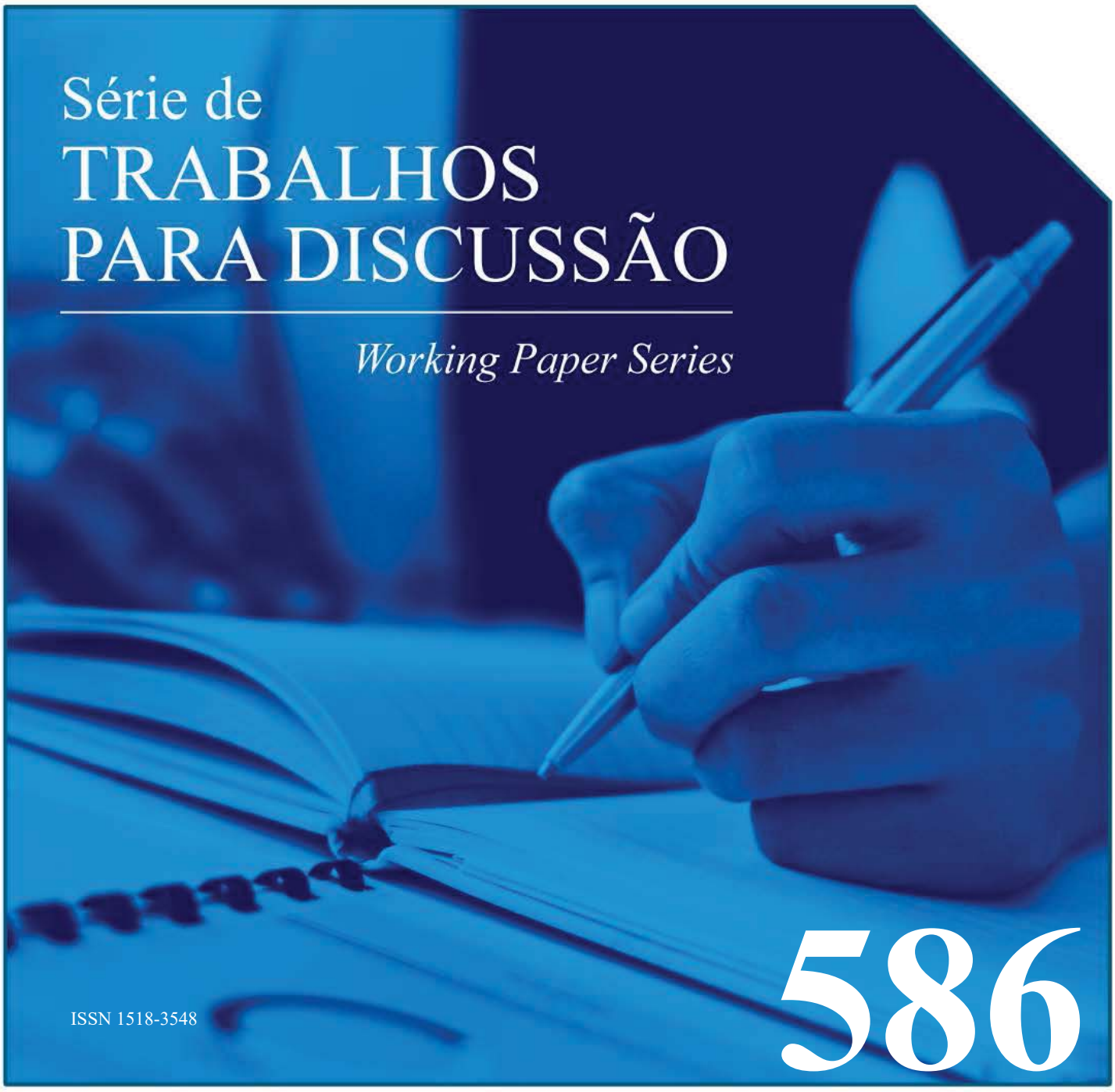


ISSN 1518-3548

Outubro 2023

Industry and Securities Market Influences on Indebtedness: evidence from a large dataset from Brazil

Flavio Alberti Docha, Luiza Betina Petroll Rodrigues

BANCO CENTRAL DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 586	Outubro	2023	p. 3-75

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Rodrigo Barbone Gonzalez

Co-editor: Eurilton Alves Araujo Jr

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Diogo Abry Guillen

The Banco Central do Brasil Working Papers are evaluated in double-blind referee process.

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-technical Summary

Academic studies with accounting microdata are largely based on data from companies that trade securities – in general, companies listed on stock exchanges. This is because, in Brazil and in most non-European countries, accounting microdata on unlisted companies are not available. Databases that combine these accounting microdata with bank and employee information from unlisted companies are even rarer.

To fill this gap in the Brazilian case, we built a new database, which we called Alexandria, gathering public and non-public, accounting and non-accounting information from more than 42 thousand accounting entities – of which more than 31 thousand non-financial companies (NFCs) – in general, medium-sized companies owned by non-residents to some extent. Accounting entities were classified as “CVM”, when participating in the securities market (trading shares or debentures), or “non-CVM”, otherwise. In all the years for which Alexandria has data, the total assets and total equity of non-CVM companies are greater than those of CVM companies, a rare feature among the databases available in Brazil.

Thus, this work presents Alexandria and shows its potential through an exploratory study of the impact of the “industry affiliation” and the “participation in the securities market” on company indebtedness (measured by the debt-to-asset-ratio) of NFCs between 2015 and 2021. The work concludes that (i) indebtedness is different between industries, with "Agriculture", "Construction" and "Real Estate Services" appearing among the least indebted; the industry difference is especially strong among non-CVM companies; (ii) participation in the securities market seems to increase indebtedness, but the evidence is weaker than that obtained for the industry affiliation; and (iii) indebtedness varies greatly between companies, even within the same industry. There are companies that operate for years with debt exceeding 100% of their assets, while others use almost exclusively equity capital. CVM companies are more uniform in terms of indebtedness levels.

Sumário Não Técnico

Estudos acadêmicos com microdados contábeis são largamente baseados em dados de empresas que negociam valores mobiliários – em geral, empresas listadas em bolsas de valores. Isso ocorre porque, no Brasil e na maioria dos países não-europeus, microdados contábeis sobre empresas não listadas não estão disponíveis. Bases de dados que combinem esses microdados contábeis com informações bancárias e de empregados de empresas não listadas são ainda mais raras.

Para cobrir essa lacuna no caso brasileiro, nós construímos uma nova base, que denominamos Alexandria, reunindo informações públicas e não públicas, contábeis e não contábeis, de mais de 42 mil entidades contábeis – das quais mais de 31 mil empresas não financeiras (ENFs) – em geral, empresas de médio porte detidas por não residentes em algum percentual. As entidades contábeis foram classificadas em "CVM", quando participam do mercado de valores mobiliários (negociando ações ou debêntures), ou "não CVM", em caso contrário. Em todos os anos para os quais Alexandria dispõe de dados, o ativo total e o patrimônio líquido total das empresas não CVM são maiores que os das empresas CVM, uma característica rara entre as bases disponíveis no Brasil.

Assim, este trabalho apresenta Alexandria e mostra seu potencial através de um estudo exploratório do impacto do setor e do uso do mercado de valores mobiliários no endividamento empresarial (medido pela relação entre passivo e ativo) das ENFs entre 2015 e 2021. O trabalho conclui que (i) o endividamento é diferente entre os setores, com "Agricultura", "Construção" e "Serviços Imobiliários" aparecendo entre os menos endividados; a diferença setorial é especialmente forte entre empresas não CVM; (ii) a participação no mercado de valores mobiliários parece aumentar o endividamento, mas as evidências são mais fracas que as obtidas para o setor; e (iii) o endividamento varia muito entre as empresas, mesmo dentro do mesmo setor. Há empresas que operam por anos com endividamento superior a 100% dos seus ativos, enquanto outras utilizam quase que somente capital próprio. Empresas CVM são mais homogêneas em termos de níveis de endividamento.

Industry and Securities Market Influences on Indebtedness: evidence from a large dataset from Brazil*

Flavio Alberti Docha **

Luiza Betina Petroll Rodrigues ***

Abstract

This article presents an unprecedented database, “Alexandria”, which consolidates a panel with more than 40 thousand Brazilian companies, of which 31 thousand are non-financial companies (NFCs). It then uses it to investigate the influence of “industry affiliation” and “participation in the securities market” on the level of indebtedness of NFCs from 2015 to 2021. The conclusions are as follows: (i) industry is important to explain the average and median level of indebtedness, especially among non-CVM companies, and "Agriculture", "Construction" and "Real Estate Services" are the least indebted industries; (ii) participation in the securities market seems to increase indebtedness, but the evidence is weaker than that obtained for the industry; and (iii) there is great heterogeneity in indebtedness, even within the same industry, especially among companies outside the securities market. These results suggest that studies based solely on listed companies fail to capture the heterogeneity of corporate debt strategies.

Keywords: database, capital structure
JEL Classification: G300, G320, M410

* The authors would like to thank (i) the company Klooks for sharing the median debt per CNAE section of its entire base, which allowed comparing and discussing possible sample biases in Alexandria; (ii) the numerous colleagues at the Central Bank of Brazil who made the construction of Alexandria viable, especially Fernando Alberto G. Sampaio Rocha; (iii) BCB/DEPEP colleagues, in particular Thiago Christiano Silva and Gabriel Garber, for their valuable contributions.

** Department of Statistics (Dstat), Central Bank of Brazil (BCB). flavio.docha@bcb.gov.br

*** Economic Department (Depec), Central Bank of Brazil (BCB). luiza.rodrigues@bcb.gov.br

1. Introduction

Academic research on companies is largely based on data from companies whose shares are traded in stock markets (hereinafter simply referred to as “listed companies”). A small number of studies include companies that, although not listed, issue other securities (such as bonds), and even fewer studies include companies outside the securities market. Often, implicitly or explicitly, it is intended that these studies reflect the reality of companies in general. However, especially in countries where the securities market – including the stock and bond markets – is not very expressive, the conclusions of these studies may not apply to the rest of companies.

This article presents **Alexandria**, an unprecedented database of Brazilian accounting entities built by the authors of this work and covering the period from 2013Q1 to 2021Q4. In its first version, Alexandria provides information on 42,143 accounting entities, of which 31,233 are non-financial companies (NFCs), including all companies that trade securities, whether listed or not, which we will call “CVM companies”[1], and thousands of companies outside the securities market, which we will call “non-CVM companies”. One of its major contributions to accounting and financial research in Brazil is that, unlike what happens with other commonly used databases, in Alexandria, non-CVM companies are more representative than CVM companies in terms of quantity, assets, shareholders' equity and outstanding credit. The second major contribution is that Alexandria also contains non-accounting data, such as information on partners, formal jobs and financial (incoming) flows.

The article uses Alexandria to carry out an exploratory study on a widely studied topic: the “capital structure”, a term that can be understood as the combination of equity and third-party capital that is being used at a given time to finance the organization's assets.

[1] The name was chosen because in Brazil the regulation and inspection of real estate securities is the responsibility of the CVM – Securities and Exchange Commission of Brazil (CVM, 2022), which is equivalent to SEC (U.S. Securities and Exchange Commission) in the US. A few securities may be traded without registration with the CVM (an example is a privately issued security).

"Debt-to-asset-ratio" – hereinafter referred as "indebtedness" – is an index frequently used to analyze the capital structure (Bressan et al., 2009; and Avelar et al., 2019)[2]. It is defined as the ratio between third-party capital and assets, that is, it is given by (Sant'Ana, 2001):

$$Indebtedness = Debt - to - asset - ratio\ (D/A) = \left(\frac{Short - Term\ Debt + Long - Term\ Debt}{Total\ assets}\right) * 100 \quad (1)$$

This work will analyze the indebtedness of non-financial companies[3] (NFCs)[4] to find out whether the indebtedness levels of NFCs are influenced by industry characteristics and/or by their participation in the securities market. In the case of this last variable, which we did not find information about in prior studies, our initial hypothesis is that the securities market tends to offer less funding costs from third-party capital. Thus, it would be expected that companies that participate in this market would desire (and achieve) a greater level of indebtedness.

Beyond this introduction, the remainder of this work is organized as follows. As most studies focus exclusively on listed companies, the second section presents a bibliographic review on the main accounting databases in the world that contain data on unlisted companies and on the indebtedness of unlisted companies. The third section analyzes the representativeness of non-financial companies (NFCs) from Alexandria in Brazil. The fourth section analyzes the indebtedness of NFCs between 2015 and 2021 using parametric and non-parametric techniques, and the fifth section presents the conclusions. Appendix A details the construction and variables of Alexandria.

---

[2] See Azevedo (2013, p. 30-33) for a discussion of the advantages and disadvantages of this index for studying the capital structure of entities. See also Parsons and Titman (2008, p. 6-7), who discuss the advantages and disadvantages of using accounting data instead of market value.

[3] To be rigorous, the work will analyze the indebtedness of formal accounting entities (that is, registered with the Federal Revenue Service of Brazil and identified by it through the CNPJ) with standardized accounting statements (in IFRS format), published and/or delivered to a public body. The accounting entities to be analyzed are of the type "formal non-financial business entity resident in Brazil" - for brevity, simply "non-financial company". The authors recognize that there are business entities that are individuals (eg: individual micro-entrepreneur) and others that are legal entities (such as joint-stock companies). Even so, we understand that changing "business entity" to "company" is a valid simplification.

[4] Numerous works in the area of finance corporations also focus on non-financial companies. Among those with this focus and that use databases that include unlisted companies are: Daskalakis et al.(2017), Demirgüç-Kunt et al.(2020), Exame (2021), Banco de Portugal (2019, p. 44) and Raalte (2021).

2. Literature review

The theoretical literature on the capital structure of companies is based on two approaches that compete with each other for the explanation of the capital structure: the Trade-off theory and the Pecking Order theory. It is not the objective of this work to deepen these theories, already well documented in Azevedo (2013, p. 15-16), Bastos e Nakamura (2009) e Marinšek (2015). However, we reproduce here a summary[5] by Campos (2009, p. 28):

> The first [theory] is called *Static Trade-off*, which determines that a company may have an indebtedness target and operate in its direction. This target may be established as a result of a cost-benefit analysis of the debt, in which the cost of bankruptcy is compared to the fiscal benefit. The second theory is that of the *Pecking Order*, corroborating that every company follows a hierarchical trend to establish its capital structure. First, the company gives preference to internal financing, i.e., using resources from its own cash flow. Should it need external financing, the logical sequence would be to enter into a debt agreement, issue debentures and convertible securities before opting to issue shares.

At an empirical level, the extensive literature on the capital structure of companies, including general debt, has not reached a consensus on the best theory (Pohlman and Iudícibus, 2010; Brito, Serrano and Franco, 2018). The influential article by Parsons and Titman (2008) summarizes the empirical literature and divides it into three segments: (i) studies that seek the determinants of capital structure; (ii) works that investigate changes in the capital structure; and (iii) works that explore the consequences of the level of indebtedness. This work is part of the first group.

The international empirical literature of this first group identified several possible determinants of capital structure: (i) deductibility of financial expense from taxes; (ii) cash flow volatility; (iii) company size; (iv) asset tangibility (liquidity); (v) ratio between the company's market value and its book value (Market-to-Book ratio); (vi) uniqueness of the product sold; (vii) industry; (viii) firm fixed effects (Parsons and Titman, 2008); (ix) profitability; (x) business risk; (xi) shares on the stock exchange (Azevedo, 2013); (xii) the personality of the CFO (Marinšek, 2015); and even (xiii) the presence of women on the company's board (García and Herrero, 2021). Azevedo (2013) and Britto, Serrano and Franco (2018) reviewed the Brazilian literature on the subject and showed most of these factors have already been tested in samples of Brazilian companies. However, most of the mentioned national and international works use a sample that only contains listed

[5] Our translation.

companies, and their conclusions about the influence of industry characteristics are not consensual[6].

As almost all studies use only data from listed companies, it is understandable that the importance of the securities market for the capital structure is a subject little explored in the literature. For this reason, this literature review often focuses on the impact of listing rather than using the impact of presence in the securities market on indebtedness. We did not find any Brazilian literature that evaluated the impact of the use of the securities market in capital structure, and the few Brazilian works that address the influence of listing are nonconsensual in their conclusions, as detailed in section 2.3.

2.1 Databases with accounting variables of unlisted companies

Among the few accounting databases that contain information on companies outside the securities market, the European databases stand out, usually focused on non-financial companies (Fernandes and Campos, 2014, p. 6).

Among these European databases, the Central de Balanços of Banco de Portugal is a database of economic and financial information on Portuguese non-financial companies (Banco de Portugal, 2021). As of 2007, the electronic accounting report "Informação Empresarial Simplificada (IES)", equivalent to the Brazilian SPED[7], allowed companies to submit information to multiple government agencies with a single statement. Thus, the database now includes almost the entire population of Portuguese companies. In 2021, this Central contained information on more than 493,000 non-financial companies.

Among the large private databases, the Orbis database, produced by the company Bureau van Dijk, deserves special mention. This database contains accounting information for 40 million companies around the world (Raalte, 2021) and has greater coverage in European developed countries. The European version of this database, "Amadeus", is the source of numerous academic works[8].

---

[6] There are also debates about the influence of other variables, but commenting on this debate would be beyond the scope of the present work.

[7] SPED (*Sistema Público de Escrituração Digital* or Public System of Digital Accounting), is an instrument that unifies the activities of receipt, validation, storage and authentication of books and documents that are part of the accounting and tax bookkeeping of businessmen and legal entities through a single digital flow of information with the Federal Revenue Service of Brazil (SPED, 2023).

[8] Bajgar et al. (2020) present a detailed study on the advantages, disadvantages and coverage of the Orbis base. Ribeiro et al. (2010) report that the OECD purchased the Orbis database and cleaned up the data to create its own database (the OECD Orbis database).

In Brazil, the largest database of accounting statements is that of the Brazilian Federal Revenue Service (RFB). In 2021, the agency received statements from around 1.19 million accounting entities (SPED, 2023)[9]. However, this information is subject to tax secrecy (Brasil, 1966, art. 198), which makes it, in practice, unavailable for research.

The second largest database of accounting entities in Brazil belongs to the company Serasa Experian (Serasa). In 2019 alone, the company collected accounting data from around 85,000 accounting entities[10]. This database is only available to customers and is less comprehensive and less detailed than the RFB's. As an advantage, it includes some balance sheets of smaller companies, which are not obliged to submit information to the RFB.

Other private companies, such as EMIS, Klooks and Balanços Patrimoniais, also collect financial statements in Brazil, usually from official gazettes. Finally, it is worth mentioning the *Maiores e Melhores* database from *Revista Exame*, compiled between 1996 and 2021 by Fipecafi (Fipecafi, 2021; Sant'Anna, 2001) and after that period by Ibmec. However, all are smaller bases than Serasa's.

2.2 Empirical studies on debt that include large databases of unlisted companies in the world and present industry and/or listing evaluation

Marinšek (2015, p. 8) used data from European unlisted companies from the Orbis base in a detailed work on the impact of debt on company performance[11]. The author concluded that factors such as asset tangibility, company size, profitability, industry and country, among others, are statistically significant to explain capital structure. For the author, it is necessary to separately analyze intercompany factors (between-firms effect, or what explains the difference between the capital structures of different companies) and intracompany factors (within-firms effect or what explains the change in capital structure from the same company):

> I showed that when comparing firms cross-sectionally by their average size (the between-firm effect) there are practically no differences in indebtedness. On the other

[9] The information refers to the number of companies that delivered digital bookkeeping (ECDs). Simply put, it can be said that the ECD is mandatory for medium and large companies (RFB, 2021c).

[10] Information received by e-mail by the authors, on March 19, 2021. The information provider was a Serasa Experian salesperson. The collection is made through balance sheets published in newspapers and through direct contact with the companies.

[11] The author used the Amadeus database (Bureau van Dijk, 2013, apud Marinšek, 2015, p. 8) with 8,777 companies from 25 European countries. Only companies with complete and consistent data in the analysis period (2005 to 2011) were kept, therefore the author warns that there may be a "survival bias".

> hand, the **within-firm increase in size reveals substantial leveraging – firms's expansions are largely financed with new debt**. I further demonstrated that an increase in growth needs additional external financing (preferring debt over new equity) and that tangibility has a strong between-firm effect, which highlights the importance of the average share of tangible assets: firms that operate with more tangible assets are able to use more debt. (Marinšek, 2015, p. 9) [our bold]

The same author decomposed the debt differences in the same year (p. 85) and concluded that the industry to which the company belongs (between firms [effect] within the same industry) explains most of this heterogeneity, in line with the extensive literature cited. The conclusion was the same using parametric and non-parametric statistical techniques (p.41).

Demirgüç-Kunt et al. (2020) also use the same basis as Marinšek (2015, p. 8) to study the capital structure of companies before and after the 2008 crisis. The authors collected information regarding about 160,000 companies in each year from 2004 to 2011, located in 47 countries[12]. The authors found strong evidence of debt reduction in both developed and developing countries, a reduction that was particularly acute for small and/or unlisted companies. However, the authors do not carry out an industry analysis.

Nehrebecka and Białek-Jaworska (2015), who analyzed indebtedness[13] with a sample of 800,000 observations of Polish companies from 1995 to 2012, showed that (i) small and medium-sized companies tend to have lower indebtedness; (ii) monetary policy had little influence on companies' debt decisions; (iii) companies incorporated in the form of limited liability companies tend to be less indebted; and (iv) the company's industry was not a significant variable to explain indebtedness.

Jaworski and Santos (2021), in a study that analyzed and compared indebtedness using a sample of around 51,000 Portuguese and Polish companies in the period from 2011 to 2019, extracted from the Orbis database, concluded that the industry, including the median indebtedness of the industry over time, is a significant variable to explain corporate indebtedness.

Den Berg (2021) analyzed a sample of 12,169 non-financial companies in the United Kingdom, obtained from the Orbis database, with data from 2014 to 2018, and

[12] The numbers refer to the data actually used, after filters applied by the authors to exclude companies with zero assets, countries with less than 20 observations, etc. About 70 Brazilian companies are used.

[13] The authors analyzed the Interest in Third-Party Capital, which is given by (Current Liabilities + Long-Term Liabilities) / (Current Liabilities + Long-Term Liabilities + Shareholders' Equity). This indicator is mathematically equivalent to general indebtedness, the subject of this work. In the case of the aforementioned article, the revaluation reserves were deducted from shareholders' equity.

concluded that profitability, tangibility, liquidity, listing, debt deductibility and company age, in addition to the industry characteristics, contribute to explaining indebtedness.

Andritzky (2003) carried out a specific study on the impact of the "industry" variable on indebtedness. The work used a sample of almost eight thousand companies from seven developed countries, with data between 1997 and 2001, to verify that the industry, in addition to the country, company size, asset tangibility, profitability and market-to-book ratio, are important to explain debt. The same author quotes Scott and Martin (1975), who used parametric and non-parametric techniques to refute previous studies and conclude that indebtedness has statistically relevant industry differences (something that will also be done in this work, in sections 4.2 and 4.3).

2.3 Empirical studies on debt in Brazil that address industry differences

The academic literature that seeks to investigate the empirical relationship between debt and industry using only data from Brazilian listed companies is numerous. Among the works that identified the industry as a statistically significant variable are: Lima and Brito (2003, apud Azevedo, 2013), Nakamura, Martin and Kimura (2004), Procianoy and Schnorrenberger (2004), Favato (2007) and Terra (2007), Bastos and Nakamura (2009) and Silva (2021)[14]. On the other hand, Gomes and Leal (2000, apud Azevedo, 2013) and Britto, Serrano and Franco (2018) found no statistically relevant industry difference in indebtedness.

Much more restricted is the Brazilian academic literature that seeks to investigate the same empirical relationship using data from listed and unlisted companies. This is somewhat curious because at least data for unlisted companies that trade securities (i.e., CVM companies that only trade securities) is relatively easily available. The same does not apply to companies outside the securities market (non-CVM companies), whose data are difficult to obtain and have lower quality[15] (in general, it is not audited – the "Maiores e Melhores" and "Valor Pro" are exceptions, because of the size of the companies

[14] Silva (2021) uses industry-specific variables, not industry dummy variables, as most studies do.
[15] Bonomo, Martins and Pint (2003), for example, studied the debt composition of Brazilian companies and the effect of exchange rates on them. The authors had data from unlisted companies from the Austin Asis database, but ended up not using data from unlisted companies because the investigation was not "fruitful". The authors attribute this failure to the fact that the unlisted statements are not audited, and therefore they decided to continue with data from listed companies only.

presented)[16]. Table 1 presents examples of studies that go beyond the universe of listed companies.

Table 1- Empirical studies on corporate debt carried out in Brazil that address industry differences using a sample that includes unlisted companies

Author(s) (year)	**Analysis period**	**Sample size and characteristics**	**Database used**	**Is the industry relevant to explain indebtedness?**
Nakamura (1992)*	1980-1989	-----	-----	Yes
Pohlman and Iudicibus (2010)	2001-2003	214 large non-financial and non-agricultural companies, of which 154 are not listed	Maiores & Melhores	Yes (1)
Azevedo (2013)	2008-2011	1,081 large non-financial and non-agricultural companies (2).	Maiores & Melhores	Yes
Correa et al. (2013)	1999-2004	389 large non-financial companies.	Maiores & Melhores (3)	Yes, for a few industries
Forte, Barros and Nakamura (2013)	1994-2006	4,400 small and medium-sized companies in the state of São Paulo (4).	Serasa Experian	Yes

Sources: compilation made by Azevedo (2013) and elaborated by the authors. Observations: (1) the variable proved to be significant in explaining short-term indebtedness. (2) About half of this sample was disregarded for the regression performed with a balanced panel.(3) Information inferred by the authors of this work from the base description; the article does not make the source explicit. (4) The initial sample consisted of 19,272 companies. This is the number after excluding financial companies, outliers and companies without full data.*

Forte, Barros and Nakamura (2013) carried out a study on indebtedness using unlisted Brazilian companies using a sample of 19,272 small and medium-sized companies in the state of São Paulo, with pieces of data between 1994 and 2006, obtained from the Serasa database (of which, 4,400 were used in the regressions). In addition to the industry affiliation, factors such as profitability, asset growth, size, risk and age of the firm, and especially indebtedness in the previous period, proved to be significant in

[16]In Brasil, large companies (with total assets greater than R$ 240 million or a gross annual revenue greater than R$ 300 million) must contract external auditors to audit their accounting statements (Brasil, 2007, art. 3). For this reason, probably all companies in the Maiores and Melhores base are audited, and almost all companies in the Valor Pro base as well. The same does not apply to the Serasa Experian database.

explaining indebtedness. Alexandria differs greatly from this database by containing (i) entities from all 27 Brazilian states; (ii) a bigger sample; (iii) in general, medium or large companies; (iv) data from a different period (2013 to 2021); and (v) extra information, such as the manager's gender and debt composition. As a disadvantage, Alexandria contains fewer accounting items, which prevents the calculation of EBITDA, for example.

