DECISIONS ON INVESTMENT AND PROFITABILITY: AN EMPIRICAL STUDY USING GENERALIZED LINEAR MIXED MODELS IN NON-FINANCIAL BRAZILIAN COMPANIES

* *Aldy Fernandes da Silva* (aldy.fsilva@gmail.com)
** *Afrânio Márcio Corrêa Vieira* (afranio.unb@gmail.com)
* *Augusto Coppi Navarro* (augustocnavarro@uol.com.br)
* *Claudio Parisi* (claudio.parisi@uol.com.br)

* FECAP – Fundação Escola de Comércio Álvares Penteado – São Paulo – Brazil
** UnB – Universidade de Brasília – Distrito Federal – Brazil

ABSTRACT
In this study we used a wide class of statistical models (generalized linear mixed models – GLMM) to examine empirically the relation between past investment and profitability (measured by ROA and Tobin´s q), considering five scenarios with different periods of investment over time. These models enabled to consider a structure of correlation for the profitability observed over time and to use a distribution of different probability for ROA and Tobin´s q coefficient. Data from the financial statements of non-financial companies listed on Brazilian Stock Exchange were collected (2001-2011), resulting in the unbalanced sample of 1,484 company-year observations. Regarding ROA, the results showed a positive relation between contemporary investment and profitability, and a negative relation between past investment and profitability. The relation of past investment with the profitability (using Tobin´s q) was positive. These relations were slightly weakening, which suggests that investments have their profitability reduced over time.

KEY WORDS: Investment. Profitability. Generalized linear mixed models. ROA. Tobin´s q coefficient.

JEL Classification: C23, G11.

† This research has received financial support from FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP: 2013/02251-0).

INTRODUCTION

Since the mid 2000s, Brazil has been consolidating its economic position in the world scenario, showing results of series of governmental measurements implemented over the last 20 years. This scenario has contributed to the development and growth of Brazilian capital market and, consecutively, to the rise of company investments. According to Luporini and Alves (2010), companies took advantage of this period of economic growth to start implementing large-scale expansionary projects by investing in new plants or increasing their already-existing capacity. Thus new demands that were stimulated by economic growth led to new investment opportunities and provided feedback to the cycle of economic growth (Giambiagi 2008).

In a globalized world, companies are involved in a competitive market environment in which: competitors act against the company, suppliers' conditions change, consumers can then switch their preferences and new technologies arise. All of these change the circumstances of competition. Faced with this situation, the results of investment can turn out to be different from what was planned and this is beginning to be reflected in the economic and financial results of companies over a period of time. In this scenario and from the standpoint of the company, investment decisions are made with the aim of adding value by obtaining a profit and positive cash flows. From the standpoint of the shareholders, the profit and positive cash flows must be revealed in the stock prices (Damodaran 2010).

There are several studies in the literature that assess the effects of investments carried out by companies from different standpoints as: the effectiveness of investments (Biddle et al. 2009; Cutillas and Sánchez 2012); their relation with the expected stock returns (McConnell and Muscarella 1985; Titman et al. 2004; Fama and French 2006); their relation with profitability or firm value (Gordon and Iyengar 1996; Echevarria 1997; Kim 2001; Li 2004; Jiang et al. 2006; Hao et al. 2011). At a first analysis of the literature, it shows that all studies are cross-sectional. Furthermore, regarding the studies that evaluated the relationship between investment carried out and profitability (or company value), some have shown difficulties in establishing this relationship: being inconclusive (as Echevarria 1997, Kim, 2001 and Jiang et al. 2006) or presenting opposed evidences (as Gordon and Iyengar 1996, 2004 and Hao Li et al 2011).

This study is based on the fact that investment decisions are made, so it must target positive returns and add value to the company. Herein, the main objective of it is to study, by means of a longitudinal form, the relationship between investment carried out and profitability of the company. We stated two hypotheses to assess this relation over time: first, using ROA as an accounting indicator of performance, and second, using Tobin´s q coefficient, as a performance indicator that reflects the expectation from investors over the company. These hypotheses were tested using a wide class of statistical models, so called generalized linear mixed models (GLMM), assuming five scenarios designed to different periods of investment lagged in time with regard to profitability. We found a positive relation between contemporary investment and profitability, and a negative one between past investment and profitability when measured by ROA. We also found a positive relation between past investment and profitability, measured by Tobin´s q coefficient. Despite, our results revealed these relations were weakening (coefficient and significance) over time.

Thus, we justified this study by the importance of the theme and the fact that it seeks to make a contribution to the capital market. In this context, the information about company investment, when associated to future profitability, becomes important to investors or to the stockbrokers, when taken into account the expectations of increased profits in the future, which are made possible by the investment, the effects result on the future dividend and share values.

In addition, we perceived a theoretical contribution in accounting field. This is due to studies that consider the effects of investments on the indicators of company profitability (Gordon and Iyengar 1996; Echevarria 1997; Kim 2001; Li 2004; Fama and French 2006; Jiang et al. 2006). This study is distinctly different, so far as it broadens the new control variables and considers the size, year’s dummy, sector of economic activity, leverage and growth opportunities of a company and it seeks, in this way, to find a better way of explaining the achieved profitability in terms of the investment undertaken. Moreover, we adopt a longitudinal approach in which it was possible to assess the effects of past investment on profitability, using five investment scenarios lagged in time (from 1 to 5 years) regarding to the profitability period (measured by ROA and Tobin´s q coefficient).

If we read only the studies which evidenced the relationship between profitability and past investment, such as Gordon and Iyengar (1996), Li (2004) and Hao et al (2011), our results, which were taken from the analysis provided by longitudinal studies, would allow to confirm the findings of Li (2004), which found a significant negative relation between the profitability and past investment. Li’s study (2004) was cross-sectional and only considered the past profitability as control variable. In a sense, our results are different from Li’s (2004), since the longitudinal study allowed considering the correlations between profitability measurements over time. Moreover, the profitability variation over time was controlled by macroeconomics variable (year's dummy) and by inherent factors to companies (past profitability, contemporary investment, size, sector of economic activity, leverage and growth opportunities). So, our study highlighted the significant relation at short term (in which the investment was lagged up to 2 years), indicating that the contemporary investment associated to the past investment decreases the company's profitability at short term (using ROA as profitability measurement). When using Tobin's q coefficient, our study also added an analysis from the perspective of the market. In this case, the found positive relation points out that contemporary investment is related to past investment, giving investors better expectances over the company, consequently, having a positive effect in the profitability.

Finally, it is very important to highlight and enhance the methodological features used in the research. We expect that the study can bring great value and contribution to the methodological field, by employing the GLMM class, in which, through an examination of profitability and investment data over a period of time, it was possible to consider the temporal correlation of these measurements in the companies and include this structure in a model of random effects. Furthermore, these models enabled to analyze the profitability data in accordance to the probability distribution that was most suited to the nature of the data (Normal distribution for the ROA and Gamma distribution for Tobin´s q coefficient). Besides, allowing a more suitable treatment on the data of profitability and past investment analysis, a greater accuracy was achieved on statistical tests carried out for the models. This approach has not been adopted in empirical studies conducted in financial accounting research so far.

We structure this paper as follows: a review of previous studies and formulation of hypotheses (Section 2); a description of methodological procedures, including variables, models and the sample (Section 3); the main results with corresponding statistical tests and analyses (Section 4) and; the summary, conclusions and suggestions for future research (Section 5).

2. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

In this section, there is an exam of the results found in the literature from the assessment of the relation between investment and profitability. Some of these studies related investment carried out by companies to the return provided by the stock valuation in the capital markets. This kind of study began in the 1980s when McConnell and Muscarella (1985) drew attention to the fact that there is little evidence of the effect of investment decisions on the market value of companies.

Titman et al. (2004), found a negative relation between capital investments and future stock returns. The authors stated that "*firms that increase their level of capital investment the most tend to achieve lower stock returns for five subsequent years and suggests that the negative relation between abnormal investments and stock returns cannot be explained by either the risks or the characteristics of the firms and are independent of the previously documented long term return reversal and secondary equity issue anomalies*".

