STATISTICAL MODELS TO MEASURE CORPORATE REPUTATION

Paola CHERCHIELLO[1]
PhD, Researcher,
University of Pavia, Italy

E-mail: paola.cerchiello@unipv.it

Abstract: *Reputation can be defined as how an entity (private or public) is perceived by each of its stakeholder groups and reputation risk as the risk that an event will negatively influence stakeholder perceptions. Since reputation involves intangible assets (public opinion, perception, reliability, merit), it is not simple to define and consequently to measure and to monitor the correlated risk. In this contribution we propose statistical models based on ordinal data aimed at measuring effectively reputation. The proposed models are applied to real data on Italian public companies taken from financial media corpora.*

Key words: *mixture of binomial and uniform r.v., ordinal models, media reputation*

1. Introduction

Reputation is a prized, and highly vulnerable, corporate asset. Risk managers debate whether reputational risk is an issue in its own right or simply a consequence of other risks. Good communication is vital to protecting against and repairing reputational damages. But good communication becomes even more important once a crisis breaks. Companies that have a communication strategy that enables them to respond quickly and effectively to "*bad news*", emerge often with their reputation enhanced.

Reputation can be defined as how an entity (private or public) is perceived by each of its stakeholder groups and reputational risk as the risk that an event will negatively influence stakeholder perceptions. It is not simple to define and consequently to measure and monitor reputation since it involves intangible assets such as: public opinion, perception, reliability, merit. By the way, it is a matter of fact, that a bad reputation can seriously affect and condition the performance of a company. In general, companies tend to act too late, that is, once the adverse event has occurred. In this way, companies do not practice risk management activities, but rather crisis management ones.

The first formal definition of reputational risk is due to Basel Committee on Banking Supervision that in 1997 stated that "*Reputational risk arises from operational failures, failure to comply with relevant laws and regulations, or other sources. Reputational risk is particularly damaging for banks since the nature of their business requires maintaining the confidence of depositors, creditors and the general marketplace*". Moreover, given the evident difficulty in defining all the relevant dimensions inherent such category risk, few strategies have been put in place in practice.

An interesting aspect embedded in reputation evaluation is the possible gap between its perception and the reality: a positive reputation of a company that is not supported by its real status leads inevitably to a dangerous risk. Several recent examples of such misalignment can be reported: Parmalat and Cirio (in Italy) and the notorious Lehman Brothers, Enron and Worldcom cases. On the other hand, media coverage plays a key role in determining a company reputation. This often occurs when a company reputation has been significantly damaged by unfair attacks from special interest groups or inaccurate reporting by the media. A detailed and structured analysis of what the media are saying is therefore important since the media shape the perceptions and expectations of all the involved actors.

Nowadays, natural language processing technologies enable companies and their business intelligence departments to scan a wide range of outlets including newspapers, magazines, TV, radio, and blogs. The above mentioned collection of textual data, aimed at measuring the reputation of an institution, motivates the development of appropriate statistical models for the analysis of such data. In order to achieve the goal, we have collaborated with the Italian market leader company in financial and economic communication, the daily newspaper Sole24ORE, that has provided us with textual data for a collection of public companies.

The paper is organized as follows: section 2 reviews the state of the art in reputation measurement, section 3 introduces our proposed methods, section 4 concerns the application of proposed models to Sole24ORE reputational data and section 5 gives the conclusions.

2. Existing Methodologies

Our proposal can be considered as an attempt to present efficient statistical methods aimed at measuring the reputation of an institution (being either private or public), through the development of appropriate statistical methodologies. From a managerial point of view, the issue is to properly measure and consequently rank the reputation level (say from high to low), thus to improve management decisions and actions directed at improving preventive controls on the correlated reputational risk. The main issue concerns the fact that the reputation of a given institution is not directly measurable, thus it is necessary to collect several different variables in order to allow an indirect measurement of the reputation dimension.

Because of the novelty of the problem, the scientific literature on the topic is limited and/or not completely shared. One of the first attempt to delimit the perimeter of reputation measurement was brought by Gabbi in 2004 that tried to offer a more operative definition of reputational risk (1). According to Gabbi, reputational risk is definable as "the set of direct economic consequences produced by the modification of the public image of an entity, such as a public company or a person with a public relevance". In a report produced by the Economist Intelligence Unit in 2005, reputational risk is defined as the "Risk of Risks" and is deemed as the most important and difficult task that senior risk managers have to face (2).

The few methodological contributions on the topic can be mainly grouped in two categories:

- Qualitative methods;
- Quantitative methods.

Among the qualitative approaches, we can list 4 indicators, often employed in actual corporate contexts:

1. *Reputation Quotient;*
2. *Reputation Index;*
3. *Fortune's Most Admired Companies;*
4. *RepTrack*.

