OBEGEF – Observatório de Economia e Gestão de Fraude

WORKING PAPERS

#19 Risk prevention of public procurement in the brazilian government using credit scoring >>

Leonardo Sales

OBEGEF
Observatório de Economia
e Gestão de Fraude

>> FICHA TÉCNICA

RISK PREVENTION OF PUBLIC PROCUREMENT IN THE BRAZILIAN GOVERNMENT USING CREDIT SCORING

WORKING PAPERS Nº 19 / 2013

OBEGEF – Observatório de Economia e Gestão de Fraude

Autores: Leonardo Sales[1]

Editor: Edições Húmus

1ª Edição: Janeiro de 2013

ISBN: 978-989-8549-64-8

Localização web: *http://www.gestaodefraude.eu*

Preço: gratuito na edição electrónica, acesso por download.

Solicitação ao leitor: Transmita-nos a sua opinião sobre este trabalho.

Paper in the International Conference *Interdisciplinary Insights on Fraud and Corruption*

[1] Brazília, Brazil. leonardojs@hotmail.com

>> ÍNDICE

>> RESUMO

Modelos de *Credit Scoring* são aplicações estatísticas utilizadas por instituições financeiras para classificar pretensos clientes quanto à possibilidade de se tornarem inadimplentes. Este trabalho visa trazer essa experiência consagrada na iniciativa privada para o contexto governamental, buscando adaptá-la e testar seu desempenho na identificação de licitantes propensos a descumprir obrigações em contratos com o governo. Compara os resultados de metodologias baseadas em diferentes técnicas estatísticas. Pretende contribuir para o controle preventivo dos riscos contratuais, tanto por parte do gestor público quanto por parte dos órgãos de controle.
Palavras-chave: Auditoria Governamental. Análise de Dados. Árvores de Decisão. Regressão Logística. *Credit Scoring*.

>> ABSTRACT

Credit Scoring models are statistical applications used by financial institutions to classify applicants as to the possibility of becoming defaulters. This work aims to bring that good experience from the private sector to the governmental context, seeking to adapt it and test its performance in identifying bidders likely to fail in the fulfillment of obligations under contracts with the government. The results of methods based on different statistical techniques are compared. We hope to contribute to the preventive control of the contractual risks, both by the public manager as by the agencies of government control.
Keywords: *Government Auditing; Data Analysis; Decision Tree; Logistic Regression; Credit Scoring.*

>> 1. INTRODUCTION

The insufficient or no provision of contracted services is one of the major problems in the Brazilian public contracts with private companies. His damages are always enhanced by the presence of the public interest purpose in these contracts.

Such situations are reflected nowadays in great challenges for both public administrators and agencies of government control.

Side of the manager, there is the need of contracts with suppliers who can meet their obligations and provide a high quality service, while ensuring the fairness of the process, in obedience to a complex and not flexible legal framework.

On the other hand, the instances of public control (public institutions of audit) have the challenge of seeking to anticipate risk situations that may impact the achievement of government objectives, among them the inadequate provision of contracted services from the private sector.

We assume that this research would be relevant to the government units of audit to understand in advance the probability of a given contract has problems.

Therefore, the focus of this paper is to prevent what we call the breach of contract, understood as the failure to provide the service. We seek to test the applicability of predictive models used by credit institutions (usually based on multivariate data analysis) in the prevention of similar cases in public agencies.

Among the models used by financial institutions, we selected the group called Credit Scoring as object of study. Credit scoring models are based on the statistical weight of a feature (registration or historical) of the company to calculate its probability of becoming defaulter. These models are widely used by such institutions and have been producing good results in the prevention of default.

The fact that these models are based on the characteristics of business is another assumption which we start believing in this study to be feasible to apply these techniques in government contracts. This is because it is possible to extract much information on the suppliers of the bases used by the federal government that are available to the control agencies.

We sought to identify which of these features contribute to an increased likelihood of suppliers fail to comply with their obligations with the government, allowing thus the risk of problems is known in advance.

We tested two statistical classification techniques used in Credit Scoring Models: Logistic Regression and Decision Tree, using as input a database refers to a sample of government contractors divided into two groups: those who have a history of breach of contract (called here "defaulter") and those with a history of compliance with obligations (called "responsible"). The performance of the techniques will be measured by comparing planned with actual ratings.

All research was set in the Office of the Comptroller General - *CGU*, which is the central organ of the internal control system of the Brazilian government[1]. We use the databases and the structure of hardware and software of the Public Expenditure Observatory - *ODP*, unit that is a reference in data analysis and production of strategic information[2].

The results are discussed from the evaluation of the accuracy of each technique to predict the situations of default and the understanding of the variables that have greater predictive power.

It is hoped that the findings of this study contribute at CGU to build a robust methodology for risk assessment contract, based on data analysis, that can be applied largely and increase the public expenditure management and its control.

1 CGU has an intelligence unit, called Public Expenditure Observatory, which makes use of information technology to monitor public spending at a distance. This work generates alerts that represent transactions (purchases, for example) suspected. More comments on the ODP can be found at http://www.cgu.gov.br/ODP.

2 ODP won the United Nations Public Service Awards 2011, in the category Advancing Knowledge Management in Government (http://unpan1.un.org/intradoc/groups/public/documents/un-dpadm/unpan048624.pdf)

>> 2. LITERATURE REVIEW

2.1. Credit Scoring

The researchers define as Credit Scoring the quantitative models adopted by financial institutions to measure and manage the risk of default by the borrower (Lewis, 1992). The correct understanding of risk helps to measure the amount of credit, its rate or term, and prevents difficulties in contracts in progress (Saunders, 2000).

In general, these models are based on multivariate statistical techniques to classify the customer in different groups, according to the default risk (Hand and Henley, 1997). For this purpose, they use as input various information, present in cadastral databases, financial or historical record of customer operations with the institution (Araújo and Carmona, 2007).