Fama and French (2006) showed that firms with higher book-to-market equity have higher than expected stock returns, when the expected profitability and investment are controlled. Moreover, given the rate of book-to-market and expected profitability, higher expected rates of investment are related to lower than expected returns. The authors said that their results are aligned with studies that concluded that firms that invest more, have lower average returns.

From a different perspective, other studies sought to relate the investment carried out with the profitability resulting from these assets within the company (with the aid of performance assessment indicators like ROA and Tobin´s q coefficient, ROE, ROI among others). Gordon and Iyegar (1996), analyzed a sample of industrial companies listed in the NYSE and AMEX, in the period 1989 to 1992, and found evidence of the positive relation between capital expenditures and ROI, and that this relation exists even where the interests of owners and managers are apparently in conflict.

In relating investment to future profitability, Echevarria (1997) selected industrial firms belonging to the Fortune group of the US, listed in the Compustat for the period 1971 to 1990 and took the period 1971 to 1980 as the base period and 1981 to 1990 as the period of profitability. The author recommends that future studies should take account of the size of companies, levels of investment, sectors of economic activity and other factors that might also be included in the assessment.

In relating investment to profitability as measured by the market value of shares, Kim (2001) conducted a study in which he selected industrial companies in the United States (listed in the Compustat Annual Tapes) for the period 1976 to 1994. After examining the data by means of a regression analysis, the results showed that, at first, there was no relation between investment and future profitability for the sample selected but after the companies had been divided between losers and winners, the investments was positively related to future profitability and negatively related to the losers.

Li (2004) analyzed the data financial statements of American companies for the period 1962 to 2002 in the cross-sectional study and classified them into the portfolios of investment groups. The author found that there is a relation between investment and future stock returns, concluding that in companies which have high free cash flow and low leverage, there is a negative relation between capital investment (long term asset accruals) and future profitability (using ROA). Moreover, the author stated that, "*when this association is analyzed in the long term, the negative coefficients of the investment get weaker. This suggests that the dampening effect from investment on future profitability becomes less severe in years in the distant past. However, the coefficients remain significantly negative 12 years ahead, then become insignificantly negative thereafter*".

In establishing a relationship between capital expenditure and profitability, measured by accounting information, Jiang et al. (2006) carried out a research which involved selecting industrial companies listed in the Taiwan Stock Exchange in the period 1992 to 2002. In the study, the first five years were used as the period of investment and the last six, as the period of profitability. They concluded that investments are positively related to the future profitability, after grouping of companies according to the level of investment.

Finally, Hao et al. (2011), relate investment growth to the value of the company, profitability (ROE) and the value of shareholder equity. The study makes it possible to conclude that past investments activities affect the value-accounting relation as a result of conservative accounting practices.

When we analyze the literature, the importance of the theme "past investment and profitability" becomes apparent not only for the academic environment but also capital market and to the company managers. However, what is being observed is still following an evolutionary pattern with different methodological proposals being carried out. Hence this study attempts to make a significant addition to this debate by exploring gaps that still exist in the literature, as well as suggesting the implementation of a more suitable statistical model for dealing with the longitudinal data of the profitability of companies.

On the basis of the analysis of previous studies in this section and, in particular, the results of research carried out by Gordon and Iyengar (1996), Echeverria (1997), Kim (2001), Li (2004), Fama and French (2006) and Jiang et al. (2006), the following hypotheses were proposed for the test:

H1: There is a statistically significant relation between profitability (measured by ROA) and past investment carried out by non-financial companies.

H2: There is a statistically significant relation between profitability (measured by Tobin´s q coefficient) and past investment carried out by non-financial companies.

The next section presents the models, variables, data collection and other methodological instruments that are employed to test the proposed hypotheses.

3. METHODOLOGY

In this section, we outline the methodological procedures employed in carrying out the research. Initially we provide an operational definition of the variables, and then, describe the statistical model used and, finally, discuss the selection of the sample and design of the research.

3.1. Operational definition of variables

We first define the operational variables in three groups: dependent variable (profitability), explanatory variables: interest (investment) and control as a means of forming our models and drawing up the sample selection criteria.

3.1.1. Proxy of profitability – dependent variable

Profitability is the net profit arising from business activities and decisions; it reflects the effectiveness of operations and shows the effects of liquidity on asset management and liabilities in the company results. Profitability can be calculated through performance measures as for example, sales margins and profit margins, return on assets, return on net worth, among others (Brigham and Houston, 2008). Indicators like ROA, ROE, ROI, and asset turnover have been used as proxy to the profitability of companies when related to levels of corporate governance, ownership concentration or even to make forecasts about future share prices, among other applications (Gordon and Iyengar 1996; Li 2004; Jiang et al. 2011).

The return on assets (ROA) is one of the most widely used profitability measurements; it is well known in the accounting literature and represents the operational return provided by all the assets of the company. As well as showing the return on investment for the whole company, it is also a key benchmark for making a comparison with third-party capital cost estimates (Weygandt et al. 2009).

Apart from the indicators for profitability calculated by accounting measures, there are indicators that use market values to measure the profitability of a company. Tobin´s q coefficient is recommended in the financial literature as a criterion that allows measure the performance of companies (Wenderfelt and Montgomery 1988; Bharadwaj et al. 1999). According to Wenderfelt and Montgomery (1988) accounting rates of return are distorted by a failure to consider differences in systematic risk, temporary disequilibrium effects, tax laws, and accounting conventions regarding R&D and advertising. The authors stated that Tobin´s q coefficient "*is a much more appealing measure than accounting returns once implicitly uses the correct risk-adjusted discount rate, imputes equilibrium returns, and minimizes distortions due to tax laws and accounting conventions*".

Two indicators were used, in this study, to measure the profitability of companies over a period of time: *i*) the ROA which shows the profitability provided by the total assets of the company (calculated annually and for each company by dividing operating results by average total assets); *ii*) the Tobin´s q coefficient which shows the performance obtained by the company´s shares in the stock market related to its total assets (calculated annually and for each company using as a basis the market share value on 31st December or the quotation immediately before, added to the short and long-term liabilities divided by the total amount of fixed assets in the balance sheet of each year, in accordance with Shin and Stulz, 2000). In this case, Tobin´s q coefficient shows a future perspective of profitability by relating the values of company´s assets with the market value of its shares and liabilities.

Finally, the profitability indicators were represented by ROA and $QTOBIN$ and were thus the two variables that this study depended on.

3.1.2 Explanatory variable of interest

Investment represents the value that the company has included in its fixed assets and which it hopes to use for future benefits. In this study, the investment index represents the value quoted as new long term investments and fixed assets and deducted the fixed asset sold or written off stated in the Sources and Uses of Funds or Cash Flow statement, divided by the average total assets. This calculation is carried out annually for each company and can be expressed as:

$$Investment\ index = \frac{Investment}{Average\ Total\ Assets} \tag{1}$$

These values considered as investments are the constant method of the sources of uses of funds or cash flow statements, registered as permanent investments, which are expected to yield returns at long term and are, thus, different from the current assets, which are for short term operations. The idea of the investment index is to associate the investment in long term with the profitability in long term, so that´s why we only consider long term investment and fixed asset. We didn´t consider changes in working capital as an investment, since it is a short term investment. Finally, we excluded companies that had negative investment for a long term. In this study, the natural logarithm of the investment index was employed and defined in (1), which will be represented by the $INVEST$ variable.

3.1.2 Control explanatory variables

The control variables sought to remove the effect of determined factors, such as the size of the company, the sector of activity where the company operates and leverage, which can influence the relationship between the investment carried out and the profitability of the company. Previous studies used past profitability as a control variable to explain the relationship between the investment carried out and future profitability (Li 2004; Jiang et al. 2006). In this way, a positive relation between past profitability and the contemporary period can be expected, since the capacity of the company to yield positive results is related to the positive results in the past. Hence, in this study, past profitability was employed as a control variable. In addition, investment in the contemporary period was used as a control variable because a positive relation is expected with contemporary profitability.

