

Longitudinal models for market reputation and risk ¹

Modelli longitudinali per il rischio reputazionale

Silvia Figini and Juan Tomas Sajago ²

Dipartimento di Statistica ed Economia Applicate “Libero Lenti”,
Università di Pavia, e-mail: silvia.figini@unipv.it

Riassunto:

L'obiettivo di questo lavoro é quello di presentare un approccio metodologico per lo studio di dati relativi al rischio reputazionale di aziende quotate in Borsa Italia. Il rischio reputazionale risulta caratterizzato sia da descrittori qualitativi, estratti ad esempio da notizie di articoli di giornale, che quantitativi. Il contributo evidenzia come integrare opportunamente dati qualitativi e quantitativi per lo studio della reputazione di una nota realtà bancaria italiana. L'approccio metodologico risulta essere generalizzabile anche per differenti settori economici e in generale per problemi di valutazione.

Keywords: Reputation Risk, Basel 2, Unstructured Data, Opinion Mining

1. Introduction

Perhaps the most critical, strategic asset that a corporation possesses is its reputation. Although corporate reputation is undoubtedly a significant and relevant corporate asset, formidable measurement challenges have effectively kept this major intangible asset out of the financial statements. We propose the creation of a reputation score that would be of broad scope and attempt to capture key dimensions and evaluate diverse organizational components including corporate strategy, financial strength and viability, organizational culture, ethics and integrity, governance processes and leadership, products/services, strategic alliances and business partnering, and innovation along with information already contained in the corporations annual report. The main contribution of this paper is to propose statistical models useful to predict reputation. Empirical evidences are given on the basis of a real data set composed of qualitative and quantitative information provided by SOLE24ORE.

2. Predictive models for reputation measurement

The first approach proposed in the financial literature to evaluate reputation is Gabbi (2003). Equations 1, 2 and 3 report the models proposed by Gabbi for reputation measurement:

$$R_{i,t} = \alpha_i + \beta_{i1}R_{Mkt,t} + \beta_{i2}R_{Banks,t} + \epsilon_{i,t}, \quad (1)$$

$$R_{i,t} = \alpha_i + \beta_{i1}R_{Mkt,t} + \beta_{i3}Repu_{i,t} + \epsilon_{i,t}, \quad (2)$$

¹This work has been supported by MUSING 2006 contract number 027097, 2006-2010.

$$R_{i,t} = \alpha_i + \beta_{i1}R_{Mkt,t} + \beta_{i2}R_{Banks,t} + \beta_{i3}Repu_{i,t} + \epsilon_{i,t}, \quad (3)$$

where for a specific company i at time t , $R_{i,t}$ is the return, $R_{Mkt,t}$ is the return in the market index (i.e. MIB30), $R_{Banks,t}$ is a hierarchical sectorial indicator (i.e. Sectorial MIB) and finally the reputation variable ($Repu_{i,t}$) that is a qualitative variable representing a measure of polarisation. More precisely $Repu_{i,t}$ is a qualitative ordinal variable derived from an opinion mining process on unstructured data (newspaper articles), while $R_{Mkt,t}$ and $R_{Banks,t}$ are based on structured data. As we can easily observe, Equation 1, 2 and 3 are special cases of multiple linear regression analysis. In order to gain improvements for reputation measurement, we propose to employ other types of models such as Generalized Additive Models (GAM) with a structure like:

$$g(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_k(x_k), \quad (4)$$

where the $f_1, f_2, \dots, f_k$ are smooth functions, β_0 is the intercept, g is a known link function and y_i is independent with some exponential family distribution. The functions $f_j(x_j)$, $j = 1, \dots, k$ may be fit using parametric or non-parametric means, thus providing the potential for better fits to data than other methods.

