

Série de
**TRABALHOS
PARA DISCUSSÃO**

Working Paper Series

ISSN 1518-3548

586

Outubro 2023

A Influência do Setor e da Participação no Mercado de Valores Mobiliários sobre o Endividamento Empresarial: evidências de uma grande base de dados brasileira

Flavio Alberti Docha, Luiz Betina Petroll Rodrigues

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 586	Outubro	2023	p. 3-76
----------------------	----------	---------	---------	------	---------

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.

Although the Working Papers often represent preliminary work, citation of source is required when used or reproduced.

The views expressed in this Working Paper are those of the authors and do not necessarily reflect those of the Banco Central do Brasil.

As opiniões expressas neste trabalho são exclusivamente do(s) autor(es) e não refletem, necessariamente, a visão do Banco Central do Brasil.

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 is 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 Brazil¹.

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%, while others use almost exclusively equity capital. CVM companies are more uniform in term of indebtedness levels.

¹ The same is true if we restrict the sample to non-financial business entities (NFEs).

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 CVM², 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%, enquanto outras utilizam quase que somente capital próprio. Empresas CVM são mais homogêneas em termos de níveis de endividamento.

² O mesmo vale se restringirmos a amostra à empresas não financeiras (ENFs).

A Influência do Setor e da Participação no Mercado de Valores Mobiliários sobre o Endividamento Empresarial: evidências de uma grande base de dados brasileira*

Flavio Alberti Docha **

Luiza Betina Petroll Rodrigues ***

Resumo

Este artigo apresenta base de dados inédita, “Alexandria”, que consolida um painel com mais de 40 mil empresas brasileiras, das quais 31 mil não financeiras (ENFs). Em seguida, utiliza-a para investigar a influência do “setor de atividade” e da “participação no mercado de valores mobiliários” no nível de endividamento das ENFs no período de 2015 a 2021. As conclusões são as seguintes: (i) o setor é importante para explicar o nível de endividamento médio e mediano, especialmente entre empresas não-CVM, e os setores "Agricultura", "Construção" e "Serviços Imobiliários" são os menos endividados; (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) existe grande heterogeneidade no endividamento, mesmo dentro do mesmo setor, em especial entre empresas fora do mercado de valores mobiliários. Estes resultados sugerem que estudos baseados unicamente em empresas listadas não conseguem captar a heterogeneidade de estratégias de endividamento empresarial.

Palavras-chave: base de dados, estrutura da capital

Classificação JEL: G300, G320, M410

Este Trabalho para Discussão não deve ser citado como representando as opiniões do Banco Central do Brasil. As opiniões expressas neste trabalho são exclusivamente do(s) autor(es) e não refletem, necessariamente, a visão do Banco Central do Brasil.

* Os autores agradecem (i) à empresa Klooks pelo compartilhamento da mediana de endividamento por seção CNAE de toda a sua base, o que permitiu comparar e discutir possíveis vieses de amostra de Alexandria; (ii) aos inúmeros colegas do Banco Central do Brasil que tornaram viável a construção de Alexandria, em especial Fernando Alberto G. Sampaio Rocha; (iii) aos colegas do BCB/DEPEP, em especial Thiago Christiano Silva e Gabriel Garber, pelas valiosas contribuições.

** Departamento de Estatísticas (Dstat), Banco Central do Brasil (BCB). Flavio.docha@bcb.gov.br

*** Departamento Econômico (Depec), Banco Central do Brasil (BCB). Luiza.rodrigues@bcb.gov.br

1. Introdução

A pesquisa acadêmica sobre empresas é largamente baseada em dados de empresas que negociam em bolsas de valores (a partir de agora, simplesmente “listadas”). Um número menor de estudos incluem empresas que, embora não listadas, emitam outros títulos mobiliários (como debêntures), e menos estudos ainda incluem empresas fora do mercado de valores mobiliários. Muitas vezes, implícita ou explicitamente, pretende-se que esses estudos reflitam a realidade das empresas. Entretanto, especialmente em países onde o mercado de valores mobiliários – incluindo o mercado de ações e o de títulos - é pouco expressivo, as conclusões desses estudos podem não se aplicar ao restante das empresas.