The size of the company can be assessed by value billing on an annual basis, the value of total assets and the total value of capital, since communicating magnitudes can be a way of making comparisons with other companies. It is expected that larger firms will have a greater capacity to invest and obtain financing than smaller firms and that they can also exert a greater influence on markets (Ehie and Olibe 2010; Stubben 2010). In this study the $SIZE$ variable was used too as a control.

The type of activity the company (sector) is involved in determines some of the features that the firms operate with, by fixing levels of profitability, investment or capacity for innovation and distribution (Li 2004; Burgstahler et al. 2006; Han et al. 2010). It is also expected that there will be a difference in the index levels for investment and profitability as a result of the sector of activity (represented by $SECTOR$). This sector can be identified by relying on constant information from the Economatica® database which identifies companies by sectors that were aggregated, since some of the original sectors obtained from Economatica® only show a few companies (for example just one or two).

Studies of earnings management have used a *Book-to-Market Equity* as a control variable to represent the growth opportunities of the company (Han et al. 2010; McNichols and Stubens 2008; Othman and Zeghal 2006). The growth opportunity variable will be represented by $GROWTH$ and was calculated annually for each company by dividing the value of the shareholders´ equity at the date of the balance sheet by the market share quotation for the same date (*Market-to-Book Equity*) or immediately before, constant of the Economatica® database. In cases where the market value quotation was not obtained, the net worth value was used. The $CRESC$ variable shows a view of the market with regard to the future results of the company and was drawn from the study carried out by Fama and French (2006).

We also included the natural logarithm of leverage (represented by LEV) as a control variable. This variable was used in the studies of Li (2004) and Ehie and Olibe (2010). In our study the leverage was calculated annually and for each company as the proportion between short and long term debts with regard to the total liabilities of the company, as a way of detecting the costly liabilities in the total liabilities of the company (Bodie and Merton 1999). Leverage, when understood as the relation between capital itself and third-party capital, shows the involvement of third party capital in financing assets. The decision regarding the involvement of third party capital is an integral part of financial decision-making and in Brazil, long-term resources are almost entirely offered by banks that are controlled by the State (BNDES, Caixa Econômica Federal – CAIXA, Banco do Brasil).

3.2. Statistical models – Generalized linear mixed models

This type of study is characterized by its longitudinal nature: performance measurements are observed or calculated for each year (semester, quarterly period or other period of time). It should be taken into account that each company has its own management practices, that the ROA or Tobin´s q coefficient indicators will reflect these practices and that these companies are relatively independent of each other. In view of these factors, it is reasonable to expect that the measures of these indicators for a company will be correlated with each other over a period of time (or autocorrelated). In case of this autocorrelation is present in the data and the usual models of linear regression are being used, the important assumption regarding “independence between the observations and the residues” will not be satisfied. If the autocorrelation is strong and depends on the direction of this correlation (whether positive or negative), it will lead to an inflation of Type I (or Type II) errors. These are involved in testing hypotheses that are related to the selection process of the statistical models and suggest factors such as erroneously significant (or ceasing to take note of important factors such as the case of the Type II error). Summing up, the use of a model with fixed and random effects (referred to as mixed effect models) improves the accuracy of the testing of hypotheses by suitably modeling correlations between the longitudinal measures of the companies.

The classic linear models that are employed in most areas of knowledge can be described (following Kitner et al. 2004) as:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{Y}$ is the vector with *n* observations of the character of interest, $\boldsymbol{X}$ is a matrix where the columns are formed by covariates exogenous (explanatory variables and/or dummy´s variables), $\boldsymbol{\beta}$ is a vector parameter with a fixed but unknown value and $\boldsymbol{\varepsilon}$ is a random error vector with elements that assume a normal distribution with an average of 0 and a variance σ^2, or rather, $e_{ij} \sim N(0, \sigma^2)$. As described in Fitzmaurice et al. (2009), the generalization of this linear model to include random effects, originated in the 1930s when statisticians realized that longitudinal measurements could be analyzed through the Split-plot ANOVA method which involves repeated measures in time like subplots. This model is described as:

$$Y_{ij} = \boldsymbol{X}_{ij}^T \boldsymbol{\beta} + b_i + e_{ij}, \quad i = 1, \cdots, N; \; j = 1, \cdots, n,$$

where Y_{ij} is the measurement of interest of the *i*-th subject (e.g. the company) in the *j*-th time, involving N subjects, each one measured n times. The covariates (exogenous variables) related to the fixed effects vector $\boldsymbol{\beta}$ can be found in the columns of the $\boldsymbol{X}_{ij}$ matrix; $b_i \sim N(0, \sigma_b^2)$ is a vector of random effects which aggregates all the non-measurable or unobserved factors related to each subject (for example, the company), which respond in a different way from one another. The $e_{ij} \sim N(0, \sigma_e^2)$ vector is a vector of errors which is common in linear models. As a result, this model induces a correlation between the observations by adopting a rigid structure for variance-covariance matrix, provided by $Var(Y_{ij}) = \sigma_b^2 + \sigma_e^2$, and $Cov(Y_{ij}, Y_{ik}) = \sigma_b^2$, at all $i \neq k$. This mixed model can be described in the most general way as in the class of Linear Mixed Models (LMM):

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Zb} + \boldsymbol{\varepsilon}, \qquad (2)$$

where, $\boldsymbol{Y}$ is the characteristic that is being modeled (ROA or Tobin´s q coefficient), $\boldsymbol{X}$ is a matrix the columns of which contain information about covariates (explanatory variables or dummy variables that represent qualitative factors); the $\boldsymbol{Z}$ matrix has factors related to their respective random effects contained in the $\boldsymbol{b}$ vector. The random effects, are assumed to be independent of the covariates of $\boldsymbol{X}$ matrix and to have a multivariate normal distribution $N(\mathbf{0}, \mathbf{D})$, with zero mean and variance-covariance $\mathbf{D}$, will adjust to the dependency structure between the observations. The $\boldsymbol{\varepsilon}$ vector is the vector of random errors with multivariate normal distribution $N(\mathbf{0}, \boldsymbol{\Sigma})$, and $\boldsymbol{\Sigma}$ is general variance-covariance matrix. Soon afterwards, this linear model shows two random effects that together with the versatility of the normal multivariate distribution (Verbeke and Molenberghs 2000), show great flexibility; as a result, various areas of knowledge make use of this class of models.

In another direction, the classic linear models were extended in a way that allows the variable answers to have other probabilistic distributions apart from the Normal distribution (e.g. Binomial, Poisson, Gamma among others) and that the relation between the linear part of the model and the average of the response variable, can be established through a nonlinear function (Nelder and Wedderburn 1972). This class is called Generalized Linear Models (GLM). It is assumed that, without loss of generality, the *i*-th observation of vector $\boldsymbol{Y}$ has a probability density function that belongs to the exponential family of distributions $f(y|\theta,\phi) = \exp\{\phi^{-1}[y\theta - \psi(\theta)] + c(y,\phi)\}$, where $\frac{\partial\psi}{\partial\theta} = \mu$, is a function of the mean. The GLM specifications are complete defining that

$$\mu_i = h(\eta_i) = h(\boldsymbol{x}_i^T\boldsymbol{\beta}),$$

being μ_i the mean of $\boldsymbol{Y}$ and $h^{-1}(\cdot)$ a *link function*, monotonic and differentiable, which allows to write $h^{-1}(\mu_i) = \eta_i$. Several statistical models are members of this class of model, such as the class linear models (which adopts a normal distribution and link function identity), the logistic or probit regression model (assuming binomial distribution and link functions logit and probit, respectively), gamma regression (assuming gamma distribution and link function log), among other models. These models are examined in detail in Lee et al. (2006).

With regard to problems similar to those we are studying in this work, the class of Generalized Linear Mixed Models – GLMM (Breslow and Clayton 1993 apud Molenberghs and Verbeke 2010) shows great flexibility in modeling this autocorrelation, and in this way, allows prognostic factors to be detected that are related to the ROA or Tobin´s q coefficient measurements. This class of models comprises other classes of models such as Panel Data Models (Wooldridge 2010) and Multilevel Models (Gelman and Hill 2006), with the added advantage of allowing the modeling of data with repeated measurements in time intervals that are not equidistant and even unbalanced, choosing different distributions for the response $\boldsymbol{Y}$.