The *Reputation Quotient* (RQ) was proposed by Harris-Fombrum and is considered a method of qualitative measure specifically built to catch the perceptions of each group of stakeholder (consumers, investors, employees, competitors) (3). The quotient is the result of an appropriate combination of 6 conceptual dimensions (emotional appeal, products and services, vision and leadership, workplace environment, social and environmental responsibility, financial performance) that are evaluated by a sample of managers in two subsequent phases: "nominations phase" and "ratings phase". The main criticism to RQ is the static measurement of the reputation and the absence of adjustable weights according to the opinions of the different stakeholder groups.

The second approach named *Reputation Index* proposed in 2003, is based on the internal evaluation of non quantitative factors produced by distinctive groups of stakeholder (4). The key factors contemplated by such index are: the leadership, the organizational culture, the innovation and the strategy. The stakeholder groups assign a score between 1 and 9, thus a weighted mean is computed where the weights are defined subjectively by the management on the basis of the experience. Finally the score is converted to a rating according to predefined subjective thresholds. It is quite evident that the previous two indexes are rather similar to each other and moreover the level of embedded subjectivity is excessive and not clearly manageable.

The *Fortune's Most Admired Companies* approach aims at ranking the 100 American companies with the best reputation. The objective is achieved by means of a very extensive survey composed of 57 questions and submitted to 100,000 executives, chiefs and financial analysts.

Finally the *RepTrack Index* created by the *Reputation Institute* is built yearly on the basis of 60,000 on-line interviews to consumers from 27 countries. Reputation Institute is the leading reference for what concerns the evaluation of corporate reputation and has a 10 years experience. The model is based on 7 'pillars' useful to create a strategic platform of communication between the company and the stakeholders. Among those dimensions we cite: citizenship, leadership, innovation. The methodology is quite standard since the interviewed are asked to give a score to each specific dimension and finally a rank of the company is produced. The strength of *RepTrack Index* lays on the large dimension of the sample, on the standardization of the results and on the long experience and gained confidence.

On the other side, there exist quantitative approaches that tend to overcome the weaknesses of the qualitative methodologies. The most recognized methodologies are listed below:

- *Intellectual Capital approach;*
- *Accounting approach;*
- *Marketing approach.*

The *Intellectual Capital approach* is based on the appropriate estimation of 5 dimensions: trademark, service marks, copyrights, authorizations and exclusive rights. Since the relative values are traceable along the balance sheets, it is feasible to calculate the amortization quotas. However, the evident limit lays on the heterogeneity of the different

balance sheets, not allowing comparison among several companies. Moreover sudden events that can seriously affect the reputation are not covered.
The second approach, named *Accounting approach*, is based on the evaluation and analysis of intangible assets, thus it is necessary to introduce criteria for fair value assessment. Roughly speaking, a kind of net reputation is computed as the result of the difference between the reputation of assets and liabilities. Finally the *Marketing approach* suggests to measure the brand of a company. The more objective method considers the royalties a company can gather by conferring its brand. By the way, the brand represents only one dimension and thus can not explain all the aspects related to the reputation concept.
As mentioned before, a different quantitative approach was proposed by Gabbi in 2004 (1). He suggests to measure the corporate reputation by means of the financial performance of the analyzed company. The underline assumption consists in the hypothesis that the market imposes an economic cost to stock value, inducing correct ethic behavior.

3. Methodological Proposal

Our proposal follows the framework showed in the previous section. In fact, in order to evaluate the corporate reputation, we propose a parametric statistical model whose estimation allows not only to describe and rank reputation, but also to predict and, therefore to prevent, reputational risks. In particular we need a parametric model suited for ordinal variables, as most reputational data is typically available in such format. In fact ordinal variables request for specific models able to exploit the discrete nature and the inner latent information contained in the data. Several methods have been proposed and employed in this context and among them we can mention the Item Response Theory (IRT) approach typically employed in the field of psychometrics.
The IRT approach makes use of mathematical functions specifying the probability of a discrete outcome on the basis of person and item specific parameters (8). Rijmen et al. have shown that IRT models can be re-phrased according to the general class of nonlinear mixed models (9). This broad category contains both generalized linear mixed models (GLLM) and generalized linear models (GLM). GLM represents the classical approach according to which the target variable (i.e. R, ranking or rating) is distributed as a random variable belonging to the exponential family (10). On the other hand in GLLM, the observations are assumed to be independent realizations from an exponential family distribution conditionally on the random effects (in addition to the covariates and fixed effects). In the light of this different formulation the latent variables, typically present in IRT models, are supposed to follow a distribution, in other words they represent the random effects (11). What we can conclude is that every time the expected value is plenty of information on the ordinal target variable, the above approaches result to be very useful and worthwhile.