The idea of using quantitative models for assessing and managing credit risk is directly related to expansion and massification of credit that occurred worldwide in recent decades, which showed the need to standardize the procedures for selection of clients and portfolio management (Marques, 2002).

Vasconcellos (2002) argues that traditional methods of assessment credit risk are based on expert opinion and are therefore subjective and individualized. Thus, do not apply satisfactorily to large credit markets. Marques (2001) adds that although these models have improved by aggregating scores (called credit ratings) still shown to be inefficient on large volumes of transactions, because they are slow and not standardized.

Therefore, the Credit Scoring models added standardization, speed and accuracy to the processes of selection and management of clients of these institutions. The structuring of systems is responsible for extending the process of analyzing credit risk, which became less dependent on subjective factors.

Another factor that favored the viability of these models was the development of computers, combined with the widespread use of computerized systems that made possible gathering and structuring large databases relating to their transactions and customers. Likewise, increasing the processing capacity of the computers also enabled the use of robust statistical techniques applicable to big volumes of data (Hand and Henley, 1997).

2.1.1. How the Credit Scoring models work

In recent papers that propose or test Credit Scoring models applied in different contexts (detailed in section 2.1.2), there is a common basis on which these systems are structured.

In general, these models are constructed from three elements: a database with information about the customers (companies or individuals), another database with negative experiences and one or more statistical techniques able to learn from this experience and predict the risk situations.

The steps of its construction generally involve: 1) the initial selection of samples of good and bad cases, creating two groups generally classified as "responsible" and "defaulter", 2) the definition of variables to be used by the model to differentiate groups, and 3) the implementation of statistical processing of learning and testing accuracy.

This last step is crucial for the classification. First, the statistical tool "learns" from the base, assigning "weights" to the variables and revealing those most able to separate the groups, then create a "formula" that define the score of each company, and finally apply this formula to each case, allowing the final classification.

The usefulness of the analysis will depend on the comparison of the ratings generated with the original classification. For this is calculated the percentage of correct classifications (accuracy of the model).

2.1.2. Applications of Credit Scoring Models

The classic application of the Credit Scoring models is the selection of tenderers to credit (Caouette et al., 1998). Beyond it, there are examples in the literature of applications of these models in other stages of the life cycle of credit. Among other times, it can be applied in prospecting for customers, when it receives the name of Response Scoring, during the execution of the contract (Behavioral Scoring) or even after a default has already consummated (Collection Scoring), to identify borrowers more likely to pay the debt (Semedo, 2009).

As an example of traditional application of Credit Scoring, we can cite Vasconcellos (2002), which proposes a methodology for analysis of lending to individuals, using as an object of study the history of operations and registration information of customers of a popular Brazilian bank. In the same vein, Gevert (2011) investigates the performance of different statistical techniques for predicting defaults on contracts of a bank.

Araujo and Carmona (2009), studying a microfinance institution, also discuss the possibility of using Credit Scoring in Behavioral Scoring method and compare the performances of two statistical techniques: Discriminant Analysis and Logistic Regression.

2.2. Statistical Techniques used in Credit Scoring models

As stated Ghodselahi (2011), Discriminant Analysis and Logistic Regression are historically used in the construction of score-cards. A favorable point for the use of these techniques is that they are easily found in the statistical packages available.

More recently, other classification techniques have been used, such as Decision Trees, Artificial Neural Networks and Support Vector Machine.

We will discuss the technical operation of Logistic Regression and Decision Trees, which are used in this study.

2.2.1. Logistic Regression

The Logistic Regression (LR) uses independent variables to classify elements in categorical variables and is based on a discriminant function constructed through of the evaluation of the capacity of each variable to distinguish the groups. Hair et al. (1998) cites the main aspects that favor the use of this technique in relation to Discriminant Analysis (another technique which, as seen, is widely used in systems of Credit Scoring): limitation of LR on the number of possible categories (two) and non-dependence of this technique to the requirement of normality of the independent variables, which also need not be numerical.

The result of the logistic function (*logit*) is binomial, assuming the values 1 or 0, indicating the presence or absence of certain characteristic.

The elements that make up the logistic function are:

α = intercept (representing the minimum value of the function);

β = coefficient of each variable that determines its ability to move the function towards 1 or 0;

x = absolute value of certain variable examined in the record;

Thus, the function assumes the following form of the equation 1:

$$P(X) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i Y_i)}}$$

According to Hair et al. (1998), the value P (x) is the probability of the function assume the value 1(indicating the presence of the studied characteristic).

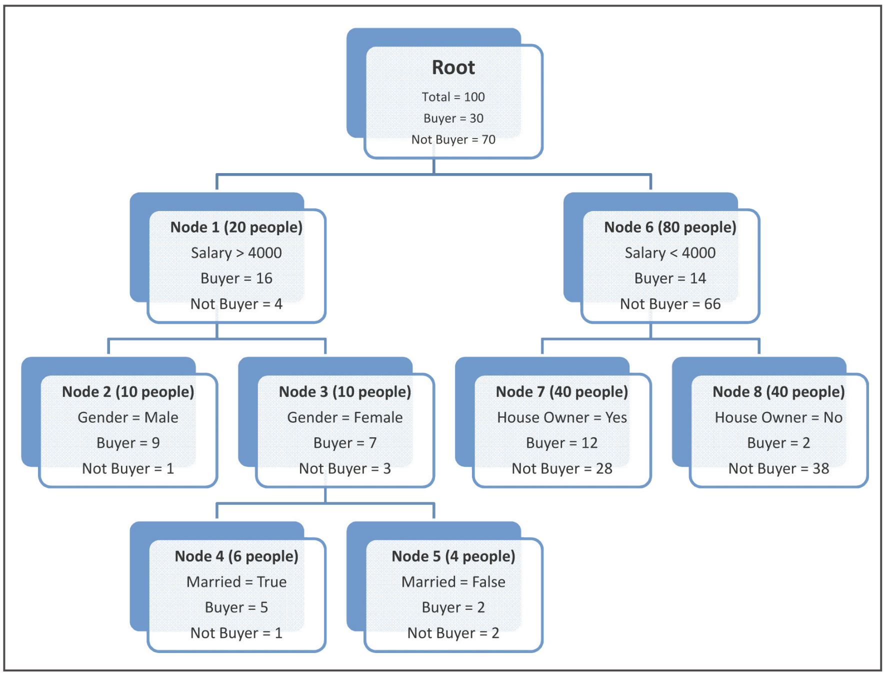
2.2.2. Decision Tree

The Decision Tree is a widespread technique to build classifiers. They consist of a simple graphical representation of knowledge acquired through learning data analysis (Shibba et al, 2005).