As there are few studies that include unlisted companies, the dummy variable "listed" appears even less than the variable "industry" in the works. Exceptions are: Pohlman and Iudícibus (2010), who concluded that the variable is not significant, and Azevedo (2013), who concluded the opposite (for the author, ceteris paribus, listed companies tend to have lower debt). We did not find Brazilian studies that evaluated the impact of participating in the securities market on capital structure.

2.4 Reasons why industry and participation in the securities market could impact corporate indebtedness

As seen in the previous sections, the industry's impacts on corporate debt are far from being consensual in the specialized literature. Among those who agree there is some effect of industry affiliation on corporate debt, there are disagreements about the reasons why it occurs.

For Pohlman and Iudícibus (2010), "the idiosyncratic characteristics of each industry, such as type of activity and degree of concentration, could affect debt policies".

For Marinšek (2015, p. 40), companies operating in the same industry have similar indicators of asset tangibility, and, consequently, a similar amount of business risk – "which importantly determines the amount of debt the capital market will provide". The author cites supplementary literature that mentions other common variables within the same industry, such as the nature of competitiveness, technological state and regulation, which also influence companies in the same industry to have similar debt patterns. Thus, the industry's median debt would be a powerful predictor of a company's debt.

Forte, Barros and Nakamura (2013), citing Frank and Goyal (2008), follow the same line and mention the industry affiliation is an important predictor of indebtedness (at least for listed companies) and can capture many omitted factors, such as industries' specific regulatory restrictions and the influence of the type of business activity on third-party financing requirements.

In fact, the “industry factor” does not seem to capture only the tangibility of assets: Raalte (2021), Marinšek (2015) and Jaworski and Santos (2021, p. 32) added the variable “tangibility” and the industry dummies remained significant.

Zonenschain (1998) understands that, in part, the fact that debt differs between industries stems from the specificity of the product produced or supplied by the company: the more specific, the lower the use of debt. Other authors, reviewed by Den Bert (2021), mention that companies in the same industry tend to have an optimal capital structure in common. This suggests what Raalte himself (2021, p.28) found in a literature review: the “industry” variable reflects many omitted variables, common to all firms in an industry, such as the degree of competition and supply and demand conditions.

Copat (2009) carried out an in-depth literature review about the reasons why the industry characteristics affects corporate indebtedness. The author criticizes the use of dummies to control for industry effects: “they are little informative, because they do not show which are the specific characteristics of a given industry that make it so that the indebtedness of the companies that operate in it be greater or smaller than that of the others” (p.20). The author suggests the use of six variables specific to industries, calculated in general based on subaccounts of the Balance Sheet and the Income Statement. However, the results showed only three of these variables were important to explain the indebtedness of listed companies in the studied countries: industry concentration, industry life cycle, and the clients’ bargaining power (p. 135).

With regard to the relationship between indebtedness and listing, Den Berg (2021) and Raalte (2021) review the literature and assess that the sign of the relationship is not clear. On the one hand, listed companies are better known, face lower interest costs on debt, and have greater bargaining power in relation to banks and other institutions. Thus, it would be expected that they would be more indebted. Similar rationale to that applied to the variable “listing” could be applied to the variable “presence in the securities market”. A CVM company, listed or not, tends to face lower funding costs, so it would be natural to expect that they would want a higher degree of debt.

On the other hand, for Den Berg (2021), the cost of issuing shares/quotas is higher for unlisted companies, as it is very difficult to sell a share of an unlisted company. In this regard, listed companies would be less leveraged, as issuing shares is cheaper. Raalte (2021) also presents studies that concluded that variables such as company size, growth, business risk and profit affect listed and unlisted companies differently. Thus, it is

possible that, among the CVM companies, the listed ones have lower debt than the unlisted ones.

2.5 Econometric techniques used in studies that test the industry and/or listing as determinants of corporate indebtedness and related difficulties

The econometric difficulties involved in identifying the determinants of corporate debt are many and have already been well documented by Parsons and Titman (2008), Bastos and Nakamura (2009), Raalte (2021) and Miniaci and Pantegini (2021). Because of these difficulties, the national and international academic literature uses different techniques.

Table 2 brings examples of works that use samples of listed and unlisted companies and summarizes the techniques used. Panel regression techniques stand out, especially GMM (Generalized Method of Moments) and fixed effects, but even among these techniques there is much debate, as can be seen in Dang et al. (2015) and Barros et al. (2020), respectively against and in favor of GMM models with instrumental variables[17].

Table 2– Studies on corporate debt using samples that include unlisted companies: econometric techniques, dependent variables and data exclusion criteria

Author	**Technique used**	**Dependent variable**	**Criteria for excluding companies from the regression /treatment of outliers**
Sant'Ana (2001)	Simple and quadratic linear regression, year by year		
Andrtzky (2003)	POLS	D/A	• Indebtedness >100%
Pohlman and Iudicibus (2010)	POLS	D/A (1); ECP(1); ELP(1)	• Companies that presented average loss in the period • Companies for which the IRPJ and CSLL, added together, have a positive value (supposed "income" from taxes)
Azevedo (2013)	Panel and multiple regression	D/A, ECP and ELP	• Companies that did not contain all the data for all years (balanced panel) • Companies with negative EBITDA (p. 37)
Correa et al. (2013)	Balanced panel, fixed effects + random effects + GMM-Sys	D/A	• Financial companies
Forte, Barros and	Unbalanced Panel - POLS and GMM-Sys –	D/A	• Indebtedness >100% and • Winsorization

[17] Gaibulloev et al. (2014 p. 261) claim this method works poorly when the coefficient of the lagged variable approaches unity, as the instrumental variables become very weak.

Nakamura (2013)			
Marinšek (2015)	Balanced panel- MLM (Multilevel linear modeling)	D/A; EF (p. 35)	• Asset less than 5 million euros; • Indebtedness >100%; • Financial companies; • 9 criteria for excluding outliers; and • Winsorization in 1%, between 0 and 86.18%
Nehrebecka and Białek-Jaworska, (2015)	Unbalanced Panel - GMM-Sys	PEG	• Indebtedness >100%
Demirgüç-Kunt et al. (2020)	Generalized least squares, or GLS (Prais-Winsten estimator).	D/A; ELP; PLP	• Firms with zero debt or assets; • Firms with less than 5 employees; • Micro enterprises; • Firms that do not have at least six consecutive years in the financial bas; • State-owned companies, financial companies; • Countries with less than 20 companies each year, companies from Taiwan, companies from countries that are tax havens; and • Companies with long-term financing before the crisis zeroed
Raalt (2021)	Unbalanced panel - POLS	D/A, ECP, ELP	• Companies with less than 5 years of data; • Financial companies; • Winsorization at 5% (removes top 2.5 and bottom 2.5%, except dummies); • Considers only large or very large companies (operating income greater than 10 million euros, or total assets greater than 20 million euros or more than 150 employees); and • Reducing the number of unlisted so that the median asset of listed and unlisted is similar.

Source: Prepared by the authors. POLS = Pooled Ordinary Least Squares. GMM-Sys = S-GMM is the estimator of Blundell R, Bond S (1998). D/A = (total liabilities)/(total assets); ECP= (current liabilities)/(total assets); ELP= (long-term liabilities)/(total assets); EF= (Financial liabilities)/(assets); PLP= (Long-term liabilities)/(Total liabilities). PEG=(total liabilities)/(total liabilities-revaluation reserves). (1) total assets have been adjusted for the effects of inflation.

Diving into this econometric discussion is beyond the scope of this work. However, the authors understood that it was important to address three essential questions for econometric studies in corporate finance: (i) the data exclusion criterion; (ii) the fact that the dependent variable is (in most studies) a ratio with logarithmic transformation; and (iii) endogeneity.

2.5.1 Data exclusion criteria

In addition to detailing the econometric techniques used by each work on the determinants of corporate indebtedness, Table 2 also highlights the different data exclusion choices used. First, most studies exclude financial companies, including

holding companies. Although this exclusion makes sense, it should be borne in mind that there are non-financial companies (NFCs) owned by these holding companies, and these NFCs owned are unlikely to be in any analysis, as the individual balance sheets of the companies owned are rarely disclosed. If we are right, the samples available for corporate finance jobs contain fewer NFCs owned by holding companies than a representative sample of NFCs would recommend.

Second, due to the difficulty of dealing with zeroed data, some works exclude companies with no debt (that is, using only equity capital in a given year)[18]. Third, due to the difficulty of dealing with negative data, many works exclude all data from companies with debt greater than 100% (hence with negative equity). Nehrebecka and Białek-Jaworska (2015, p. 9) say the exclusion of this data is one of the main limitations of their study. The present work diverges in part from the literature because it includes in its analysis both NFCs with zero debt and NFCs with debt greater than 100%.

Fourth, probably in order to exclude misstatements and other outliers, some studies discard extreme data by the winsorization method (Dang et al, 2015; Forte, Barros and Nakamura, 2013; Raalte, 2021; Marinšek, 2015, p. 64). This study followed this practice only in parametric analysis (Section 4.3).

Fifth, in addition to the afore mentioned exclusions, studies using balanced panels forcibly exclude all observations from companies that do not have data for all years, thus generating a "survival bias" (Demirgüç-Kunt et al; 2020; Marinšek; 2015). In other words, balanced panels exclude new companies (which start their activities after the beginning of the period of analysis), companies that stopped operating in the middle of the period of analysis (including due to excessive debt and bankruptcy), and those which, for other reasons, do not have a financial statement for all years in the base[19]. As this study shows in Section 4.3, the use of balanced panels can alter these conclusions.

[18] The exclusion of companies with zero liabilities is not a problem for studies that only use data from companies that participate in the securities market, as practically all of them have some liability. It is even likely that the company's decision to participate in the securities market reflects an interest in increasing the use of third-party capital.

[19] The exclusion of companies for the formation of a balanced panel can be an especially big problem in the analysis of companies with a short and predetermined duration, such as special purpose companies formed for the construction and sale of real estate projects (generally framed in group 681 in ISIC/CNAE classification – see IBGE,2021, for more information on this classification).

2.5.2 Dependent variable as a logarithm of a ratio

The variable of interest in this work is a ratio – total liabilities/total asset, or debt-to-asset (D/A) – which always assumes non-negative values.

According to an influential article by Firebaugh and Gibbs (1985), the issue of using a ratio as a dependent variable in an econometric regression has raised great debates in the literature. The authors point out, however, that this use can be made without major adjustments to the model as long as the intercept is zero (here, if the liabilities are zero when the assets are zero, which is true in almost all cases due to the accounting identity[20]).

Bartlett and Partnoy (2020), however, suggest that studies that use a ratio as a dependent variable should add some additional variables to avoid bias. Thus, suppose "x_i" is any variable that is supposed to explain the D/A ratio. The authors state that, instead of using an equation such as 2, one should use equation such as 3, which allows the possibility that the explanatory variable "x_i" is related to liabilities, regardless of assets.

$$\frac{Total\ liabilities_i}{Total\ Assets_i} = \beta_0 + \beta_1 x_i + \epsilon_i \quad (2)$$

$$\frac{Total\ liabilities_i}{Total\ Assets_i} = \beta_0 + \beta_1 x_i + \underbrace{\beta_2 \left(\frac{1}{Total\ Assets_i}\right)}_{\substack{not\ needed\ if\ total\ liabilities\ equal\ 0 \\ when\ total\ assets\ equal\ 0}} + \beta_3 \left(\frac{x_i}{Total\ Assets_i}\right) + \nu_i \quad (3)$$

Another important point in the corporate debt literature is the use of logarithmic transformation of the dependent variable, which is subdivided into two questions: (i) whether the use of the logarithmic specification is theoretically appropriate; and (ii) what changes are needed in the model specification to accommodate the logarithmic transformation.

Regarding the first issue, the logarithmic transformation is generally used to deal with data with outliers, with positive asymmetry and/or with a lot of dispersion. Furthermore, the transformation allows interpreting the coefficients as elasticities or

[20] There is only one case in which assets are zero but liabilities are not: when equity is negative and exactly equal to third-party liabilities in absolute terms, that is, (current liabilities + non-current liabilities) = -(shareholders' equity). We understand it is such a rare case that it can be ignored.

semi-elasticities. However, this decision must be careful, as the logarithmic transformation is not always the best choice and brings costs to the researcher, as it makes it difficult to interpret the regression results and test hypotheses (Lee, 2020; and Feng et al., 2014).

Regarding the second issue – necessary changes in the model specification to accommodate the logarithmic transformation – Bartlett and Partnoy (2020) state that this transformation, applied to a ratio, avoids the problem of the variable omitted from equation 2 (since equation 4 is mathematically equivalent to equation 5).

$$ln\left[\frac{Total\ liabilities_i}{Total\ Assets_i}\right] = \beta_0 + \beta_1 x_i + \epsilon_i \quad (4)$$

$$ln[Total\ liabilities_i] = \beta_0 + \beta_1 x_i + 1 \cdot ln[Total\ Assets_i] + \epsilon_i \quad (5)$$

However, according to Bartlett and Partnoy (2020), equations 4 and 5 assume a one-to-one relationship between ln(Liability$_i$) and ln(Asset$_i$), which is rarely true. Thus, the authors suggest using equation 6 (which includes a control, the logarithm of the denominator as the dependent variable) instead of equation 4.

$$ln\left[\frac{Total\ liabilities_i}{Total\ Assets_i}\right] = \beta_0 + \beta_1 x_i + \beta_2 \cdot ln[Total\ Assets_i] + u_i \quad (6)$$

Finally, as mentioned in Section 2.5.1, many authors who econometrically modeled the determinants of corporate debt chose to exclude from the sample companies with zero debt or leverage greater than one, also because the logarithmic transformation requires non-null values. Bellégo, Benatia and Pape (2019) state this procedure adds selection bias or, at the very least, changes the scope of the study. The literature points to alternatives used to avoid deleting data, but none is mentioned as ideal:

- add a constant before doing the logarithmic transformation, so that the dependent variable Y_{it} is transformed into $\tilde{Y}_{it} = \ln(Y_{it} + c)$ (Feng et al., 2014)[21].

  We used this alternative in the parametric tests, with c = 1.

[21] See Lee (2020), Feng et al.(2014) for a detailed discussion of the disadvantages of constant addition. Bellégo, Benatia and Pape (2019) calculate the endogeneity bias brought about by this choice and comment that this procedure also reduces standard errors.

- use the IHS transformation, in which the dependent variable Y_{it} is transformed into $\tilde{Y}_{it} = \left\{ ln\left[\theta Y_{it} + \sqrt{\theta^2 Y_{it}^2 + 1} \right] \right\} \div \theta$ (where generally $\theta = 1$) (Bellégo, Benatia and Pape, 2019)[22];
- use a Poisson model (Bellégo, Benatia and Pape, 2019).
- use the Box-Cox transformation, where $\tilde{Y}_{it} = \begin{cases} \frac{(Y_{it}+\lambda_2)^{\lambda_1}-1}{\lambda_1} \; if \; \lambda_1 \neq 0 \\ \log(Y_{it}+\lambda_2) \; if \; \lambda_1 = 0 \end{cases}$ (where normally $\lambda_2 = 0$) (Atkinson et al., 2021).

2.5.3 Endogeneity

Endogeneity, which describes the situation in which an explanatory variable is correlated with error, usually occurs in the form of omitted variables, measurement error, simultaneity (Wooldridge, 2010, p. 50) or significant degree of persistence[23] (dynamic endogeneity). Barros et al. (2020) describe how all these forms appear in corporate finance panels. We add here only the complications associated with databases that contain companies outside the securities market, which refer to the omission of variables and measurement errors.

Databases that only contain companies that trade securities typically contain more variables than databases that include companies that do not. As a result, it is natural to think the problem of omission of variables is even more important in studies that use the second type of database. Because of this problem, many of these studies significantly reduce the initial sample and build models only with companies that contain all the desired line items[24]. This strategy, while understandable, creates another problem (selection bias).

Regarding measurement errors, the authors of this work believe the liabilities of non-CVM companies are normally measured more accurately than the assets. This would also occur, in the authors' opinion, because debts are periodically recalculated according

[22] See Aïhounton and Henningsen (2019) for a detailed discussion of the bias brought about by the IHS (inverse hyperbolic sine) transformation and its relevance depending on the measurement unit.
[23] Most of the authors mentioned in sections 2.2 and 2.3 verified that the debt of the previous period is an important variable to explain the current debt (which could happen, for example, because the company cannot get rid of debts quickly or because it maintains a level of indebtedness that it considers appropriate). However, the inclusion of the lagged dependent as an explanatory variable violates the hypothesis of strict exogeneity (Wooldridge, 2010, p. 256).
[24] This is the case of Forte, Barros and Nakamura (2013), Marinšek (2015, p. 37) and Raalte (2021).

to contracts, while assets, especially long-term assets, are subject to less accurate accounting valuation in the case of smaller companies:

- CVM companies and large non-CVM companies are subject to compliance with accounting rules that often require valuation of their assets at fair value or at recoverable value.
- Smaller non-CVM companies, on the other hand, tend to record the asset based on amortized cost or historical cost methods.

Assuming a (common) scenario in which the historical cost is lower than the market price of the asset, the indebtedness of smaller non-CVM companies would *appear higher than it really is*, since the numerator (liabilities) is updated periodically (with new debts and interest on previous debts), while part of the denominator (assets) has its value always reduced by depreciation. However, we found no reference in the literature on this topic.

Barros et al. (2020) perform Monte Carlo simulations with panels of companies to verify how the different sources of endogeneity alter the regression coefficients. The authors concluded that the OLS, fixed effects and random effects estimators may be inconsistent in the presence of endogeneity problems which are very likely in the context of corporate finance. On the other hand, the GMM estimators (in particular the GMM-Sys), if well specified, are capable of adequately dealing with these problems.

3. Non-financial companies (NFCs) analyzed in this work: representativeness

As already mentioned, we have built a database – which we called Alexandria – on formal entities located in Brazil, each identified by its company identification number, "CNPJ". Alexandria contains a rich combination of registration, accounting, partners composition, credit (including external debt) and incoming bank flows information. The financial statements, which represent the main novelty of the database, were compiled from different sources, with lengthy manual verification to reduce inconsistencies. The database structure and its variables, including those that were not used in the empirical exercise of this work, are presented for the first time in Annex A.

Non-financial companies (NFCs) represent the majority of entities contained in Alexandria and are the focus of the exploratory study of this work. This chapter is

intended to compare these NFCs from Alexandria with the population of NFCs from Brazil to assess their representativeness.

3.1 Scope of the empirical exercise: annual data from non-financial companies (NFCs) from 2015 to 2021

Alexandria NFCs will be divided into twelve groups according to industry and securities market registration, as per Table 3:

Table 3– Non-financial companies (NFCs): definition of the industries used in this work and name of the variables that identify the groups

Industry code used in this work	**Industry description**	**CNAE/ ISIC sections included**	**NFC subgroup name**	
			Registered with the CVM	**Not registered with the CVM**
Agr	"Agriculture", including Agriculture and Livestock	A	*CVM_Agr*	*NCVM_Agr*
Ind	"Extractive and manufacturing", including • Mining and quarrying; • manufacturing; • electricity; • gas and water supply.	B to E	*CVM_IND*	*NCVM_IND*
Constr	"Construction"	F	*CVM_Constr*	*NCVM_Constr*
Com	Commerce: "Trading and repair of vehicles"	G	*CVM_Com*	*NCVM_Com*
Imob	"Real-estate activities"	L	*CVM_Imob*	*NCVM_Imob*
Serv	"Services", except financial, domestic, real-estate and public administration	H to S, except K, L and O	*CVM_SERV*	*NVM_SERV*

Source: prepared by the authors. For a detailed description of the CNAEs classification, which corresponds to the ISIC international classification at the level of sections, see IBGE (2021). The reader should note that the NFCs consider only CNPJ8 with a legal nature starting at 2, which identifies business entities, but excludes all individual rural producers. The "legal nature" 221 was also withdrawn, as it identifies non-resident business entities.

Only annual financial statements will be considered. This definition includes both the statements effectively closed in December and those that end in a different period, generally to coincide with the harvest year. In addition, only data from 2015 to 2021 will be considered[25].

[25] The period was chosen due to the significant increase in sample size from 2015 onwards (due to the five-year Census and regulatory changes that made the Accounting and Partners Statement mandatory – see section A.5.1).

3.2 Number of NFC companies in Brazil and Alexandria

Between 2015 and 2021, the number of active CNPJ8s in Brazil was between 17.3 and 20.8 million, the vast majority of which are non-financial microenterprises. However, a large number of these active entities were, in practice, inactive. For example, of the 19.1 million active NFCs in 2020,

- only 10.7 million paid some federal tax during 2020 (RFB, 2021b);
- only 6.6 million registered any incoming bank flow during 2020 and/or had any bank, foreign or securities liabilities at the end of 2020[26].

Even considering only the entities that had some effective economic activity, the minority produces financial statements, as explained in Section 3.1. In 2021, the RFB received 1.19 million bookkeeping documents (SPED, 2023). Assuming 95% of companies are NFCs, we can conclude that, of the 19.6 million NFCs active in 2021, no more than 1.13 million produce financial statements.

Of those NFCs that produce financial statements (at most 1.13 million, as estimated above), a small number, 27,573, are in Alexandria[27]. Even so, this is a larger sample than that used in any other study on corporate debt in Brazil, as shown in Section 2.3.

3.3 Representativeness of Alexandria in terms of volume: incoming bank flows and credit flows.

By construction, Alexandria captures only companies that produce accounting statements[28], and therefore the vast majority of micro and small companies are outside this base. Table 4 shows that between 2015 and 2021 (period of interest of the present work), the median assets of NFCs in Alexandria ranged between R$ 7.83 and R$ 46.88 million (column A), which corresponds to a medium enterprise (large enterprises have assets superior to R$ 240 million, according to Brasil, 2007).

Moreover, Table 4 also shows the median of assets for the group of NFCs for which Alexandria had data about their net profits (column B), as only this group was used in the parametric evaluation in section 4.3. As observed in the data in column B, these

[26] See Annex A, item A.8 and A.7 for the definition of the variables “incoming bank flows” and the various types of liabilities compiled in Alexandria.
[27] Please note that the number of entities in Alexandria is much greater: 42,143. Of these, 42,068 are corporate entities, and of these, 31,233 are NFCs. Of these NFC, 27,573 were active at the end of 2020.
[28] Regarding this aspect, see the annex, item A.1.

NFCs are still, in general, of medium size[29], but are much larger than NFCs for which Alexandria does not have data about net profits (column C).

Table 4– Median assets and number of NFCs with data about assets in Alexandria, by availability or non-availability of data about profits in each base date

Year	Median assets of NFCs (R$ million)			Number of NFCs		
	(A) With asset data by date	(B) Of which present data on profits	(C) Of which do not present data on profits	(D) With asset data by date	(E) Of which present data on profits	(F) Of which do not present data on profits
2015	7.83	14.14	0.80	18,724	15,237	3,487
2016	37.46	149.77	5.22	10,949	5,778	5,171
2017	46.88	172.26	5.50	10,634	6,029	4,605
2018	42.01	205.61	5.55	11,406	5,836	5,570
2019	15.87	244.22	5.11	17,830	5,630	12,200
2020	11.12	25.44	3.08	21,035	13,810	7,225
2021	12.79	347.18	5.58	18,629	4,950	13,679

Source: Alexandria. Observe that the median of assets in 2015 and 2020 was lower than that of other years, and that the number of NFCs, including those with data about profits, is greater. This occurs because in these years the Five-Year Brazilian Foreign Capitals Census took place; they capture data from all companies owned in some percentage by non residents, and Alexandria used all these statements. See section A.5.1.

To estimate the importance of NFCs in Alexandria, compared to the total number of active NFCs on each base date, this section used indicators for which the total values (referring to all NFCs) are known:

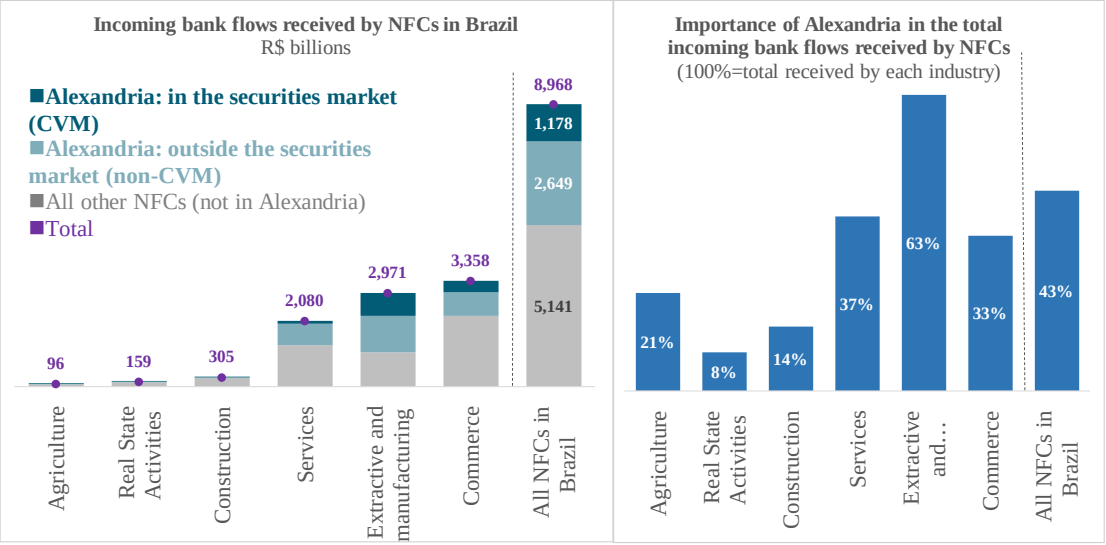
- Incoming bank flows received by the NFCs (variable described in item A.8 of the annex)
- Total credit to NFCs (variable described in item A.7 of the annex)
  - Of which, credit from the National Financial System to NFCs (variable described in item A.7.1 of the annex)

Figure 1 shows that, of the R$ 8.97 trillion in incoming bank flows that entered the bank account of all NFCs throughout 2020, 43% was received by NFCs included in Alexandria. The representativeness of Alexandria varies by industry: only 8% of flows from "real-estate activities" are received by NFCs from Alexandria, while in the "extractive and manufacturing industry", this number is 63%[30].