In the situation, in which the Normal distribution is not appropriate, the GLMM is one of the alternative analyses. In the description of this model, random effects are included in the linear predictor η_i of the GLM:

$$h^{-1}\{\mu(\boldsymbol{Y}_i)\} = h^{-1}\{E(\boldsymbol{Y}_i|\boldsymbol{b}_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)\} = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \tag{3}$$

where, $h^{-1}(\cdot)$ is a link function which ensures continuity and suitable linearization of fixed effects $\boldsymbol{\beta}$ and random effects $\boldsymbol{b}$. The random effect $\boldsymbol{b}$ is assumed to follow normal multivariate distribution with vector mean $\mathbf{0}$, and variance-covariance matrix **D**. Wherefore, measurements that show asymmetric distribution like Tobin´s q coefficient can be modeled with distributions that adjust to their shape without the need of data transformations, like box-cox type (Kitner et al. 2004), simplifying interpretation and forecasting, as well as ensuring greater precision by adjusting the data in their original scale. In spite of general structure that the matrices $\mathbf{D}$ and $\boldsymbol{\Sigma}$ accommodate, in this study only a random effect will be employed and when adjusted to this random effect, the errors will be assumed to be conditionally independent. Thus, the vector $\boldsymbol{b}$ consists of $b_i \sim N(0, \sigma_b^2)$ and the matrix $\boldsymbol{\Sigma}$ consists of $\mathbf{I}\sigma_e^2$. Different structures for the variance-covariance matrices **D** and $\boldsymbol{\Sigma}$ can be found in Verbeke and Molenberghs (2000, 94-101) and SAS Institute Inc. (2011, 253-273).

In the special case of linear mixed models (LMM), where an identity link function has been adopted, the expression in (3) is simplified by $\mu(\boldsymbol{Y}_i) = E(\boldsymbol{Y}_i|\boldsymbol{b}_i, \boldsymbol{X}_i, \boldsymbol{Z}_i) = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i$, which have a interpretation in terms of the population means (Fitzmaurice et al. 2009, 19-20, 86-88).

The estimate of the vector $\boldsymbol{\beta}$ of GLMM, can be obtained using maximum or restricted likelihood method based on Laplace Approximation or Gaussian Adaptive Quadrature (SAS Institute Inc. 2011, 284-299). The estimate of the vector $\boldsymbol{b}$ of random effects an obtained through Empirical Bayes technique (Verbeke and Molenberghs 2000, 78-81). Finally, assumptions of symmetry, normality, homoscedasticity and other diagnostic influences can be checked about residuals for the LMM models (SAS Institute Inc. 2011, 347-352). These measures must be used with care for non-normal cases of the GLMM because few results are known for the diagnostic analyses.

3.3. Sample selection and design of research

The research hypotheses were evaluated by investigating a sample of non-financial companies listed in the Brazilian Stock Exchange – BM&FBovespa during the period 2001 to 2011 that initially consisted of 185 companies in the study period (11 years). Data for all companies in the sample were drawn from individual and consolidated financial statements found in the Economatica® database and the Securities and Exchange Commission of Brazil (CVM) site.

The choice of the period of analysis (2001-2011) was supported by other research work that uses two periods of study: the period of investment and the period of profitability (Kim 2001; Li 2004; Jiang et al. 2006; Arslan 2008). According to Jiang et al. (2006), before a company can finalize an investment project, it needs several years to reach a conclusion and reap the benefits of the investments and several years are also needed before the projects can be abandoned. The authors also state that the division into two periods – one for investment and one for profitability – may not exactly be in accordance with the reality of each company but makes it possible to measure the capital invested in a previous period and the profitability of a subsequent period.

The following companies were rejected from the original list: those where it was not possible to identify the investment values; those that showed a failure to invest in at least 50% of the period under study; and those that showed a continuous negative net worth. This led to a final sample comprising 137 companies that were used to construct the statistical models.

To mitigate the effects of outliers in the sample, we winsorized the variables: LEV, $CRESC$ and $INVEST$, using 0.5% and 99.5% percentiles. The final unbalanced sample thus consisted of 1,484 of company-year observations.

The design of the research adopted the following form to develop and construct the statistical models: *i*) five scenarios were initially set up with different periods of investment and profitability; *ii*) in each scenario the relation between contemporary profitability and past investment was analyzed, while also taking into account the control variables. In scenario 1, the relation between contemporary profitability (denoted by ROA_t or $QTOBIN_t$) with investment in the immediately preceding time period ($INVEST_{t-1}$). In scenario 2, the relation was with the lagged investment of 2 time units, or rather, $t-2$ years. In the same way, in scenarios 3, 4 e 5, there was a lagged investment in 3, 4 and 5 years; *iii*) two models for future profitability were constructed for all the scenarios – one for the profitability measured by the ROA_t, and the other for the profitability measured by $QTOBIN_t$.

As described above, the planning of the research enabled the relation between contemporary profitability and past investment to be established, by varying these measures throughout the study, with different time lags. Thus for example, in scenario 1, the profitability of the year 2011 was related to the investment for the year 2010 and that of 2010 was related to the investment of 2009 and so on, until the profitability of 2002 was related to the investment for 2001.

The operational definitions of the variables and investment are shown in Chart 1, in the same way as the control variables, with their respective index of time reference.

CHART 1 HERE

Finally, the linear predictor, as described in (3), was established on the basis of the definition of the sample and the planning of the research for the contemporary profitability measured by the ROA_t:

$$h^{-1}\{\mu(\boldsymbol{Y_t})\} = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 SIZE_t + \beta_2 LEV_t + \beta_3 GROWTH_t + \beta_4 ROA_{t-l} + \beta_5 INVEST_t + \beta_6 INVEST_{t-l} + b_i, \quad (4)$$

where, $j = 1, \dots, 9$ and $k = 1, \dots, 13$, are the coefficient indexes α and γ, and represent the number of categories of the dummy´s variables $YEAR_t$ and $SECTOR$, which vary in accordance with the model constructed for each scenario. The l index, which varies from 1 to 5, denotes the time-lag (in years) for the ROA_{t-l} and $INVEST_{t-l}$ variables, which are used in the model as past profitability and past investment, respectively. In the case where contemporary profitability $\boldsymbol{Y_t}$ is measured by Tobin´s q coefficient ($QTOBIN_t$), we use $QTOBIN_{t-l}$ as past profitability. Then, the linear predictor was established as:

$$h^{-1}\{\mu(\boldsymbol{Y_t})\} = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 SIZE_t + \beta_2 LEV_t + \beta_3 GROWTH_t + \beta_4 QTOBIN_{t-l} + \beta_5 INVEST_t + \beta_6 INVEST_{t-l} + b_i. \quad (5)$$

The models described in (4) and (5) were estimated using the regression methodology for longitudinal data (GLMM), according to Verbeke and Molenberghs (2000), Fitzmaurice et al. (2009) and SAS Institute Inc. (2011).

4. RESULTS

In this section we outline the results of the data analysis. Initially, we conducted an exploratory analysis to show the behaviour of the variables used in the models. Following this the results of the longitudinal regression models (unbalanced data) were used to evaluate the hypotheses of the researches (**H1** and **H2**).

4.1. Exploratory analysis

Table 1 provides a descriptive statistical account of the investments and profitability of the companies in the sample over a period of time (2001-2011). It can be seen how the investments carried out by the 137 firms in the sample have evolved in each year. Although the data are shown as historical values, the first observation that can be made is that there is a fall in investment for the period 2002-2003 (the initial period of the 1st Lula´s mandate) with regard to 2001. This same fall is also reflected in the investment indices after the year 2008 (as a result of the world economic crisis). As regards the profitability indicators (ROA and $QTOBIN$) more rapid changes can be seen in the values of Tobin´s q coefficient over a period of time. This suggests that an indicator that takes account of the market value of companies can reveal more rapidly the way the profitability of a company evolves than an indicator of profitability accounting.

TABLE 1 HERE

Evidence is required to support the conclusion that if the data analysis is conducted on the basis of the average of the indicators for the period of investment and profitability (as carried out by Kim 2001; Jiang et al. 2006; Hao et al. 2011), a linear regression model would not be able to incorporate the variations of the profitability indicators over a period of time, as observed in Table 1.