Cortes et al. (2002) compares the decision tree to a flowchart composed of the following elements: 1) nodes, which represent questions concerning attributes (predictor variables), 2) branches, which indicate answers to these questions, and 3) leaves (which are nodes without branch) that denote the distribution of records according to the response variable (dependent).

Figure 1, taken from Cortes et al. (2002), exemplifies a Decision Tree applied to classify customers into buyers or not buyers, based on the variables salary, sex, marital status and residence property.

Figure 1: Decision Tree for prospective buyers (Côrtes et al., 2002)



In this classification, the technique found the variables that best divide the two groups, calculating the value (concerning the variable) more effective in separating them. This generated a hierarchy of divisions that would be used to classify each new records analyzed. Imagine that a new potential client have salary of 3000 and did not own their own home. In this case, following the series of divisions (branches of the tree) and passing through the nodes 6 and 8, he would be classified as non-buyer.

2.2.3. Applications in academic research

Logistic Regression pervades many academic studies that propose models of Credit Scoring, among them we could highlight Araujo and Carmona (2009), that compares the LR with a Discriminant Analysis model and propose risk management for a microfinance institution, and Semedo (2009), which investigates the usefulness of the technique in market companies in Cape Verde.

Vasconcellos (2002) uses decision trees to construct a model of credit risk management of individuals in a Brazilian bank.

More recent studies seek to use other techniques. Semedo (2009) compares the performance of logistic regression with the technique of Artificial Neural Networks. Ghodeslahi (2011) and Gevert et al (2011) test the technique of Support Vector Machine, comparing their degree of agreement with other techniques such as Neural Networks and Logistic Regression.

2.3. Government Biddings

The whole process of buying goods or hiring services in the federal government is under the rules of the Law n°. 8666/1993, entitled Bidding Law. Other normative complement this law as the law n °. 10,520 in 2005, establishing a modality of reverse auction, and the Complementary Law N °. 123, which provides privileges for micro and small companies in bidding. Law n°. 8.666/93 details the phases of the bidding process, allowed procedures for bidding, types of contracts, aspects of the qualification of companies and also provides administrative and criminal sanctions to be applied to suppliers in case of noncompliance.

We highlight the following parts of the Bidding Law which defines the administrative penalties:

> *Seção II*
> *Das Sanções Administrativas*
> *Art. 87. Pela inexecução total ou parcial do contrato a Administração poderá, garantida a prévia defesa, aplicar ao contratado as seguintes sanções:*
> *I - advertência;*
> *II - multa, na forma prevista no instrumento convocatório ou no contrato;*
> *III - suspensão temporária de participação em licitação e impedimento de contratar com a Administração, por prazo não superior a 2 (dois) anos;*
> *IV - declaração de inidoneidade para licitar ou contratar com a Administração Pública enquanto perdurarem os motivos determinantes da punição ou até que seja promovida a reabilitação perante a própria*

> *autoridade que aplicou a penalidade, que será concedida sempre que o contratado ressarcir a Administração pelos prejuízos resultantes e após decorrido o prazo da sanção aplicada com base no inciso anterior.*
> *(Brazil. Law n°. 8.666/1993)*

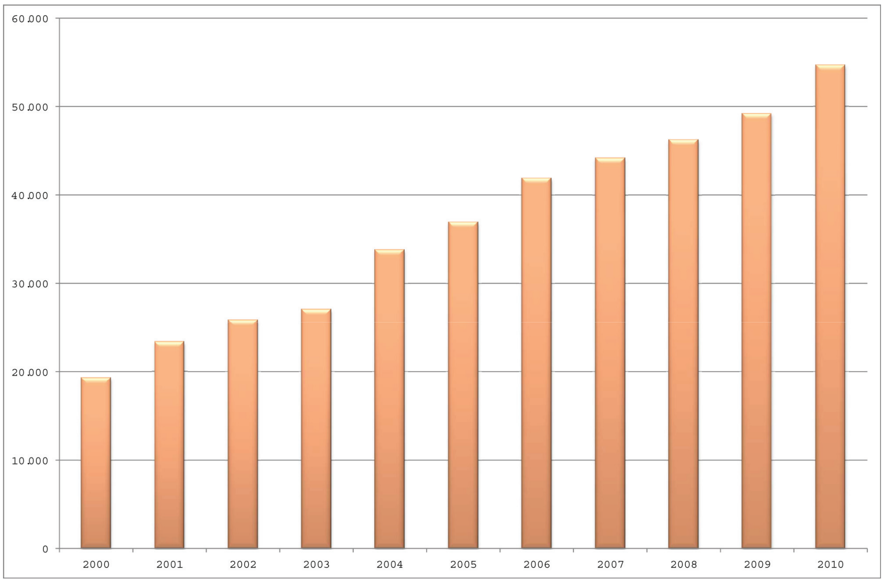
The article 87 states the punishments that the companies would face if they fail to comply with the commitments entered into with the government.