[29] Except in 2019 and 2021, probably because of inflation. The definition of large companies was defined in 2007 by the law on nominal values (Brasil, 2007).

[30] In the authors' opinion, this different representativeness may reflect different incentives in each industry. It is possible, for example, that the "extractive and manufacturing industry" have significant economies of scale and incentives to attract foreign capital and to be more present in the securities market. For this reason, industrial companies would generally be larger and more present in BCB's and Economatica's database,

Figure 1– Incoming bank flows received by NFCs in Brazil in R$ billion (2020) and importance of Alexandria



Source: Authors' preparation based on BCB and Alexandria data. Note: "Alexandria: not registered with CVM" NFCs are active NFCs on December 31, 2020, with no CVM data on that date and with at least one balance sheet available between 2013Q1 and 2021Q4 (condition for being in Alexandria). That is, all NFCs in Alexandria are considered, even if the balance sheet is not specifically available for 2020.

Following similar rationale, the representativeness of non-financial companies (NFCs) in Alexandria was analyzed using credit data. Alexandria companies accounted for 75% of the total outstanding credit to NFCs[31] at the end of 2020. Considering only the R$ 1.4 trillion[32] in credit owed to the National Financial System (SFN) by NFCs at the end of 2020, Alexandria companies accounted for 49%.

In summary, although the number of NFCs in Alexandria is small compared to the total number of Brazilian NFCs, and even compared to the number of NFCs with a balance sheet (1.13 million maximum in 2021, according to our estimate in section 3.2), the representativeness in the economic activity is quite expressive.

---

which would make it so that Alexandria could capture more from this industry (see Annex A.5.2). More studies are needed to corroborate these hypotheses.

[31] This high percentage can be explained, largely, because the Alexandria sample contains all enterprises detained by non-residents, and these companies have greater ease of external funding. This results also seems to support the conclusion of Freire et al. (2021), that "companies with a greater level of internationalization have greater debt." However, further studies might be needed to verify the presence of causality between these variables.

[32] This value is very similar to that reported in the SGS 22047 series - Credit operations outstanding by type of borrower - Private sector - Legal entities - R$ (million). The difference is that this series includes credit granted to holding companies, explicitly excluded from NFCs.

It is interesting to note that the importance of “CVM NFCs” in Alexandria is smaller than the importance of “non-CVM NFCs”, both in terms of incoming bank flows and in terms of domestic bank credit (“SFN credit”)[33]. Undoubtedly, this will be one of Alexandria's great contributions to accounting and finance research in Brazil.

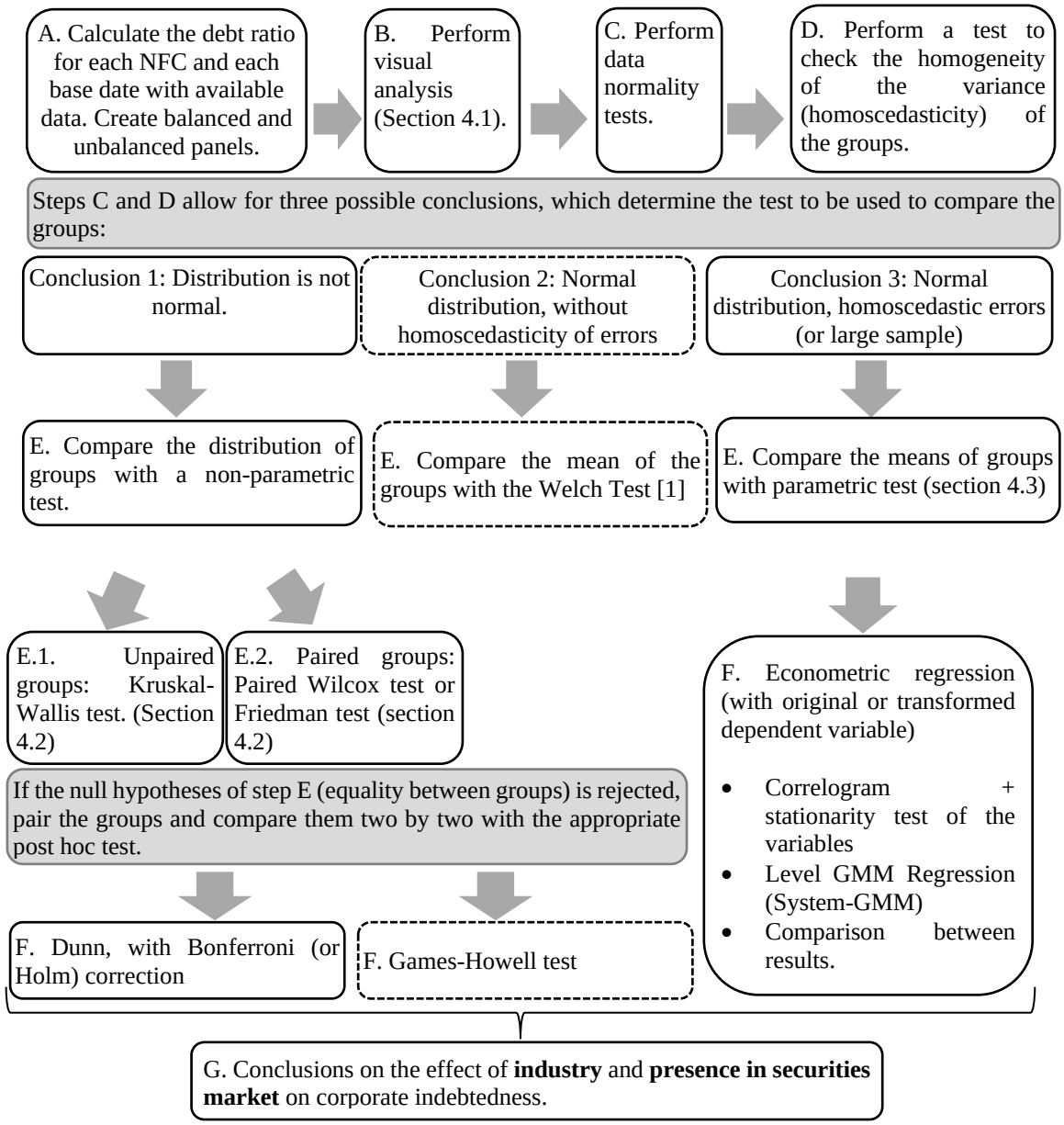
4. Panel of Non-financial Companies (NFCs): Differences in the Industry Debt- to- Asset Ratio

This section will maintain the outline of Section 3 and continue to analyze the annual statements of non-financial companies (NFCs) by industry, in the period from 2015 to 2021.

The core of this section is to see whether there are statistically significant differences in indebtedness according to industry and presence in the securities market. The steps of this analysis are summarized in Figure 2. Steps that weren't necessary (as the tests led to another path) are in dashed lines.

[33] However, the greater importance of non-CVM companies in Alexandria is not verified in the concept of “total credit”: in December 2020, “CVM” NFCs in Alexandria owed R$ 1.43 trillion in total credit, while “non-CVM” companies in Alexandria owed R$ 1.22 trillion. Non-CVM companies outside of Alexandria owed the remaining R$ 0.87 trillion.

Figure 2- Non-financial companies (NFCs) in Alexandria: steps of debt-to-asset ratio analysis



Source: prepared by the authors. [1] If the asymmetry is strong, prefer the Brown-Forsythe ANOVA test.

4.1 Visual, normality and homoscedasticity analysis

The first noteworthy element regarding companies' indebtedness, according to the literature, is that it changes very little over time, regardless of the industry. In fact, a large group of authors have observed that indebtedness in a given year is similar to that of the year before (see, for example, Nehrebecka and Białek-Jaworska, 2015). Because many companies have debts whose payment can take longer than 12 months, this observation can simply mean that the company has not yet finished paying the debt it had the previous

year. However, it is also possible that this stability reflects individual factors of the firm (such as the deliberate choice to maintain a certain level of indebtedness, for example). To try to separate the accounting effect between one year and another from supposedly more long-lasting characteristics, we used a longer timeframe. We considered the debt of each NFC in 2015 and 2020 and classified the debt of each one in relation to the industry it operates in as "low", "normal", and "high". The result, presented in Table 5, shows that of the 8,206 NFCs for which Alexandria had indebtedness data both in 2015 and 2020, most did not change classification between 2015 and 2020. In other words, those whose debt was considered low in 2015 maintained this level in 2020, and the same was true of the other classifications. The median variation was 0.16 percentage points, which means that those with a debt-to-asset ratio of 50% in 2015 increased to 50.16% in 2020. The result reinforces the perception that companies – regardless of industry – tend to maintain their standard of indebtedness over time, and that the companies in Alexandria reflect the pattern already observed in other databases.

Table 5 – Debt ratio of non-financial companies (NFCs) in 2015 and 2020, classified in relation to the indebtedness of their industry.

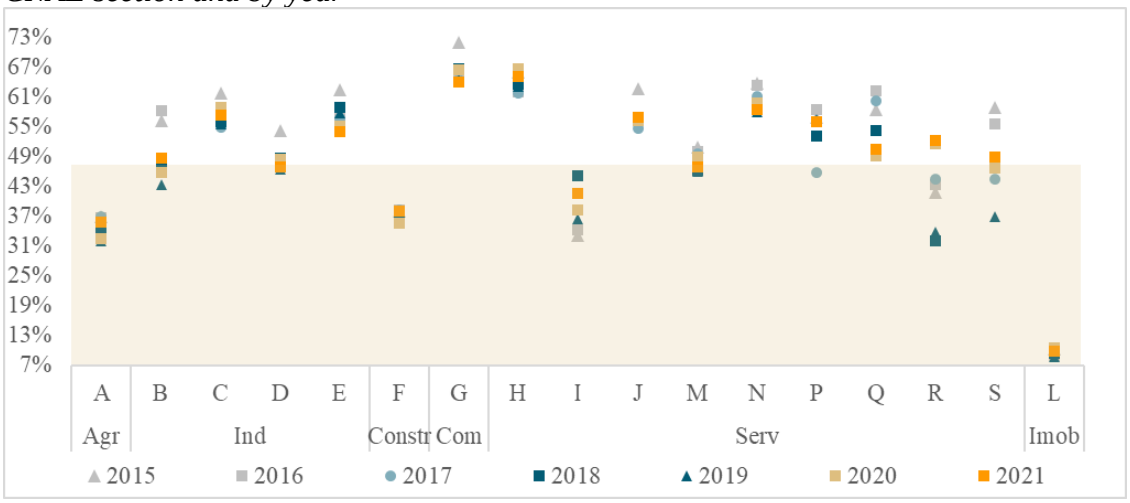
	Debt classification in 2020			**Total number of NFCs**
Debt classification in 2015	*Low* for the industry (lower than the first quartile)	*Normal* for the industry (between the first and third quartiles)	*High* for the industry (greater than the third quartile)	
Low for the industry (lower than the first quartile)	1,154	642	97	**1,893**
Normal for the industry (between the first and third quartiles)	696	3,185	646	**4,527**
High for the industry (greater than the third quartile)	125	538	1,123	**1,786**
Total	**1,975**	**4,365**	**1,866**	**8,206**

Source: Alexandria.

Going back to the static analysis, Figure 3 presents the median indebtedness of the CNAE sections of the NFCs of Alexandria, without breakdown by presence in the securities market. This first visual analysis suggests the industry is an important variable to explain the level of indebtedness. In particular, three CNAE/ISIC sections had a median

indebtedness of less than 40% in all years analyzed: "A" (Agriculture), "F" (Construction) and "L" (Real estate activities)[34]. At the other end, two CNAE/ISIC sections had a median indebtedness above 60% in all years: "G" (Commerce) and "H" (Transport). In general terms, the median values are slightly higher than the medians of the Klooks company database (to which the authors had access), but the industry ordering (which industry is more – or less- indebted) is very similar[35]. The important difference is section G (Commerce): in the Klooks base, the median indebtedness is much lower than in the Alexandria base, and therefore in this private database sections H and E, and not G and H, are the most indebted.

Figure 3–Non-financial companies (NFCs) in Alexandria: median indebtedness, by CNAE section and by year



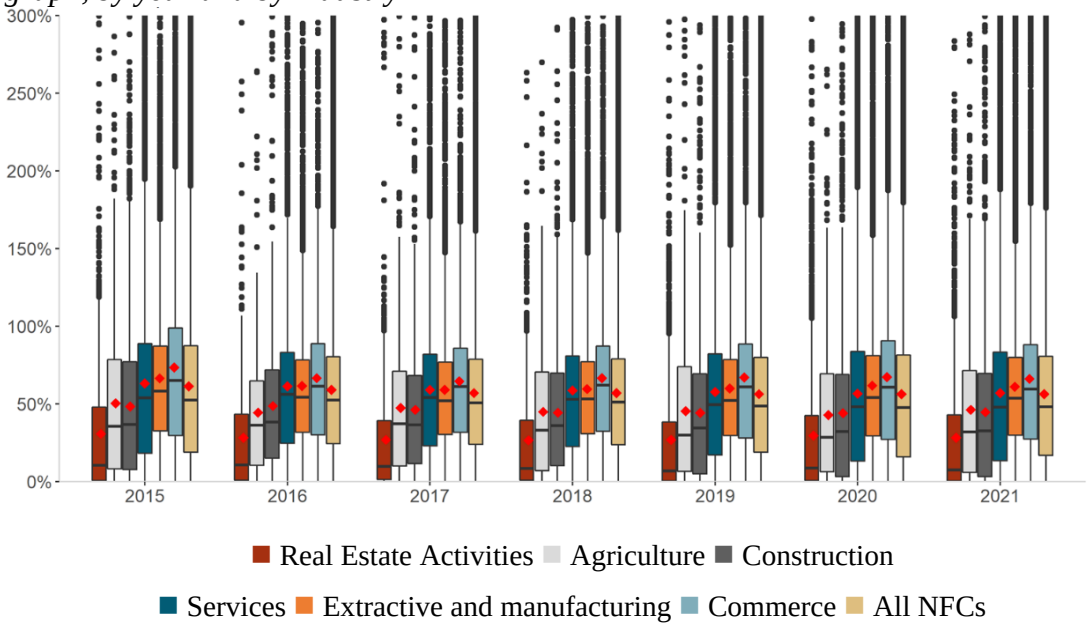
Source: Alexandria.

Figure 4 aggregates the CNAE/ISIC sections into industries (without separating by presence in the securities market) and presents the distribution of industry debt in box-plot format. Visual analysis suggests that (i) the sample shows significant variability in debt and a large number of outliers (representing heavily indebted companies)[36]; (ii) in

[34] These CNAE/ISIC sections coincide with the groups "Agriculture", "Construction" and "Real Estate Activities" presented in section 3.1.
[35] Annex B compares the median debts of NFCs in Alexandria and the Klooks base.
[36] The existence of a significant number of companies with indebtedness above 100%, even much higher than that, is not exclusive to Alexandria. Marinšek (2015, p. 82) cites several authors who found huge differences in debt among the firms in their samples. Banco de Portugal (2019b, p. 32-33) informs that in 2017 and 2018 around 26% of Portuguese non-financial companies had negative equity (therefore they had a debt-to-asset ratio bigger than 100%) - a situation more common in micro-enterprises and in industries

all the years analyzed, the lowest, mean and median indebtedness were those in the real estate activities industry, while the greater mean and median indebtedness were in the commerce industry[37].

Figure 4– Non-financial companies (NFCs) of Alexandria: debt-to-asset ratio box-plot graph, by year and by industry



Source: Alexandria. The vertical axis has been cropped at 300% for ease of viewing.

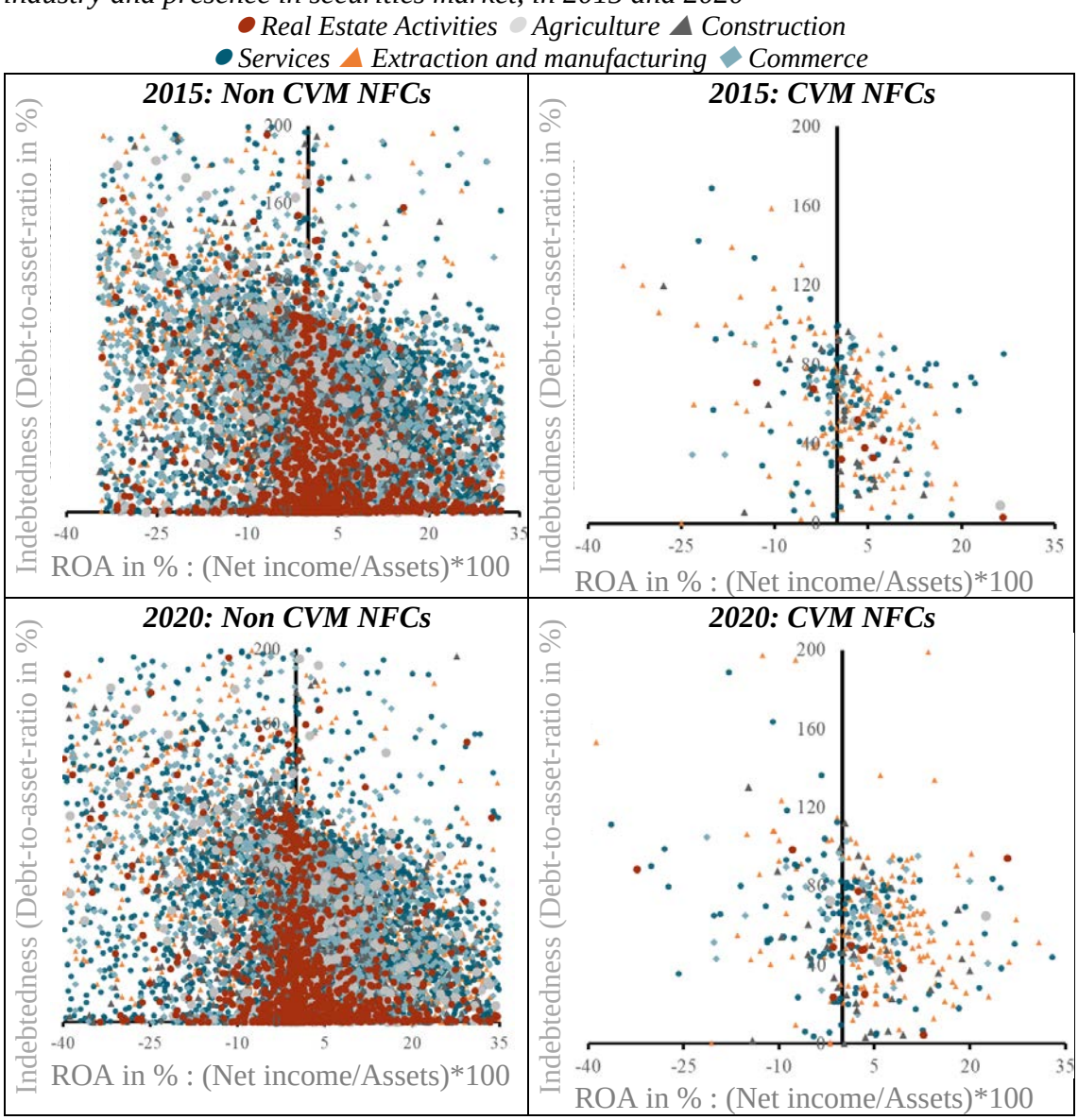
Figure 5 presents the indebtedness and profitability values (measured in ROA – the net-income-to-assets ratio) in 2015 and 2020, by industry and presence of the securities market. The result suggests a slightly negative relationship, i.e., in general, more profitable companies present lower indebtedness that enterprises with loss – a result in line with the Pecking Order theory, according to which firms prefer internal financial resources (Bastos and Nakamura, 2009). However, it is impossible not to note (i) the expressive dispersion of data, which decreases correlation between both variables; and (ii) the large number of companies in the "Real Estate Activities" industry that operate with debt-to-asset ratio equal to zero and outside the securities market (maybe, precisely because they do not have the conditions or the interest in capturing third-party capital).

"commerce" and "services". Urionabarrenetxea et al. (2016, p. 1), in an article that used the Amadeus database, report that in 2012 almost 20% of European companies had negative capital.

[37] The balanced panel presents similar conclusions regarding industry medians but presents less outliers.

This group of companies is not captured in studies that exclude companies that are debt-free, such as that by Demirgüç-Kunt et al. (2020).

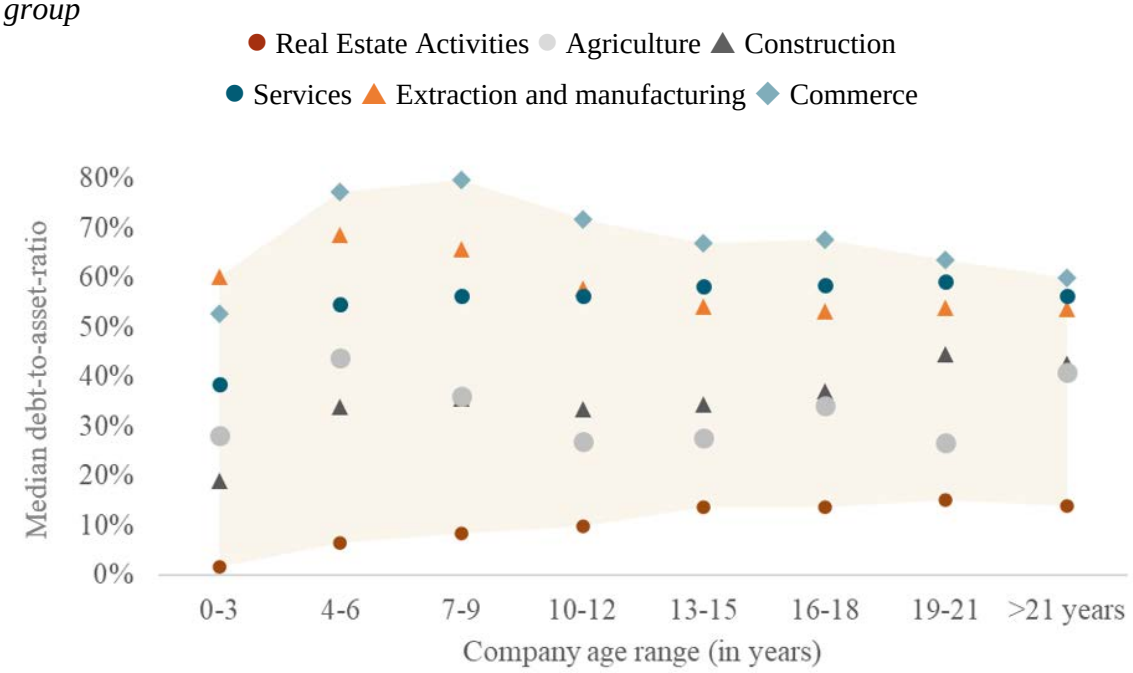
Figure 5- Alexandria non-financial companies (NFCs): Indebtedness versus ROA, by industry and presence in securities market, in 2015 and 2020



Source: Alexandria. The vertical axis has been cropped at 200% for ease of viewing. The reader should note that the sample of NFCs with profit data is smaller than the total number of NFCs with indebtedness data, which is why this graph presents a smaller sample of companies than most of the graphs in this section 4.1.

Figure 6 presents the median indebtedness by age group of the company. The visual analysis suggests that (i) all the industries increased their debt starting in the third year of age, but (ii) some industries, notably commerce and extraction and manufacturing, slowly reduced their indebtedness starting in their ninth year.

Figure 6– Alexandria non-financial companies (NFCs): median debt by company age group



Source: Alexandria. (i) Data refer to the period from 2015 to 2021, therefore the same company may appear more than once in the same age range.

Moving on to the analysis by industry and by presence in the securities market together (tests with two factors[38]), the CNAE sections were aggregated into the ten groups described in Section 3.1. The number of CVM NFCs in the "Agriculture" industry ("CVMAgr" group) is very small (4 companies in 2021), and therefore it was necessary to disregard it in some analyses.