The behavior of the profitability indicators in the contemporary period of the study (ROA_t and $QTOBIN_t$) can be visualized in Figure 1. The return on assets distribution (ROA_t) has a form of a symmetrical distribution around an average equal to 0.1040 and standard deviation equal to 0.0913. While Tobin´s q coefficient ($QTOBIN_t$) shows a strong asymmetry distribution with average equal to 0.9909 and standard deviation equal to 0.8516.

FIGURE 1 HERE

Finally, the Table 2 shows a correlation matrix for the variables – which will be used in models – that are based on the Spearman correlation (a non-parametric measure of correlation does not imply the existence of a linear relationship between the variables under study). The correlations indicate that the returns on assets in the contemporary period (ROA_t) are positively correlated with $INVEST_t$, $SIZE_t$ and $GROWTH_t$, while ($QTOBIN_t$) Tobin´s q coefficient is positively correlated with $INVEST_t$, $SIZE_t$, LEV_t and $GROWTH_t$.

TABLE 2 HERE

4.2. Results of regression

Various GLMM models were evaluated to examine the existence of a relation between contemporary profitability (measured by the ROA_t and $QTOBIN_t$ indicators) and past investment, in accordance with the five scenarios described in Section 3.3 and the linear predictors provided in (4) and (5). The adjustment of the models was carried out by SAS® 9.3 Statistical Software using the maximum likelihood method based on Adaptive Quadrature. The GLIMMIX procedure (SAS Institute Inc. 2011, Chapter 3) enabled the adjustment of the GLMM to be carried out with a wide range of discrete and continuous statistical distributions. As well as the distributions belonging to the exponential family (Binomial, Poisson, Negative binomial, Normal, Gamma and Inverse gaussian among others), this procedure allows other distributions to be defined that are outside the exponential family to carry out the adjustment of the GLMM (Lognormal, Beta, Multinomial, and t distribution among others). Finally, in addition to specifying the suitable distribution probability for the nature of the data, it was possible to select a link function which relates the mean value of the dependent variable (profitability) to the linear predictor (covariates).

The Normal (Gaussian) distribution with an identity link function was used for the construction of regression models for profitability measured by the ROA_t. The choice of this distribution was determined by the analysis in Figure 1 where the return on assets (ROA_t) shows a symmetrical behavior around the mean. The models developed for the five scenarios set out in Section 3.3 are reported in Table 3 together with the adjusted measurements of the models (AIC and BIC information criteria) and the variance estimations: residual (σ_e^2) and the random effect due to the companies (σ_b^2).

Before describing the interpretation of the results of the models reported in Table 3, it should be stressed that in all the scenarios, the $SIZE_t$ variable did not have a statistical significance and was thus excluded from all the models that were evaluated. The same occurred to the models of profitability measured by Tobin´s q coefficient ($QTOBIN_t$) reported in Table 4.

We start the interpretation of the results of Table 3 with “Model Fit Statistics”. In each of the five scenarios shown, information is given about the size of the sample and the AIC and BIC information criteria. Clearly, the size of the sample declined in the time from Scenario 1 to Scenario 5. With regard to AIC and BIC information criteria (Litell et al. 2006; SAS Institute Inc. 2011, 167), these are measurements used to estimate the adjustment of the regression models, the parameters of which are calculated by some kind of likelihood function. The AIC criterion is called Akaike´s Information Criteria and the BIC, the Schwarz´ Bayesian Criterion. Both types of information criteria are used for the selection of the models. The less the value of the criteria, the better is the adjustment of the model selected. Hence, within each scenario, it is presented the model with the best adjusted based on the set of explanatory variables proposed in the linear predictor described in (4). Clearly in Scenario 1, the AIC and BIC values would be less with the exclusion of the insignificant $SECTOR$ variable. The same would occur in Scenarios 2 and 5 with the exclusion of insignificant variables. Again, the AIC and BIC values are reduced from Scenario 1 to the other scenario, owing to a loss of information in the models resulting from the reduction in the size of the sample. The great importance of information criteria are that they can allow a comparison to be made between two or more statistical models, which is not the case in the scenarios shown in Table 3, since there is a concern with assessing the relation between the control variables and investment in profitability.

We have to analyze the information contained in the "Covariance Parameter Estimates" to assess whether the generalized linear mixed model (with random effects) is really more appropriate to the adjustment of the data. This information shows the residual variance estimates (σ_e^2) and the random effect (σ_b^2). Linked to each variance estimate is shown the standard error (s.e.), which allows the construction of confidence intervals. When the variance measurement that is linked to the companies (σ_b^2) and used for the Scenario 1 model, and its standard error are examined, it can be noted that this variance is relatively small. This suggests that for this scenario, the intra-class correlation (between the measures of profitability within the companies over a period of time) is negligible and thus a classic linear model could be used that treats each repeated measure of the profitability as being independent. However, the same thing is not determined for the other scenarios and the variances related to their effect on the companies, show that there is a correlation between the repeated measures and that these affect the accuracy of the estimates and linked testing of the hypotheses. This result points to the importance of the use of the GLMM class of models in the analysis of these data.

TABLE 3 HERE

The Gamma distribution with a logarithmic link function was selected for the construction of the regression models used for the profitability measured by Tobin´s q coefficient ($QTOBIN_t$). The choice of this distribution was made by means of the analysis in Figure 1 where the distribution of profitability ($QTOBIN_t$) shows a behavior with right asymmetry. In this case, the logarithmic link function enables the profitability $QTOBIN_t$ to be related to the linear predictors by means of the expression $\boldsymbol{\mu}(QTOBIN_t) = \exp(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i)$. This relation has the advantage of assuring the positive nature of the profitability measurements calculated by Tobin´s q coefficient.

The models developed for the five scenarios outlined in Section 3.3, are shown in Table 4, together with the adjustment measures of the models (AIC and BIC criteria) and the estimates of residual variance and the random effects of the models.

The same interpretation given to the results in Table 3 for the importance of the use of the GLMM class of models in the data analysis of profitability can be extended to the results shown in Table 4. However, it should be noted now that in all the five scenarios prepared for profitability measured by $QTOBIN_t$, the use of models with random effects of the GLMM class is of great importance (variance of random effects related to companies are significant). This is mainly owing to the fact that this enables a probability distribution to be used that is more appropriate to the nature of the $QTOBIN_t$ data, in this case, a Gamma distribution.

Finally, we can interpret the results of the relations between the explanatory variables and profitability of the companies by examining the coefficients and the p-value of the fixed effects models carried out in five scenarios, shown in Table 3 and 4. The first observation that should be made is about the linear predictor $h^{-1}\{\mu(\boldsymbol{Y_t})\}$ of the profitability models measured by ROA_t e $QTOBIN_t$. The linear predictor for the models of profitability measured by ROA_t is composed of the following variables: $YEAR_t$, $SECTOR$, LEV_t, $GROWTH_t$, ROA_{t-l}, $INVEST_t$ and $INVEST_{t-l}$. The $YEAR_t$ variable was used to control the macroeconomic effects in the profitability data (ROA_t and $QTOBIN_t$). The $SECTOR$ variable was significant in scenarios two to five, whereas the only case when the past profitability ROA_{t-l}, did not show significance was in the scenario when $l = 5$.

In the models that use $QTOBIN_t$ like profitability, a quadratic relationship was identified with the LEV_t and $GROWTH_t$ variables (when carrying out the exploratory data analysis). Then the linear predictor for the models was composed of the following variables: $YEAR_t$, $SECTOR$, LEV_t, LEV_t^2, $GROWTH_t$, $GROWTH_t^2$ $QTOBIN_{t-l}$, $INVEST_t$ and $INVEST_{t-l}$. The $SECTOR$ variable was not significant in the Five scenarios employed, whereas LEV_t and LEV_t^2 only had statistical significance in the scenario when $l = 5$. The past profitability measured by $QTOBIN_{t-l}$, did not show significance in the scenario when $l = 3$ to 5.