Among them we list just a few: proportional ordinal models, IHG based on inverse hypergeometric random variable and BIT based on shifted binomial random variable (12)(13)(14). IHG has been used as statistical model for rank data in order to link the expressed ranks to the main features of the raters thus to explain the sequential choice of several objects (15). However the last two models present drawbacks due to the insufficient flexibility towards the empirical ranking (or rating) distributions and to the inability in modelling all the components guiding the choice process. Mixture models, born to fit responses rates coming from heterogeneous sub-population, deserve particular attention and will be the basis of our parametric approach. Our proposal moves from D'Elia and Piccolo paper of 2005 that presents a model based on a mixture of two different probabilistic structures (16).

In order to compare on a fair ground our proposal with ranking models, we shall also consider a simple non parametric model able to produce rankings in a simple and statistically coherent way. The use of non-parametric methods may be necessary when data presents a ranking (or a rating) but no clear numerical interpretation, as typically happens when coping with the assessment of the reputation of a given institution. In this context emerges the scorecard approach that develops the so-called Self Assessment, based on the experience of a number of internal "experts" of an institution who usually correspond to different areas of activities or processes. An internal procedure of self assessment can be periodically carried out through questionnaires, submitted to such experts. The collected questionnaires give information on which risks are perceived as most important by the chosen experts for a future given period. Once interviews are collected, the aim is to assign an ordinal “rating” to each risk event, based on the distribution of the opinions. Giudici in 2007 proposed to employ the median as a location measure for each distribution, and the normalized Gini index as an indicator of the "consensus" on such location measure (17). That method results in three rating measures for each event, expressed using the conventional risk letters: A for low risk, B for medium risk, C for higher risk and so on. While the median is used to assign a "single letter" measure, the Gini index is used to double or triple the letter, depending on its value.(18) For example: if the median of the frequency distribution of a certain risk type is "yearly", corresponding to the lowest risk category, a letter A is assigned. Then, if all interviewed experts agree on that evaluation (e.g. the Gini index is equal to zero), A is converted to AAA; if instead the Gini index corresponds to maximum heterogeneity A remains A. Intermediate cases will receive a double rating of AA. The same approach can be followed with regards to each question of a questionnaire, leading to a complete scorecard that can be used for intervention purposes. On the other hand, for visualization purposes, colours can then be associated to letters, using a "traffic-light" convention: green corresponds to A, yellow to B, red to C and so on. In this paper we shall apply scorecard models to the available textual reputational data. Different models of operational risk measurement, that take both opinions and data into account, may be used also in reputational risk modeling, (19)(20)(21).

We now focus on our main proposal, that is, a parametric model based on the mixture of two random variables (a shifted Binomial r.v. and an uniform r.v.) able to model effectively ordinal data. Mixture models represent a type of density model which comprise a finite number of component functions, either discrete or continuous. These component functions are combined to provide a multimodal density and they are useful in affording greater flexibility and precision in modeling the underlying statistics of sample data.

As we said above, a mixture model suited to describe ordinal variables arising from questionnaires was introduced by D’Elia and Piccolo (2005) and is based on the appropriate combination of a shifted binomial and an uniform random variable (16). More precisely the main idea of such mixture, named CUB, is to model the latent components (i.e. feeling and uncertainty) that guide questionnaires respondents during the choice process. Moreover CUB r.v. can assume different structures depending on the presence of absence of covariates to be inserted into the model (22).

The above mentioned latent components are defined as follows:

- ***Feeling***: a continuous latent variable expressing the general personal mood towards the item under analysis;

- ***Uncertainty***: a continuous latent variable expressing the personal indecision towards the item under analysis and due to several correlated factors such as the knowledge, the interest, the time spent.

Assume that *Ri* is a random variable that describes the rate (or rank) attributed by an individual to an item, on the basis of *m* possible alternatives (the rating scale). Thus the CUB model can be described formally as follows.

Let $Ri \sim CUB(\pi, \xi)$ indicate a mixture of an uniform and a binomial random variable with parameters π, ξ and m. We assume that π is the mixture weight, $(1 - \pi)$ expresses the uncertainty, m is the rate scale, $(1 - \pi)/m$ the uncertainty share (i.e. normalized uncertainty on the basis of the rate scale) and $(1 - \xi)$, as parameter of the shifted binomial r.v., represents the feeling.