The sections III and IV will be particularly relevant in this work. They define the companies that comprise the National Registry of Companies Not-Reputable or Suspended - *CEIS*, treated in section 2.4.1.

2.3.1. Integrated System of General Services Administration - *SIASG*

As predict Law N°. 8.666/93, the process of procurement and hiring in the federal government is done into the Integrated System of General Services Administration - *SIASG*. Every purchase or contract is registered in the system since the opening of proceedings until the issue of commitment. During the execution of contracts public managers can use the system to assign penalties for suppliers, if any breach of obligations is detected.

Figure 2 - Quantitative contracts registered in SIASG, 2000 to 2010



Source: Integrated System of General Services Administration - SIASG

Existing since 1994, *SIASG* recorded more than 400,000 contracts and nearly 4 million purchases.

Figure 2 shows the evolution of the quantity of contracts published between 2000 and 2010.

2.4. Control over government procurements

As shown in Figure 2, only in 2010 more than 50,000 contracts were published, and we can see a linear trend of growth. This large volume of contracts is under the government audit/control agencies supervision, in particular the Office of the Comptroller-General - *CGU*, which is the central unit of the Internal Control System of the Federal Government[3].

The CGU, in its manual of audit procedures, provides that, when auditing a governmental unit, each audit unit must select the contracts following criteria of materiality, criticality, relevance and operational capability of the area to be audited. Then, the contracts have to be separated into three groups: layoffs, non-requirement of bid tender and other modalities.

CGU also acts preventively. In 2009 created a systematic of previous analyzing bidding notices, that look for some indications that the purchase could be being targeted (attempt to favor one bidder). Usually, the targeting is related to the requirements that the supplier must complete to be hired.

2.4.1. Administrative penalties to suppliers

In 2009, CGU created (fulfilling one of the goals of the National Strategy Against Corruption and Money Laundering – *ENCCLA*), the National Registry of Companies Not-Reputable or Suspended - *CEIS*[4]. The purpose of the register is to consolidate the list of companies that suffered sanctions by agencies and government entities of the several federal spheres of Public Administration (Municipalities, States and Federal).

These penalties (Declaration of Not-Reputable and Suspension to compete in bidding) are provided in the Law 8.666/93 and arise from non-execution or partial execution of the contract and the existence of proven fraud at any stage of the contract or the procurement itself.

The CEIS currently has more than 5,000 records.

3 The Federal Law nº. 10.180/2001, assigned to the role of CGU central organ of the Internal Control System of the Federal Executive.

4 http://www.portaldatransparencia.gov.br/ceis/SaibaMais.asp

>> 3. REASONS AND OBJECTIVES

We understand that the construction of statistical methods of data analysis, focusing on management and risk prevention on a large scale, will innovate how the CGU operates in audit tenders and contracts.

There are gaps in the process of risk assessment of government procurement by this agency of control. This is because the analysis of procurement processes always occur in a punctual manner (even when it is preventive, as the previous analysis of bidding notices), in a number of processes limited to the operational capability of the audit units and the fragmentation of work on the structure of the government. The very subjective nature of assessments, such as the study of bidding notices, allows that only few of them were actually analyzed.

This context of ad-hoc analysis depending on expertise also exist in banking institutions, and has motivated the credit risk management departments to develop Credit Scoring models. Nowadays, such models work supporting the activity of the professional, who develop subjective analysis only for contracts of great materiality, attributing to the tools the great mass of statistical analysis.

This paper attempts to take the first step towards a methodology based on statistical and data analysis that allows prediction risk situations in government hires and could be applied to 100% of processes, supporting the auditor to prioritize their actions. It focuses on one of the key elements of the problems encountered in these contracts: the supplier.

The specific goals of this research are: 1) Test the applicability and the performance of Credit Scoring models to predict the breach of contract caused by the supplier, and 2) Understand the characteristics of suppliers that contribute to themselves to be more prone to the condition of defaulting.

>> 4. METHODOLOGICAL CONSIDERATIONS

In this research, we tested the performance of two statistical classification techniques, often used in Credit Scoring models: Logistic Regression and Decision Tree.

To apply in the context of public biddings, we use a database that represents two distinct groups: one of companies supposedly likely the breach of contract and other with companies considered less prone to this type of situation. The database was also divided into groups for knowledge and test. The next section will explain the criteria used to build the database.

We tried, with the use of classification techniques, to identify and understand the variables that could best separate these two groups. Finally, we compare the results of each technique, calculating each hit percentage.

4.1. Criteria for classification of groups

We divided the suppliers into two groups, which we call "Responsible" and "Defaulter".

To be classified as "defaulter" is enough to the company to be registered in the CEIS, described in Section 2.4.1. The fact that the supplier was registered in the CEIS means that it has committed serious flaws in the execution of any contract.

To be classified as "responsible" not only is necessary to the company not be registered in the CEIS, but also it had to have a minimum of five contracts concluded with the government over the past five years.

The next section describes the criteria for selection of cases for analysis.

4.2. Construction of the main database

The database used contains 2000 companies, 1000 is considered "responsible" and 1000 "defaulter". Each group contains 500 firms selected for knowledge and 500 for testing databases.

The 1000 companies in the group "defaulter" were selected randomically from the database of the CEIS, after exclusion the following cases: individuals, government, businesses with no contract with the federal government and companies punished exclusively by municipalities or states. After these exclusions, our universe was reduced from 4100 to 1222 companies.

The construction of the group "responsible" was much simpler. First, since as we started from a very large universe of companies (according to the criteria described in the previous section), we selected those who signed contracts this year (2011). We accept here that this action would reduce the universe without bias the model. From the new universe, of 4011 companies, we selected 1000, also randomically.