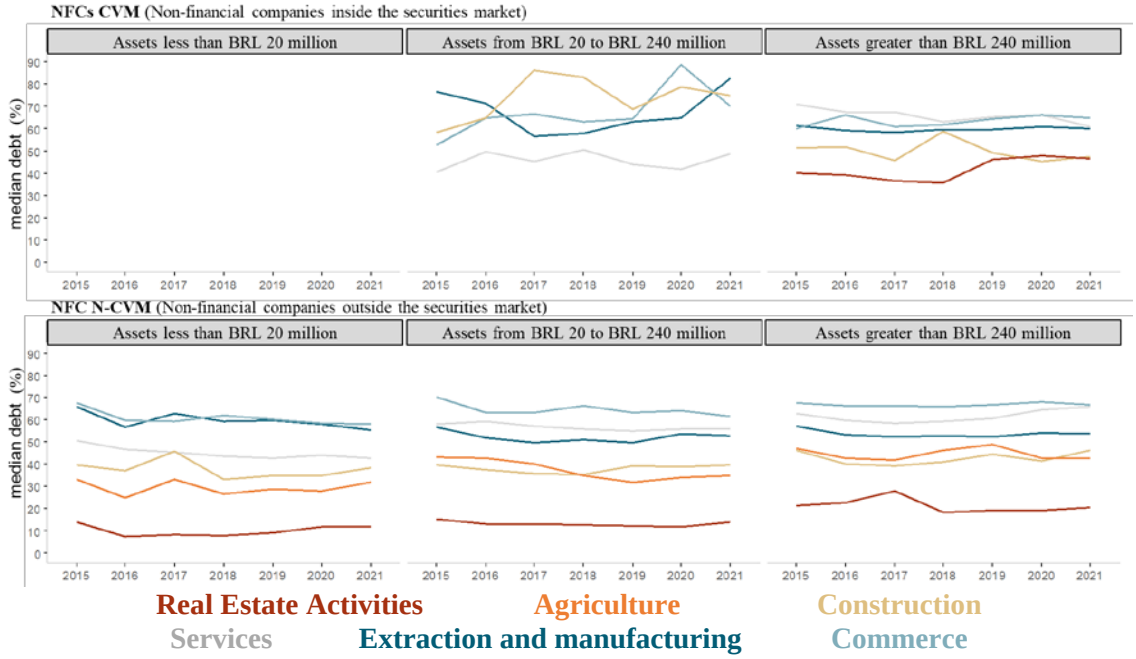
Figure 7 shows the median of the debt ratio separated by industry, presence in the securities market and size. Visual analysis suggests that:

- On comparing companies of the same industry and size, CVM companies are in general more indebted than non-CVM companies.
- Among non-CVM companies
  - "Real Estate Activities" uses less third-party capital than all the other industries in all the years.
  - "Non CVM agriculture" and "Non CVM construction" present similar median indebtedness (a little higher than "Non CVM Real Estate Activities", but lower than the other industries).

[38] Two-factor tests to be used (see section 5.2): 1. Non-parametric Mann-Whitney test pairing samples from pairs of groups; 2. Non-parametric Kruskal-Wallis test with post-hoc Dunn analysis.

- Size of companies (measured here by total assets)
    - impacts the indebtedness of non-CVM companies: the larger the company, the greater its level of indebtedness (in line with what most academic papers find[39]). This conclusion is not so clear only in the extractive and manufacturing industry.
    - does not seem to impact the indebtedness of CVM companies as clearly (there are industries that presented an increase of indebtedness with increased assets, but other industries showed the opposite).

Figure 7– Non-financial companies (NFCs): Median debt-to-asset ratio (%) by asset size, presence in the securities market and industry



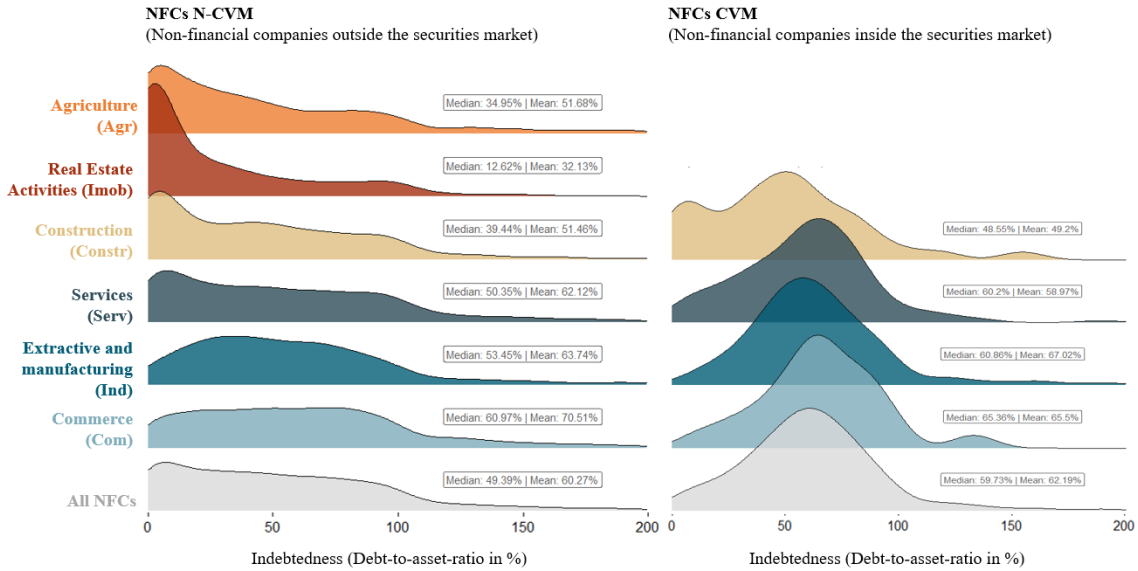
Source: Alexandria. Due to the small number of companies, the following groups were not presented: (i) CVM NFCs of Agriculture (CVMAgr); (ii) CVM companies in the Real Estate Activities industry with assets between R$ 20 million and R$ 240 million; (iii) CVM companies with assets of up to R$ 20 million.

To allow an analysis beyond the median, Figure 8 shows the distribution of the frequency of indebtedness in 2021 according to the groups. Especially among non-CVM companies, the visual analysis shows great differences in the distributions and does not suggest normal distribution. In other words, CVM companies present a more concentrated distribution around the median, non-CVM companies present more disperse indebtedness

[39] Section 5.3 presents a list of variables considered significant by the literature to explain corporate debt.

(greater variance). In accounting and statistical terms, this means that the visual analysis does not suggest a single or dominant pattern of capital structure.

Figure 8- Non-financial companies (NFCs): debt frequency distribution, by industry and presence in the securities market (2021)



Source: Alexandria. The horizontal axes have been cropped at 200% for ease of viewing.

Next, several normality tests and two variance homogeneity tests were performed (steps C and D of Figure 2). The results (detailed in Table 6) show the null hypothesis of normality should be rejected, but not that of homogeneity of variances[40].

[40] The results of Table 6 show that Levene's and Barlett's tests are discrepant. Levene's test is a relatively insensitive procedure to deviations from normality, which makes it a robust test, as in the absence of normality its actual size is close to the level of significance set for a wide variety of probability distributions. On the other hand, Bartlett's test is sensitive in relation to the hypothesis of data normality, being robust if the variables have an approximately normal distribution (ALMEIDA, ELIAN e NOBRE, 2008). Thus, we prefer the conclusion of Levene's test: we cannot reject the hypothesis of homogeneity of variances.

Table 6- Debt-to-asset-ratio of non-financial companies (NFCs), except CVMAgr and CVMImob: Tests of normality and homogeneity of variances

	Tests of normality (p-value) [Is distribution normal?]						Tests homogeneity variances (p-value) [Is variance homogeneous?]	
	Anderson-Darling	Lilliefors	Shapiro-Wilk	Cramer-von Mises	D´Agostino-Pearson	Jarque Bera	Levene	Bartlett
2015	7208.89 (0.00) [No]	0.5021 (0.00) [No]	0.01 (0.00) [No]	1555.28 (0.00) [No]	242.31 (0.00) [No]	41.18×10^9 (0.00) [No]	0.61 (0.77) [Yes]	49943.21 (0.00) [No]
2016	4223.61 (0.00) [No]	0.5025 (0.00) [No]	0.00 (0.00) [No]	911.13 (0.00) [No]	190.19 (0.00) [No]	39.68×10^9 (0.00) [No]	0.14 (1.00) [Yes]	52977.04 (0.00) [No]
2017	4102.42 (0.00) [No]	0.5019 (0.00) [No]	0.00 (0.00) [No]	885.01 (0.00) [No]	190.30 (0.00) [No]	49.67×10^9 (0.00) [No]	0.23 (0.99) [Yes]	98539.02 (0.00) [No]
2018	4400.02 (0.00) [No]	0.5015 (0.00) [No]	0.00 (0.00) [No]	949.24 (0.00) [No]	198.32 (0.00) [No]	61.23×10^9 (0.00) [No]	0.22 (0.99) [Yes]	103247.79 (0.00) [No]
2019	6876.52 (0.00) [No]	0.4998 (0.00) [No]	0.00 (0.00) [No]	1483.49 (0.00) [No]	254.98 (0.00) [No]	193.98×10^9 (0.00) [No]	0.19 (0.99) [Yes]	77999.69 (0.00) [No]
2020	8105.23 (0.00) [No]	0.5040 (0.00) [No]	0.01 (0.00) [No]	1748.58 (0.00) [No]	248.54 (0.00) [No]	23.99×10^9 (0.00) [No]	0.42 (0.91) [Yes]	63283.92 (0.00) [No]
2021	7178.51 (0.00) [No]	0.5013 (0.00) [No]	0.00 (0.00) [No]	1548.74 (0.00) [No]	237.61 (0.00) [No]	29.42×10^9 (0.00) [No]	0.79 (0.61) [Yes]	102640.20 (0.00) [No]
Concl.	{No}	{No}	{No}	{No}	{No}	{No}	{Yes}	{No}

Sources: Ghasemi and Zahediasl (2012) and authors' calculations. Notes: (i) For the Shapiro-Wilk test, when n > 5,000, a random sample was extracted due to the test limitation for samples between 3 and 5,000 observations. (ii) The robust Levene test is the Brown-Forsythe F test. (iii) [Yes] indicates we do not reject the null hypothesis of normal distribution or homogeneity of variances, with 95% confidence; [No] indicates we reject the same hypothesis. (iv) The last line is the "conclusion" line, which presents the applicable conclusion for most years analyzed, that is, whether in most years the null hypothesis was accepted {Yes} or rejected {No}; (v) for the variance homogeneity tests we used the ten groups, which resulted in 9 degrees of freedom; (vi) we removed the CVMAgr and CVMImob groups given the reduced sample size, which would bring a low power of the tests performed.

The rejection of the hypothesis of normality suggests the use of non-parametric tests, which are presented in Section 4.2. However, as this suggestion is not consensual in the literature, we also performed the parametric tests, which are presented in Section 4.3.

4.2 Comparison of indebtedness medians using non-parametric tests

As it was not possible to conclude that the variances of the indebtedness data were normal, and as the groups had different sizes, the Kruskal-Wallis non-parametric test for independent samples was adopted (step E.1 of Figure 2), following Marinšek (2015, p. 42). The test, carried out with the 11 groups of NFCs (all except CVMAgr), led to the conclusion that there was at least one group with a different distribution from the others,

in all the years analyzed[41]. For this reason, Dunn's post hoc analysis was performed for each of the 55 pairs formed by the 11 groups, with Bonferroni's significance[42] correction to verify any differences between the medians (step F of Figure 2). It is noteworthy that this procedure is sensitive to differences in sample distribution.

To generate more robust conclusions, the non-parametric Mann-Whitney test was also performed for independent and paired samples (step E.2 of Figure 2). In each of the 55 pairs, the smallest group size determines the random sample size to be created for the largest group. For example, if the smaller group size is n = 100, 1000 random samples from the larger group of the pair were created, all of size n = 100, thus allowing 1000 paired comparisons. As a result, it is possible to analyze in how many comparisons the null hypothesis (medians are equal) should be accepted.

Figure 9 shows, for each year and for each pair of groups, the result of Dunn's test, together with its conclusion in square brackets: "=" or "≠". When the Mann-Whitney test ("MH-MC") disagreed with the conclusion of the Dunn test, a sign ("!") was added. For example, in the first column (2015), line "CVMCom-CVMConstr", we compare the median of the debt ratio of the 28 CVM commerce companies and the 30 CVM construction companies. The cell displays the value "0.71 [=]", which should be interpreted as follows:

- Dunn's test did not reject the null hypothesis that the medians are equal (hence the "=" sign).
- the Mann-Whitney test was performed, that is, 1,000 comparisons were made between the 28 commerce CVM companies and a random sample of 28 CVM construction companies. In more than 50% of the cases, the hypothesis of median equality was accepted.

[41] Some authors state that the Kruskal-Wallis test compares the median of the groups, but without homoscedasticity of variances, it is more correct to state that the null hypothesis of this test is stochastic homogeneity (VARGHA and DELANEY, 1998, p. 175). If we adopt the conclusion of the Levene test presented in Table 6 (that the variance is homogeneous between the groups), we can adopt the simplest explanation: the Kruskal-Wallis test showed at least one of the nine groups of NFCs has a median different from the others.

[42] The Bonferroni correction, in its simplest form, means that instead of using a critical value of 5% for each test, the value of 5% divided by the number of groups tested is used. One effect is that the Bonferroni correction inflates the type II error, that is, we fail to identify differences that may exist.

- therefore, in this case, both the Dunn test and the Mann-Whitney test resulted in the same conclusion (indebtedness medians of CVMCom and CVM Constr are equal). Therefore, the cell is colored green.

Figure 9- Indebtedness ratio of non-financial companies (NFCs): comparison of the median of pairs using the Dunn test with Bonferroni correction and the Mann-Whitney test.

Color legend: results suggest that groups have medium indebtedness ■ equal ■ different ■ could not define (tests diverge)

Groups		Pair	2015	2016	2017	2018	2019	2020	2021	Conclusion
CVM companies only		CVMCom - CVMConstr	0.71 [=]	0.79 [=]	0.11 [=]	-0.08 [=]	1.07 [=]	1.46 [=]	1.64 [=]	{ = }
		CVMCom - CVMInd	-0.2 [=]	0.43 [=]	0.14 [=]	0.28 [=]	0.78 [=]	0.44 [=]	0.05 [=]	{ = }
		CVMCom - CVMServ	0.1 [=]	0.45 [=]	0.37 [=]	0.46 [=]	0.94 [=]	0.95 [=]	1.02 [=]	{ = }
		CVMConstr - CVMInd	-1.14 [=]	-0.62 [=]	-0.01 [=]	0.37 [=]	-0.61 [=]	-1.41 [=]	-2.03 [=]	{ = }
		CVMConstr - CVMServ	-0.8 [=]	-0.55 [=]	0.23 [=]	0.55 [=]	-0.39 [=]	-0.86 [=]	-1.03 [=]	{ = }
		CVMInd - CVMServ	0.52 [=]	0.07 [=]	0.39 [=]	0.31 [=]	0.31 [=]	0.81 [=]	1.54 [=]	{ = }
		CVMCom - CVMImob	0.41 [=]	0.81 [=]	1.05 [=]	0.82 [=]	1.04 [=]	0.55 [=]	0.68 [=]	{ = }
		CVMConstr - CVMImob	-0.12 [=]	0.23 [=]	0.97 [=]	0.87 [=]	0.27 [=]	-0.49 [=]	-0.51 [=]	{ = }
		CVMInd - CVMImob	0.6 [=]	0.65 [=]	1.1 [=]	0.75 [=]	0.68 [=]	0.35 [=]	0.73 [=]	{ = }
		CVMServ - CVMImob	0.39 [=]	0.61 [=]	0.93 [=]	0.61 [=]	0.55 [=]	0.02 [=]	0.11 [=]	{ = }
CVM X N-CVM	Same industry	CVMCom - NCVMCom	-0.29 [=]	0.76 [=]	0.59 [=]	0.6 [=]	1.33 [=]	0.61 [=]	0.46 [=]	{ = }
		CVMConstr - NCVMConstr	1.81 [=]	2.62 [=]	3.83** [≠]	4.17** [≠]	3.33 [=] (!)	1.82 [=]	1.77 [=]	{ = }
		CVMInd - NCVMInd	1.53 [=]	3.44* [≠]	3.41* [≠]	3.57* [≠]	4.42** [≠]	2.22 [=] (!)	4.4** [≠]	{ ≠ }
		CVMServ - NCVMServ	1.42 [=]	1.48 [=]	1.3 [=]	1.72 [=]	2.45 [=]	1.03 [=]	0.54 [=]	{ = }
		CVMImob - NCVMImob	2.57 [=] (!)	2.63 [=] (!)	2.73 [=]	3.15 [=] (!)	3.57* [≠]	3.22 [=] (!)	3.24 [=] (!)	{ = }
	Different industries	CVMCom - NCVMAgr	2.34 [=]	3.69* [≠]	3.34 [=] (!)	3.63* [≠]	4.79** [≠]	3.73** [≠]	3.6* [≠]	{ ≠ }
		CVMCom - NCVMConstr	2.72 [=] (!)	3.51* [≠]	3.89** [≠]	4.12** [≠]	4.81** [≠]	4** [≠]	4** [≠]	{ ≠ }
		CVMCom - NCVMInd	0.4 [=]	1.84 [=]	1.67 [=]	1.97 [=]	2.92 [=] (!)	1.56 [=]	2.16 [=]	{ = }
		CVMCom - NCVMServ	0.82 [=]	1.24 [=]	1.11 [=]	1.46 [=]	2.46 [=]	1.68 [=]	1.45 [=]	{ = }
		CVMCom - NCVMImob	4.83** [≠]	5.38** [≠]	6.11** [≠]	6.63** [≠]	7.65** [≠]	6.74** [≠]	6.29** [≠]	{ ≠ }
		CVMConstr - NCVMAgr	1.43 [=]	2.84 [=] (!)	3.27 [=] (!)	3.69** [≠]	3.39* [≠]	1.65 [=]	1.56 [=]	{ = }
		CVMConstr - NCVMCom	-1.31 [=]	-0.33 [=]	0.45 [=]	0.69 [=]	-0.16 [=]	-1.41 [=]	-1.77 [=]	{ = }
		CVMConstr - NCVMInd	-0.59 [=]	0.82 [=]	1.55 [=]	2.05 [=]	1.39 [=]	-0.52 [=]	-0.17 [=]	{ = }
		CVMConstr - NCVMServ	-0.16 [=]	0.17 [=]	0.98 [=]	1.55 [=]	0.94 [=]	-0.42 [=]	-0.85 [=]	{ = }
		CVMConstr - NCVMImob	3.99** [≠]	4.63** [≠]	6.11** [≠]	6.66** [≠]	6.19** [≠]	4.4** [≠]	4.1** [≠]	{ ≠ }
		CVMInd - NCVMAgr	5.65** [≠]	6.29** [≠]	5.7** [≠]	5.68** [≠]	6.64** [≠]	5.82** [≠]	5.36** [≠]	{ ≠ }
		CVMInd - NCVMCom	-0.18 [=]	0.74 [=]	0.95 [=]	0.61 [=]	1 [=]	0.27 [=]	0.8 [=]	{ = }
		CVMInd - NCVMConstr	7.04** [≠]	6.64** [≠]	7.4** [≠]	7.21** [≠]	7.35** [≠]	6.9** [≠]	6.84** [≠]	{ ≠ }
		CVMInd - NCVMServ	2.58 [=]	1.93 [=]	2.14 [=]	2.46 [=]	3.38* [≠]	2.46 [=]	2.86 [=] (!)	{ = }
		CVMInd - NCVMImob	12.15** [≠]	9.81** [≠]	11.16** [≠]	11.31** [≠]	11.78** [≠]	12.46** [≠]	9.91** [≠]	{ ≠ }
		CVMServ - NCVMAgr	4.15** [≠]	5.5** [≠]	4.73** [≠]	4.9** [≠]	5.72** [≠]	4.28** [≠]	3.71** [≠]	{ ≠ }
		CVMServ - NCVMCom	-0.8 [=]	0.52 [=]	0.33 [=]	0.15 [=]	0.48 [=]	-0.8 [=]	-1.14 [=]	{ = }
		CVMServ - NCVMConstr	5.1** [≠]	5.58** [≠]	6** [≠]	6.07** [≠]	6.12** [≠]	4.95** [≠]	4.6** [≠]	{ ≠ }
		CVMServ - NCVMInd	0.58 [=]	2.7 [=] (!)	2.34 [=]	2.65 [=] (!)	3.29 [=] (!)	0.83 [=]	1.82 [=]	{ = }
		CVMServ - NCVMImob	9.24** [≠]	8.54** [≠]	9.47** [≠]	9.89** [≠]	10.25** [≠]	9.65** [≠]	7.76** [≠]	{ ≠ }
		CVMImob - NCVMAgr	1.02 [=]	1.52 [=]	0.98 [=]	1.35 [=]	1.85 [=]	1.57 [=]	1.6 [=]	{ = }
		CVMImob - NCVMCom	-0.66 [=]	-0.46 [=]	-0.86 [=]	-0.59 [=]	-0.4 [=]	-0.28 [=]	-0.5 [=]	{ = }
		CVMImob - NCVMConstr	1.24 [=]	1.34 [=]	1.26 [=]	1.56 [=]	1.74 [=]	1.67 [=]	1.7 [=]	{ = }
		CVMImob - NCVMInd	-0.22 [=]	0.22 [=]	-0.2 [=]	0.21 [=]	0.52 [=]	0.25 [=]	0.49 [=]	{ = }
		CVMImob - NCVMServ	0.04 [=]	-0.16 [=]	-0.54 [=]	-0.08 [=]	0.25 [=]	0.32 [=]	0.07 [=]	{ = }
N-CVM companies only		NCVMAgr - NCVMCom	-10.6** [≠]	-7.34** [≠]	-6.36** [≠]	-6.58** [≠]	-7.24** [≠]	-8.66** [≠]	-5.57** [≠]	{ ≠ }
		NCVMAgr - NCVMConstr	1.24 [=]	-0.67 [=]	0.78 [=]	0.56 [=]	-0.41 [=]	0.31 [=]	0.1 [=]	{ = }
		NCVMAgr - NCVMInd	-7.97** [≠]	-5.12** [≠]	-4.35** [≠]	-4.16** [≠]	-4.62** [≠]	-6.32** [≠]	-3.18 [=] (!)	{ ≠ }
		NCVMAgr - NCVMServ	-6.38** [≠]	-6.52** [≠]	-5.51** [≠]	-5.12** [≠]	-5.42** [≠]	-6.09** [≠]	-4.23** [≠]	{ ≠ }
		NCVMAgr - NCVMImob	9.09** [≠]	3.26 [=] (!)	5.01** [≠]	5.06** [≠]	4.49** [≠]	7.09** [≠]	3.47* [≠]	{ ≠ }
		NCVMCom - NCVMConstr	18.71** [≠]	8.72** [≠]	9.53** [≠]	9.51** [≠]	8.83** [≠]	13.55** [≠]	7.89** [≠]	{ ≠ }
		NCVMCom - NCVMInd	5.88** [≠]	5.21** [≠]	4.67** [≠]	5.5** [≠]	6.01** [≠]	5.4** [≠]	6** [≠]	{ ≠ }
		NCVMCom - NCVMServ	10.05** [≠]	2.2 [=]	2.14 [=]	3.29 [=] (!)	4.02** [≠]	6.34** [≠]	3.29 [=] (!)	{ ≠ }
		NCVMCom - NCVMImob	32.64** [≠]	12.13** [≠]	14.01** [≠]	14.38** [≠]	13.94** [≠]	26.78** [≠]	11.12** [≠]	{ ≠ }
		NCVMConstr - NCVMInd	-14.91** [≠]	-5.98** [≠]	-7.27** [≠]	-6.68** [≠]	-5.63** [≠]	-10.32** [≠]	-4.76** [≠]	{ ≠ }
		NCVMConstr - NCVMServ	-12.68** [≠]	-7.85** [≠]	-8.74** [≠]	-7.88** [≠]	-6.65** [≠]	-10.12** [≠]	-6.21** [≠]	{ ≠ }
		NCVMConstr - NCVMImob	11.28** [≠]	4.53** [≠]	5.1** [≠]	5.42** [≠]	5.7** [≠]	9.58** [≠]	4.07** [≠]	{ ≠ }
		NCVMInd - NCVMServ	4.07** [≠]	-3.63* [≠]	-2.99 [=] (!)	-2.51 [=] (!)	-2.17 [=] (!)	0.75 [=]	-3.06 [=] (!)	{ - }
		NCVMInd - NCVMImob	29.24** [≠]	10.07** [≠]	12.34** [≠]	12.26** [≠]	11.57** [≠]	24.05** [≠]	8.71** [≠]	{ ≠ }
		NCVMServ - NCVMImob	27.52** [≠]	11.5** [≠]	13.51** [≠]	13.19** [≠]	12.32** [≠]	24.3** [≠]	9.83** [≠]	{ ≠ }

*Sources: Alexandria and authors' calculations. Notes: (i) the figure shows the Z statistic values of Dunn's post hoc test with Bonferroni adjustment for each sample pair; (ii) * 5% significance; ** 1% significance. (iii) [=] indicates that we cannot reject the null hypothesis that the median indebtedness of the two groups, in that year, is the same. [≠] indicates that we reject, at a significance level of 5%, the same null hypothesis; (v) (!) indicates that most of the Mann-Whitney simulations diverged from the Dunn test conclusion; (vi) {=} indicates that in most years analyzed the null hypothesis is not rejected; (vii) {≠} indicates that in most of the analyzed years the null hypothesis is rejected.*

The results of Figure 9 in relation to the differences between the medians of the debt ratios of the NFCs suggest that, within the same year:

- There is no difference between the median indebtedness of the different industries of companies registered with the CVM (that is why the first lines of the "only CVM companies" are in green).
- The presence in the securities market affects the level of indebtedness of companies in the extractive and manufacturing industry (because the median debt of CVM companies in these industries is different from non-CVM ones – see group "same industry" in the Figure, column "Conclusion")
    - For the other industries, in contrast, presence in the securities market did not seem to affect the median level of indebtedness.
- Among companies outside the securities market (see the lines "only non CVM companies"), the difference by industry is very clear: each industry has a different level of indebtedness than all the others, with the exception of agriculture and construction, whose median was equal in statistical terms, and the extraction and manufacturing and services industries, for which it was not possible to reach a conclusion because the tests diverge.
- The reader should note the apparent contradiction between the conclusions: the median indebtedness of CVM industries of services and commerce are equal to each other, and these are respectively equal to the Non-CVM industries of services and commerce. However, these last two are different from each other. This apparent contradiction (which also happen in other industries) can occur in the Dunn and Mann Whitney tests and must be evaluated using other approaches, as we will do later.