The model is:

$$R_i \sim CUB(\pi, \xi)$$
$$P(Ri = r) = \pi \binom{m-1}{r-1} (1 - \xi)^{r-1} \xi^{m-r} + (1 - \pi)\, 1/m$$

where:

$\pi \in [0,1]$, $\xi \in [0, 1]$, $r = 1, \ldots, m$, $i = 1, \ldots, n$ whit n the number of observations and m fixed and known by data construction. The expected value is:

$$E[R] = \pi (m - 1) (1/2 - \xi) + (m + 1)/2$$

It is interesting to observe that if both π and ξ tend to 1 then the mean value approaches 1 that for a rating problem represents the lowest value on the measurement scale that goes from 1 to m. In order to improve the performance of this structure, an extension of the CUB model with covariates (henceforth CUB(p,q)) has been proposed in Piccolo and D'Elia (23).
This model is as follows:

$$R_{ik} \sim CUB(\pi_{ik}, \xi_{ik}) \quad i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, K;$$
$$\pi_{ik} = F(y_i, z_k ; \beta, \delta);$$
$$\xi_{ik} = G(w_i, z_k ; \gamma, \eta);$$

where K is the number of items, y_i and w_i represent the subject covariates, z_k the items covariates for π and ξ, while F and G are the linking functions (typically logistic). Finally the values $[\beta, \gamma]$ are the subjects regression coefficients and $[\delta, \eta]$ are those for the items. The above mixtures show good performance on several real data set because they are able to model different empirical distributions. Moreover, the possibility to insert covariates into the mixture, appropriately chosen among the available variables, allows to model efficiently several real situations.(24)(25)

It is important to notice that for some real contexts, as the reputational one, covariates are not available by construction or for inaccessibility of further information. Thus, in order to manage a limited information framework, we propose a generalization that is particularly suited in the context of reputation measurement. Instead of taking into account only one shifted binomial random variable, we suggest to employ a mixture with 2 shifted binomial and 1 uniform random variables (henceforth Mixture Uniform Binomial Binomial, CUBB) as follows:

$P(R_i = r) = \pi_1(\binom{m-1}{r-1}(1 - \xi_1)^{r-1}\,\xi_1^{m-r}) + \pi_2\,(\binom{m-1}{r-1}(1 - \xi_2)^{r-1}\,\xi_2^{m-r}) + \pi_3\,(1/m);$

$b_1(r\,;\,\xi_1) = (\binom{m-1}{r-1}(1 - \xi_1)^{r-1}\,\xi_1^{m-r})$

$b_2(r\,;\,\xi_2) = (\binom{m-1}{r-1}(1 - \xi_2)^{r-1}\,\xi_2^{m-r})$

$\pi_3 = 1 - \pi_1 - \pi_2;$

where m is again the rate scale, ξ and π_1, π_2 are respectively the binomial parameters and the weight coefficients. We are now going to derive the expected value and the variance of a CUBB model.

Let μ_{B1}, μ_{B2}, μ_U indicate the mean and σ^2_{B1}, σ^2_{B2}, σ^2_U the variance of a shifted binomial and an uniform distributions:

$Mean : \mu_{Bi} = \xi i + m(1 - \xi_i);\;\; Variance : \sigma^2_{B\,i} = (m - 1)(1 - \xi_i)$

$Mean : \mu_U = (m + 1)\,/2\,;\;\; Variance : \sigma^2_U = (m^2 - 1)\,/\,12$

Their expression can be introduced into the general form of a mixture model mean which is:

$E[R] = \pi_1\,\mu_{B1} + \pi_2\,\mu_{B2} + \pi_3\mu_U$

and after some calculations we have:

$E[R] = (\pi_1\,\xi_1 + \pi_2\,\xi_2) + m(\pi_1\,(1 - \xi_1) + \pi_2\,(1 - \xi_2)) - (\pi_1 + \pi_2)\,(m + 1)\,/\,2 + (m + 1)\,/\,2$

Multiplying and dividing by $\pi_1 + \pi_2$ we have:

$E[R] = (\pi 1 + \pi 2)\,[(\,\pi_1\xi_1 + \pi_2\xi_2)\,/\,(\pi_1 + \pi_2) + m\,(1 - (\pi_1\xi_1 + \pi_2\xi_2)\,/\,(\pi_1 + \pi_2))] - (\pi_1 + \pi_2)\,(m + 1)\,/\,2 + (m + 1)\,/\,2$

Finally the expected value of a CUBB distribution turns out to be:

$E[R] = (\pi_t)(m - 1\,)(1/2 - \xi_t\,) + (m + 1)\,/2$

where $\pi_t = \pi_1 + \pi_2$ and $\xi_t = (\pi_1\xi_1 + \pi_2\xi_2)\,/\,(\pi_1 + \pi_2)$

Note that the above mean is the same as for a CUB model where the binomial parameter is ξ_t and the weight is π_t.

Therefore, as done for the CUB, we can consider the feeling as $1 - \xi_t$ (this can be named total feeling) and the uncertainty share as $(1 - \pi_t)\,/\,m$. Obviously if $\xi_1 = \xi_2$, the feeling is constant and the CUBB coincides with CUB; moreover if one out of the two weights ($\pi 1$ or $\pi 2$) is equal to zero, CUB ≡CUBB.