Table 1 describes the construction of the base:

Table 1 - Procedure for construction of the base

	Group “Defaulters”	Group “Responsible”
Initial universe of research	All registered punished in CEIS (4,100 companies)	All supliers with no record in CEIS, with at least five contracts completed in the last five years (21,743 companies)
Selection criteria	Excluded: individuals, government, businesses with no contract with the federal government and companies punished exclusively by municipalities or states.	Selected: companies with contracts signed in 2011.
New universe	1,222 companies	4,011 companies
Sampling	Random	Random
Main database	1,000 companies	1,000 companies
Learning database	500 companies	500 companies
Test database	500 companies	500 companies

4.3. Statistical techniques selected

To build the classification models we used Logistic Regression and Decision Trees techniques.

We decided to use logistic regression because this technique is one of the most found in models in the Credit Scoring and, moreover, is less restrictive than the assumptions of the database used, nor requiring of predictor variables the normality of distributions neither the exclusive presence of numeric values.

This flexibility is also present in the Decision Tree technique. According Meira et al. (2009), several variables in these models, numerical or categorical, could be worked simultaneously without damage to its performance and reliability. Moreover, this technique can graphically represent the functionality of the prediction model, which is an advantage over regression models.

To enable the models construction we used the software STATISTICA, from Statsoft, Inc.

4.4. Definition of variables

The response variable used, called the dependent variable, is the group ("responsible" and "defaulter"), and will be obtained by the models through the analysis and weighting of the independent variables (predictors).

The predictors variables used in the models were selected according to their availability in the database that CGU has access, and also according to the consistency of the observed records (no missing data or unreliable, for example).

Table 2, below, describes each predictor variable chosen:

Table 2 – Predictor variables

Variable	Description	Source
Days until 1st contract	Number of days elapsed between the opining of the company and the first contract	SIASG and database of the Federal Revenue
Campaign donations in 2010	Amount donated to political campaigns in 2010	Brazilian High Court of Elections
Occurrences in monitoring made by CGU	Cases in which the company was selected in some trail monitoring of CGU	CGU[1]
No. of partners	Number of shareholders of the company	Database of the Federal Revenue
No. of employees	Number of employees	Database of Ministry of Social Security
Median age of partners	Median age of the shareholders of the company	Database of the Federal Revenue
Median age of employees	Median age of employees	Database of the Federal Revenue and of the Ministry of Social Security
Median salary of partners	Median salary of the shareholders of the company	Database of the Federal Revenue and of the Ministry of Social Security
Median salary of employees	Median salary of employees	Database of Ministry of Social Security

4.5. Estimation of the models

4.5.1. Logistic Regression

The logistic regression function was built in two stages. First, using the knowledge database, all selected variables were used and had their coefficients of discrimination (using the Maximum Likelihood Estimation method[5]) calculated to identify those with more discriminating power. Then, only the variables that were statistically significant (at a significance level of 5%) were used together in the construction of the function.

After having the best four coefficients (relative to the variables from the first round) recalculated, we had the enough parameters to build the classification function. This function was applied to each of 1000 records of the test database, given us the result according to the classification model.

To validate the results, as is usual in such work, we built a classification matrix that compares each prediction to the actual value of the dependent variable, thus achieving a hit rate (percentage).

The significance level of accuracy was tested according to the Press's Q test[6].

4.5.2. Decision Tree

Using the same variables and the same test database, the decision tree was built using the method of recursive partitioning, which is traditional for this technique.

The partitioning algorithm used was CART - Classification and Regression Trees[7], which divides the distribution of each variable in order to improve the separation of groups. This division is based on the Gini coefficient (Hoffman, 1980). Partitioning creates the conditions ("if, then") for each variable and distributes them in a tree hierarchically configured according to their discriminating power.

Once constructed the tree, the software itself classified the test database records.

Again, this classification was validated by a comparison matrix with the actual values, when we have obtained a new hit rate.

---

5 According to Portugal (1995), the Maximum Likelihood method is a procedure, as well as the Least Squares, which allows the estimation of statistical parameters and testing hypotheses.

6 Details on the use of Press 'Q test on Hair Jr. et al. (1998). The equation is described in Section 6.1.

7 The CART is a binary partitioning algorithm often used in the construction of Decision Trees. For more information, see Berry et al suggest. (1997).

>> 5. RESULTS

5.1. Logistic Regression

The first step of the logistic regression returned the following parameters for the variables:

Table 3 - Estimated parameters in the first stage of logistic regression.

Variable	β	Wald test	P-value
Days until 1st contract	0,000015	0,31086	0,577153
Campaign donations in 2010	0,000004	2,87676	0,089867
Occurrences in monitoring made by CGU	0,000110	1,08392	0,297822
No. of partners	0,018996	1,03650	0,308637
No. of employees	0,000056	0,35190	0,553037
Median age of partners	-0,038611	28,78896	0,000000
Median age of employees	0,017014	6,05829	0,013841
Median salary of partners	-0,000179	8,34594	0,003866
Median salary of employees	-0,000654	11,71519	0,000620

We highlight those variables that were significant according to the Wald test (Hair Jr. et al - 1998, states that there is significance for the variables with p-value less than 0.05).

The next step, in order to optimize the fit of the model, it was rebuilt using only the significant variables (the median of age and salary of members and employees). The second step resulted in new parameters for the selected variables, as described in Table 4:

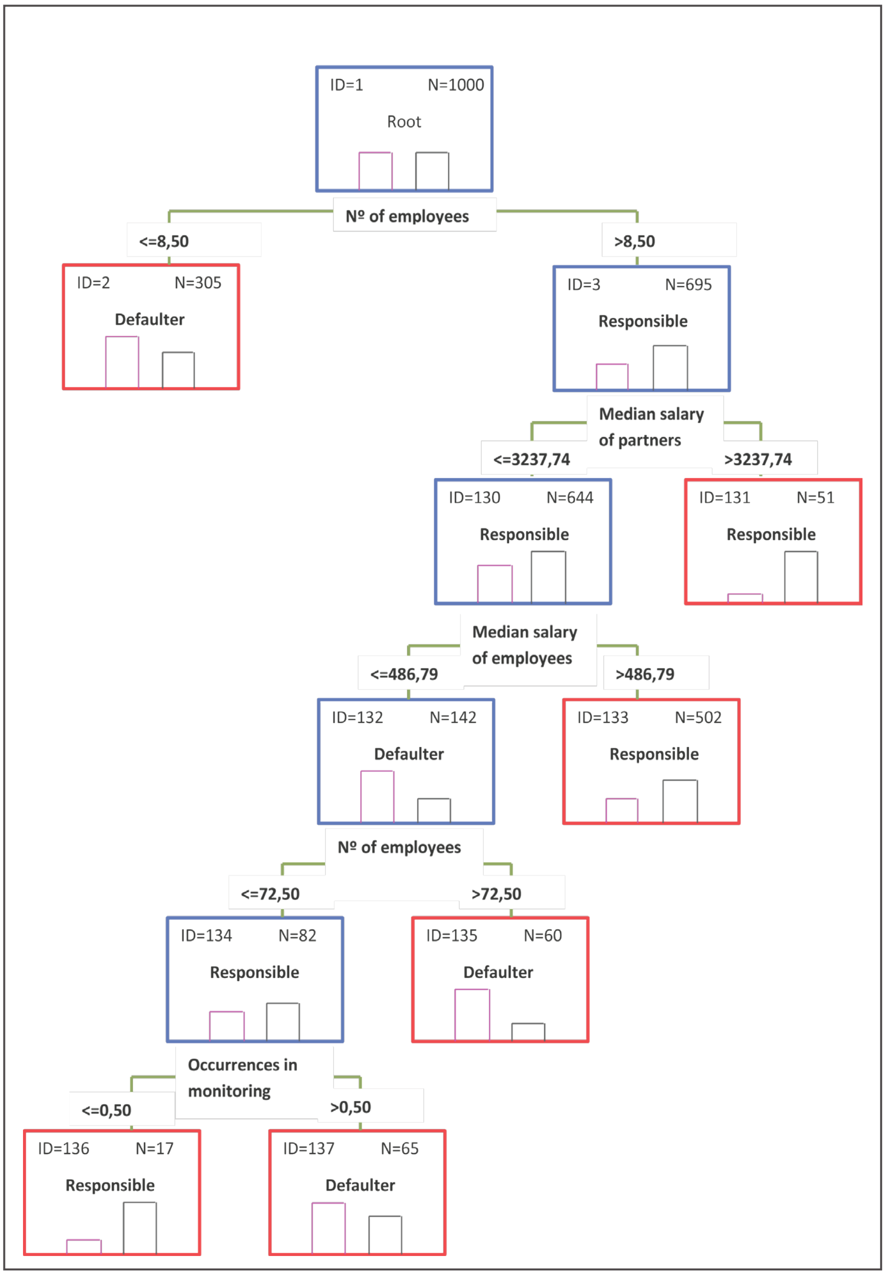
Table 4 - Parameters estimated in the second stage of logistic regression.

Variable	β	Wald test	P-value
Median age of partners	-0,037118	32,86038	0,000000
Median age of employees	-0,014772	4,71279	0,029939
Median salary of partners	-0,000174	7,64116	0,005705
Median salary of employees	-0,000654	11,76702	0,005705

Thus, the logit function was assembled using the coefficients of the four variables besides of the intercept coefficient (equal to 2.688939).

As already mentioned, the coefficients β determine the influence of the predictor variable on the dependent variable (group "Defaulter" or "Responsible"). The four main predictors were the age of partners, the age of employees, the salaries of partners and the salaries of employees. The negative sign of the coefficient in all of them indicates that the variable inversely contributes to the likelihood of the value of the Logit function being equal to 1 (presence of the value "Defaulter"). That is, the smaller the values of variables, the closer of 1 will be the result of a function. It can be argued, for example, that the

Figure 3 - Results - Decision Tree



lower the age of the partners the more likely the company is classified as a "Responsible", using the same reasoning to the other three predictors.

Besides the "direction" of influence, we must also analyze its "intensity", characterized by the modulus of the coefficient. The closer the first module is of 1, the greater the intensity of the influence of the predictor variable. In this particular case, it is clear the increased capacity of the characteristics of age (of partners and employees) in effect on classification.

5.2. Decision Tree

The tree generated from the learning base is presented in Figure 3.

As shown in the figure above, the decision tree represents a set of rules that determine the classification of each record. These rules will be applied to the base test to validate the model.

The first division of the tree is determined by the variable of number of employees. The algorithm determined as the cutoff the value of 8.5. Below this value the company is classified as "Responsible". Above it is a new division. Here, we find that the algorithm has separated 305 companies (understood as "Defaulters") and went in search of the other´s classification.

The second division involves the median salaries of partners, of which R$ 3,237.74 is the cutoff value. Above that, the company receives a rating of "Responsible" and new divisions below occur.

The tree goes down by splitting up until the last condition that divides the 82 remaining companies in "Responsibles" to a number of instances of monitoring less than 0.5, and "Defaulters" for occurrences above this value.

>> 6. VALIDATION OF MODELS

6.1.Logistic Regression

The scoring matrix of Table 5 shows the degree of adjustment of the Logit model:

Table 5 - Classification Matrix - Logistic Regression.

Accuracy Matrix	Classification of the model	
Real classification	Defaulter	Responsible
Defaulter	68%	32%
Responsible	43%	57%
Total accuracy	62,5%	

The Press' Q test can be used to measure the significance of the model. The formula is as follows (equation 2):

$$Q = \frac{[N-(nk)]^2}{N(k-1)}$$

In this equation, "N" is the number of records in the test database, "n" is the number of correct classifications and "k" is the number of groups.