Este artigo apresenta **Alexandria**, uma base de dados inédita de entidades contábeis brasileiras construída pelos autores deste trabalho e que cobre o período de 2013T1 a 2021T4. Em sua primeira versão, Alexandria traz informações sobre 42.143 entidades contábeis, das quais 31.233 são empresas não financeiras (ENFs), incluindo todas as empresas que negociam valores mobiliários, listadas ou não, que chamaremos “empresas CVM”³, e milhares de empresas fora do mercado de valores mobiliários, que chamaremos “empresas não-CVM”. Uma de suas grandes contribuições à pesquisa contábil e financeira no Brasil é que, ao contrário do que ocorre com outras bases comumente utilizadas, em Alexandria as empresas não-CVM são mais representativas que empresas CVM em quantidade, ativo, patrimônio líquido e em saldo de crédito. A segunda grande contribuição é que Alexandria contém também dados não contábeis, como informações sobre os sócios, empregos formais e fluxos financeiros (de recebimento).

O artigo utiliza Alexandria para fazer um estudo exploratório sobre um tema largamente estudado: a “estrutura de capital”, termo que pode ser entendido como a combinação de capital próprio e de terceiros que está sendo utilizada em um determinado momento para financiar os ativos da organização.

³ O nome foi escolhido porque no Brasil a regulação e fiscalização dos valores mobiliários é de responsabilidade da CVM - Comissão de Valores Mobiliários (CVM, 2022), equivalente à SEC (*U.S. Securities and Exchange Commission*) nos EUA. Alguns poucos valores mobiliários podem ser negociados sem registro na CVM (um exemplo é um título mobiliário emitido privadamente).

O “endividamento geral” - a partir de agora simplesmente “endividamento” - é um índice frequentemente utilizado para analisar a estrutura de capital (Bressan et al, 2009, e Avelar et al, 2019)⁴. Ele é definido como a relação entre capital de terceiros e ativo, ou seja, é dado por (Sant’Ana, 2001):

$$\text{Endividamento geral}(EG) = \left(\frac{\text{Passivo circulante} + \text{Passivo não circulante}}{\text{Ativo total}} \right) * 100 \quad (1)$$

Este trabalho analisará o endividamento de empresas não financeiras⁵ (ENFs)⁶ para verificar se o nível de endividamento é influenciado pelo o setor e/ou pela participação no mercado de valores mobiliários. No caso desta última variável, sobre a qual não encontramos estudos anteriores, nossa hipótese inicial é que o mercado de valores mobiliários tende a oferecer menores custos de captação de capital de terceiros. Assim, seria de se esperar que as empresas que participam desse mercado desejam (e conseguem) um grau maior de endividamento.

Além da presente introdução, o restante do trabalho está organizado da seguinte forma. Como a maioria dos estudos foca exclusivamente em empresas listadas, a segunda seção apresenta uma revisão bibliográfica sobre as principais bases contábeis do mundo que contêm dados de empresas não listadas e sobre endividamento de empresas não listadas. A terceira seção analisa a representatividade das entidades empresariais não financeiras (ENFs) de Alexandria no Brasil. A quarta seção analisa o endividamento das ENFs entre 2015 e 2021 utilizando técnicas paramétricas e não paramétricas, e a quinta

⁴ Veja Azevedo (2013, p. 30-33) para uma discussão das vantagens e desvantagens desse índice para estudar a estrutura de capital das entidades. Veja também Parsons e Titman (2008, p. 6-7), que discute as vantagens e desvantagens de usar dados contábeis em vez de valor de mercado.

⁵ Para ser rigoroso, o trabalho analisará o endividamento de entidades contábeis formais (ou seja, registradas junto à Receita Federal do Brasil e por ela identificadas pelo CNPJ) com demonstrativos contábeis padronizados (no formato IFRS), publicados e/ou entregues a algum órgão público. As entidades contábeis a serem analisadas são do tipo “entidade empresarial formal não financeira residente no Brasil” - por concisão, simplesmente “empresa não financeira”. Os autores reconhecem que existem entidades empresariais que são pessoa física (ex: microempreendedor individual) e outras que são pessoa jurídica (como as sociedades por ações). Mesmo assim, entendemos que trocar “entidade empresarial” por “empresa” é uma simplificação válida.