For what concerns the variance we first calculate the second moment of a mixture which is (26):

$E[R^2] = \pi_1\,E[B^2_1\,] + \pi_2 E[B^2_2\,] + \pi_3 E[U^2]$

therefore substituting the mean and the variance of each single distribution we get:

$E[R^2] = \pi_1(\sigma^2 B_1 + \mu^2 B_1) + \pi_2(\sigma^2 B_2 + \mu^2 B_2) + (1 - \pi_1 - \pi_2)(\sigma^2_U + \mu^2_U)$

and, thus the variance is:

$Var[R] = \pi_1\sigma^2 B_1 + \pi_2\sigma^2 B_2 + (1 - \pi_1 - \pi_2)\sigma^2_U + \pi_1(1 - \pi_1)(\mu_{B1} - \mu_U)^2 + \pi_2(1 - \pi_2)(\mu_{B2} - \mu_U)^2 - 2\pi_1\pi_2(\mu_{B1} - \mu_U)(\mu_{B2} - \mu_U)$.

From the previous expression note that, if one of the two weights (π_i) is zero, then we obtain the variance of the CUB model.

In order to apply this new mixture, first of all we shall derive the estimators of all the involved parameters by means of an EM algorithm, later on we comparatively apply CUB and CUBB models on the same data set.

We finally remark that a covariate-dependant CUBB could be introduced, similarly to what done with CUB. Moreover, the previous distribution can be generalized to a mixture among 1 uniform and p shifted binomials. Indeed this is a theoretical model as, in practice, to fit it $2p$ parameters would need to be estimated and this would require $m > 2p$.

4. Application

As we have already said in section 1, media coverage plays a key role in determining a company's reputation. A detailed and systematic analysis of what the media are saying is especially important because the media shape the perceptions and expectations of all the involved actors. Natural language processing technologies enable these services to scan a wide range of outlets, including newspapers, magazines, TV, radio, and blogs. In order to enable the application of the scorecard approach and of the CUB-CUBB models in this context, we have collaborated with the Italian market leader company in financial and economic communication, "IlSole24ORE" and with DFKI a German Research Center for Artificial Intelligence in the framework of the European research project MUSING. The objective is to evaluate the corporate reputation of 40 Italian companies listed as Blue Chips in the Italian Stock market, on the basis of newspaper articles delivered by "IlSole24Ore". The German center DFKI has analyzed over 1Gb of newspaper articles, dating from January 1st 2009 to June 31 2009, that have been used to train and validate the opinion mining classifier. The opinion mining (OM) tool created by DFKI, through sophisticated natural language processing, is able to capture how verbs, nouns, and other language structures interact.

In essence, the meaning and context of information, not just the words themselves, are extracted from unstructured documents. The opinion mining tool executes a sentiment classification, that is to determine the attitude (a judgment or an evaluation) of a speaker or a writer with respect to a given topic. The OM results pursue data structured according to the following ordinal scale:

1: very good news;
2: good news;
3: neutral news;
4: bad news;
5: very bad news.

Such output represents our variable of interest, which we use to assess the scorecard approach and the mixture model. For what concerns the non parametric approach, first of all we have to choose the position index to represent the reputation variable and secondly the heterogeneity index. As position index, we are obliged to employ the median since the variable is ordinal. As discussed in section 3, there exist several heterogeneity indexes employable with categorical data, among them we cite: Gini, Leti, Shannon, Heterogeneity-concentration indexes and the Concentration ratio.(26) Our choice is the Gini index, the most known; to further justify such decision, we have calculated the correlation index among all those indexes on the basis of simulated data. The result shows positive correlations (very next to 1) among the indexes, indicating their concordance. This means that using other indexes rather than the Gini one, would not bring to meaningfully different results. According to the previous choices, the median can be quantified in A, B, C, D or E (since we have 5 modalities as output of the OM tool) and the final rating depends on the value of the Gini index. The final rating can thus be labelled according to the following codes: AAA, AA, A; B, BB, BBB; C, CC, CCC; D, DD, DDD; E, EE, EEE. We have fifteen different categories of risk rating, that can be represented, for example, by means of a pie chart.(18) As we have already said, the analysis has been applied to the reputation variable produced by OM tool on the basis of 40 Italian companies. For sake of simplicity and representability, such companies have been divided into 12 groups according to the core business: Insurance, Bank, Public Utility, Industrial Services, Food, Automotive, Construction, Energy, Media, Home services, Telco, Spare Time. The relative ratings, obtained through the explained methodology, can be seen in Table 1.