4.3 Comparison of debt-to-asset ratio averages using parametric test (econometric model)

The rationale "if the distribution is not normal, perform a non-parametric test", which led to the tests performed in Section 4.2, is not consensual in the literature. Fagerland (2012) warns that the Wilcoxon-Mann-Whitney tests (including Kruskal-Wallis) produce p-values lower than the t-test, on average, which can alter conclusions. Therefore, the author suggests that, for large samples, the t test should be used, even if

the distribution is strongly asymmetric. Vargha and Delaney (1998, p. 181) state that, when the variance between groups differs substantially, the Kruskal-Wallis test may prove inadequate, especially in the case of groups with different sizes (the case of the groups analyzed in Section 4.2). Finally, Ghasemi and Zahediasl (2012) mention that, for sufficiently large samples, the violation of the hypothesis of normality does not cause major problems and parametric tests can be used, by virtue of the Central Limit Theorem (step F of Figure 2).

Following this literature, the authors also compared the average debt-to-asset ratio of each group using panel econometric regressions.

Although the objective of this work is to test only the significance of industry and presence in the securities market to explain corporate indebtedness, the model was built with other variables, recommended in the literature (see section 2), for robustness. Table 7 summarizes the variables tested.

Regarding explanatory variables, it was not possible to test all variables commonly used in studies. In particular, the impact of the variables (i) tangibility of assets[43], (ii) tax deductibility[44] and (iii) ratio between market value and book value (Market-to-Book ratio)[45], often cited as relevant variables, was not tested, for lack of the necessary line items in Alexandria. Therefore, our study possibly undergoes the omitted variable problem, and it is possible that the industry variable captures part of the tangibility effect.

At the theoretical level, we believe that the variable "tax deductibility" is not relevant in Brazil, as only companies in the "real profit" taxation system can deduct interest payments from income tax. These companies represent only 16% of the companies that produce balance sheets (Sped, 2023), but include 100% of CVM companies.

[43] Examples of studies that include unlisted companies and that conclude that the variable "tangibility of assets" is relevant to explain indebtedness: Marinšek (2015) and Den Berg (2021).
[44] Den Berg (2021) conducts a study that includes unlisted companies and concludes that the variable "tax deductibility" is relevant to explain indebtedness. Demirgüç-Kunt et al. (2020) reach the same conclusion for Portugal and Poland.
[45] Jaworski and Santos (2021, p. 40) used data from unlisted companies, but estimated the variable "market/book ratio" from cash flow.

Table 7- Variables tested in the econometric models of this work

Variables tested	**Some previous studies that used/tested the same variable (a) or a transformation of it.**
Dependent variable	
$\boldsymbol{LOG_{DA}} = log[D/A + 1]$, namely D/A= $= \left(\frac{Short - Term\ Debt + Long - Term\ Debt}{Total\ assets}\right) * 100$ *The variable was Winsorized*[46]	Raalte (2021, p. 26); Nehrebecka e Białek-Jaworska (2015); Den Berg (2021); Forte, Barros and Nakamura (2013)
Independent variables	
Indebtedness from the previous year (dependent variable lag): **LAG_LOG_D/A**	Nehrebecka and Białek-Jaworska (2015); Correa et al.(2013); Forte, Barros and Nakamura (2013)
Company size: **LOG_Assets**=log(Total assets)	Marinšek (2015, p. 37); Raalte (2021, p. 25); Forte, Barros and Nakamura (2013); Andritzky (2003); Pohlman and Iudícibus (2010) [b]
Profitability in the base year: **ROA**= $\left(\frac{Net\ income}{Total\ assets}\right) * 100$ *The variable was winsorized.*	Marinšek (2015, p. 37); Raalte (2021, p. 26); Forte, Barros and Nakamura (2013); Correa et al.(2013); Andritzky (2003) [c]
Profitability in the base year: **ROE**= $\left(\frac{Net\ income}{Net\ equity}\right) * 100$	Forte, Barros and Nakamura (2013)
Age of the company at the base date, in years: **LOG_AGE**=log(age of the company)	Jaworski and Santos (2021, p. 40); Forte, Barros and Nakamura (2013); Den Berg (2021)
GROWTH: $\left(\frac{Total\ assets[t] - Total\ Assets[t-1]}{Total\ Assets[t-1]}\right) * 100$	Marinšek (2015, p. 37); Raalte (2021, p. 30); Forte, Barros and Nakamura (2013)
STATE-OWNED: dummy that assumes the value 1 if the company is owned, in some proportion, by the Brazilian Government in 2022.	Marinšek (2015, p. 40); Nehrebecka and Białek-Jaworska (2015); Azevedo (2013), Pohlman and Iudícibus (2010)
PART_NRES: share of non-residents in companies' net equity on the reference date. It varies from 0 to 1.	Nehrebecka and Białek-Jaworska (2015); Azevedo (2013), Pohlman and Iudícibus (2010); Correa et al.(2013)
LISTED: dummy that takes the value 1 if the company is listed on the Brazilian stock exchange (B3) at the end of the base year and 0 otherwise.	Raalte (2021, p. 25); Den Berg (2021); Azevedo (2013); Demirgüç-Kunt et al.(2020) and Marinšek (2015, p. 39)
CVM: dummy that takes the value 1 if the entity participates in the Brazilian securities market and 0 otherwise. Includes all listed.	We did not find literature with this variable.
FAMILY: dummy that takes the value 1 if the company is a limited liability company owned by two or more people with the same last name.	Gottardo e Maria Moisello (2014)

[46] The winsorization maintained many NFCs with indebtedness superior to 100% in the sample, contrary to what many previous works have done (section 2.5.1). This decision was made by the authors after manual analysis of hundreds of balance sheets published in newspapers, which allowed for the verification that there are a considerable number of companies that operate at levels higher than 100% of indebtedness, including for many years. In other words, in many cases this is not a statement error.

WOMAN_MAN: dummy that takes the value 1 if at least one of the managers was a woman.	García e Herrero (2021) [d]
INDUSTRY: dummies that assume the value 1 if the company belongs to the industry and 0 otherwise. The industries considered were the six industries described in Table 3	Marinšek (2015, p. 41); Raalte (2021, p. 25); Forte, Barros and Nakamura (2013); Azevedo (2013); Andritzky (2003)
YEAR: dummies that take the value 1 if the data belong to the year in question and 0 otherwise.	Den Berg (2021); Nehrebecka and Białek-Jaworska (2015); Raalte (2021); Forte, Barros and Nakamura (2013)
CAPITAL: dummy that assumes the value 1 if the company's headquarters is located in the capital of some federation unit and 0 otherwise.	We did not find literature with this variable.

Source: prepared by the authors. [a] Only studies whose sample included unlisted companies were mentioned; [b] Correa et al. (2013), Azevedo (2013) and Den Berg (2021) use the natural logarithm of net revenue as a measure of company size, and Demirgüç-Kunt et al. (2020) used a dummy (small company or not). As Alexandria does not contain the net revenue of all companies, we chose to use the natural logarithm of total assets, including the possibility of bias mentioned in section 2.5.2. The mentioned studies used the same metric. [c] some cited authors measure profitability as operating profit/assets, that is, EBIT/assets. As operating profit is not available in Alexandria, we used ROA.[d] This is the only work cited in the table that does not use unlisted companies, as we did not find any work that investigated the variable "gender" using a sample that included unlisted companies.

The descriptive statistics of the variables used are in Annex D. The variables were stationary and the correlogram (Annex E) helped to select the best variables.

The correlogram also showed the strong persistence of indebtedness, seen by the high correlation of indebtedness and its lagged value – a result that is largely in line with the national and international literature (see Table 7 and Correa et al., 2013). This fact and the likely problem of omitted variables suggest endogeneity (see Section 2.5.3) and led the authors to discard POLS-types regressions, fixed effects and random effects.

The econometric method chosen, due to its robustness and ability to deal with the problem of endogeneity, was the level GMM regression (GMM-Sys), in line with what was suggested by Barros et al. (2020). The final equation was:

$$\ln\left(\frac{D_t}{A_t}+1\right) = \ln\left(\frac{Liabilities_t}{Assets_t}+1\right)$$
$$= \beta_0 + \beta_1\left(\frac{Liabilities_{t-1}}{Assets_{t-1}}+1\right) + \beta_{2-5} industry\ dummies + \beta_6 \quad (7)$$
$$* presence\ in\ CVM + \beta_7\ state-owned$$
$$+ \beta_7\ limited\ liability\ family\ company + \beta_8 \ln(Assets_t)$$
$$+ \beta_9 PartNres + \beta_{10} * d\_capital + \beta_{11}\ woman\_man + \beta_{12} \ln(age)$$

The regression results are detailed in Table 8.

Table 8- Estimation of indebtedness models using the Sys-GMM method in balanced panel and unbalanced panel of non-financial companies (NFCs) that participate (CVM) or not (N-CVM) in the Brazilian securities market with annual data from 2015 to 2021.

	Unbalanced panel			**Balanced panel**		
	(A) NFCs	**(B) CVM**	**(C) N-CVM**	**(D) NFCs**	**(E) CVM**	**(F) N-CVM**
lag(log D/A, 1)	**0.7918*****	**0.8203*****	**0.8665*****	**0.9187*****	**0.8416*****	**0.9065*****
	(0.0359)	(0.0583)	(0.0385)	(0.0211)	(0.0450)	(0.0239)
Services	**0.0038**	**-0.0531****	**0.0073*****	**0.0019**	**-0.0111**	**0.0050**
	(0.0026)	(0.0212)	(0.0025)	(0.0037)	(0.0106)	(0.0046)
Commerce	**0.0048**	**-0.0008**	**0.0043**	**0.0051***	**-0.0093**	**0.0074****
	(0.0032)	(0.0102)	(0.0030)	(0.0028)	(0.0060)	(0.0034)
Construction	**-0.0137****	**-0.0821*****	**-0.0078**	**-0.0041**	**-0.0146**	**-0.0048**
	(0.0056)	(0.0210)	(0.0051)	(0.0043)	(0.0142)	(0.0048)
Agriculture	**-0.0132***		**-0.0082**	**-0.0066**		**-0.0079**
	(0.0076)		(0.0055)	(0.0046)		(0.0051)
Real Estate Activities	**-0.0401*****	**-0.1648*****	**-0.0270*****	**-0.0114***	**-0.0328**	**-0.0151****
	(0.0082)	(0.0406)	(0.0073)	(0.0059)	(0.0240)	(0.0067)
ROA	**-0.1471*****	**-0.1056***	**-0.1319*****	**-0.1905*****	**-0.2384*****	**-0.2402*****
	(0.0132)	(0.0554)	(0.0334)	(0.0388)	(0.0449)	(0.0395)
Data in CVM	**0.0223*****			**-0.0073**		
	(0.0057)			(0.0079)		
State-owned	**0.0069**	**0.0735*****	**0.0027**	**-0.0138*****	**-0.0091**	**-0.0184*****
	(0.0065)	(0.0113)	(0.0059)	(0.0044)	(0.0105)	(0.0055)
Family	**0.0021**		**0.0064**	**0.0032**		**0.0048**
	(0.0058)		(0.0052)	(0.0048)		(0.0054)
log(asset)	**-0.0012**	**-0.0822*****	**0.0007**	**0.0058**	**-0.0014**	**0.0093***
	(0.0009)	(0.0151)	(0.0014)	(0.0040)	(0.0096)	(0.0049)
ParticNres	**0.0020**	**0.0895****	**-0.0032**	**-0.0088***	**-0.0045**	**-0.0137****
	(0.0031)	(0.0368)	(0.0033)	(0.0050)	(0.0242)	(0.0061)
Headquarters in the State Capital	**-0.0046****	**0.0809*****	**-0.0058*****	**-0.0040**	**0.0103**	**-0.0068****
	(0.0021)	(0.0164)	(0.0022)	(0.0026)	(0.0100)	(0.0029)
Woman in management	**-0.0023**	**0.0062**	**-0.0004**	**-0.0012**	**0.0010**	**-0.0011**
	(0.0021)	(0.0051)	(0.0020)	(0.0012)	(0.0046)	(0.0014)
log(age)	**0.0017**	**0.0015**	**0.0003**	**0.0035*****	**-0.0018**	**0.0028****
	(0.0015)	(0.0088)	(0.0014)	(0.0010)	(0.0053)	(0.0011)
nobs	53,891	4,162	49,307	20,097	2,151	17,208
Instruments	E [2:99]	E, R, A [2:99]	E, R [2:99]	E, R, A [3:99]	E, A [3:99]	E, R, A [3:99]
Hansen-Sargan	22.9208	31.3635	33.4197	36.0678	25.1214	36.0450
[p-value]	*[0.2408]*	*[0.1436]*	*[0.0562]*	*[0.0541]*	*[0.1216]*	*[0.0408]*
AR1	-11.58	-2.87	-11.56	-14. 69	-4.60	-13.34
[p-value]	*[0.0000]*	*[0.0040]*	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*
AR2	-1.58	1.23	-1.75	1.45	1.50	1.69
[p-value]	*[0.1149]*	*[0.2188]*	*[0.0798]*	*[0.1460]*	*[0.1349]*	*[0.0914]*
Wald Coef {GL}	6,610.94 {15}	361.40 {12}	7,702.21 {14}	7,327.56 {15}	1,170.76 {12}	5,676.27 {14}
[p-value]	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*
Wald Time {GL}	17.07 {5}	10.75 {5}	27.74 {4}	28.57 {4}	15.03 {4}	15.85 {4}
[p-value]	*[0.0044]*	*[0.0565]*	*[0.0000]*	*[0.0000]*	*[0.0046]*	*[0.0032]*

*Notes: (i) Analysis of the sample's annual variables; (ii) Errors robust to heteroscedasticity in parentheses using firm-level error clustering; (iii) * 10% significance; ** 5% significance; *** 1% significance; (iv) AR(1) and AR(2) verify the presence of serial correlation of first order and second order in the residuals of first difference (v) in B and E regressions, companies in the CVM agro group were disregarded because they represented a small sample (between 3 and 4 companies, depending on the year).; (vi) We do not show the time and intercept dummies for brevity. (vii) in regressions with an unbalanced panel, the number of CVM companies, added to the number of non-CVM companies, is lower than the total because there are companies that were CVM companies for a period and then ceased to be CVM companies, or vice versa, and therefore could not be included in the last two balanced panels; (viii) the winsorization technique was used at a 1% level for the debt-to-asset ratio and ROA variables; (ix) variables in lag used as instrumen–s: E - ln(debt+1) winsorized; R - ROA winsorized; A - ln(asset); values between [ ]´s are the lags of the variables used as instruments.*

Table 9 summarizes the main conclusions to be made from the regressions. In sum, (i) industry affiliation is a significant variable to explain indebtedness in all regressions with unbalanced panel. In balanced panels, industries only differed in indebtedness among non-CVM companies; and (ii) with all else constant, the presence in the securities market increases the level of indebtedness of NFCs in the unbalanced panel, but not in the balanced panel[47]

Table 9– Main conclusions from the regressions in Table 8 (considering 5% significance coefficient)

	Unbalanced panel			**Balanced panel**		
	(A) NFCs	**(B) CVM**	**(C)N-CVM**	**(D) NFCs**	**(E) CVM**	**(F) N-CVM**
Does the industry impact mean indebtedness?	Yes	Yes	Yes	No(*)	No	Yes
If yes, which are the most indebted industries? (**)	Extraction and manufacturing, Commerce, Services	Extraction and manufacturing, Commerce	Services	---	---	Commerce
If yes, which is the least indebted industry?	Real Estate Activities	Real Estate Activities	Real Estate Activities	---	---	Real Estate Activities
Does presence in the securities market increase mean indebtedness?	Yes	---	---	No	---	---

Obs: (*) The answer changes to "yes" when the 10% significance coefficient is considered, and in regressions made for robustness in Annex E. (**) When more than one industry is mentioned, the regression did not show statistical difference between the indebtedness of these industries, that is, they were tied as the most indebted.

In order to test the robustness of the models, we estimated alternative models that contemplate a shorter period (2015 to 2019 and 2017 to 2021). We also estimated separate models for each industry. These robustness tests, available respectively in Annex E and F, did not change our main conclusions.

Although it is not the focus of this work to analyze the other variables, it is interesting to mention: first, that the previous year's indebtedness is the most significant variable to explain the NFC's current indebtedness. Second, corporate earnings (measured by ROA[48]) are negatively correlated with debt, a result that is in line with the literature[49].

[47] Other studies will be necessary to verify whether the cost of debt differs between CVM and non-CVM companies.
[48] Estimates made by the authors using ROE showed a less significant effect.
[49] One possible interpretation is that, when an enterprise has profits, it prefers to reduce indebtedness. Observe that the visual analysis suggested great heterogeneity (see Figure 5).

Third, all things being equal, the size of the company (measured by assets) does not impact indebtedness. Fourth, variables frequently mentioned as significant in the literature to explain indebtedness, such as the fact of a company being state-owned or its age, proved to be significant, but not in all regressions and sometimes with contradictory signs.

Furthermore, the share of non-residents in net equity did not prove significant to explain indebtedness in the main regressions (A and D) – and, in the regressions in which some impact was detected, the signals were contradictory. This result was no different when we changed the percentage of participation for a dummy, as some works have done. It also did not alter when we changed "percentage of non-residents" for "direct investment company", a subgroup of the first. Because Alexandria contains practically all the companies directly held by non-residents in Brazil and the classification of these as receivers of direct investment or not, this result is especially strong and goes against previous literature (including for the Brazilian case, as in Correa et al, 2013).

Finally, it is important to mention that the authors tested three variables little explored in the Brazilian literature: the presence of women in management, the fact that the company is limited liability and family-owned, and the fact that the company's headquarters are in a state capital. However, only the latter variable showed a significant impact on indebtedness (and with a negative sign, contrary to what we expected).

5. Conclusions

5.1 Effects of the "industry affiliation" and "presence in the securities market" on the indebtedness of Brazilian NFCs

Alexandria is an invaluable database for academic study in the areas of Accounting and Economics in Brazil. The base contains a sample of companies whose importance in the national economy is clear (for example, because they represent about half of all SFN credit taken by NFCs and just under half of all incoming bank flows). Alexandria makes room for debates that were previously impossible and allows verifying whether conclusions using data from listed companies are applicable to unlisted companies inside and outside the securities market.

The analysis of corporate indebtedness, carried out in this work, is an exploratory study that exemplifies the potential of Alexandria.

Most part of analyses carried out suggest the industry affiliation affects indebtedness. However, there were subtle differences in the conclusions of each analysis.

The visual and non-parametric analyses suggest the industries "commerce" and "services" tend to have a higher median indebtedness than the others. At the other end, the "Construction" and "Agriculture" industries appear with relatively low median debt, and "Real Estate Activities" has even lower indebtedness.[50]

The parametric (econometric) analyses show commerce companies tend to have a greater mean indebtedness, but in some regressions, other industries appeared tied with commerce as the most indebted. On the other extreme, the real estate activities industry was shown to have the lowest mean indebtedness.

The conclusion about the impact of the presence in the securities market was very clear in the non-parametric assessment ("presence increases median indebtedness), but not in the parametric assessment ("presence increases mean indebtedness" only in the unbalanced panel).

Both for the analysis of the industry and for the analysis of the presence of securities market, the non-parametric analyses, which have a greater sample, showed much more marked differences. It seems to us, in particular, *that large companies and CVM companies – more concentrated in the balanced panel – are much more similar among themselves, regardless of industry*. The smaller companies are much more heterogeneous, and the industry profile is more striking.

In line with the international literature, the analyses suggest the industry variable captures many variables not observed by the econometrician, such as differences in competitiveness, business cycle and regulation, which are often associated with each industry. It does not seem to us, however, that tangibility is such an important factor among non-CVM companies: one of the industries with less tangible assets, commerce, seems to be one of the most indebted.

It is thus concluded that the studies about company indebtedness using only listed companies, and often using only parametric analyses, must be analyzed taking into

[50] These three industries with lower indebtedness than the others have in common longer business cycles. However, especially for agriculture, this result (low indebtedness) can be surprising given the offer of subsidized credit in this industry. Our sample is large (between 288 observations and 493 observations per year for this industry), but still it is a topic that deserves further studies – in general this industry does not appear in the works we found.

account their limitations, as the results change much according to the sample and technique used. Regardless, most of the analyses made here suggest that: (i) companies that participate in the securities market are slightly more indebted than others; (ii) indebtedness varies significantly among industries; and (iii) companies do not alter much their own level of indebtedness over time, but differ greatly among themselves, even within the same industry. This reduces the explanatory power of analyses that use only the mean, as is the case of some works.

5.2 Study limitations

Evidently, this preliminary study has limitations, to be offset in further studies. First, the sample of NFCs from Alexandria *is not random*, and therefore probably does not adequately represent the population of NFCs producing financial statements.

In particular, the results for the commerce industry must be looked at with care, because despite the large sample, the median of Alexandria's debt was higher than the median of another database analyzed by the authors (Klooks).

Second, *the measurement error of accounting variables* is probably not uniformly distributed in the sample: large CVM and non-CVM companies (more concentrated in the balanced panel) have more accurate accounting and audited financial statements, but this is far from being the rule among medium-sized non-CVM companies. Thus, measurement errors can affect the conclusions. We believe, in particular, that the assets of medium-sized non-CVM companies are, in general, recorded at a value below the real value (since, according to accounting rules, these are often measured according to historical cost). If our assumption is correct, this group *is less indebted than suggested by assets statements*[51], *which could make our conclusion - that presence in the securities market increases indebtedness - stronger*.

Third, the fact that Alexandria does not have earnings information for all non-financial companies impairs the comparison between parametric and non-parametric analyses: while parametric (econometric) analyses used only companies with this data,

[51] This is because indebtedness is calculated from the ratio between values recorded in the balance sheet: $\frac{Recorded\ Liabilities}{Recorded\ Assets}$. If the assets of a non-CVM company are under-evaluated, but not its liabilities, real indebtedness of the company is lower than the accounting data suggest. If, for example, the market value of the asset is equal to 120% of the registered asset, we would have real indebtedness = $\frac{Recorded\ Liabilities}{1.2 * Recorded\ Assets}$ = $\frac{1}{1.2} \cdot \frac{Recorded\ Liabilities}{Recorded\ Assets} \cong 0.83 \cdot$ indebtedness as measured by the Balance Sheet.

non-parametric analyses used a larger sample. It is possible that this limitation explains why indebtedness of agricultural and construction companies appears much lower than the other industries only in the non-parametric analysis: because the sample used in this analysis is larger and includes more medium-sized companies. From this point of view, we understand that the non-parametric analysis is more accurate to explain the differences in industry indebtedness.

Finally, in its 1.0 version, Alexandria does not have some line items considered important by other authors to explain indebtedness, especially the tangibility of assets and tax deductibility.

6. References

AÏHOUNTON, G. and HENNINGSEN, A. Units of Measurement and the Inverse Hyperbolic Sine Transformation, IFRO Working Paper, No. 2019/10, University of Copenhagen, Department of Food and Resource Economics (IFRO), Copenhagen. Available at: https://www.econstor.eu/bitstream/10419/211087/1/IFRO_WP_2019_10.pdf. Accessed on May 11, 2022.

ALMEIDA, A.; ELIAN, S. N.; NOBRE, J. S. Modificações e alternativas aos testes de Levene e de Brown e Forsythe para igualdade de variâncias e médias [ALMEIDA, A.; ELIAN, S.N.; NOBRE, J. S. Modifications and alternatives to the Levene and Brown and Forsythe tests for equality of variances and means]. Revista Colombiana de Estadística, v. 31, p. 241-260, 2008. Available at: http://repositorio.ufc.br/handle/riufc/41465. Accessed on December 22, 2021.

ALMEIDA et al. Substituição da captação dos dados do CAGED pelo eSocial: implicações para as estatísticas de emprego formal [Replacement of CAGED data capture by eSocial: implications for formal employment statistics]. IPEA, 2020. Available at http://repositorio.ipea.gov.br/bitstream/11058/10212/1/BMT_69_substituicao_da_capacitacao.pdf. Accessed on July 16, 2021.

ANDRTZKY, J. Leverage, Industries, and Countries: Evidence from Seven Countries. Nov. 2003. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=469942. Accessed on May 06, 2022.

ANS. Demonstrações Contábeis. 2021. Available at: https://www.gov.br/ans/pt-br/acesso-a-informacao/perfil-do-setor/dados-e-indicadores-do-setor/demonstracoes-contabeis. Accessed on December 16, 2021.

ATKINSON, A. et al. The box-cox transformation: review and extensions. 2021. Statistical Science, 36 (2). 239 - 255. ISSN 0883-4237. Available at: http://eprints.lse.ac.uk/103537/. Accessed on May 18, 2020.

AVELAR, E. A.; SOUZA, A. A. e AMARAL, H. F. Determinantes do endividamento em cooperativas médicas sob a regulação da Agência Nacional de Saúde Suplementar (ANS) [AVELAR, E. A.; SOUZA, A. A. and AMARAL, H. F. Determinants of indebtedness in medical cooperatives regulated by the National Supplementary Health Agency (ANS)]. 2019. In: Revista de Gestão em Sistemas de Saúde. Available at: https://periodicos.uninove.br/revistargss/article/view/11537/7948. Accessed on January 10, 2021.

AZEVEDO, Alexandre Bonfim de. Fatores determinantes do endividamento de empresas por setor no Brasil [Determining factors of corporate indebtedness by industry in Brazil]. 2013. Masters Dissertation. Available at: https://repositorio.insper.edu.br/handle/11224/757. Accessed on January 10, 2022.

BAJGAR, Matej et al. Coverage and representativeness of Orbis data. OECD Science, Technology and Industry Working Papers, 2020/06. Available at: https://www.oecd-ilibrary.org/docserver/c7bdaa03-en.pdf?expires=1614629919&id=id&accname=guest&checksum=CA1AFFDD06D086D931025D8966E3DF10. Accessed on: March 1, 2021.

BARROS, L. et al. Endogeneity in panel data regressions: methodological guidance for corporate finance researchers. In: Rev. Bras. Gest. Neg. São Paulo v.22, Special Issue. 2020 p. 437-461. Available at: https://www.scielo.br/j/rbgn/a/gPYpsvqrBQKDLgLqH8LHNPz/?lang=en&format=pdf. Accessed on June 16, 2022.