⁶ Inúmeros trabalhos na área de finanças corporativas também focam nas empresas não financeiras. Entre os que têm esse foco e usam bases de dados que incluem empresas não listadas, estão: Daskalakis et al (2017), Demirgüç-Kunt et al (2020), Exame (2021), Banco de Portugal (2019, p. 44) e Raalte (2021).

seção apresenta as conclusões. Merece ser mencionado também o anexo (metodológico) A, que detalha a construção e as variáveis de Alexandria.

2. Revisão da literatura

A literatura teórica sobre a estrutura de capital das empresas está embasada em duas correntes principais que competem entre si pela explicação da estrutura de capital: a teoria do *Trade-off* e a teoria do *Pecking Order* (Teoria do Equilíbrio das Fontes). Não é objetivo deste trabalho aprofundar essas teorias, já bem documentadas em Azevedo (2013, p. 15-16), Bastos e Nakamura (2009) e Marinšek (2015). Mas reproduzimos aqui o resumo de Campos (2009, p. 28):

A primeira [teoria] é chamada de *Static Trade-off*, que determina que a empresa possua uma meta de endividamento e caminhe em sua direção. Essa meta seria estabelecida como resultado do confronto entre o custo e o benefício da dívida, em que o custo de falência se contraporiria ao benefício fiscal. A segunda teoria é denominada *Pecking Order*, corroborando que toda empresa segue uma tendência hierárquica ao estabelecer sua estrutura de capital. Primeiramente, a empresa daria preferência ao financiamento interno, ou seja, usando recursos do próprio caixa. Caso necessite de financiamento externo, a sequência lógica seria a de contrair dívidas, emitir debêntures e títulos conversíveis, antes de optar pela emissão de ações.

No nível empírico, a extensa literatura sobre a estrutura de capital das empresas, incluindo o endividamento geral, não chegou a um consenso sobre a melhor teoria (Pohlman e Iudícibus, 2010; Brito, Serrano e Franco, 2018). O influente artigo de Parsons e Titman (2008) faz um resumo da literatura empírica e a divide em três segmentos: (i) trabalhos que buscam os determinantes da estrutura de capital; (ii) trabalhos que investigam as mudanças na estrutura de capital e (iii) trabalhos que exploram as consequências do nível de endividamento. Este trabalho faz parte do primeiro grupo.

A literatura empírica internacional desse primeiro grupo identificou diversos possíveis determinantes da estrutura de capital: (i) dedutibilidade do gasto financeiro dos impostos, (ii) volatilidade do fluxo de caixa; (iii) tamanho da empresa; (iv) tangibilidade do ativo (liquidez); (v) razão entre o valor de mercado da empresa e seu valor patrimonial (*Market-to-Book ratio*); (vi) singularidade do produto vendido; (vii) setor; (viii) efeitos fixos da firma (Parsons e Titman, 2008); (ix) lucratividade; (x) risco empresarial, (xi) ações em bolsa (Azevedo, 2013), (xii) a personalidade do CFO (Marinšek, 2015) e até (xiii) a presença de mulheres no conselho da empresa (García e Herrero, 2021). Azevedo (2013) e Britto, Serrano e Franco (2018) revisaram a literatura brasileira sobre o tema e mostraram que a maioria destes fatores já foi testada em amostras de empresas brasileiras. Entretanto, a maioria dos trabalhos nacionais e internacionais mencionados utiliza

amostra que contém apenas empresas listadas, e suas conclusões sobre setores não são consensuais⁷.

Como quase todas os estudos utilizam apenas dados de empresas listadas, é compreensível que a importância do mercado de valores mobiliários para a estrutura de capital seja um tema pouco explorado na literatura. Por esse motivo, esta revisão de literatura muitas vezes aborda o impacto da listagem, em vez da participação no mercado de valores mobiliário, sobre o endividamento. Não encontramos trabalhos brasileiros que avaliaram o impacto do uso do mercado de valores mobiliários na estrutura de capital, e os poucos trabalhos brasileiros que abordam a influência da listagem não são consensuais em suas conclusões, como detalharemos na seção 2.3.

2.1 Bases de dados com variáveis contábeis de empresas não listadas

Entre as poucas bases de dados contábeis que contém informações de empresas fora do mercado de valores mobiliários, destacam-se as centrais de balanços europeias, normalmente focadas em sociedades não financeiras (Fernandes e Campos, 2014, p. 6).