Table 1. Results from scorecard approach

Insurance	**Bank**	**Public Utilities**	**Industrial Service**	**Food**	**Car**
E	***C***	***D***	***E***	***C***	***C***
Construction	**Energy**	**Media**	**Home Service**	**Telco**	**Spare Time**
C	***D***	***C***	***C***	***D***	***D***

From Table 1 we can conclude that the ratings are not particularly high: they all range from C to E. From an interpretative point of view, such results are not positive: the reputation of all the listed groups of companies is far from being positive. The articles, analyzed in the given horizon time, report judgments that are neutral or negative. In order to strengthen the non parametric analysis based on the proposed rating index, we have applied the CUB mixture model to evaluate the latent components (feeling and uncertainty). Table 2 shows the obtained results.

Table 2. Parameters from estimated CUB model

Company Groups	**1-π**	**1-ξ**	**Diss**
Insurance	0.59	0.00	0.11
Bank	0.22	0.38	0.15
Public Utilities	0.37	0.22	0.17
Industrial Service	0.41	0.05	0.07
Food	0.47	0.35	0.28
Car	0.43	0.53	0.09

Construction	0.62	0.59	0.32
Energy	0.09	0.24	0.03
Media	0.00	0.33	0.37
Home service	0.76	0.34	0.29
Telco	0.38	0.24	0.18
Spare Time	0.07	0.20	0.02

In Table 2 we report the companies groups, the values of the latent parameters ξ (feeling) and π (uncertainty). The last parameter Diss[2] (27), reports the dissimilarity index, useful to evaluate the goodness of the model fitting. According to the context analysis, the feeling parameter assumes particular and different meanings: degree of perception, index of selectiveness/awareness, measure of concern, threshold of pain, subjective probability. In our framework we propose to interpret feeling as the reputation awareness, that is the level of consciousness expressed throughout the newspaper articles with regards to the reputation of a given company. The higher is the value, the bigger is the consideration and importance given to the company. On the other hand the level of uncertainty explains well the firmness associated to the given reputation rate. High values for the uncertainty component suggest the presence of a not clear judgment with regard to the company.

Table 3. Parameters CUB vs CUBB (in bold we report best Diss values)

Company Groups	$\xi 1$	$\xi 2$	$\pi 1$	$\pi 2$	DissCUBB	DissCUB
Bank	1.00	0.52	0.17	0.83	**0.06**	0.15
Car	0.47	0.47	0.32	0.25	0.01	0.09
Construction	0.42	0.42	0.22	0.16	0.31	0.32
Energy	0.76	0.76	0.67	0.24	0.03	0.03
Food	1.00	0.46	0.30	0.70	**0.12**	0.27
Home Service	1.00	0.42	0.28	0.61	**0.19**	0.29
Insurance	1.00	0.58	0.42	0.08	0.11	0.11
Media	0.00	0.01	0.08	0.29	**0.08**	0.37
Public Utilities	1.00	0.59	0.29	0.54	**0.07**	0.17
Industrial Service	1.00	0.64	0.49	0.33	**0.02**	0.07
SpareTime	0.82	0.39	0.92	0.08	0.02	0.02
Telco	1.00	0.59	0.25	0.54	**0.01**	0.18