BARTLETT, R. and PARTNOY, F. The Ratio Problem. 18, 2020. Available at https://ssrn.com/abstract=3605606 e em https://www.law.nyu.edu/sites/default/files/The%20Ratio%20Problem.pdf. Accessed on May 07, 2022.

Banco Central do Brasil (BCB). Crédito ampliado ao setor não financeiro. 2018. In: Relatório de Economia Bancária. P. 56-61. Available at: https://www.bcb.gov.br/content/publicacoes/relatorioeconomiabancaria/reb_2018.pdf. Accessed on: December 28, 2021.

_____. Relatório de Investimento Direto – 2019 [Direct Investment Report - 2019]. Nov. 2019. 2019b. Available at: https://www.bcb.gov.br/content/publicacoes/relatorioidp/RelatorioID2018/RID_2019.pdf. Accessed on: July 22, 2020.

_____. RDE-IED - Manual do Declarante [RDE-IED - Declarant's Manual]. May 2020. 2020a. Available at: https://www.bcb.gov.br/content/estabilidadefinanceira/rde/manuais_RDE/Manual-RDE-IED.pdf. Accessed on June 01, 2020.

_____. Censo de Capitais Estrangeiros no País - Manual do Declarante 2020 [Brazilian Foreign Capitals Census - Declarant Manual 2020]. 2020b. Available at: https://www.bcb.gov.br/content/estabilidadefinanceira/cambiocapitais/Manual2020.pdf. Accessed on: July 7, 2020.

_____. Sistema de Pagamentos Instantâneos (SPI) [Instant Payment System (SPI)]. 2021a. Available at: https://www.bcb.gov.br/estabilidadefinanceira/sistemapagamentosinstantaneos. Accessed on October 08, 2021.

_____. O que é Pix? [What is Pix?]. 2021b. Available at: https://www.bcb.gov.br/estabilidadefinanceira/pix. Accessed on October 08, 2021.

_____.Relatório de Estabilidade Financeira [_____.Financial Stability Report]. April 2021. 2021c. Available at: https://www.bcb.gov.br/publicacoes/ref/202104. Accessed on: June 30, 2023.

_____. Special Tables - Gross external debt and short-term external debt by residual maturity. 2022. Available at: https://www.bcb.gov.br/content/estatisticas/Documents/Tabelas_especiais/DivExt_T.xls. Accessed on: Mar 26, 2022.

BANCO DE PORTUGAL. Sector tables and company and sector tables. Fev. 2019. In: Estudos da Central de Balanços – 36. [Studies of the Central Balance Sheet – 36]. Available at https://www.bportugal.pt/sites/default/files/flexpapers/2024100/index.html.

_____. Sector analysis of non-financial corporations in Portugal | 2018. 2019b. Estudos da Central de Balanços, v. 40 [Central Balance Sheet Studies, v. 40]. Available at: https://www.bportugal.pt/sites/default/files/anexos/pdf-boletim/estudos_da_cb_40_2019.pdf. Accessed on: Jan 06, 2021.

_____. Central de Balanços. 2021. Available at: https://www.bportugal.pt/page/central-de-balancos?mlid=1045. Accessed on: Oct 21, 2021.

BARROS, L., et al. Endogeneity in panel data regressions: methodological guidance for corporate finance researchers. Rev. Bras. Gest. Neg. (Online) 22 (spe), 2020. Available at: https://doi.org/10.7819/rbgn.v22i0.4059. Accessed on Feb 25, 2022.

BASTOS, D. and NAKAMURA, W. Determinants of the capital structure of publicly-traded companies in Brazil, Mexico and Chile in the period 2001-2006. Revista Contabilidade & Finanças 20 (50), Ago 2009. Available at: https://www.scielo.br/j/rcf/a/3LpqmV393TZF7wTddnT7NBq/?lang=pt&format=html#. Accessed on Feb 10, 2022.

BELLÉGO, C.; BENATIA, D. and PAPE, L-D. Dealing with Logs and Zeros in Regression Models. CREST - Série des Documents de Travail n° 2019-13. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3444996. Accessed on: Feb 25, 2022.

BLUNDELL, Richard; BOND, Stephen. Initial conditions and moment restrictions in dynamic panel data models. Journal of econometrics, v. 87, n. 1, p. 115-143, 1998. Available at: https://www.ucl.ac.uk/~uctp39a/Blundell-Bond-1998.pdf. Accessed on October 07, 2022.

BONOMO, Marco. MARTINS, Betina e PINTO, Rodrigo. Debt composition and exchange rate balance sheet effect in Brazil: a firm level analysis. In: Emerging Markets Review. Volume 4, 4, Dez. 2003, Pp. 368-396. Available at: https://www.sciencedirect.com/science/article/pii/S156601410300061X. Accessed on: Feb 24, 2022.

BRASIL. Law n. 11,638, of December 28, 2007. Amends and revokes provisions of Law No. 6,404, of December 15, 1976, and Law No. 6.385, of December 7, 1976, and extends provisions related to the preparation and disclosure of financial statements to large companies. Available at http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/lei/l11638.htm. Accessed on March 29, 2021.

_____. Law n. 9,656, of June 03, 1998. Provides for private health care plans and insurance. Available at http://www.planalto.gov.br/ccivil_03/leis/l9656.htm. Accessed on Feb 16,2021.

_____. LEI Nº 5.172, DE 25 DE OUTUBRO DE 1966 (Código Tributário Nacional) [LAW No. 5,172, OF OCTOBER 25, 1966 (Brazilian National Tax Code)]. Available at http://www.planalto.gov.br/ccivil_03/leis/l5172compilado.htm. Accessed on March 10, 2021.

_____. Lei Nº 14.286, de 29 de dezembro de 2021 [Law No. 14,286, of December 29th, 2021]. Dispõe sobre o mercado de câmbio brasileiro [...]. Available at https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14286.htm. Accessed on July 1st, 2023.

BRESSAN, Valéria Gama Fully et al. Análise dos determinantes do endividamento das empresas de capital aberto do agronegócio brasileiro [Analysis of the debt determinants of publicly traded companies in Brazilian agribusiness]. Revista de Economia e Sociologia Rural 47(1), Mar. 2009. Available at https://www.scielo.br/j/resr/a/vKR79rd78fjhBqgDrhYwdpr/?lang=pt#.

BRITO, Paulo; SERRANO, André; FRANCO, Víthor. Determinantes da Estrutura de Capital de Empresas Brasileiras de Capital Aberto em Período de Crise [Determinants of the Capital Structure of Publicly Traded Brazilian Companies in a Period of Crisis]. In: Revista Ambiente Contábil, Vol. 10. n. 2, jul./dez. 2018, ISSN 2176-9036. Available at http://www.atena.org.br/revista/ojs-2.2.3-08/index.php/Ambiente/article/view/3347/2587. Accessed on Feb 10, 2022.

CAMPOS, Claudio. Confronto das Teorias de *Pecking Order* e *Trade-off*: Evidências com base nas companhias Brasileiras Abertas [Confrontation of Pecking Order and Trade-off Theories: Evidence based on Brazilian Public Companies]. São Paulo: 2008. Masters Dissertation. Available at https://dspace.mackenzie.br/handle/10899/23547. Accessed on May 21, 2023.

COPAT, Rafael. Fatores específicos da empresa, do setor e do país: quais deles são os direcionadores-chave da estrutura de capital na América Latina?. 2009. Masters Dissertation. Available at https://www.lume.ufrgs.br/bitstream/handle/10183/16145/000697442.pdf. Accessed on April 16, 2023.

CORREA, C. et al. A estrutura de capital das maiores empresas brasileiras: análise empírica das teorias de Pecking Order e trade-off, usando panel data [The capital structure of the largest Brazilian companies: empirical analysis of Pecking Order and trade-off theories, using panel data]. Finanças Estratégicas. RAM, Rev. Adm. Mackenzie 14 (4). Ago 2013. Available at: https://doi.org/10.1590/S1678-69712013000400005. Accessed on May 12, 2022.

CVM. O que é a CVM [What is CVM]. 2022. Available at: https://www.investidor.gov.br/menu/Menu_Investidor/a_cvm/ACVM.html#:~:text=A%20comiss%C3%A3o%20de%20Valores%20Mobili%C3%A1rios,o%20mercado%20de%20valores%20mobili%C3%A1rios. Accessed on: Sep 14, 2022.

DANG, V. et al. In search of robust methods for dynamic panel data models in empirical corporate finance. In: Journal of Banking & Finance, Volume 53, 2015, P. 84-98, ISSN 0378-4266. Available at https://doi.org/10.1016/j.jbankfin.2014.12.009 e https://www.sciencedirect.com/science/article/pii/S0378426614003902. Accessed on May 18, 2022.

DEMIRGÜÇ-KUNT, Asli; PERIA, Maria Soledad Martinez e TRESSEL, Thierry. The global financial crisis and the capital structure of firms: Was the impact more severe among SMEs and non-listed firms?. Journal of Corporate Finance, Volume 60, 2020, 101514, ISSN 0929-1199. Available at: https://doi.org/10.1016/j.jcorpfin.2019.101514. Accessed on December 20, 2021.

DEN BERG, Bas. Firm-specific determinants of capital structure: Evidence on UK listed and unlisted firms. Masters Dissertation. University of Twente. 2021. Available at: http://essay.utwente.nl/85557/1/vandenBerg_MA_BMS.pdf. Accessed on: Fev 20, 2022.

EXAME. Melhores e Maiores – metodologia [Best and Biggest – methodology]. Available at https://mm.exame.com/metodologia/. Accessed on Jul. 21, 2021

FAGERLAND, Morten W. t-tests, non-parametric tests, and large studies--a paradox of statistical practice?. BMC Med Res Methodol. 2012 Jun 14;12:78. doi: 10.1186/1471-2288-12-78. PMID: 22697476; PMCID: PMC3445820. Available at: https://pubmed.ncbi.nlm.nih.gov/22697476/. Accessed on January 4, 2022.

FORTE, D.; BARROS, L. e NAKAMURA, Wilson. Determinants of the Capital Structure of Small and Medium Sized Brazilian Enterprises. ANPAD. BAR, Rio de Janeiro, v. 10, n. 3, art. 6, pp. 347-369, jul-set. 2013. Available at: https://www.scielo.br/j/bar/a/XVC8cq8TYXQ9zzPrj79CnHj/?format=pdf&lang=en. Accessed on Feb 15, 2022.

FENG, C. et al. Log-transformation and its implications for data analysis. IN; Shanghai Arch Psychiatry. 2014 Apr; 26(2): 105–109. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4120293/. Accessed on: May 10, 2022.

FERNANDES, Pedro Onofre and CAMPOS, Luiz Fernando de Barros. Fundamentos para uma Central de Balanços Brasileira: o uso da XBRL [Foundations for a Brazilian Central Balance Sheet: the use of XBRL]. Nov. 2014. In: CONTECSI USP - International Conference on Information Systems and Technology Management - ISSN 2448-1041. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.517.2033&rep=rep1&type=pdf. Accessed on Oct 20, 2021.

FIPECAFI. Melhores e maiores [Best and Biggest]. 2021. Available at: https://fipecafi.org/Fundacao/FundacaoMelhoresMaiores. Accessed on Oct 04, 2021.

FIREBAUGH, G. e GIBBS, J. User's Guide to Ratio Variables. In: Article in American Sociological Review · October 1985. Available at: https://www.jstor.org/stable/2095384. Accessed on: May 07, 2022.

FREIRE, Janayna; CARVALHO, Eduardo; and CALEGÁRIO, Cristina. Análise da Relação Grau de Internacionalização versus Indicadores Financeiros de Multinacionais Brasileiras [Analysis of the Relation Degree of Internationalization versus Financial Indicators of Brazilian Multinational Companies]. In:

Revista de Negócios, v. 26, n. 2, p. 39-56, Abril, 2021. Available at: https://bu.furb.br/ojs/index.php/rn/article/view/9710/5123. Accessed on: Oct 18, 2021.

GAIBULLOEV, K. et al. Dynamic Panel Analysis under Cross-Sectional Dependence. Political Analysis (2014) 22:258–273. doi:10.1093/pan/mpt029. Available at: https://www.cambridge.org/core/services/aop-cambridge-core/content/view/88101DECF0F2804EF7E8AE2EEC53127F/S1047198700013723a.pdf/div-class-title-dynamic-panel-analysis-under-cross-sectional-dependence-div.pdf. Accessed on: 19. 18, 2022.

GARCÍA, J. and HERRERO, Begoña. Female directors, capital structure, and financial distress. In: Journal of Business Research, Volume 136, 2021, P. 592-601, ISSN 0148-2963, https://doi.org/10.1016/j.jbusres.2021.07.061. Accessed on Feb 20, 2022.

GHASEMI, Asghar and ZAHEDIASL, Saleh. Normality tests for statistical analysis: a guide for non-statisticians. International journal of endocrinology and metabolism vol. 10,2 (2012): 486-9. doi:10.5812/ijem.3505. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3693611/. Accessed on January 13, 2022.

GOTTARDO, Pietro, and MARIA MOISELLO, Anna. The capital structure choices of family firms: Evidence from Italian medium-large unlisted firms. Managerial Finance 40.3 (2014): 254-275. Available at https://www.emerald.com/insight/content/doi/10.1108/MF-03-2013-0065/full/html. Accessed on June 29, 2023.

IBGE. Bases de dados»metadados»MTE»relação anual de informações sociais – RAIS. 2020?. Available at: https://ces.ibge.gov.br/base-de-dados/metadados/mte/relacao-anual-de-informacoes-sociais-rais.html#:~:text=Metodologia%3A,%C3%A9%20preservada%20a%20informa%C3%A7%C3%A3o%20original. Accessed on Jul 16, 2021.

_____. Concla – Comissão Nacional de Classificação [Concla – National Classification Commission]. 2021. Available at: https://concla.ibge.gov.br/busca-online-cnae.html?view=estrutura. Accessed on July 30, 2021.

IMF. External Debt Statistics – Guide for compilers and users. 2014. Available at http://tffs.org/pdf/edsg/ft2014.pdf. Accessed on Dec 30,2021.

___. Balance of Payments and International Investment Position Manual - Sixth Edition (BPM6). 2009. 371 p. ISBN 978-1-58906-812-4. Available at: https://www.imf.org/external/pubs/ft/bop/2007/pdf/bpm6.pdf. Accessed on: May 23, 2020.

JAWORSKI, J. and SANTOS, M. Coutinho dos. Which Determinants Matter for Capital Structure? Evidence from Polish and Portuguese Nonfinancial Firms. 2021. European Review of Business Economics I(1): 27-60; DOI: https://doi.org/10.26619/ERBE-2021.01.2. Accessed on Feb 19, 2022.

LEE, D. K. Data transformation: a focus on the interpretation. In: Korean J Anesthesiol. 2020 Dec; 73(6): 503–508. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7714623/. Accessed on May 10, 2022.

MARINŠEK, Denis. The Impact of Indebtedness on a Firm's Performance: Evidence from European Countries. Doctoral thesis. Ljubljana, jan. 2015. Available at http://www.cek.ef.uni-lj.si/doktor/marinsek.pdf. Accessed on Apr 4, 2021.

MOTTA, Elaine Cristina de Oliveira. Juros sobre o Capital Próprio – Função e efeitos na tributação da renda [Interest on Equity – Function and effects on income taxation]. São Paulo: FGV, 2021. Masters Dissertation. Available at: https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/30672/Elaine.dissertacao_versao%20final%20%281%29.pdf?sequence=5&isAllowed=y. Accessed on Jul 19, 2021.

NAKAMURA, W.; MARTIN, D. and KIMURA, H. Indicadores Contábeis como Determinantes do Endividamento das Empresas Brasileiras [NAKAMURA, W.; MARTIN, D. and KIMURA, H. Accounting

Indicators as Determinants of the Indebtedness of Brazilian Companies]. In: Encontro Nacional dos Programas de Pós-Graduação em Administração, 28, 2004, Curitiba, Anais... [National Meeting of Post Graduation Programs in Business Administration, 28, 2004, Curitiba, Anals...] Curitiba: EnANPAD, 2004. Available at: http://www.anpad.org.br/admin/pdf/enanpad2004-ccg-1177.pdf. Accessed on: Feb 10, 2022.

NEHREBECKA, Natalia and BIAŁEK-JAWORSKA, Anna. Determinants of Polish Companies' debt Financing Preferences. 2015. Social Sciences 87(1). Available at: https://doi.org/10.5755/j01.ss.87.1.12319. Accessed on: Jan 7, 2021.

NEVES et al. Qual é o Efeito da Garantia Real na Exigência de Capital para Risco de Crédito no Brasil? [What is the Effect of Collateral on Capital Requirement for Credit Risk in Brazil?]. XXXI Encontro da ANPAD – Rio de Janeiro – 22 a 26 de setembro de 2007 [XXXI ANPAD Meeting – Rio de Janeiro – September 22 to 26, 2007]. Available at http://www.anpad.org.br/admin/pdf/FIN-C720.pdf. Accessed on Jul. 16, 2021.

Normas Brasil. Resposta à Consulta nº 8791 DE 11/04/2016 [Brazil Standards]. Response to Consultation No. 8791 OF 04/11/2016]. 2016. Available at https://www.normasbrasil.com.br/norma/resposta-a-consulta-8791-2016-sp_319569.html. Accessed on Oct 21, 2021.

Parsons, Christopher A. and Titman, Sheridan, Empirical Capital Structure: A Review (2008). Foundations and Trends in Finance, Vol. 3, No. 1, pp. 1–93, 2008. Available at: https://ssrn.com/abstract=2131759. Accessed on Feb 10, 2022.

POHLMANN, M. and IUDÍCIBUS, S. The relationship between the income taxation and the capital structure of large companies in Brazil. In: Revista Contabilidade & Finanças, 21 (53), ago. 2010. Available at: https://www.scielo.br/j/rcf/a/DS8SPscdbQ8R4CdLZwZtwLv/abstract/?lang=pt. Acesso em 10 fev. 2022.

PROCIANOY, Jairo Laser and SCHNORRENBERGER, Adalberto. A Influência da Estrutura de Controle nas Decisões de Estrutura de Capital das Companhias Brasileiras [The Influence of Control Structure on Capital Structure Decisions of Brazilian Companies]. Rio de Janeiro: RBE, 58(1):121-146 jan/mar 2004. Available at https://www.scielo.br/j/rbe/a/PBWn7Bmw85DZsZsRKRCndxv/abstract/?lang=pt. Access on Dec 17, 2021

RAALTE, Floris van. The impact of firm-specific determinants on the capital structure of Dutch listed and non-listed companies. 2021. Master's thesis, University of Twente. Available at: http://essay.utwente.nl/87321/1/Van%20Raalte_BA_Behavioural%2C%20Management%20and%20Social%20Sciences.pdf. Accessed on Feb 19, 2022.

Receita Federal do Brasil (RFB). Dados Públicos CNPJ [CNPJ Public Data]. 2021a. Available at: https://www.gov.br/receitafederal/pt-br/assuntos/orientacao-tributaria/cadastros/consultas/dados-publicos-cnpj . Accessed on July 16, 2021.

_____. SIC_Quantidade_de_CNPJs_em_docs_pagos_com_NJ.xlsx. 2021b. Answer, in April 26th, 2021 to que question 18870.001442/2021-22.

_____. RFB Normative Instruction No. 2003, of January 18, 2021. 2021c. Available at http://normas.receita.fazenda.gov.br/sijut2consulta/link.action?visao=anotado&idAto=114965#2223871. Accessed on: July 21, 2021.

RIBEIRO, Samuel, MENGHINELLO, Stefano and BACKER, Koen De. The OECD ORBIS Database: Responding to the Need for Firm-Level Micro-Data in the OECD. OECD Statistics Working Papers 2010/01. Available at: https://www.oecd-ilibrary.org/economics/the-oecd-orbis-database_5kmhds8mzj8w-en. Accessed on: March 25, 2021.

SANT'ANA, José Augusto de Almeida. A relação entre endividamento e rentabilidade nas empresas brasileiras de 1996 a 2000 [The relationship between debt and profitability in Brazilian companies from 1996 to 2000]. 2001. Masters Dissertation. Available at:

https://repositorio.unb.br/bitstream/10482/15203/1/2001_JoseAugustodeAlmeidaSantAna.pdf. Accessed on December 21, 2021.

SILVA, Veronica Favato. Determinantes da estrutura de capital na América Latina e nos Estados Unidos [Capital structure determinants in Latin America and the United States]. Universidade Federal de Uberlândia, Uberlândia, 2007. Masters Dissertation. Available at: https://repositorio.ufu.br/handle/123456789/12034. Accessed on Feb 10, 2022.

SILVA, Mailliw. Crédito Direcionado e Desenvolvimento Econômico: uma avaliação da política de crédito brasileira [Targeted Credit and Economic Development: an assessment of Brazilian credit policy]. Masters Dissertation. 2018. Available at: https://www.teses.usp.br/teses/disponiveis/12/12138/tde-12062018-111324/publico/CorrigidaMailliw.pdf. Accessed on July 16, 2021.

SILVA, Marcos Roberto Alves da. Um Painel Setorizado das Variáveis Determinantes da Estrutura de Capital de Empresas Brasileiras [A Sectorized Panel of the Determining Variables of the Capital Structure of Brazilian Companies]. In: Brazilian Journal of Business, v. 3, n. 1, 2021. Available at: https://brazilianjournals.com/index.php/BJB/article/view/26314/20866. Accessed on: Feb 10, 2022.

SPED – Sistema Público de Escrituração Digital [SPED – Public Digital Bookkeeping System]. Estatísticas ECD Junta [Joint ECDS Statistics]. 2023. Available at https://www.sped.fazenda.gov.br/estatisticascontabil/Default.aspx?ano=2021. Accessed on April 15, 2023.

TERRA, Paulo Renato Soares. Estrutura de capital e fatores macroeconômicos na América Latina [Capital structure and macroeconomic factors in Latin America]. R. Adm., v. 42, p. 192-204, abr/maio/junho 2007. Available at: https://www.revistas.usp.br/rausp/article/view/44435/48055. Accessed on December 17, 2021.

URIONABARRENETXEA, Sara, SAN-JOSE, Leire and RETOLAZA, Jose-Luis. Negative equity companies in Europe: theory and evidence. Business: Theory and Practice 17(4) 2016, 307–316. Available at: https://doi.org/10.3846/btp.17.11125. Accessed on January 06, 2022.

VARGHA, A. and DELANEY, H. The Kruskal-Wallis Test and Stochastic Homogeneity. Journal of Educational and Behavioral Statistics. Summer 1998, Vol. 23, No. 2, pp. 170-192. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.899.3263&rep=rep1&type=pdf. Accessed on: Jan 05, 2021.

VALORPRO. Serviço de Informações em Tempo Real [Real Time Information Service]. 2021. Available at: https://valorpro.globo.com/#/downloads. Accessed on: July 19, 2020.

WOOLDRIDGE, Jeffrey M. Econometric analysis of cross section and panel data. 2ª edição. The MIT Press, 2010. 1064 pages.

ZONENSCHAIN, Claudia Nessi. Estrutura de Capital das Empresas no Brasil [Capital structure of companies in Brazil]. Revista do BNDES, Rio de Janeiro, V. 5, N. 10, P. 63-92, Dez. 1998. Available at: https://web.bndes.gov.br/bib/jspui/handle/1408/13505. Accessed on: Feb 24, 2002.

Annex A - Description of the Alexandria base

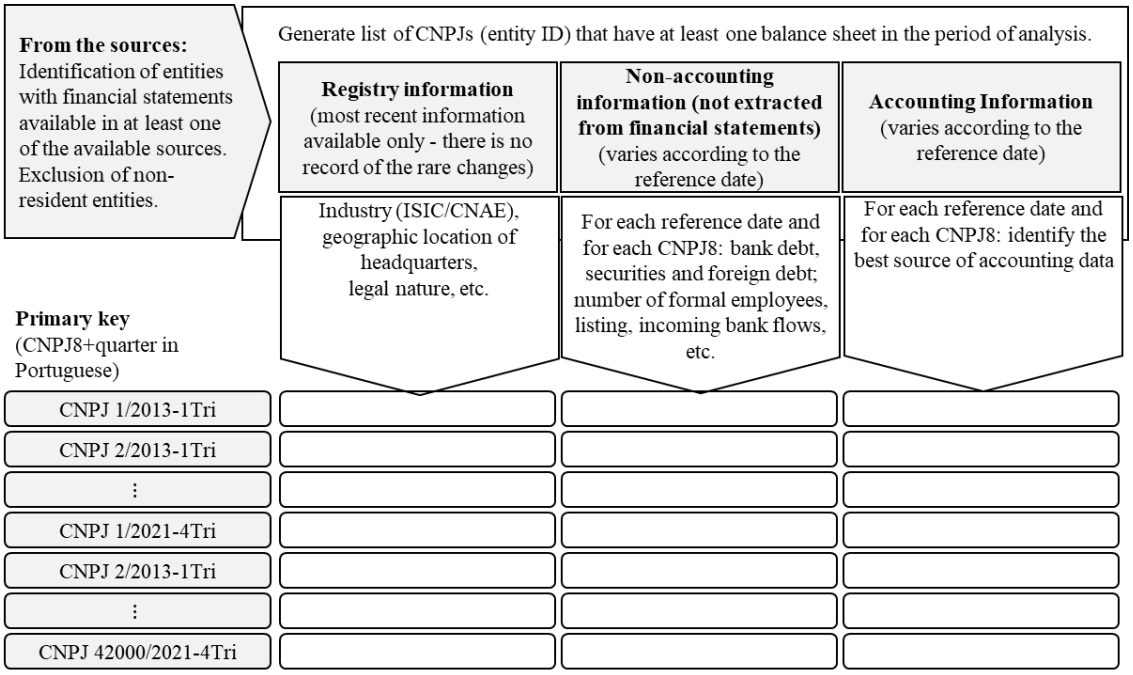
The Alexandria database, built by the authors of this work, is a database of formal accounting entities residing in Brazil – mostly medium-sized non-financial companies (NFCs). In its 1.0 version, it contains around 18,000 annual financial statements per year, on average (of which, 13,300 from NFCs), and information covering the period from 2013Q1 to 2021Q4, being expanded on a quarterly basis. The period of analysis of version 1.0 precedes Law no. 14.286 of December 29, 2021 (Brasil, 2021) – "new foreign exchange framework" – which changed data collection of foreign capital in the country by the Central Bank of Brazil (BCB). Therefore, the sample of Alexandria did not undergo methodological changes as a result of this law.