When the coefficients and variables are analyzed, it can be seen that LEV_t shows a negative sign for both the profitability modes, which suggests that the greater the leverage, the lower is the profitability. With regard to the variable for growth opportunities ($GROWTH_t$), defined as market-to-book equity, this provides a view of the market with regard to the future results of the company. In this way, it is anticipated that the greater the expectation of growth, the greater the profitability.

As expected from previous studies (Li 2004; Jiang et al. 2006), past profitability measured by ROA_{t-l} showed a positive relation with contemporary profitability ROA_t. However, when the past profitability was measured by $QTOBIN_{t-l}$, this positive relationship was observed until the $l = 2$ time-lag and became insignificant from the period of the $l = 3$ time-lag. This suggests that profitability in the distant past loses significance when measured by a market indicator. The contemporary investment $INVEST_t$.which is also used as a control variable, showed a significant positive relation with contemporary profitability in all scenarios of the study ($l = 1$ to 5).

Regarding the past investment $INVEST_{t-l}$ controlled by $YEAR_t$, $SECTOR$, LEV_t, $GROWTH_t$, ROA_{t-l}, and $INVEST_t$, the results show a significant negative relation with ROA_t until the time-lag $l = 2$ years, becoming insignificant from the period of the time-lag $l = 3$. The same occurs with $QTOBIN_t$, although there is a significant positive relation until the time-lag $l = 2$ years.

TABLE 4 HERE

SUMMARY AND CONCLUSIONS

Considering the importance of investment carried out by companies for economic development and the capacity of investment to bring about wealth for the company and shareholders (by creating new jobs, increasing the volume of consumption, and creating new investment opportunities), the purpose of this study was to examine the relation between the investment carried out and the profitability of companies. In view of this, we associate the profitability with the investment lagged in time considering five different scenarios. The study also employed two profitability measurements, ROA and Tobin´s q coefficient.

On the basis of the results of this study, the first observation that can be made is the robust result shown by the regression models, where two statistical models (using ROA and Tobin´s q coefficient) showed that when carrying out studies about profitability and investment, a wide range of control variables should be taken into account, such as the year, sector, leverage and growth opportunities. This was made possible by implementing the proposal to conduct the data analysis in a longitudinal form, using a new class of statistical models. For this reason, it is very important to highlight and enhance the methodological features used in the research.

When we interpret the results of this research, the model based on ROA was an empirical demonstration that different sectors have different requirements for investment and profitability and in the case of the model based on Tobin´s q coefficient, the sector of economic activity did not exert an influence in the explanation of the profitability of companies. We observe a negative association between leverage and ROA, as well the same relation to leverage and Tobin´s q coefficient. These finds show a typical situation of Brazilian companies, once it is cultural that firms have a higher dependence on the banks capital and consequently this kind of resource creates financial expenses, which effects are the decrease in profits. Even, with regard to the control variables, it could be shown that both contemporary investment and past profitability are important factors in determining the contemporary profitability of the company, which suggests that current investment and corporate earnings reflect competitive advantages for companies.

It is important to discuss the relation between past investment and the profitability that was identified in a negative way in the model, in which profitability was measured by ROA. Beyond that, our results revealed that this relation was weakening (coefficient and significance) over time. Our understanding is that in short term (time-lag from 1 and 2 years), the contemporary investment associated with past investment reduces the profitability of company, but in long term, the past investment is not important to explain the contemporary profitability. Then, this is an indication that the company must make new investments to maintain its rate of profitability.

The results for Tobin´s q coefficient are very similar to those obtained by ROA indicator of profitability. However, in this case, we found a positive relation between past investment and contemporary profitability (in the first two time-lags of the year). Our understanding is that when the contemporary investment is associated to past investment, investors have better expectations over the company, bringing a positive effect in profitability.

In this way, since investment decisions are important for the company, it is hoped that this study can make a positive contribution to decision-making about investment. It should also be of value to the stock market and its investors by providing confirmation through market measurements (Tobin´s q coefficient) that profitability is positively related to past and present investment, past profitability and growth opportunities.

Future researches can use our results and approaches to replicate these statistical models in empirical studies with observed data over time.

REFERENCES

Arslan, O. 2008. Overinvestment and investment inefficiency: a study in UK firms. *International Journal of Economic Perspectives* 2: 134-149.

Bharadwaj, A. S., S. G. Bharadwaj, and B. R. Konsynski. 1999. Information technology effects on firm performance as measured by Tobin's q. *Management Science* 45: 1008-1024.

Biddle, G. C., G. Hilary, and R. S. Verdi. 2009. How does financial reporting quality relate to investment efficiency? Available at SSRN: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1146536>.

Bodie, Z., and R. C. Merton. 1999. *Finance*. Prentice Hall.

Breslow, N. E., and D. G. Clayton. 1993. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88: 9-25.

Brigham, E. F., and J. F. Houston. 2008. *Fundamentals of financial management*, 6th ed. South-Western College Pub.

Burgstahler, D., L. Hail, and C. Leuz. 2006. The importance of reporting incentives: earnings management in European private and public firms. *The Accounting Review* 81: 983-1016.

Cutillas, M. F., and J. P. Sánchez. 2012. Financial reporting quality, debt maturity and investment efficiency. Working Paper WP-EC 2012-07. Instituto Valenciano de Investigaciones Económicas – Ivie. Available at http://www.ivie.es/downloads/docs/wpasec/wpasec-2012-07.pdf

Damodaran, A. 2010. *Applied corporate finance*, 3rd ed. John Wiley & Sons.

Dechow, P. M., and D. J. Skinner. 2000. Earnings management: reconciling the views of Accounting academics, practitioners, and regulators. *Accounting Horizons* 14: 235-250.

Echevarria, D. P. 1997. Capital investment and the profitability of Fortune 500 Industrials: 1971-1990. *Studies in Economics and Finance* 18: 3-35.

Ehie, I. C., and K. Olibe. 2010. The effect of R&D investment on firm value: an examination of US manufacturing and service industries. *International Journal Productions Economics* 128: 127-135.

Fama, E. F., and K. R. French. 2006. Profitability, investment and average returns. *Journal of Financial Economics* 82: 491-518.

Fitzmaurice, G., M. Davidian, G. Verbeke, and G. Molenberghs. 2009. *Longitudinal Data Analysis*. Boca Raton: Chapman & Hall / CRC.

Gelman, A., and M. Hill. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University.

Giambiagi, F. 2008. 18 anos de política fiscal no Brasil: 1991/2008. *Revista de Economia Aplicada – USP (RP)* 12: 535-580.

Gordon, L. A., and R. J. Iyengar. 1996. Return on investment and corporate capital expenditures: empirical evidence. *Journal of Accounting and Public Policy* 15: 305-325.

Han, S., T. Kang, S. Salter, and Y. K. Yoo. 2010. A cross-country study on the effects of national culture on earnings management. *Journal of International Business Studies* 41: 123-141.

Hao, S., Q. Jin, and G. Zhang. 2011. Investment growth and relation between equity value, earnings, and equity book value. *The Accounting Review* 86: 605-635.

Jensen, M. C. 1986. Agency cost of free cash flow, corporate finance and takeovers. *American Economic Review* 76: 323-329.

Jiang, C., H. Chen, and Y. Huang. 2006. Capital expenditures and corporate earnings evidence from the Taiwan Stock Exchange. *Managerial Finance* 32: 853-861.

Kim, S. 2001. The near-term financial performance of capital expenditures: a managerial perspective. *Managerial Finance* 27: 48-62.

Kitner, M. H., M. H. Nachtsheim, and J. Neter. 2004. *Applied linear regression models*, 4. ed. New York: McGraw-Hill.

Kothari, S. P., A. J. Leone, and C. E. Wasley 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39: 163-197.

Lee, Y., J. A. Nelder, and Y. Pawitan. 2006. *Generalized linear models with random effects: unified analysis via H-likelihood*. Chapman and Hall/CRC.

Li, D. 2004. The implications of capital investment for future profitability and stock returns: an over-investment perspective. Thesis. Berkeley: *University of California*. Available at http://www2.accountancy.smu.edu.sg/research/seminar/pdf/Donglin.pdf.