Entre essas centrais europeias, a Central de Balanços do Banco de Portugal é uma base de dados de informação econômica e financeira sobre as empresas não financeiras portuguesas (Banco de Portugal, 2021). A partir de 2007, o reporte contábil eletrônico “Informação Empresarial Simplificada (IES)”, equivalente ao SPED⁸ brasileiro, permitiu às empresas enviar informações a diversos órgãos do governo com uma declaração única. Assim, a base passou a incluir quase toda a população de empresas portuguesas. Em 2021, essa Central continha informações de mais de 493 mil empresas não financeiras.

Entre as grandes bases privadas, merece destaque a *Orbis database*, produzida pela empresa Bureau van Dijk. Essa base contém informações contábeis de 40 milhões de empresas ao redor do mundo (Raalte, 2021) e tem maior cobertura em países

⁷ Também existe debate sobre a influência das outras variáveis, mas comentar esse debate fugiria ao tema deste trabalho.

⁸ O SPED, Sistema Público de Escrituração Digital, é um instrumento que unifica as atividades de recepção, validação, armazenamento e autenticação de livros e documentos que integram a escrituração contábil e fiscal dos empresários e das pessoas jurídicas mediante fluxo único digital de prestação de informações junto à Receita Federal do Brasil (SPED, 2020).

desenvolvidos europeus. A versão europeia dessa base, “Amadeus”, é fonte de numerosos trabalhos acadêmicos⁹.

No Brasil, a maior base de dados de demonstrativos contábeis é a da Receita Federal do Brasil (RFB). Em 2021, o órgão recebeu demonstrativos de aproximadamente 1,19 milhão de entidades contábeis (SPED, 2023)¹⁰. Entretanto, essas informações estão sujeitas a sigilo fiscal (Brasil, 1966, art. 198), o que as torna, na prática, indisponíveis para pesquisa.

A segunda grande base de dados de entidades contábeis no Brasil pertence à empresa Serasa Experian (Serasa). Somente em 2019, a empresa coletou dados contábeis de cerca de 85 mil entidades contábeis¹¹. Essa base está disponível apenas a clientes e é menos abrangente e menos detalhada que a da RFB. Como vantagem, inclui alguns balanços patrimoniais de empresas de menor porte, desobrigadas de apresentar a informação à RFB.

Outras empresas privadas, como a EMIS, a Klooks e a Balanços Patrimoniais, também coletam demonstrativos financeiros no Brasil, normalmente de diários oficiais. Por fim, vale mencionar a base *Maiores e Melhores da Revista Exame, compilada entre 1996 e 2021 pela Fipecafi* (Fipecafi, 2021; Sant’Anna, 2001) e após esse período pelo Ibmecc. Entretanto, todas são bases menores que a da Serasa.

2.2 Estudos empíricos sobre endividamento que incluem grandes bases de dados de empresas não listadas no mundo e apresentam avaliação setorial e/ou de listagem

Marinšek (2015, p. 8) utilizou dados de empresas não listadas europeias da base Orbis em obra detalhada sobre o impacto do endividamento no desempenho das empresas¹². O autor concluiu que fatores como a tangibilidade dos ativos, o tamanho da

⁹ Bajgar et al. (2020) apresentam estudo detalhado sobre as vantagens, desvantagens e cobertura da base Orbis. Ribeiro et al. (2010) informam que a OCDE comprou a base Orbis e fez um trabalho de limpeza dos dados para criar uma base própria (a base OCDE Orbis).

¹⁰ A informação se refere à quantidade de empresas que entregaram a ECD (Escrituração Contábil Digital). Simplificadamente, pode-se dizer que a ECD é obrigatória para algumas empresas de médio porte e para todas as empresas de grande porte (RFB, 2021c).

¹¹ Informação recebida por email pelos autores, em 19 de março de 2021. O provedor da informação foi um vendedor da Serasa Experian. A coleta é feita através de balanços publicados em jornais e por contato direto com as empresas.