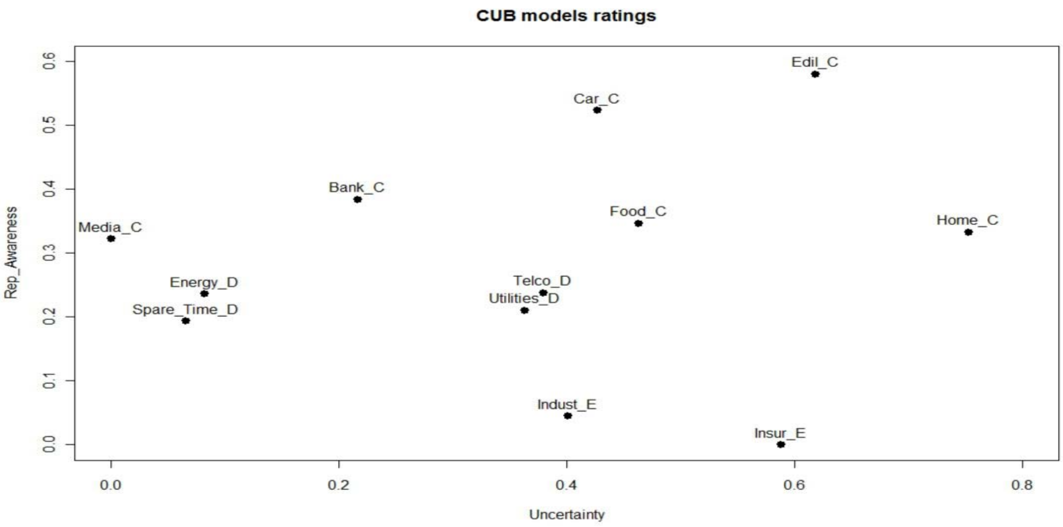

Figure 1. Graphical Representation of CUB model results

Figure 1 offers a clear picture of the combination reputation awareness/uncertainty for each company group. For sake of clarity in the figure we have added to each company group the letter produced by the rating index. It emerges clearly that there is variability in the results, in particular high values of uncertainty are expressed for at least three groups (Construction, Home and Insurance). On the contrary, far from being uncertain are the evaluations for other three groups: Media, Energy and Spare Time. Another interesting aspect is related to the level of reputation awareness; in fact the only two groups labelled with the worse ranting (E) present also very low values of feeling, the groups with label D have medium feeling and finally groups labelled C are characterized by the highest values. Thus we can conclude that there exists a direct relation between the scorecard rating and the estimated feeling awareness, while the estimated uncertainty helps us to evaluate the reliability of the reputation evaluations. Going further with the analysis of the result, we notice that the values of Diss index are rather high. We remark that a general practical rule suggests that Diss should not be greater than 0.10. Such consideration suggests to insert covariates into the CUB model in order to improve the performance. However, as we have already said, covariates are not always available or employable, thus we propose to employ the CUBB mixture proposed in this paper, whose results are reported in table 3. The last two columns of table 3 contain the performance measure (Diss) respectively for CUB and CUBB. As the reader can notice, CUBB is able to improve the performance in many cases (see bold numbers).

Table 4. Results from simulation

Rating Class	Average Feeling
A	0.95
B	0.76
C	0.49
D	0.24
E	0.19

Moreover the fact that in two cases either CUB or CUBB present Diss values greater than the threshold (0.10) depends on the small number of training observations (25 for Construction and 30 for Home service).

In order to verify with a parallel run approach, the coherence between Rank index and CUB model, we have run a simulation study, generating five times 100 different frequency distributions produced by an hypothetic opinion mining tool. Table 4 reports the summary of the simulation; more precisely, simulated data is ordered by the scorecard measure rating classes (e.g. A, B, C, D or E). For each class we have calculated the mean value of estimated feeling parameter (aka reputation awareness). From formula 2 we know that the location index of the CUB models is a monotonic function of the feeling parameter and therefore, we expect concordance between scorecard location ratings (based on the median) and CUB (or CUBB) location estimates (based on the complement of the feeling since we are coping with ratings). From table 2 note in fact correspondence between rating class and the mean feeling parameters.

5. Concluding Remarks

In this paper we have proposed statistical models aimed at measuring effectively reputation of an institution. The scientific literature does not offer an agreed method of the topic given the complexity in defining and measuring the related dimensions. We decided to focus on a specific aspect of the reputation: media reputation. In fact a detailed and structured analysis of what the media are saying is important since the media shape the perceptions and expectations of all the involved actors. Moreover the media can aliment the misalignment between the perception of an event and the reality.

In order to cope with the media reputation measurement, we have proposed a parametric model, that extends the CUB mixture proposed by D'Elia and Piccolo in 2005, particularly useful when covariates are not available. We have compared our model with a simple non parametric ranking model, taken from the operational risk management literature. Our proposed method is a powerful tool, easy to interpret and clearly representable by means of a graphical device. This helps in taking under control the reputation level and to communicate the results to a management board.

The power of the proposed integrated approach has been proved employing a real media-based reputational data. To further confirm the validity of the model, we have run a simulation study. An extension of this paper will investigate the opportunity to employ, in an integrated way, other types of ordered data, as for example those arising from stakeholders questionnaires.

References

1. Gabbi, G., **Definizione, misurazione e gestione del rischio reputazionale degli intermediari bancari**, il Mulino, 2004
2. Economist Intelligence Unit, **Reputation: Risk of Risks**, available online at: http:/www.ies21.edu.ar/jornada/descargas/risk-reputation-report.pdf, 2005
3. Fombrum, C.J and Gardberg, N., **Who tops in corporate reputation? Corporate reputation review**, 3(1), 2000, pp. 13–17
4. Cravens, K., Oliver, E. and Ramamoorti, S. (2003). **The reputation index: Measuring and managing corporate reputation**. European Management Journal 21(2), pp. 201-212