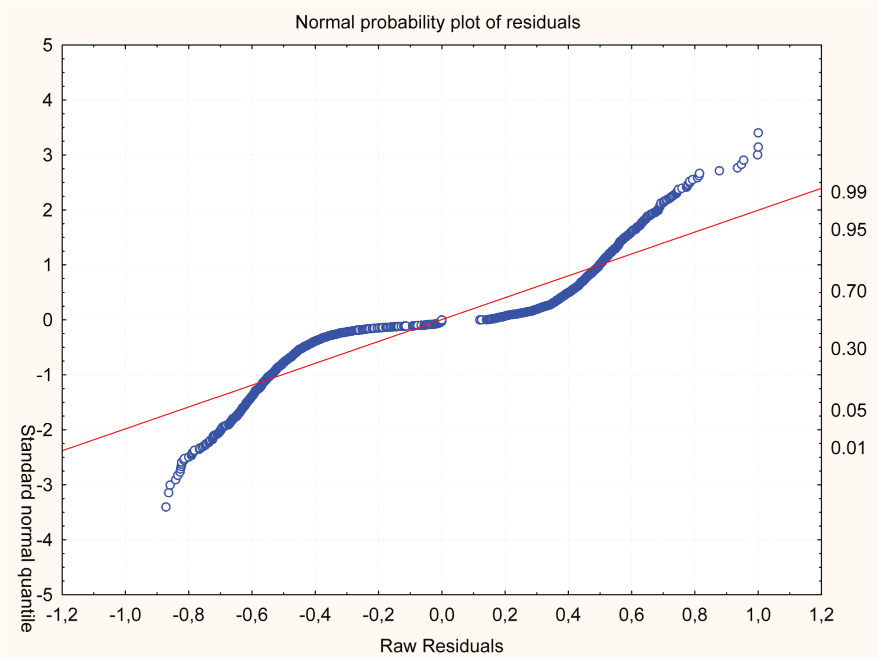
We obtained the value of 62.5 with the application of the test, which is a result considered by Hair Jr et al. (1998) as the minimum acceptable to differentiate the correctness of the model than would be obtained by the criterion of chance (randomness).

This result indicates that the model adds value to the classification, having acceptable predictive capabilities, but needs improvement. A positive point is that, as can be seen in the classification matrix, the performance to predict "Defaulter" companies was 68% better than that obtained in the classification of "Responsible", which is more interesting for the predictive control of procurements.

Another way to evaluate the consistency of a Logit model is the dispersion curve of the residuals, reproduced in Figure 4.

The shape of the curve indicates the approach of residual dispersion to the normal curve, indicating that the model is suitable for data.

Figure 4 - Logit function. Dispersion curve of Residuals



6.2. Decision Tree

Table 6 details the classification matrix obtained in the implementation of the tree on the basis of test:

Table 6 - Classification Matrix - Decision Tree.

Accuracy Matrix	Classification of the model	
Real classification	Defaulter	Responsible
Defaulter	56%	44%
Responsible	30%	70%
Total accuracy	63%	

The Press' Q test, applied again, showed an index of 67.6, slightly above those obtained with the logistic regression but still of medium efficacy.

>> 7. CONCLUSIONS

In order to achieve the overall goal of this work, to contribute to building a predictive model for risk management in government contracts with private companies, we tested the applicability of statistical techniques for selection of clients used by financial institutions.

The two models used for classification set out from an identical set of variables, selecting the most capable of separating two groups of government contractors: those most likely to fail to execute the contracts (called "Defaulters") and the least likely ("Responsibles"). The criterion of maximum likelihood of logistic regression prioritized the characteristics of salary and age of partners and employees. Whereas the algorithm CART of Decision Tree selected also the features of salary of partners and employees, but not included in the final model aspects connected to their age. Instead, the features of number of employees and occurrences in monitoring trails were added.

It is evident the prevalence of issues linked to the "size" of companies such as major discriminating of good and bad contractors in both models (quantity, age and salary of employees and partners). This indicates that the basis used for selecting the group "Defaulter," the Register of non-Reputable and Suspended Companies - *CEIS*, is more populated with companies of smaller "size", which may indicate that these companies have made commitments to the government without having enough structure and experience to meet them.

Besides these features, appears in the tree model the variable of occurrences recorded in audits made by CGU (monitoring trails), which adds a different look, more connected to the findings of audits.

Possible contributions of the study

Understanding the ability of variables to determine the potential default of suppliers is, in our view, the main and immediate contribution of this study to the context of public audit, as well as the need for managers of public contracts to foresee risky situations.

As a possibility in the medium term, our study demonstrated the feasibility of using multivariate statistical techniques, commonly used for evaluation of credit with financial institutions, for the prediction of government managers problems with their suppliers. Based on the use of two statistical techniques and only nine variables, we can achieve a performance more representative than mere randomness.

However, we understand the models constructed here as a starting point, not as a solution already finished. The prediction results were only slightly above the acceptable and indicate a great way to go in search of effective forecasts. We foresee that gains in accuracy can be obtained from the test of more techniques, and, especially, the increase of more variables.

Regarding the use of classification models, such as Credit Scoring, we must emphasize that these cannot be used by government in exactly the same way that they are used by financial institutions. This is because the public manager cannot be based on a predictive model to reject a vendor that wants to be hired by the government. The requirements for participation in bidding contests and employment are only those defined by law.

We understand, however, that information coming from such prediction is interesting for the government in two respects. First, for managers, it is interesting to know their contractors and the risks they offer to the performance of the services under their responsibility. This way, the management effort may be focused on those contracts that are being conducted by suppliers potentially problematic.

Second, regarding the audit activity, beyond the obvious usefulness for the selection of contracts to be audited, there is the capability of the models focused on suppliers being expanded to a larger process of risk management in the provision of government services.