¹² O autor utilizou a base de dados Amadeus (Bureau van Dijk, 2013, apud Marinšek, 2015, p. 8) com 8.777 empresas de 25 países europeus. Somente empresas com dados completos e consistentes no período de análise (2005 a 2011) foram mantidos, portanto o autor alerta que pode haver um “viés de sobrevivência”.

empresa, a lucratividade, o setor e o país, entre outros, são estatisticamente significativos para explicar a estrutura de capital. Para o autor, é necessário analisar separadamente fatores entre empresas (*between-firms effect*, ou o que explica a diferença entre as estruturas de capital de diferentes empresas) e intraempresas (*within-firms effect*, ou o que explica a mudança da estrutura de capital da mesma empresa):

As firmas, quando comparadas transversalmente por seu tamanho médio (o efeito entre firmas), praticamente não apresentam diferenças no endividamento. Por outro lado, **o aumento intra-firma em tamanho revela alavancagem substancial - expansões da empresa são em grande parte financiados com novas dívidas.** [Além disso, um] aumento no crescimento precisa de financiamento externo adicional (preferindo dívida em vez de novo patrimônio) e a tangibilidade tem um efeito mais forte entre as empresas, o que destaca a importância da participação média dos ativos tangíveis: as empresas que operam com ativos mais tangíveis são capazes de usar mais dívidas. (Marinšek, 2015, p. 9) [tradução com pequenas adaptações e negrito nossos]

O mesmo autor decompôs as diferenças de endividamento em um mesmo ano (p. 85) e concluiu que o setor ao qual a empresa pertence (*between firms [effect] within the same industry*) explica a maior parte dessa heterogeneidade, em consonância com ampla literatura citada. A conclusão foi a mesma usando técnicas estatísticas paramétricas e não paramétricas (p.41).

Demirgüç-Kunt et al (2020) também utilizam a mesma base de Marinšek (2015, p. 8) para estudar a estrutura de capital de empresas antes e após a crise de 2008. Os autores coletaram informações referentes a cerca de 160 mil empresas em cada ano de 2004 a 2011, localizadas em 47 países¹³. Os autores encontraram forte evidência de redução do endividamento em países desenvolvidos e em desenvolvimento, redução essa que foi particularmente aguda para empresas pequenas e/ou não listadas. Entretanto, os autores não fazem análise setorial.

Nehrebecka e Białek-Jaworska (2015), que analisam o endividamento¹⁴ com amostra de 800 mil observações de empresas polonesas no período de 1995 a 2012, mostraram que (i) empresas pequenas e médias tendem a ter endividamento menor, (ii) a política monetária influenciou pouco as decisões de endividamento das empresas, (iii) empresas constituídas sob a forma de empresas de responsabilidade limitada tendem a ser

¹³ Os números referem-se aos dados efetivamente utilizados, após filtros aplicados pelos autores para excluir empresas com ativo nulo, países com menos de 20 observações etc. Cerca de 70 empresas brasileiras são utilizadas.

¹⁴ Os autores analisaram a Participação de Capitais de Terceiros, que é dada por (Passivo Circulante + Exigíveis a longo prazo) / (Passivo Circulante + Exigíveis a longo prazo + Patrimônio líquido). Esse indicador equivale matematicamente ao endividamento geral, tema deste trabalho. No caso do artigo mencionado, as reservas de reavaliação foram deduzidas do patrimônio líquido.

menos endividadas; (iv) o setor da empresa não foi variável significativa para explicar o endividamento.

Jaworski e Santos (2021), em estudo que analisou e comparou o endividamento utilizando amostra de cerca de 51 mil empresas portuguesas e polonesas no período de 2011 a 2019, extraídas da base Orbis, concluíram que o setor, inclusive o endividamento mediano do setor ao longo do tempo, é uma variável significativa para explicar o endividamento empresarial.

Den Berg (2021) analisou amostra de 12.169 empresas não financeiras do Reino Unido, obtidas da base Orbis, com dados de 2014 a 2018, e concluiu que lucratividade, tangibilidade, liquidez, listagem, dedutibilidade da dívida e idade da empresa, além do setor, contribuem para explicar o endividamento.