3. Application

We describe our reputation models by means of an example, that concerns the Italian financial company Italease. The data set is composed of 4 variables (three quantitative features and one qualitative). The data available is composed of 860 observations starting from June 15 2005 to October 31 2008. The quantitative data used comes from the Italian financial market and this information corresponds to the price index for Italease, Mib 30 and the sectorial index. It is important to review how each quantitative 'variable' is constructed: $\text{Return}_{\text{variable}} = \frac{\text{Variable}_t - \text{Variable}_{t-1}}{\text{Variable}_{t-1}}$ where 'variable' refers to the daily closing price of Italease, Mib 30, and Sectorial Mib (in this case the banking sector). In order to understand the motivation of our choice and be able to speak about reputation for Italease, it is important to remember that Italease is an Italian bank specialized on leasing formed by popular banks who had signed a pact of stability. Furthermore, it should be remembered that the company in June of 2007 began a deep process of reorganization and have made more internal changes. Concerning the qualitative variable for reputation measurement, as pointed out before, is derived on the basis of an opinion mining procedure that classifies the newspaper information collected in SOLE24ORE. Opinion mining is a methodology that, according to a predefined dictionary, gives a value to every word and provides consideration to the different types of word used in articles: such as adjectives, adverbs, complements, prepositions and others. As final result, the qualitative variable $Repu$ reports the following labels: 0 (worst news), 1 (bad news), 2 (neutral news), 3 (good news) and 4 (great news). We remark that there is only one value for $Repu$ for each date (t). Based on the data at hand, we move our attention on predictive models for reputation and more precisely, as described in Section 2, we compare linear regression models (Gabbi, 2003) with GAM. The results of the Gabbi models are not really good in terms of key performance indicators as reported in Table 1. The quantitative variables are significant, but the reputation variable derived from the opinion mining process shows a lower impact on reputation forecasting. The models in Table 1 are based on backward, forward and stepwise selection respectively. In order to improve the results a set of GAM are proposed:

Process & equation Variable	$Model_1$ (1)	$Model_2$ (2)	$Model_3$ (3)
Intercept	0.00024 (0.817)	-0.00874 (0.0293)*	-0.0087 (0.0275)*
Mib30	0.8604 (2.23e-11)***	1.3608 (2e-16)***	0.8507 (3.36e-11)***
Sectmib	0.6144 (2.62e-07)***		0.615682 (2.31e-07)***
Repu		0.00461 (0.0216)*	0.00465 (0.0185)*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			
R-Squared	0.2695	0.2511	0.2742
F-statistic	158.1 (2.2e - 16)	143.7 (2.2e - 16)	107.8 (2.2e - 16)

Table 1: Gabbi Multivariate Linear Models: Results

1. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}, x_{2t}) + \epsilon_i$
2. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}) + \beta_2 f_2(x_{2t}) + \epsilon_i$
3. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}) + \beta_2 f_2(x_{2t}) + \beta_3 Repu + \epsilon_i$
4. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}, x_{2t}) + \beta_2 Repu + \epsilon_i$
5. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}) + \beta_2 f(x_{2t}, Repu) + \epsilon_i$
6. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}, Repu) + \beta_2 f(x_{2t}) + \epsilon_i$
7. $E(y_t) = \alpha + \beta_1 f_1(x_{1t}, x_{2t}, Repu) + \epsilon_i$

where y_t , x_{1t} , and x_{2t} are the values of Italease, Mib30 and Sectorial Mib returns. In order to select the best model, we report in Table 2 the anova analysis and we select as best models, Model 1, 4, and 7. We select as best model, Model 4 which explains 29.8%

Model	Resid. Df	Resid. Dev	Df	Deviance
1	847.39415	0.82451		
2	846.47796	0.82054	0.91619	0.00397
3	844.96110	0.81516	1.51685	0.00539
4	846.64006	0.81991	-1.67896	-0.00475
5	842.89305	0.80341	3.74701	0.01650
6	843.08550	0.80389	-0.19245	-0.00049
7	849.05361	0.82666	-5.96812	-0.02277

Table 2: The Anova analysis for GAM models

of the total variance.

	Estimate	Std. Error	t value	$Pr(> t)$
(Intercept)	-0.009125	0.003918	-2.329	0.0201 *
Repusole	0.004547	0.001961	2.319	0.0206 *
	edf	Ref.df	F	p-value
$s(Mib30, Sectmib)$	11.36	11.86	28.55	2e-16 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
R-sq.(adj)	0.288	Dev expl	29.8	
GCV score	0.00098371	Scale est.	0.00096843	

Table 3: *Generalized Additive Model with an smoothing parameter for Mib30 and Sectorial Mib*

4. Conclusions

In this contribution we have presented a framework for reputation measurement. The statistical models that we have proposed to measure reputation could be improved, on the basis of better qualitative information and new data. Further research is needed in the area, especially concerning the integration between qualitative and quantitative data.

References

- Gabbi G. (2003) *Definizione, misurazione e gestione del rischio reputazionale degli intermediari bancari*, Mulino.
- Hastie T. and R.J. T. (1990) *Generalized additive models*, Chapman and Hall.