In the view of the authors of this work, Alexandria is a relevant contribution to Brazilian academic studies in Accounting, Finance and Economics, for the following reasons:

- Alexandria is a database of significant size that brings together medium and large companies in Brazil. It is larger in size than the main database used in academic work on unlisted companies (the Maiores e Melhores database).
    - Alexandria also breaks new ground by including an unique database of over 4,300 quarterly statements since 2016.
- Alexandria crosses accounting data with confidential and proprietary data from the BCB and other government bodies, available to the BCB via an agreement. No other database with accounting data in Brazil has this feature.

Figure 10 illustrates the steps in the construction of Alexandria, to be detailed in the next subsections. In summary, the authors took all available sources with accounting statements and thereby generated a list of accounting entities. From these sources, they obtained the financial statements. From other sources, they obtained information on registry, incoming bank flows, number of employees, etc.

Figure 10– Schematic representation of the construction of Alexandria



Source: prepared by the authors.

A.1 Scope of Alexandria

Alexandria is a database that gathers information from formal accounting entities resident in Brazil that have at least one balance sheet available[52] between 2013 and 2021. It uses as a primary key the first 8 digits of the CNPJ - the "CNPJ8"[53] - combined with the base date. Thus, for the CNPJ8 “01234567” and for the base date “2013-1st quarter, the primary key would be “01234567/2013-1Tri” (because the database is in Portuguese).

The criteria chosen to participate in Alexandria (resident entity with the main lines items of Balance Sheet) were purposefully broad, to allow Alexandria 1.0 and its later versions to host entities of different types. As the choice of CNPJ8s for Alexandria was not random (it was based on the availability of accounting information), the analyses must take this possible bias into account.

[52] Although Alexandria provides accounting information for the main line items of the Balance Sheet and Income Statement for the Year, in addition to two line items for the Statement of Changes in Equity, not all line items are always available for all entities. Thus, the minimum criterion used was the availability of information on total assets, total liabilities and total equity on at least one base date.

[53] Formal entities are recognized in Brazil by their identification number with the Federal Revenue Service of Brazil (RFB), here called “CNPJ14” because it has 14 digits. For establishments or subsidiaries of the same entity, the first 8 digits of a CNPJ - "CNPJ8" - are the same. Thus, the use of “CNPJ8” means the entity is being considered with all its subsidiaries and establishments.

As Alexandria is based on accounting statements, there are certain entities that are unlikely to be captured, as they are not obliged to produce them. Examples of these entities are: (i) micro-enterprises or small businesses (RFB, 2021b, art. 3); (ii) immune and exempt legal entities (such as churches and unions) that earned (in a broad concept) annual revenue of less than R$ 4.8 million (RFB, 2021c) and (iii) individual rural producers (Normas Brasil, 2016). If these entities decide to produce an accounting statement even though they are not obliged to do so, they may be part of Alexandria, but this is a rare case. This means, in practice, that Alexandria generally contains medium and large companies (simply put, those with revenue equal to or greater than R$ 4.8 million).

A.2. Entities in Alexandria: valid base dates and availability of information

Alexandria includes information from the first quarter of 2013 to the last quarter of 2021 – a total of 36 quarters or base dates. However, not all entities are active on all base dates. For each CNPJ8, the "opening date" and "closing date" (if any) were taken and the valid base dates were defined for each entity. When the CNPJ8 is written off, the closing date is inferred as the last change date in the CNPJ record.

For each CNPJ8 and for each valid base date, Alexandria contains a series of information about the entity. However, not all information is available for all valid base dates. In particular, the financial statements are concentrated on the base date of the fourth quarter and are generally annual information, as the minority of companies produces quarterly statements. Furthermore, many times the balance sheet is simply not available for a valid base date (because it is not included in any of the sources that Alexandria uses).

The 31,233 CNPJ8s of NFCs listed in Alexandria have, in total, 936,873 valid base dates, which means that each CNPJ8 tracked was active for an average of 29.99 quarters. Most CNPJ8s were active on all valid base dates[54].

Of the 936,873 valid base dates, Alexandria has financial statements on 166,958 of them - of which 119.839 are annual statements. This means that Alexandria averages 3.8 annual financial statements from each NFC.

Therefore, in econometric terms, Alexandria is an unbalanced panel, as not all entities are present on all base dates, and some variables, notably financial statements, are

[54] For reasons of space and scope, comments in this section A.2 are restricted to NFCs, although Alexandria contains other accounting entities.

not available even for all valid base dates (Wooldridge, 2010, p. 828)[55]. It is, moreover, a short panel, in which the number of base dates (36) is much lower than the number of observations (31,233, considering only NFCs).

A.3 Types of information in Alexandria

Alexandria is a database that compiles and organizes information from other databases. The compiled databases can be divided into two large groups.

The first large group corresponds to census databases, that is, those that provide information for the entire population of CNPJ8s and – in the case of information that changes over time – for all valid dates. This is the case of the external debt database (see section A.7.3): as virtually all external debt with resources entering Brazil must be registered in the BCB's RDE-ROF module[56], all valid dates contain some information (even if it be the information that entity "x", on base date "y", has external liabilities equal to zero).

The second large group corresponds to sample databases, that is, those that provide information for some valid dates and for some CNPJ8. This is the case for formal employee information (section A.9) and accounting information (section A.5).

A.4 Census database: registration with the Federal Revenue Service of Brazil

Alexandria uses the Federal Revenue of Brazil (RFB) register to obtain registration data of the current situation of the entity. These data are open to the public and represent all formal entities (with CNPJ8). The detailed explanation of each variable is in RFB (2021a).

The data only present the current status of the entity. Thus, for example, if the entity changed its headquarters from a municipality in the analysis period, Alexandria will consider only the last municipality (which is in the register)[57]. However, in the case of partners and managers, there is information on the entry date, which allows the analysis of partners over time.

---

[55] Wooldridge (2010, p. 809 and 828) comments on the econometric consequences of an unbalanced panel in which the reason for a variable or data being missing may be correlated with the error.

[56] The RDE-ROF System is now called the SCE-Crédito system because of the new normative framework over foreign capital in the country (Brasil, 2021). We use the former name here because, as already mentioned, we captured the variables before the changes brought on by this law.

[57] Generally, all CNPJ14 have the same legal nature and carry out the same activity (CNAE/ISIC). In the few cases where this does not happen, Alexandria classifies the entity according to the headquarters registration data.

The authors used information from partners and administrators to identify the gender of managers at each valid reference date. We identified the managers as being all adults (18 to 85 years old) alive on the reference date, present on the corporate board (or in the administration) before or on the reference date, who have been formally appointed as directors[58]. In the absence of this information, all alive adults from the corporate board were considered managers. In the (rare) case of absence of this second information, all individuals in the corporate structure were considered managers. This information allowed the construction of a variable (changeable over time) that identifies whether or not the entity has at least one woman in management.

For the limited liability companies, the authors also created a family company identifier. All companies owned by two or more individuals with the same family name were identified as family companies, regardless of the percentage of this group in the net equity. In addition, companies owned by legal entities that are family businesses were also considered family businesses (because they are owned indirectly by families).

A.5 Sample databases: accounting statements (RDE-IED, Brazilian Foreign Capitals Census, Economatica, Valor Pro and ANS)

The great advantage of Alexandria is the availability of the main line items of the financial statements of listed and unlisted companies. This innovation required an effort to unite sources that contained, individually and for some base dates, some entities, but which together represent a more expressive group of entities and base dates.

We collected only data from individual, not consolidated balance sheets, due to the greater availability of data.

The sources of financial statements collected by Alexandria are four: BCB (subdivided into RDE-IED and Brazilian Foreign Capitals Census[59]), Economatica, ValorPro and ANS. These sources are described in the next subsections.

[58] It was simply considered that administrators, board members, directors and CEOs are "managers".
[59] The RDE-IED System was renamed SCE-IED as a result of the new regulatory framework on foreign capital in the country (Brasil, 2021). Here we use the former name because, as already mentioned, we captured the variables before the changes brought about by this law.

A.5.1 Source of financial statement data and information on non-resident members: BCB (RDE-IED and Brazilian Foreign Capitals Census)

The BCB, as it is responsible for collecting external sector statistics, collects accounting and non-accounting information on companies owned by non-residents through two systems: RDE-IED and Brazilian Foreign Capitals Census. Because they are the main source of Alexandria balance sheets and because they are little explored in the literature, these systems deserve more attention.

All companies[60] and investment funds resident[61] in Brazil that are owned by non-residents in some percentage must declare summarized accounting information to the BCB – either in the Census of Foreign Capital, in the RDE-IED or in both. The frequency of declaration depends mainly on the size of the entity. In any case, it is a census database[62]. As a result, Alexandria is a base that contains more non-resident owned companies than wholly resident owned companies.

A.5.1.1 The Brazilian Foreign Capitals Census in the Country

The Brazilian Foreign Capitals Census, in its Annual and Five-Year versions, is a mandatory electronic survey for legal entities owned by non-residents and/or with a total outstanding balance of short-term commercial credits above a certain amount[63]. The survey captures accounting, corporate, external debt data, among others (BCB, 2019b, p. 48-51).

A.5.1.2 The RDE-IED

The RDE-IED is a mandatory registration module for investment made by non-residents in companies residing in the country. Despite the name “IED” (which refers to the direct investment position according to the BPM5 concept - today the name would be IDP in Portuguese), the RDE-IED does not only capture direct investment in the country, nor does it capture all direct investment (BCB, 2020a).

---

60 Individuals, including individual entrepreneurs, cannot receive investment from non-residents. For this reason, we do not use the generic term “accounting entity” here, but “company” (business accounting entity that is a legal entity).

61 Readers unfamiliar with the concept of residency may refer to IMF (2009, chap. 4, item E).

62 It stopped being a census survey when Law no. 14.286/2021 was enacted, especially due to the implementation of statement floors for the identification of the pair between receiver and non-residing investor, but this does not affect the sample of this study (Brasil, 2021).

63 The outstanding balance of short-term and long-term trade credits began to be captured exclusively by the SCE-Crédito system (former RDE-ROF) starting at the base date of December 31, 2021 (Brasil, 2021).

In the case of companies, the RDE-IED is a wide-ranging system: it captures all share participation held by non-resident investors in unlisted companies and some share participation in listed companies.

- As the voting criterion is not taken into account in the registration of the RDE-IED, this system captures, for these companies, both direct investment in the country (IDP) and portfolio investment.
- The RDE-IED captures the share participation of non-resident investors in the capital of resident companies. Thus, the portion of foreign investment that arrives through the capital market is not registered in the RDE-IED, but in the RDE-Portfolio module, as these investors do not pay in the capital nor are they part of the corporate structure of the receiving company.

The scope of the RDE-IED does not apply to investment funds, which are not required to make declarations in this module. Thus, their direct and portfolio investment positions are only captured by the Brazilian Foreign Capitals Census (BCB, 2020b, p. 7).

Resident companies declaring the RDE-IED must fill in economic and financial information in the system, information that was incorporated into Alexandria.

Whenever there is a change to company's share structure, the declaring company must complete an Accounting and Partners Statement (the Quadro Societário, or QS). This statement contains the following variables: total assets, shareholders' equity and total paid-in capital (of all partners, resident or not). Unfortunately, it does not contain net income data.

Companies with assets and shareholders' equity of less than R$250 million must also fill out the QS annually, even when there is no change to company's share structure. The annual QS, mandatory since 2016, is a periodic statement and source of thousands of balance sheets for Alexandria on each base date.

Companies with assets or equity equal to or greater than R$ 250 million must fill out the Economic and Financial Statement (DEF) quarterly and do not need to fill out the QS annually.

DEF is of high quality, being one of the rare sources of quarterly data for unlisted companies in the country. In addition to the data reported in the QS, the DEF contains information on net income and dividend distribution.

Table 10 summarizes the obligation to fill in accounting information in the Brazilian Foreign Capitals Census and RDE-IED. The accounting and corporate information stored in the two systems is largely comparable[64].

Table 10– Obligation to complete accounting information in the Brazilian Foreign Capitals Census and RDE-IED according to the achievement of the criterion on December 31 of the base year

System		**Legal entities (other than investment funds) headquartered in the country, with direct share participation of non-residents in their share capital**	**Resident investment funds with non-resident shareholders**
Brazilian Foreign Capitals Census	**Five-year period**	All	All
	Yearly	Shareholders' equity equal to or greater than US$ 100 million	
RDE-IED	**DEF (quarterly)**	Assets or shareholders' equity greater than R$ 250 million	----
	QS (annual)	All except those who declared DEF	-----

Source: BCB (2019b, p. 48-51). In some cases, legal entities headquartered in the country with a total outstanding balance of short-term trade credits (payable within 360 days) granted by non-residents must also complete the Brazilian Foreign Capitals Census, but only need to complete accounting information if they meet any of the above requirements.

A.5.2 Source of financial statement data: Economatica / CVM

The Economatica database provides subscribers with accounting and financial information on companies in 45 countries – the vast majority of which are listed. It is widely used in academic studies and by market analysts. Alexandria used information from all 1,152 companies residing in Brazil and available at Economatica. These companies were registered with the CVM to trade securities and are the companies that this work calls “CVM companies”. The reader should note that a company can be “CVM” in one quarter and cease to be CVM in the following quarter, and vice versa.

A.5.3 Source of financial statement data: Valor Pro

ValorPro (2021) is an information service from the economic group Globo. It has a business database with 7,360 listed and unlisted companies that registered accounting information with the CVM or published accounting statements in the official gazette. It prioritizes large companies, with annual net revenue above R$ 100 million. It is also a good quality source and used in previous literature (MOTTA, 2021, p. 48).

[64] Especially from 2017 onwards, the RDE-IED accounting data are of excellent quality, as the system was updated to a web interface in January 2017. This update maintained the previous data, but created a more user-friendly interface for the declarant, in addition to simplifying several business rules. As a result, it facilitated registration (for the declarant), data monitoring (for the BCB team) and, consequently, the quality of the final data.

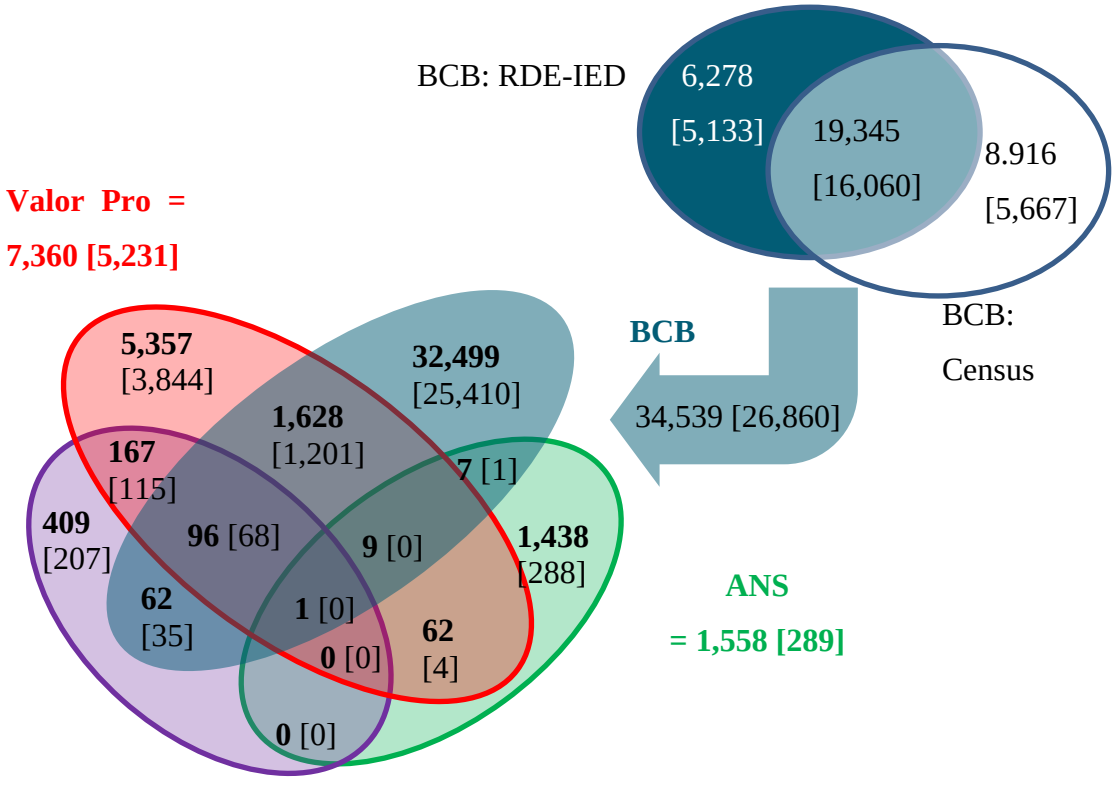
A.5.4 Source of financial statement data: ANS

The National Agency of Supplementary Health (ANS) supervises legal entities governed by private law that operate health care plans. The agency has been compiling and publishing quarterly financial information for small entities (Brasil, 1998 and ANS, 2021) for many years, defined as those with less than 20,000 beneficiaries. Avelar et al.(2019) have already used this database to investigate determinants (including regulatory ones) of the indebtedness of medical cooperatives.

A.5.5 Union of different sources of accounting data

Figure 11 presents a Venn diagram with the number of CNPJ8 with accounting data available in each source (described in sections 3.5.1 to 3.5.4). The data between square brackets informs the number of CNPJ8 that are non-financial companies [NFCs]. One of the conclusions that can be drawn from Figure 2 is that the "BCB" source is the most important for the construction of the Alexandria base: of the more than 42,000 legal entities in Alexandria, 32,499 have accounting information only in the "BCB" source.

Figure 11– Venn diagram representing the number of CNPJ8s [of which, NFCs] with accounting information present in each source



Source: prepared by the authors. Numbers reflect position on December 12, 2022.

It can be seen that in Figure 11 there is a large number of statements that are available from more than one source. When there was data from more than one source for the same base date, a source selection process was necessary. This process included a detailed manual comparison and checking of data in official gazettes (in the case of corporations). In addition, as one of the authors works in the area responsible for monitoring the quality of the RDE-IED data, it was possible to question declarants who presented incorrect data and request the correction of this information. This resulted in an excellent quality database, especially for data from 2018 or later, and especially for larger entities**.**

Just to give you an idea, Alexandria has accounting information for 21,035 NFCs on the base date of December 2020. Only 14.6% of them have total assets bigger than R$ 250 million. However, this small group represents 94.3% of Alexandria's total assets at this base date. The number must be interpreted with care, as there are some cases of Alexandria entities holding others. Even so, it already gives an idea of the importance of this group of entities and justifies the focus on the quality of these data.

The authors of this work established a quality hierarchy of accounting data sources, based on the presence or absence of audited data and extensive manual validation. Thus, if the company has, on a given base date, financial statements whose source is Economatica, a well known company in the market and specialized in data, this data prevails in Alexandria, regardless of existing in another source.

The second highest quality base is Valor Pro. Thus, if your data conflicts with data reported in another source (except Economatica), it will prevail. Then we consider it to be of better quality, in this order: RDE-IED, ANS and Brazilian Foreign Capitals Census. As the corporate charts (QS) of the RDE-IED do not capture information on profit, gross revenue and dividends, this data was sought from the Brazilian Foreign Capitals Census when available.

After selecting the best source of the financial statements for each base date and each CNPJ8, the main line items of the best quality financial statements referring to that base date are obtained, as per the selection explained above.

It is important to note that, for the same CNPJ8, different sources may occur depending on the base date. For example, a company owned by non-residents may have declared the Brazilian Foreign Capitals Census in December 2015. In 2016, it became listed, a situation in which the source used by Alexandria becomes Economatica.

After structuring the base, the data was exhaustively compared to identify outliers and errors. Special attention was given to entities with total assets bigger than R$ 10 billion.

A.6 Census database: shareholding of non-residents in companies

In addition to accounting data, the BCB captures, through the Brazilian Foreign Capitals Census and RDE-IED surveys (section A.5.1), shareholder's composition in companies and investment funds wholly or partially owned by non-residents on census basis until 2021.

The combination of the two sources, Brazilian Foreign Capitals Census and RDE-IED, generated unique information on the participation of investors in companies over the years. This information, combined with corporate data available at the Federal Revenue of Brazil base, became quarterly data at Alexandria.

A.7 Census database: total credit to the non-financial sector

The total credit to the non-financial sector is a statistic produced by the BCB in accordance with the international methodological standard that computes the debt of non-financial companies against (i) institutions of the National Financial System (SFN), (ii) bondholders and (iii) non-resident entities and individuals (BCB, 2018).

Alexandria used the microdata that originated this statistic and captured debt information from the CNPJ8s already included in Alexandria, separating the total outstanding credit of each CNPJ8 by type of debt and by term. This last separation was made to compare this information with current liabilities and non-current liabilities available in the balance sheet. Table 11 summarizes the three types of debt captured by this statistic, the system that captures the microdata, and the variables created in Alexandria.

Table 11 – Components of total credit lent to the non-financial sector, microdata collection system and variables generated for Alexandria

Type of debt registered in the system		**Loans and financing granted by SFN financial institutions**	**Public and private debt securities**	**Credits granted by non-residents (external debt)**
System that collects microdata [see Section]		**SCR (BCB)** [A.7.1]	**B3 Registration** [A.7.2]	**RDE-ROF (BCB)** [A.7.3]
Variables generated for Alexandria from microdata	**Short term (ST -expiration within 12 months)**	Short term SCR liability	Liabilities in short-term securities	• ST intercompany declared • ST intercompany inferred, • Other ST
	Long-term (LT - maturity equal to or greater than 12 months)	Long-term SCR liability	Liabilities in long-term securities	• LT intercompany declared • LT intercompany inferred, • Other LT

Sources: Prepared by the authors based on BCB (2018).

A.7.1 Debt against institutions of the national financial system: source "SCR"

The BCB's Credit Information System (SCR) gathers all the individualized information on customer credit agreements from a certain amount (which in 2021 was equal to R$ 200). It began in 2002 and its data are the source of numerous academic works – Neves et al. (2007) and Silva (2018) are just a few examples.

The variables obtained from the SCR by Alexandria, for each base date from 2013Q4 (inclusive) and for each CNPJ8, are as follows:

- "Short-term SCR liabilities": represents CNPJ8 liabilities against all institutions of the National Financial System, to be paid within 12 months.
- "Long-term SCR liabilities": represents CNPJ8 liabilities against all institutions of the National Financial System, to be paid over a period of more than twelve months.

When the entity does not have external debt or issue securities, the sum of the two variables above should be identical to that reported under the line items "Loans and financing" of Liabilities in the Balance Sheet.

A.7.2 Liabilities in securities issued in the domestic market: source "B3 registration"

This database compiles securities issued in the domestic market by non-financial legal entities and by public sector entities, excluding those held by non-residents, which are classified under external debt.

The variables obtained for each base date from 2013Q4 (inclusive) and for each CNPJ8 are as follows:

- "Liabilities in short-term securities": represents the CNPJ8 liability issued in securities, to be paid within 12 months. The maturity structure is estimated by the BCB.
- "Liabilities in long-term securities": represents the remaining liability issued in securities (to be paid over a period of more than twelve months).

A.7.3 Credits granted by non-residents: source "RDE-ROF"

The BCB compiles the data associated with each external debt transaction, including those between companies of the same economic group in the RDE system, module RDE-ROF. The data are of good quality, as almost all payments to external creditors require currency conversion, which is reported via the Exchange system to the BCB. The only external liabilities that are not recordable in the RDE-ROF are:

- Short-term trade credits (such as import financing and export prepayment). This liability is compiled in the Brazilian Foreign Capitals Census (section 3.5.1.1).
  - However, operations originally contracted with a short payment period and which, when refinanced, reach a payment period of more than 360 days, must be registered in the RDE-ROF.
- External liabilities that did not enter the country, that is, which were taken abroad and remained there (for example, to pay some commitment of the company abroad). This liability is compiled in the Brazilian Foreign Capitals Census (see section 3.5.1.1) and is uncommon, even among large companies[65].

For the purposes of this study, the outstanding balances of each entity in Alexandria were compiled, on each base date, and the amounts were converted into reais

[65] Just to give the reader an approximate idea, the BCB informs that the Brazilian gross external debt of the private sector was US$ 465.9 billion in December 2020 (BCB, 2022). This amount is compiled by the RDE-ROF system, therefore it does not include debt not entered. The companies declaring the Census of Foreign Capital reported liabilities whose values were classified as "not entered in the country" of around US$ 164 billion, with two companies accounting for 88.9% of this amount. Part of this amount may be owed to a subsidiary of the reporting company headquartered abroad, which is why not necessarily all of this amount will appear on the individual balance sheet of the resident company.

at the exchange rate at the time. The compiled amount is theoretically slightly less than or equal to that recorded as an external liability under "liabilities" in the balance sheet of the debtor entity (resident entity). This is because the RDE-ROF only calculates the principal of each debt – without accrued interest[66] – and because there are some external liabilities captured only by the Brazilian Foreign Capitals Census.

External liabilities recorded in the RDE-ROF received two classifications: by relationship between creditor and debtor and by the term of each installment.

The classification according to the relationship between creditor and debtor separated the liabilities into "intercompany" and "other". Alexandria's definition of "intercompany" lending is very similar to that used in external sector statistics (IMF, 2009, §6.26), except that these exclude loans taken between affiliated financial institutions (IMF, 2009, §6.28). That is, in Alexandria, all liabilities between affiliated companies are classified as intercompany in Alexandria, regardless of the industry of the creditor or borrower[67].

The loan is considered "intercompany" in two situations:

- Declared: when the liability is registered in the RDE-ROF and the declarant informs that the creditor holds 10% or more of the voting power.
- Inferred: When the debtor appears in the RDE-IED (section 3.2.1.2) as held by the creditor with paid-in capital greater than zero.