Andritzky (2003) realizou estudo específico sobre o impacto da variável “setor” no endividamento. O trabalho utilizou amostra de quase oito mil empresas de sete países desenvolvidos, com dados entre 1997 e 2001, para verificar que o setor, além do país, tamanho da empresa, tangibilidade dos ativos, lucratividade e *market-to-book ratio*, são importantes para explicar o endividamento. O mesmo autor cita Scott e Martin (1975), que usaram técnicas paramétricas e não paramétricas para refutar estudos anteriores e concluir que o endividamento tem diferenças setoriais estatisticamente relevantes (algo que também será feito neste trabalho, nas seções 4.2 e 4.3).

2.3 Estudos empíricos sobre endividamento no Brasil que abordam diferenças setoriais

A literatura acadêmica que busca investigar a relação empírica entre endividamento e setor utilizando apenas dados de empresas listadas brasileiras é numerosa. Entre os trabalhos que identificaram o setor como variável estatisticamente significativa, estão: Lima e Brito (2003, apud Azevedo, 2013), Nakamura, Martin e Kimura (2004), Procianoy e Schnorrenberger (2004), Favato (2007) e Terra (2007), Bastos e Nakamura (2009) e Silva (2021)¹⁵. Por outro lado, Gomes e Leal (2000, apud Azevedo, 2013) e Britto, Serrano e Franco (2018) não encontraram diferença setorial estatisticamente relevante no endividamento.

¹⁵ Silva (2021) utiliza variáveis específicas do setor, e não variáveis dummy de setor, como a maioria dos estudos.

Muito mais restrita é a literatura acadêmica brasileira que busca investigar a mesma relação empírica utilizando dados de empresas listadas e não listadas. Isso é um tanto curioso porque pelo menos os dados de empresas não listadas que negociam valores mobiliários (ou seja, empresas CVM que só negociam títulos) estão disponíveis com relativa facilidade. O mesmo não se aplica às empresas fora do mercado mobiliário (empresas não-CVM), cujos dados são de difícil obtenção e têm menor qualidade¹⁶. (em geral não são auditados – as bases “Maiores e Melhores” e “Valor Pro” são exceções, por causa do porte das empresas apresentadas¹⁷). O Quadro 1 apresenta exemplos de estudos que extrapolam o universo das empresas listadas.

Quadro 1 - Estudos empíricos sobre endividamento empresarial realizados no Brasil que abordam diferenças setoriais utilizando amostra que inclui empresas não listadas

Autor(es) (ano)	Período de análise	Tamanho e características da amostra	Base de dados utilizada	Setor é relevante para explicar endividamento?
Nakamura (1992)*	1980-1989	----	----	Sim
Pohlman e Iudícibus (2010)	2001-2003	214 empresas não financeiras e não agropecuárias de grande porte, das quais 154 não listadas	Maiores & Melhores	Sim (1)
Azevedo (2013)	2008-2011	1.081 empresas não financeiras e não agropecuárias de grande porte (2).	Maiores & Melhores	Sim
Correa et al (2013)	1999-2004	389 empresas não financeiras de grande porte.	Maiores & Melhores (3)	Sim, para alguns poucos setores
Forte, Barros e Nakamura (2013)	1994-2006	4.400 pequenas e médias empresas do estado de São Paulo (4).	Serasa Experian	Sim

Fontes: compilação feita por Azevedo (2013) e elaboração dos autores. Observações: (1) a variável se mostrou significativa para explicar o endividamento de curto prazo. (2) Cerca de metade dessa amostra foi desconsiderada para a regressão feita com painel balanceado. (3) Informação inferida pelos autores*

¹⁶ Bonomo, Martins e Pint (2003), por exemplo, estudaram a composição da dívida de empresas brasileiras e o efeito das taxas de câmbio sobre elas. Os autores dispunham de dados de empresas não listadas da base Austin Asis, mas acabaram desistindo de usar os dados das não listadas, porque a investigação não foi “frutífera”. Os autores atribuem esse fracasso ao fato de os demonstrativos de não listadas não serem auditados, e por isso decidiram continuar apenas com dados de empresas listadas.

¹⁷ No Brasil, empresas de grande porte (com ativo total superior a R\$ 240 milhões ou receita bruta anual superior a R\$ 300 milhões) devem contratar auditoria externa para auditar as demonstrações contábeis. (Brasil, 2007, art. 3º). Por isso, provavelmente todas as empresas da base Maiores e Melhores são auditadas, e quase todas as empresas da base Valor Pro também. O mesmo não se aplica à base serasa Experian.

deste trabalho a partir da descrição da base; o artigo não deixa a fonte explícita. (4) A amostra inicial era de 19.272 empresas. Esse é o número após excluir empresas financeiras, outliers e empresas sem os dados completos.