References

ARAÚJO, E.A., CARMONA, C.U.M. 2007. Desenvolvimento de Modelos Credit Scoring com Abordagem de Regressão Logística para a Gestão da Inadimplência de uma Instituição de Microcrédito. Contabilidade Vista & Revista. Vol.18, n°. 3: 107-131.

BEASENS, B., VAN GESTEL, T., VIAENE, S., STEPANOVA, M., SUYKENS, J., VANTHIENEN, J. 2003. State-of-the-art classification Benchmarking algorithms for credit scoring. The Journal of the Operational Research Society. Vol. 54, n°. 6: 627 - 635.

BERRY, M. J. A.; LINOFF, G. 1997. Data Mining Techniques - For Marketing, Sales and Custumer Support. Wiley Computer Publishing, Indianapolis.

BRASIL. 1988. Constituição da República Federativa do Brasil de 1988. Available in: http://www.planalto.gov.br/ccivil_03/constituicao/constituiçao.htm. Accessed on: 10/20/2011.

BRASIL. 2006. Lei Complementar nº 123 de 14 de dezembro de 2006. Institui o Estatuto Nacional da Microempresa e da Empresa de Pequeno Porte. Available in: http://www.planalto.gov.br/ccivil_03/leis/LCP/Lcp123.htm. Accessed on: 10/20/2011.

BRASIL. 2001. Lei nº 10180 de 06 de fevereiro de 2001. Organiza e disciplina os Sistemas de Planejamento e de Orçamento Federal, de Administração Financeira Federal, de Contabilidade Federal e de Controle Interno do Poder Executivo Federal, e dá outras providências. Available in: http://www.planalto.gov.br/ccivil_03/leis/LEIS_2001/L10180.htm. Accessed on: 10/20/2011.

BRASIL. 2002. Lei nº 10520 de 17 de julho de 2002. Institui, no âmbito da União, Estados, Distrito Federal e Municípios, nos termos do art. 37, inciso XXI, da Constituição Federal, modalidade de licitação denominada pregão, para aquisição de bens e serviços comuns, e dá outras providências. Available in: http://www.planalto.gov.br/ccivil_03/leis/2002/L10520.htm. Accessed on: 10/20/2011.

BRASIL. 1993. Lei nº 8666 de 21 de junho de 1993. Referente à regulamentação do art. nº 37, inciso XXI, da Constituição Federal, instituindo normas para licitações e contratos da administração pública. Available in: http://www.planalto.gov.br/ccivil_03/leis/L8666cons.htm. Accessed on: 10/20/2011.

CAOUETTE, J., ALTMAN, E., NARAYANAM, P. 2009. Gestão do Risco de Crédito: O Grande Desafio dos Mercados Financeiros Globais. 2° edição. Qualitymark, Rio de Janeiro.

CÔRTES, S.C., PORCARO, R.M., LIFSHITZ, S. 2002. Mineração de Dados: Funcionalidades, Técnicas e Abordagens. PUC-RioInf, Rio de Janeiro.

GEVERT, V.G., SILVA, A.C.L., GEVERT, F., ALVES, V.T. 2007. Modelos de Regressão Logística, Redes Neurais e Support Vector Machine (SVM´s) na Análise de Crédito a Pessoas Jurídicas. Revista de Ciências Exatas e Naturais. Vol. 12, n°. 2.

GHODSELAHI, A. 2011. A Hybrid Support Vector Machine Ensemble Model for Credit Scoring. International Journal of Computer Applications. Vol. 17: 1 - 5.

HAIR, J. F., TATHAM, R. L., ANDERSON, R. E., BLACK, W. 1998. Multivariate Data Analysis. 5° edição. Prentice Hall, New Jersey.

HAND, D.J., HENLEY, W.E. 1997. Statistical Classification Methods in Consumer Credit Scoring: a review. Journal of the Royal Statistical Society. Vol. 160, part 3: 523-541, Series A.

HOFFMAN, R. 1980. Estatística para Economistas. Pioneira, São Paulo.

LEWIS, E. 1992. An Introduction to Credit scoring. Fair Isaac, San Rafael.

MARQUES, L. F. B., KLOECKNER, G. O. 2002. Gerenciamento do risco de crédito: um modelo de mercado vs. normas do Banco Central do Brasil. Master thesis. Universidade Federal do Rio Grande do Sul, Porto Alegre.

MEIRA, C.A.A., RODRIGUES, L.H.A., MORAES, S.A. 2009. Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente. Pesquisa Agropecuária Brasileira. Vol. 44, n°. 3: 233-242.

PORTUGAL, M.S. 1995. Notas Introdutórias Sobre o Princípio da Máxima Verossimilhança: Estimação e Testes de Hipóteses. DECON/UFRGS, Porto Alegre.

SAUNDERS, A. 2000. Medindo o Risco de Crédito – Novas Abordagens para o Value at Risk e Outros Paradigmas. Qualitymark, Rio de Janeiro.

SEMEDO, D.P.V. 2009. Credit Scoring: Aplicação da regressão logística vs redes neuronais artificiais na avaliação do risco de crédito no mercado Cabo-verdiano. Master thesis. Instituto Superior de Estatística e Gestão de Informação da Universidade de Lisboa, Lisboa.

SHIBBA, M.H., SANTOS, R.L., QUINTANILHA, J.A., KIM, H.Y. 2005. Classificação de imagens de sensoriamento remoto pela aprendizagem por árvore de decisão: uma avaliação de desempenho. pp. 4319-4326. In Anais do XII Simpósio Brasileiro de Sensoriamento Remoto. INPE, São José dos Campos.

THOMAS, L.C., EDELMAN, D.B., CROOK, J.N. 2002. CrediStcoring and Its Applications. Monographs on Mathematical Modelingand Computation. Society for Industrial and Applied Mathematics, Philadelphia.

VASCONCELLOS, M.S. 2002. Proposta de Método para Análise de Concessões de Crédito a Pessoas Físicas. Master thesis. Universidade de São Paulo, São Paulo.