The intercompany loan can, in many situations, be considered equivalent to equity. As this work analyzes the indebtedness of business entities in relation to third parties, it would be interesting to separate them from the others. However, we do not have information on domestic intercompany loans (that is, between companies residing in the same economic group), only on external intercompany loans[68].

[66] The option not to calculate interest was made to simplify the calculations. The authors understand that this simplification does not cause great harm, because: (i) international interest rates are very low, and companies generally pay interest over the course of the loan (i.e., there is little interest accrued and not paid on each base date); especially in the case of intercompany loans, there are cases of interest-free loans and cases of interest never paid (an implicit or explicit forgiveness of interest agreed in the contract).

[67] Researchers who want to seek comparability between the intercompany loan from the external statistics and the Alexandria data can simply exclude intercompany loan balances where both the borrower and the lender are financial institutions. For NFCs, the focus of this article, the definition coincides with the methodological standard of the BPM6 (IMF, 2009, §6.26).

[68] For methodological comments on measuring intercompany borrowing in external debt, see IMF (2014, §3.20 and §4.3).

Considering the total credit balance of each NFC, we find excellent data: for the 165,754 base dates for which information on liabilities and total credit is available, in only 3,601 cases (2.1%) the value exceeds the liability total (thus indicating an error, which may be (i) failure of the declarant to update the record on any of the information or (ii) wrong declaration of liabilities).

A.8 Census database from 2019Q1: incoming bank flows (DOC, TED, PIX, debit card and export)

The incoming bank flows represent the amounts received by the CNPJ8 in the following ways (BCB 2021c): (i) slips, (ii) Available Electronic Transfers (TEDs), (iii) debit cards, (iv) credit cards, (v) Exports and (vi) Transfers via Pix (BCB, 2021b) made through the payment system snapshots (SPI), which identifies the parts (BCB, 2021a).

Alexandria takes the variable "incoming bank flows" for each entity and for each quarter. Data is only available from 2019 (inclusive) and only for non-financial companies[69].

Incoming bank flows have been used by BCB (2021c) as a proxy for entity billing. This is an analysis to be confirmed (and future studies using Alexandria may help with this). In any case, it is important to point out that incoming bank flows do not necessarily represent billing, much less indicate the total billing, since companies can receive by other means, not captured by this variable: (i) cash, (ii) check, (iii) intra-bank transfers (except via TED and identified Pix), (iv) Pix outside the SPI system, (v) DOCs, (vi) TECs and (vii) invoices of water, energy, gas, cable TV and other bills, whose payment is not made by any of the forms included in the flows (and described above).

Thus, there are active business entities that do not record any incoming bank flows for long periods. In addition to being able to receive by other means, as mentioned above, there are cases in which the entrepreneur uses his personal account to deal with the company's cash (common in the case of individual entrepreneurs). However, the authors understand that this will hardly happen with a medium or large business entity.

Even with these observations on the variable, it is undeniable that this is precious information, which allows, at the very least, to investigate the level of activity and financial inclusion of the company, among other possibilities.

[69] That is, this base does not contain incoming flows received by companies with CNAE/ISIC 64 (except 64.62, for which data are available), 65 (except 655) or 66.

A.9 Sample database: formal jobs (RAIS, CAGED and eSocial)

RAIS, Caged and their successor, eSocial, are state repositories of information on the formal employment relationships of resident legal entities (IBGE, 2020?; Almeida et al., 2020).

Although RAIS, Caged and eSocial declarations are mandatory for legal entities, in practice many do not declare them, especially those without formal employees – which is why a sample base was considered. However, IBGE (2020?) estimates that RAIS has coverage of “97% of the total organized sector of the economy”.

The employment variables captured by Alexandria, for each CNPJ8 and for each base date, refer to (i) the number of employees and (ii) the average wage, broken down both by gender and education.

A.10 Summary of sources

Figure 12 illustrates the main information that Alexandria has about each business entity, in each base date, and classifies them according to the source used. This is a summary, as Alexandria has more than 120 variables.

The summary makes it clear that, although Alexandria collected only a few items of the entities' liabilities, it was possible to complete the information with internal data from the BCB. Thus, liabilities could be broken down into liabilities against external creditors, against SFN creditors and in securities. The rest of the liability was obtained by residual and includes liabilities against suppliers, with payroll, against the government, among others.

Figure 12– Information about each entity in Alexandria, in each base date.
CNPJ XXXXXX, Base date: YYYY/Q

Balance sheet	
1) Total assets [A]	2) Total Liabilities [A] 2.1) Current liabilities [A] 2.1.1) External creditors [B] – includes intercompany 2.1.2) SFN creditors [C] 2.1.3) Securities held by residents [D] 2.1.4) Other current liabilities (*=2.1-2.1.1-2.1.2-2.1.3*) 2.2) Non-current liabilities [A] 2.2.1) External creditors [B] – includes intercompany 2.2.2) SFN creditors [C] 2.2.3) Securities held by residents [D] 2.1.4) Other non-current liabilities (*=2.2-2.2.1-2.2.2-2.2.3*)
	2.3) Equity [A] 2.3.1) Share capital [A]
Income Statement for the Year (IS)	
3) Net Revenue [A] 4) Net income [A]	
Other information	
5) Formal Employees, by gender and education [E] 6) Incoming bank flows [F] 7) Held entities [G] 8) Gender of the administrator [G] 10) Corporate data: family business (G); participation of non-residents in share capital and voting power [H]	

Sources used to obtain each information [Yes=available on all quarterly base dates]:
[A] Brazilian Foreign Capitals Census (BCB), RDE-IED (BCB), Valor Pro or Economatica (according to the best source); [B] RDE-ROF (BCB) [Yes]; [C] SCR (BCB) [Yes from 2013Q4]; [D] Cetip/B3 [Yes, from 2013Q4]; [E] RAIS/CAGED/eSocial; [F] BCB [Yes, as of 2019]; [G] CNPJ database (Federal Revenue of Brazil); [H] Brazilian Foreign Capitals Census (BCB) and/or RDE-IED (BCB) [Yes].

Source: prepared by the authors.

Appendix B – Median indebtedness of NFCs in the Alexandria base versus the Klooks base (without removing outliers; in %)

CNAE Section	Alexandria Base									Klooks Base		
	2013	2014	2015	2016	2017	2018	2019	2020	2021	2018	2019	2020
A	37.9	37.3	37.4	37.6	38.5	34.8	32.6	30.9	34.5	25.4	25.8	26.4
B	43.2	44.6	57.1	56.1	50.2	46.4	41.9	43.1	45.2	34.5	39.9	34.5
C	57.6	57.7	61.9	56.6	54.3	55.4	55.4	57.4	56.5	48.4	50.6	54.7
D	54.5	55.8	58.8	53.1	49.0	51.8	49.7	52.3	51.1	49.2	47.1	52.1
E	56.2	58.1	58.9	56.0	55.3	55.5	56.1	54.5	54.1	54.9	57.9	58.6
F	48.9	47.7	38.6	39.7	38.0	37.4	36.7	34.7	35.0	30.2	32.6	35.5
G	65.5	64.9	71.7	65.1	63.2	64.7	64.3	65.1	62.7	30.9	36.0	44.4
H	67.3	69.6	70.1	67.1	65.2	65.3	63.8	67.0	66.6	51.7	55.5	60.5
I	63.4	52.2	33.9	34.2	46.7	47.5	41.1	34.8	37.8	36.0	39.5	44.0
J	49.5	51.3	61.3	55.7	53.6	55.6	56.1	54.0	54.6	48.2	49.6	49.4
L	40.6	35.2	11.7	11.5	10.2	9.5	7.9	9.5	8.7			
M	55.5	49.2	51.1	50.9	48.0	42.5	45.0	45.9	44.3	31.7	35.1	36.8
N	65.1	64.4	65.1	65.7	63.1	61.2	60.5	60.4	58.9	33.7	39.1	51.2
P	62.6	57.1	57.0	58.4	55.8	46.6	54.8	50.6	53.4	28.1	41.8	48.0

Q	67.3	68.2	65.0	66.4	63.2	56.5	52.5	49.6	47.7	46.7	53.6	53.6
R	42.8	47.2	43.9	42.3	38.7	31.6	31.1	40.5	52.2	27.9	35.6	48.6
S	61.1	65.4	62.8	62.5	50.4	46.0	37.1	44.3	36.6	26.1	23.7	33.0
NFCs	58.4	58.3	57.0	55.2	52.8	53.4	51.5	51.6	51.4	39.7	43.2	48.9

Source: Alexandria base and Klooks base. Sections K, O, T and U were not presented because they do not belong to the NFCs, as defined in section 3.1.

Annex C – Mean and median indebtedness of NFCs by industry (in %)

	Mean (ignores amounts greater than 300)						Median					
	Ind	Serv	Com	Constr	Agr	Imob	Ind	Serv	Com	Constr	Agr	Imob
2015	66.36	63.00	73.94	47.80	50.75	29.84	60.34	57.95	72.11	38.03	37.04	10.38
2016	60.68	60.34	67.08	46.93	44.06	26.30	54.69	56.76	65.21	38.38	36.68	9.09
2017	59.99	59.79	65.75	46.59	46.04	26.90	53.55	55.89	64.57	38.31	37.07	9.91
2018	60.01	59.21	68.76	44.53	45.05	27.40	54.50	54.98	66.70	37.33	34.41	9.40
2019	61.13	58.04	69.23	44.79	45.64	27.69	53.89	53.59	64.53	36.39	32.13	8.70
2020	62.81	57.49	69.35	45.35	43.99	30.95	56.46	54.63	66.44	35.65	32.53	10.59
2021	62.73	58.74	68.17	47.16	49.36	29.42	55.54	54.01	64.07	37.99	35.82	9.91
ENFs	62.15	59.39	69.29	46.23	46.61	28.94	55.72	55.35	66.39	37.21	35.04	9.80

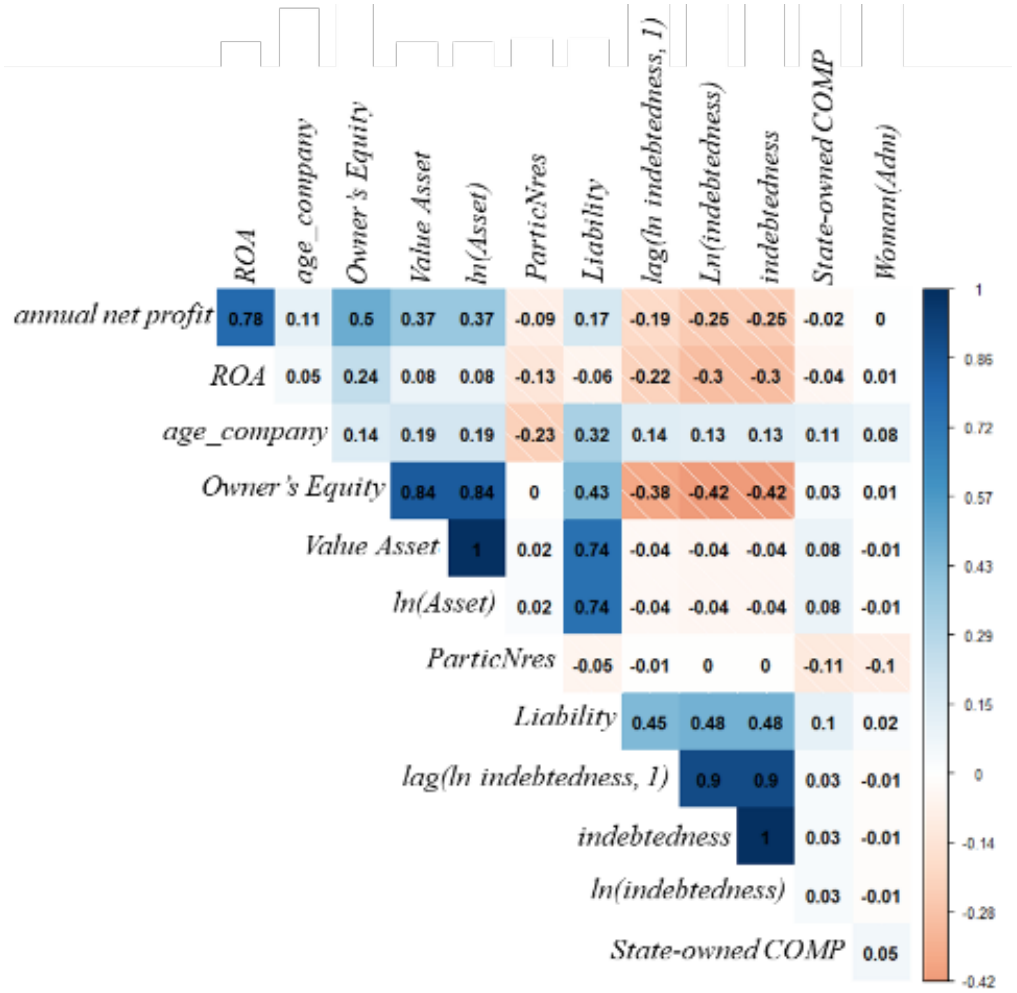
Annex D – Descriptive statistics of the variables used in section 4.3

Number of observations used between 2015 and 2021, including: 100,944

Variable	**Average**	**Standard deviation**	**Minimum value**	**Perc. 25**	**Median**	**Perc. 55**	**Maximum value**
Total Assets (R$ million)	1,648.58	26,959.45	0	7.074	71.496	403.291	1,988,646.00
Total liabilities (R$ million)	1,133.179	24,107.94	-12.548	1.022	19.211	142.915	1,854,420.00
Net equity value (R$ million)	515.404	5,419.95	-64,485.81	0.996	24.136	183.567	892,952.50
Annual net income (R$ million)	44.892	1,294.914	-44,212.19	-0.506	0.392	13.252	290,820.00
log(indebtedness+1) (%) (a)	45.291	49.79	0.00	14.524	39.24	58.667	345.51
Indebtedness (Debt-to-asset ratio, in %) (a)	102.266	336.935	0	15.582	48.033	79.792	3,065.613
ROA (%) (a)	-10.215	86.94	-700.559	-3.365	1.781	9.906	108.503
Log(Asset) (R$ million)	3.737	3.451	-18.644	1.956	4.27	6	14.503
Nres Share Participation (%)	42.548	47.125	0	0	0.4	100	100

Observation: (a) Descriptive statistics using winsorization technique at 1% level.

Annex E - Correlogram of the variables mentioned in Table 7



Annex F –Time Robustness Tests

Table 12–Estimation of indebtedness models using the Sys-GMM method in balanced panel and unbalanced panel of non-financial companies (NFCs) in different periods.

	Unbalanced panel					Balanced panel				
	(A) NFCs 2015 to 2021 (1)	(A.1) NFCs 2015 to 2019 (2)	Δ	(A.2) NFCs 2017 to 2021 (3)	Δ	(D) NFCs 2015 to 2021 (4)	(D.1) NFCs 2015 to 2019 (5)	Δ	(D.2) NFCs 2017 to 2021 (6)	Δ
lag(log D/A, 1)	**0.7918*** **	**0.7771*** **	▼	**0.8339*****	▲	**0.9187*** **	**0.8563*** **	▼	**0.9086*****	▼
	(0.0359)	(0.0497)		(0.0390)		(0.0211)	(0.0247)		(0.0322)	
Services	**0.0038**	**0.0058***	▲	**0.0014**		**0.0019**	**0.0122****	▲	**0.0028**	
	(0.0026)	(0.0031)		(0.0025)		(0.0037)	(0.0050)		(0.0035)	
Commerce	**0.0048**	**0.0083****	▲	**0.0031**		**0.0051***	**0.0125*****	▲	**0.0066****	▲
	(0.0032)	(0.0040)		(0.0033)		(0.0028)	(0.0036)		(0.0033)	
Construction	**-0.0137****	**-0.0126***	▼	**-0.0095***	▲	**-0.0041**	**-0.0031**		**-0.0023**	
	(0.0056)	(0.0069)		(0.0056)		(0.0043)	(0.0057)		(0.0046)	
Agriculture	**-0.0132***	**-0.0140**	▲	**-0.0087**	▲	**-0.0066**	**-0.0071**		**-0.0083***	▲
	(0.0076)	(0.0095)		(0.0067)		(0.0046)	(0.0063)		(0.0051)	
Real Estate Act.	**-0.0401*** **	**-0.0382*** **	▲	**-0.0384*****	▲	**-0.0114***	**-0.0138***	▼	**-0.0147****	▼
	(0.0082)	(0.0101)		(0.0084)		(0.0059)	(0.0077)		(0.0071)	
ROA	**-0.1471*** **	**-0.1465*** **	▲	**-0.1335*****	▲	**-0.1905*** **	**-0.2038*** **	▲	**-0.0840**	▼
	(0.0132)	(0.0164)		(0.0146)		(0.0388)	(0.0735)		(0.0650)	
Data in CVM	**0.0223*** **	**0.0295*** **	▲	**0.0151****	▼	**-0.0073**	**-0.0139**		**0.0016**	
	(0.0057)	(0.0076)		(0.0061)		(0.0079)	(0.0113)		(0.0065)	
State-Owned	**0.0069**	**0.0111**		**0.0006**		**-0.0138*** **	**-0.0179*** **		**-0.0109****	
	(0.0065)	(0.0080)		(0.0061)		(0.0044)	(0.0065)		(0.0051)	
Familiar	**0.0021**	**-0.0005**		**0.0101**		**0.0032**	**0.0151****		**0.0014**	
	(0.0058)	(0.0069)		(0.0065)		(0.0048)	(0.0075)		(0.0049)	
Log(asset)	**-0.0012**	**-0.0020***		**0.0010**		**0.0058**	**0.0114***		**0.0027**	
	(0.0009)	(0.0011)		(0.0009)		(0.0040)	(0.0059)		(0.0035)	
ParticNres	**0.0020**	**-0.0005**		**-0.0003**		**-0.0088***	**-0.0215*** **		**-0.0037**	
	(0.0031)	(0.0041)		(0.0030)		(0.0050)	(0.0068)		(0.0050)	
Headquarter	**-0.0046****	**-0.0047***		**-0.0045****		**-0.0040**	**-0.0068***		**-0.0023**	
In State Capital	(0.0021)	(0.0026)		(0.0021)		(0.0026)	(0.0038)		(0.0026)	
Woman in	**-0.0023**	**-0.0039**		**-0.0010**		**-0.0012**	**-0.0012**		**-0.0002**	
management	(0.0021)	(0.0025)		(0.0022)		(0.0012)	(0.0015)		(0.0014)	
Log(age)	**0.0017**	**-0.0005**		**0.0022**		**0.0035*** **	**0.0028****		**0.0034*****	
	(0.0015)	(0.0017)		(0.0016)		(0.0010)	(0.0013)		(0.0013)	
nobs	53,891	35,132		34,427		20,097	15,631		15,631	
Instruments	E [2:99]	E [2:99]		E [2:3]		E. R. A [3:99]	E. R. A. I		E. R. A. I	
Hansen-Sargan	229.2080	244.8990		147.6770		360.6780	31.1580		277.2850	
[p-value]	*[0.2408]*	*[0.1067]*		*[0.5417]*		*[0.0541]*	*[0.0711]*		*[0.1480]*	
AR1	-11.58	-7.89		-8.78		-14. 69	-11.14		-12.69	
[p-value]	*[0.0000]*	*[0.0000]*		*[0.0000]*		*[0.0000]*	*[0.0000]*		*[0.0000]*	
AR2	-1.58	-1.58		0.67		1.45	-1.88		1.86	
[p-value]	*[0.1149]*	*[0.1132]*		*[0.5009]*		*[0.1460]*	*[0.0604]*		*[0.0625]*	
Wald Coef {GL}	6,610.94	4,069.33 {15}		6,446.53		7,327.56 {15}	4,288.53 {15}		5,592.19 {15}	
[p-value]	*[0.0000]*	*[0.0000]*		*[0.0000]*		*[0.0000]*	*[0.0000]*		*[0.0000]*	
Wald Time {GL}	17.07 {5}	7.86 {3}		2.69 {3}		28.57 {4}	41.76 {3}		3.34 {3}	
[p-value]	*[0.0044]*	*[0.0490]*		*[0.4415]*		*[0.0000]*	*[0.0000]*		*[0.3421]*	

*Notes: (i) Analysis of the sample's annual variables; (ii) Errors robust to heteroscedasticity in parentheses using firm-level error clustering; (iii) * 10% significance; ** 5% significance; *** 1% significance; (iv) AR(1) and AR(2) verify the presence of serial correlation of first order and second order in the residuals of first difference (v) We do not show the time and intercept dummies for brevity. (vi) the winsorization technique was used at a 1% level for the debt-to-asset ratio and ROA variables; (vii) variables in lag used as instruments: E - ln(debt+1) winsorized; R - ROA winsorized; A - ln(asset); values between [ ]´s are the lags of the variables used as instruments.*

Anexo G – Industry robustness test

Table 13–Robustness analysis of debt models using the Sys-GMM method on unbalanced panels of non-financial companies (ENFs) with breakdown by sector.

	Extractive and manufacturing	Services	Commerce	Construction	Agriculture	Real Estate Activities
lag(log D/A, 1)	**0.8987 *****	**0.6824 *****	**0.9528 *****	**0.8423 *****	**0.9070 *****	**0.8384 *****
	(0.0461)	(0.0691)	(0.0541)	(0.0955)	(0.0565)	(0.0833)
ROA	**-0.1130 *****	**-0.1599 *****	**-0.1319 *****	**-0.1123 ****	**-0.1842 *****	**-0.0150**
	(0.0251)	(0.0187)	(0.0332)	(0.0437)	(0.0465)	(0.0116)
Data in CVM	**0.0108**	**0.0094**	**-0.0108**	**0.0331**	**0.0327**	**0.0861 ****
	(0.0076)	(0.0098)	(0.0078)	(0.0202)	(0.0497)	(0.0401)
State-owned	**0.0014**	**0.0026**	**-0.0104**	**0.0044**		**0.0425**
	(0.0076)	(0.0132)	(0.0165)	(0.0424)		(0.0335)
Family	**0.0055**	**-0.0002**	**0.0068**	**-0.0067**	**-0.0045**	**-0.0034**
	(0.0071)	(0.0152)	(0.0071)	(0.0102)	(0.0128)	(0.0158)
Log(asset)	**0.0004**	**-0.0018**	**0.0023 ***	**-0.0043**	**-0.0005**	**0.0028**
	(0.0012)	(0.0016)	(0.0014)	(0.0037)	(0.0030)	(0.0031)
ParticNres	**-0.0046**	**0.0058**	**-0.0051**	**-0.0076**	**-0.0118**	**0.0081**
	(0.0030)	(0.0075)	(0.0053)	(0.0141)	(0.0144)	(0.0150)
Headquarter	**-0.0031**	**-0.0049**	**-0.0027**	**-0.0070**	**0.0045**	**-0.0081**
in state capital	(0.0029)	(0.0050)	(0.0038)	(0.0090)	(0.0110)	(0.0110)
Woman in	**-0.0014**	**0.0000**	**-0.0013**	**-0.0064**	**0.0075**	**0.0079**
management	(0.0024)	(0.0046)	(0.0044)	(0.0111)	(0.0089)	(0.0089)
log(age)	**0.0014**	**0.0004**	**0.0017**	**0.0063**	**-0.0028**	**0.0012**
	(0.0018)	(0.0035)	(0.0041)	(0.0056)	(0.0084)	(0.0060)
Intercept	**0.0261**	**0.1919 *****	**-0.0222**	**0.1225 ***	**0.0571**	**-0.0105**
	(0.0361)	(0.0541)	(0.0458)	(0.0709)	(0.0672)	(0.0627)
2017	**0.0106 *****	**-0.0083**		**0.0013**		**-0.0139**
	(0.0026)	(0.0055)		(0.0104)		(0.0094)
2018	**0.0123 *****	**-0.0149 ****	**0.0063**	**0.0155**		**-0.0191 ***
	(0.0029)	(0.0060)	(0.0053)	(0.0102)		(0.0102)
2019	**0.0144 *****	**-0.0057**	**0.0050**	**0.0116**	**0.0014**	**-0.0144**
	(0.0027)	(0.0057)	(0.0056)	(0.0110)	(0.0091)	(0.0108)
2020	**0.0202 *****	**-0.0036**	**-0.0047**	**0.0057**	**0.0085**	**-0.0218 ***
	(0.0028)	(0.0061)	(0.0060)	(0.0112)	(0.0083)	(0.0117)
2021	**0.0149 *****	**-0.0072**	**-0.0021**	**0.0212**	**0.0091**	**-0.0089**
	(0.0036)	(0.0064)	(0.0073)	(0.0129)	(0.0135)	(0.0122)
nobs	23,525	15,807	6,325	3,446	931	1,852
Instruments	E [2:99]	E [2:99]	E [3:99]	E. R [2:99]	E [4:99]	E. R [2:99]
Hansen-Sargan	173.1740	182.0540	168.2210	166.6630	118.6160	168.4380
[p-valor]	*[0.2397]*	*[0.1976]*	*[0.2076]*	*[0.5461]*	*[0.4569]*	*[0.5339]*
AR1	-9.39	-5.79	-5.60	-3.04	-1.68	-2.51
[p-valor]	*[0.0000]*	*[0.0040]*	*[0.0000]*	*[0.0023]*	*[0.0921]*	*[0.0121]*
AR2	-1,6	-1,2	-2,06	1,15	1,96	0,34
[p-valor]	*[0.1087]*	*[0.2308]*	*[0.0396]*	*[0.2492]*	*[0.0495]*	*[0.7307]*
Wald Coef {GL}	1870,25 {10}	959,89 {10}	1673,2 {10}	298,58 {10}	648,54 {9}	873,45 {10}
[p-valor]	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*	*[0.0000]*
Wald Time {GL}	55.00 {5}	9.06 {5}	5.23 {4}	8.37 {5}	1.18 {3}	4.29 {5}
[p-valor]	*[0.0000]*	*[0.1067]*	*[0.2641]*	*[0.1372]*	*[0.7590]*	*[0.5080]*

Notes: the same of Table 12.