countries, because it allows them to see what measures they have to take in order to ameliorate their ranks and situation. A polytomous logistic regression permitted us also to see to which extent we could predict the countries’ belonging class using as regressors the three partial indexes.

### Is it possible to model the final score with exogenous variables?

We tried to explain the freedom of the press index by freeing ourselves from RWB’s questionnaire and taking into account other characteristics of countries, more particularly the GDP per capita, religion, political regime, respect of workers’ rights... We created our own data base thanks to online data on World Databank’s and PNUD’s websites. Statistical processes permitted us to select the most relevant variables in order to model the final score:  $Final\ score_i = 40.7 - 10.6 * Speech_i - 8.1 * Physintcr_i - 7.9 * Politycr_i - 7.9 * Worker_i + \epsilon_i$  with  $\epsilon$  a random error term, and:

Speech	Physintcr	Politycr	Worker
Index of freedom of speech and (0 if this freedom is transgressed, 1 otherwise)	Index or torture, sentence without trial, political imprisonment and disappearing (0 if the government does not do anything against this, 1 otherwise)	Numerical index about governance authorities and political regime (from -10 for dictatorships, to 10 for strong democracies)	Index of workers’ rights – strike, minimum legal age, hours of work, minimum wage... – (0 if these rights are not respected, 1 otherwise)

Therefore an increase of one point in the physical integrity index – an improvement in related rights – implies a decrease of 8 points in the final score and so the country gets closer to the top countries in the ranking. Moreover a country where freedom of speech and association is respected has, other things being equal, a final score reduced by 10.64 points compared with a country where this freedom is violated. This model allowed us also to show that variables such as GDP are less significant since they were not selected by the model. In other words a rich country does not necessarily have a good ranking: we had a premonition about that when we analyzed the ranking’s evolution, in particular according to France’s mediocre ranking (44<sup>th</sup>), behind much poorer countries and whose press situation is very encouraging (South Korea – 42<sup>nd</sup>, Latvia – 30<sup>th</sup>). Yet we regretted that we could not prove the statistical relevance of other variables of which we thought they could bring information, such as the situation of women or the freedom of religion. However our model explains 69.3% of the countries’ final score’s variations.

Then we checked whether our model was correct with statistical tests: we confirmed that every theoretical hypothesis was validated, and more precisely that there was no perfect collinearity between the explanatory variables and that our error term had a constant variance (homoscedasticity assumption).

Lastly we tested our model on 4 countries randomly selected at the beginning of our analysis and isolated in another data base: Finland, Iran, Gabon and Canada. Our model turned out as a rather good explanation of the score, since the error terms were not high. However we noticed that our model works less and less well as the country is badly ranked: in other words we can go in depth into the previous result according to which differences between the top countries can hardly be seen “to the naked eye” but have quite predictable press situations, while badly ranked countries have separated scores, but in order to explain their scores’ variations we may resort to more diverse and comprehensive variables.