Luporini, V., and J. Alves. 2010. Investimento privado: uma análise empírica para o Brasil. *Revista de Economia e Sociedade – UNICAMP* 19: 449-475.

McConnell, J. J., and C. Muscarella. 1985. Corporate capital expenditure decisions and the market value of the firm. *Journal of Financial Economics* 14: 399-422.

McNichols, M. F., and S. R. Stubben. 2008. Does earnings management affect firms' investment decisions? *The Accounting Review* 83: 1571-1603.

Molenberghs, G., and G. Verbeke. 2010. *Models for discrete longitudinal data*. New York: Springer Verlag.

Nelder, J. A., and R. W. M. Wedderburn. 1972. Generalized linear models. *Journal of the Royal Statistical Society, Serie A* 135: 370-384.

Othman, H. B., and D. Zeghal. 2006. A Study of earnings-management motives in the Anglo-American and Euro-Continental accounting models: the Canadian and French cases. *The International Journal of Accounting* 41: 406-435.

Littell, R. C. et al. 2006. *SAS for mixed models*, 2nd ed. NC: SAS Institute Inc.

SAS Institute Inc. 2011. *SAS/STAT 9.3 User´s guide: the mixed modeling*. NC: SAS Institute Inc.

Shin, H., and R. M. Stulz. 2000. *Firm value, risk and growth opportunities*. Cambridge: National Bureau of Economic Research, July 2000, NBER Technical Working Paper 7808.

Stubben, S. 2010. Discretionary revenues as a measure of earnings management. *The Accounting Review* 85: 695-717.

Titman, S., K. C. John, and F. X. Wei. 2004. Capital investments and stock returns. *The Journal of Financial and Quantitative Analysis* 39: 677-700.

Verbeke, G., and G. Molenberghs. 2000. *Linear mixed models for longitudinal data*. New York: Springer Verlag.

WFE – World Federation of exchanges. 2011. *WFE database*. Available at: http:www.wordexchanges.org/statistics.

Wenderfelt, B., C. A. Montgomery. 1988. Tobin's q and the importance of focus in firm performance. *The American Economic Review* 78: 246-250.

Weygandt, J. J., P. D. Kimmel, and D. E. Kieso. 2009. *Financial Accounting*, 7th ed. John Wiley & Sons.

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*, 2nd ed. The MIT Press.

Chart 1: Summary table with an operational definition of the variables under study.

Variable	**Dependent variables**
ROA_t	Represents the return on assets in the period of contemporary profitability t (the future). It is the first proxy for profitability.
$QTOBIN_t$	Represents Tobin´s q coefficient in the period of contemporary profitability t (the future). It is the second proxy for profitability.
Variable	**Control explanatory variables**
$INVEST_t$	Represents the investment carried out by the company in the contemporary period t in a logarithmic scale.
$YEAR_t$	Represents the dummy variables of the year in contemporary periods which will be used to control macroeconomic fluctuations over a period of time.
$SECTOR$	Represents the dummy variables of the sector (14 sectors: Food and drinks; Commerce; Electro electronics; Electric power and sanitation; Plastic products industry; Industrial machinery and equipment; Paper and cellulose; Oil, gas and mining; Chemicals; Steel and metal works industry; Telecommunications; Textile industry; Transport and services; Vehicles and spare parts).
$SIZE_t$	Represents the size of the company in the contemporary period t.
LEV_t	Represents the leverage of the company in the contemporary period t.
$GROWTH_t$	Represents the opportunities for the growth of the company in the contemporary period t.
ROA_{t-l}	Represents the return on assets with a time-lag of *l* years $(t-l)$, where the l index varies from 1 to 5.
$QTOBIN_{t-l}$	Represents Tobin´s q coefficient with a time-lag of *l* years $(t-l)$, where the l index varies from 1 to 5.
Variable	**Explanatory variable of interest**
$INVEST_{t-l}$	Represents the investment carried out (in a logarithmic scale) by the company with a time-lag of *l* years $(t-l)$, where the l index varies from 1 to 5.

Table 1: Investment and profitability of the sample companies.

Period (Years)	Average Total Asset (in R$ billion)	Investment Index (average)	ROA (average)	$QTOBIN$ (average)
2001	60.4832	0.0853	0.0962	0.6138
2002	71.6889	0.0656	0.0981	0.6142
2003	78.8657	0.0658	0.1094	0.7372
2004	85.9208	0.0699	0.1349	0.9984
2005	94.4575	0.0791	0.1101	1.0037
2006	110.9516	0.0878	0.0965	1.1996
2007	121.8915	0.0885	0.1032	1.4077
2008	149.4785	0.1022	0.1187	0.9203
2009	158.8679	0.0686	0.0915	1.1766
2010	197.0447	0.0685	0.0971	1.1318
2011	219.0301	0.0674	0.0868	1.0585

Source: the author.

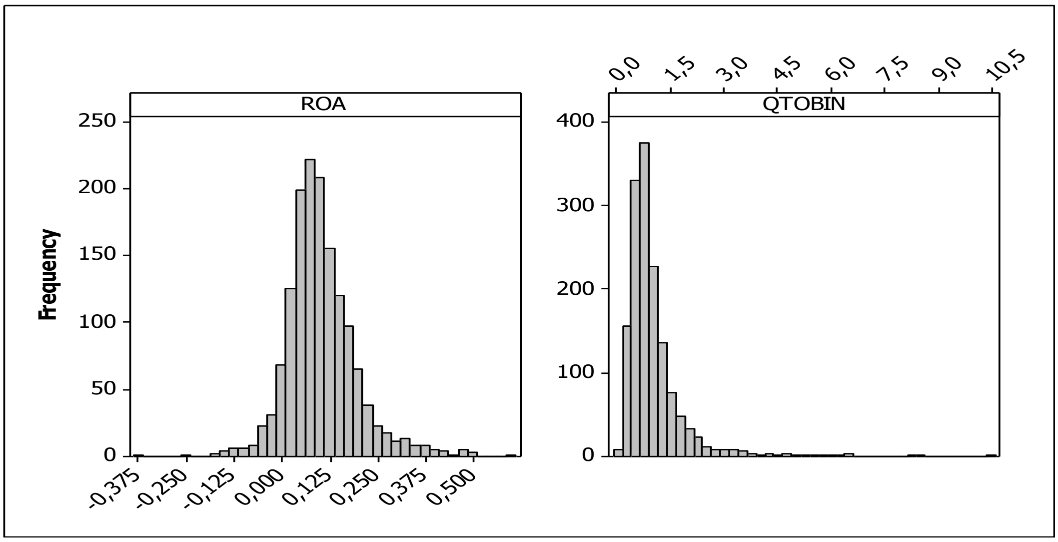

Figure 1: Histogram showing indicators of profitability (ROA and Tobin´s q coefficient).

Table 2: Spearman´s correlation matrices between the variables in the models.

Variables	ROA_t	$QTOBIN_t$	$INVEST_t$	$SIZE_t$	LEV_t
$QTOBIN_t$	0.392 †0.000	-			
$INVEST_t$	0.170 †0.000	0.203 †0.000	-		
$SIZE_t$	0.065 *0.013	0.226 †0,000	0.166 †0.000	-	
LEV_t	-0.005 0.839	0.270 †0.000	0.164 †0.002	0.281 †0.000	-
$GROWTH_t$	0.350 †0.000	0.876 †0.000	0.168 †0.000	0.210 †0.005	0.200 †0.000

Levels of significance: ' * ' 5% ' † ' 1%.

Table 3: Generalized linear mixed models for ROA_t using a Gaussian distribution and identity link function.