5. Alexander, C. **The Bayesian methods to measuring operational risks,** Mastering Risk, Vol.2, FTPrentice Hall, London, 2001
6. Cruz, M. **Modelling, Measuring and Hedging Operational Risk**. John Wiley and Sons, 2002
7. Siegel and Castellan, **Nonparametric statistics for the behavioral sciences**, 2nd edition. McGraw-Hill, London, 1988
8. Lazarsfeld, P. F. and Henry, N. W. **Latent structure analysis**. Boston: Houghton Mifflin, 1968
9. Rijmen, F., Tuerlinckx, F., De Boeck, P. and Kuppens, P. **A Nonlinear Mixed Model Framework for Item Response Theory**. Psychological Methods, 8(2), 2003, pp. 185-205
10. McCullagh, P. and Nelder, J.A. **Generalized Linear Models, 2nd edition**, Chapman and Jall, London, 1989
11. Agresti, A., Booth, J. G., Hobert, J. P. and Caffo, B. **Random effects modelling of categorical response data**, Sociological Methodology, 30, 2000, pp. 27–80
12. Agresti, A. **Categorical Data Analysis**, 2nd edition, J. Wiley and Sons, New York, 2002
13. Guenther, W.C. **The inverse hypergeometric - a useful mode,**. Statist., Neerlandica, 29, 1975, pp. 129–144
14. D'Elia, A. **A Shifted Binomial Model for Ranking**. Statistical Modelling-XV International Workshop on Statistical Modelling, Servicio Editorial de la Universidad del Pais Vasc, 2000, pp. 412–416
15. D'Elia, A. **Modelling ranks using the Inverse Hypergeometric distribution**, Statistical Modelling: an International Journal, 3, 1, 2003, pp. 65–78
16. D'Elia, A. and Piccolo, D. **A mixture model for preference data analysis**. Computational Statistics and Data Analysis, 49, 2005, pp. 917–934
17. Giudici, P. **Governo dei Rischi: Il ruolo dei modelli statistici**. Isituto Lombardo (Rend. Lett.), Edizioni Universitarie di Lettere Economia Diritto, 141, 2007, pp. 361–376
18. Cerchiello, P., Dequarti. E., Giudici, P. and Magni, C. **Scorecard models to evaluate perceived quality of academic teaching**. Statistica Applicazioni, 2, 2010, pp. 145-156
19. Bonafede, E. C., Cerchiello, P. and Giudici, P. **Statistical models for Business Continuity Management**, Journal of Operational Risk 2, 4, 2007, pp. 79–96
20. Argentero, P., Cerchiello, P., Giudici, P. **Models for questionnaire data in injury risk management**, Advances and Applications in Statistics, 10, 1, 2010, pp. 77-100
21. Cerchiello, P., Bonafede, E. C. **A Proposal to Fuzzify Categorical Variables in Operational Risk Management**, Studies in Classification, Data Analysis, and Knowledge Organization, 4, 2010, pp. 129-135
22. Piccolo, D. **Observed information matrix for MUB models**, Quaderni di Statistica, 8, 2006, pp. 3–78
23. Piccolo, D. and D'Elia, A. **A new approach for modelling consumers' preferences** food Quality and Preferences, 19, 2007, pp. 247–259
24. Iannario, M. and Piccolo, D. **University teaching and students perception: models and evidences of the evaluation process**, Proceedings of DIVAGO Meeting, University of Palermo, Springer, forthcoming, 2009
25. Cerchiello, P., Iannario, M., Piccolo, D. **Assessing risk perception by means of ordinal models**, Proceeding of 2008 MAF Meeting, Springer, 2009, pp. 65-73
26. Johnson, N.L., Kotz, S. and Kemp, A.W. **Univariate discrete distributions**, J. Wiley and Sons, New York, 1989
27. Leti, G. **Sull'entropia. Su un indice del Gini e su altre misure dell'eterogeneita' di un collettivo**, Metron, XXIV, 1-4, 1965, pp. 332–378
28. Simonoff, J. S. **Analyzing catgorical data**, New York, Springer, 2003

---

[1] Paola Cerchiello has BS in Economics at the University of Pavia, 2002; Degree IUSS in 2002; PhD in Statistics at the University Milan-Bicocca in 2005. She has spent a research period abroad: in particular at the Rutgers University, New Jersey. She is researcher in statistics at University of Pavia and she is lecturer of: statistics and statistical data analysis. Author of publications in the area of methodological statistics, ordinal models, textual analysis and statistical models for operational risk management. She is still involved in a research project granted by The Minister of Economic Development on "*Tecnologies for Made in Italy*" (@abilita, 2010-2012). She is involved

in the European network COBEREN on consumer behaviors. She has participated to several projects: (P.R.I.N. 2009-2010) on "*Multivariate Methods for the evaluation of the quality and of the risks in academic context*", (P.R.I.N., 2007-2008) on "*Multivariate statistical methods for the evaluation of public services*", (F.I.R.B, 2006-2009) on "*Data mining methods for small and medium enterprises*", (European Integrated Project M.U.S.I.N.G., 2006-2010) on "*Data mining models for advanced business intelligence applications*". She has also participated to the territorial project "*Antica Pieve di Mazzo*" and to the European network 'Young' for comparative studies on consumer behaviour'.

[2] Diss = $\frac{1}{2}\sum_{r=1}^{m}|p_o(r)-p_e(r)|$ where po(r) is the observed relative frequency, pe(r) is the frequency expected by the model and m is the rank scale length