Fixed Effects	Scenario 1 (l = 1) Coefficient	p-value	Scenario 2 (l = 2) Coefficient	p-value	Scenario 3 (l = 3) Coefficient	p-value	Scenario 4 (l = 4) Coefficient	p-value	Scenario 5 (l = 5) Coefficient	p-value
$Intercept$	0.08043	<0.0001†	0.1134	<0.0001†	0.1816	<0.0001†	0.1642	<0.0001†	0.1553	<0.0001†
$YEAR_t$	-	<0.0001†	-	<0.0001†	-	<0.0001†	-	0.0002†	-	<0.0001†
$SECTOR$	-	0.2304	-	0.0873+	-	0.0964+	-	0.0887+	-	0.0788+
LEV_t	-0.0616	<0.0001†	-0.1121	<0.0001†	-0.1355	<0.0001†	-0.1360	<0.0001†	-0.1483	<0.0001†
$GROWTH_t$	0.0189	<0.0001†	0.0289	<0.0001†	0.0323	<0.0001†	0.0296	<0.0001†	0.0334	<0.0001†
ROA_{t-l}	0.5828	<0.0001†	0.2970	<0.0001†	0.1245	0.0004†	0.0944	0.0129+	0.0342	0.4119
$INVEST_t$	0.0082	0.0027†	0.0095	0.0004†	0.0101	0.0005†	0.0096	0.0019†	0.0091	0.0067†
$INVEST_{t-l}$	-0.0078	0.0044†	-0.0095	0.0007†	-0.0039	0.2024	0.0000	0.9968	-0.0016	0.6775
Model Fit Statistics										
	$N =$	1,217	$N =$	1,094	$N =$	966	$N =$	839	$N =$	714
	$AIC =$	-3,255.29	$AIC =$	-2,830.52	$AIC =$	-2,465.64	$AIC =$	-2,096.69	$AIC =$	-1,758.44
	$BIC =$	-3,167.69	$BIC =$	-2,745.84	$BIC =$	-2,383.88	$BIC =$	-2,017.85	$BIC =$	-1,682.52
Covariance Parameter Estimates										
$Cov\ Parm$	**Estimate**	**s. e.**	**Estimate**	**s. e.**	**Estimate**	**s. e.**	**Estimate**	**s. e.**	**Estimate**	**s. e.**
$Companies\ (\sigma_b^2)$	0.000132	0.000123	0.001218	0.000253	0.002229	0.000384	0.002568	0.000439	0.002884	0.000481
$Residual\ (\sigma_e^2)$	0.003724	0.000179	0.003543	0.000168	0.003373	0.000170	0.003413	0.000188	0.003358	0.000203

Levels of significance: '+' 10% '*' 5% '†' 1%.

Generalized linear mixed models to:

Scenario 1: $\mu(ROA_t) = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 GROWTH_t + \beta_3 ROA_{t-1} + \beta_4 INVEST_t + \beta_5 INVEST_{t-1} + b_i$.

Scenario 2: $\mu(ROA_t) = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 GROWTH_t + \beta_3 ROA_{t-2} + \beta_4 INVEST_t + \beta_5 INVEST_{t-2} + b_i$.

Scenario 3: $\mu(ROA_t) = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 GROWTH_t + \beta_3 ROA_{t-3} + \beta_4 INVEST_t + \beta_5 INVEST_{t-3} + b_i$.

Scenario 4: $\mu(ROA_t) = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 GROWTH_t + \beta_3 ROA_{t-4} + \beta_4 INVEST_t + \beta_5 INVEST_{t-4} + b_i$.

Scenario 5: $\mu(ROA_t) = intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 GROWTH_t + \beta_3 ROA_{t-5} + \beta_4 INVEST_t + \beta_5 INVEST_{t-5} + b_i$.

The assumptions of normality and homoscedasticity about residuals were satisfied (according diagnostics in SAS Institute Inc., 2011, p. 347-352).

Table 4: Generalized linear mixed models for Tobin´s q coefficient ($QTOBIN_t$) using a Gamma distribution and logarithmic link function.

	Scenario 1 (*l* = 1)		**Scenario 2 (*l* = 2)**		**Scenario 3 (*l* = 3)**		**Scenario 4 (*l* = 4)**		**Scenario 5 (*l* = 5)**	
Fixed Effects	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
$Intercept$	-0.5079	<0.0001†	-0.4337	<0.0001†	-0.3742	0.0003†	-0.3960	0.0002†	-0.2753	0.0058†
$YEAR_t$	-	<0.0001†	-	<0.0001†	-	<0.0001†	-	<0.0001†	-	<0.0001†
$SECTOR$	-	0.9480	-	0.8824	-	0.8853	-	0.9187	-	0.8548
LEV_t	0.7600	<0.0001†	0.8466	<0.0001†	0.5496	0.0062†	0.4473	0.0269*	0.1882	0.3779
LEV_t^2	-0.8566	0.0025†	-1.2764	<0.0001†	-0.8045	0.0075†	-0.5235	0.0807+	-0.5133	0.1014
$GROWTH_t$	0.4971	<0.0001†	0.5207	<0.0001†	0.5308	<0.0001†	0.5271	<0.0001†	0.5423	<0.0001†
$GROWTH_t^2$	0.0153	0.0096†	0.0222	0.0011†	0.0351	<0.0001†	0.0284	<0.0001†	0.0276	0.0014†
$QTOBIN_{t-l}$	0.1041	<0.0001†	0.0334	0.0045†	-0.0142	0.1900	-0.0136	0.2040	-0.0025	0.8227
$INVEST_t$	0.0099	0.3109	0.0289	0.0022†	0.0247	0.0065†	0.0154	0.0763+	0.0182	0.0409*
$INVEST_{t-l}$	0.0258	0.0084†	0.0192	0.0479*	0.0089	0.3619	-0.0021	0.8162	-0.0023	0.8253
Model Fit Statistics										
	$N =$	1,217	$N =$	1,094	$N =$	966	$N =$	839	$N =$	714
	$AIC =$	-476.09	$AIC =$	-340.32	$AIC =$	-379.32	$AIC =$	-456.95	$AIC =$	-402.11
	$BIC =$	-382.65	$BIC =$	-249.80	$BIC =$	-291.72	$BIC =$	-372.27	$BIC =$	-320.35
Covariance Parameter Estimates										
$Cov\ Parm$	**Estimate**	**s. e.**	**Estimate**	**s. e.**	**Estimate**	**s. e.**	**Estimate**	**s. e.**	**Estimate**	**s. e.**
$Companies\ (\sigma_b^2)$	0.05226	0.007200	0.05883	0.008075	0.06834	0.009171	0.07392	0.009819	0.06289	0.008461
$Residual\ (\sigma_e^2)$	0.03961	0.001702	0.03819	0.001745	0.03076	0.001512	0.02356	0.001262	0.02069	0.001226

Levels of significance: '+' 10% '*' 5% '†' 1%.

Generalized linear mixed models to:

Scenario 1: $\mu(QTOBIN_t) = \exp\,(intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 LEV_t^2 + \beta_3 GROWTH_t + \beta_4 GROWTH_t^2 + \beta_5 ROA_{t-1} + \beta_6 INVEST_t + \beta_7 INVEST_{t-1} + b_i)$.

Scenario 2: $\mu(QTOBIN_t) = \exp\,(intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 LEV_t^2 + \beta_3 GROWTH_t + \beta_4 GROWTH_t^2 + \beta_5 ROA_{t-2} + \beta_6 INVEST_t + \beta_7 INVEST_{t-2} + b_i)$.

Scenario 3: $\mu(QTOBIN_t) = \exp\,(intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 LEV_t^2 + \beta_3 GROWTH_t + \beta_4 GROWTH_t^2 + \beta_5 ROA_{t-3} + \beta_6 INVEST_t + \beta_7 INVEST_{t-3} + b_i)$.

Scenario 4: $\mu(QTOBIN_t) = \exp\,(intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 LEV_t^2 + \beta_3 GROWTH_t + \beta_4 GROWTH_t^2 + \beta_5 ROA_{t-4} + \beta_6 INVEST_t + \beta_7 INVEST_{t-4} + b_i)$.

Scenario 5: $\mu(QTOBIN_t) = \exp\,(intercept + \alpha_j YEAR_t + \gamma_k SECTOR + \beta_1 LEV_t + \beta_2 LEV_t^2 + \beta_3 GROWTH_t + \beta_4 GROWTH_t^2 + \beta_5 ROA_{t-5} + \beta_6 INVEST_t + \beta_7 INVEST_{t-5} + b_i)$.

The assumptions of normality and homoscedasticity about residuals were satisfied (according diagnostics in SAS Institute Inc., 2011, p. 347-352).