Forte, Barros e Nakamura (2013) fizeram um estudo sobre endividamento utilizando empresas brasileiras não listadas utilizando uma amostra de 19.272 pequenas e médias empresas do estado de São Paulo, com dados entre 1994 e 2006, obtida da base Serasa (das quais, 4.400 foram usadas nas regressões). Além do setor, fatores como lucratividade, crescimento do ativo, tamanho, risco e idade da firma e principalmente o endividamento no período anterior mostraram-se significantes para explicar o endividamento. Alexandria difere bastante dessa base por conter (i) entidades de todas os 27 estados brasileiros; (ii) uma amostra maior; (iii) em geral, empresas de porte médio ou grande; (iv) dados de um período diferente (2013 a 2021) e (v) informações extras, como o gênero do administrador e a composição da dívida. Como desvantagem, Alexandria contém menos rubricas contábeis, o que impede por exemplo o cálculo do EBITDA.

Como são poucos os estudos que incluem empresas não listadas, a variável dummy “listada” aparece ainda menos que a variável “setor” nos trabalhos. São exceções: Pohlman e Iudícibus (2010), que concluiu que a variável não é significativa, e Azevedo (2013), que concluiu o contrário (para o autor, *ceteris paribus*, empresas listadas tendem a ter endividamento menor). Não encontramos trabalhos brasileiros que avaliaram o impacto da participação no mercado de valores mobiliários na estrutura de capital.

2.4 Razões pelas quais o setor e a participação no mercado de valores mobiliários poderiam impactar o endividamento empresarial

Como visto nas seções anteriores, os impactos do setor sobre o endividamento empresarial estão longe de ser consensuais na literatura especializada. Entre os que concordam que há algum efeito setorial no endividamento empresarial, há divergências sobre as razões pelas quais ele ocorre.

Para Pohlman e Iudícibus (2010), “as características idiossincráticas de cada setor, tais como tipo de atividade e grau de concentração, afetariam as políticas de endividamento”.

Para Marinšek (2015, p. 40), as empresas que operam no mesmo setor têm indicadores semelhantes de tangibilidade de ativo, e, conseqüentemente, uma quantidade semelhante de risco de negócios – o que determina de maneira importante a quantidade de dívida que os mercados de capitais fornecerão. O autor cita literatura suplementar que

menciona outras variáveis em comum dentro do mesmo setor, como natureza da competitividade, estado tecnológico e regulação, que também influenciam empresas do mesmo setor a ter padrões semelhantes de endividamento. Assim, o endividamento mediano do setor seria um poderoso preditor do endividamento de uma empresa.

Forte, Barros e Nakamura (2013), citando Frank e Goyal (2008), vão na mesma linha e mencionam que o setor é um importante preditor do endividamento (pelo menos para empresas listadas) e pode captar muitos fatores omitidos, como restrições regulatórias específicas do setor e a influência do tipo de atividade comercial nos requisitos ao financiamento de terceiros.

De fato, o “fator setorial” não parece captar apenas a tangibilidade dos ativos: Raalte (2021), Marinšek (2015) e Jaworski e Santos (2021, p. 32) adicionaram a variável “tangibilidade” e as *dummies* setoriais continuaram significantes.

Zonenschain (1998) entende que, em parte, o fato de o endividamento diferir entre setores decorre da especificidade do produto produzido ou fornecido pela empresa: quando mais específico, menor seria o uso de endividamento. Outros autores, revisados por Den Berg (2021), mencionam que as empresas do mesmo setor tendem a ter uma estrutura de capital ótima em comum. Isso sugere o que o próprio Raalte (2021, p.28) encontrou em revisão da literatura: a variável “setor” reflete muitas variáveis omitidas, comuns a todas as firmas de um setor, como o grau de competição e as condições de oferta e demanda.

Copat (2009) faz revisão aprofundada da literatura sobre as razões de o setor afetar o endividamento empresarial. O autor critica o uso de *dummies* para controlar os efeitos setoriais: “são muito pouco informativas, pois não mostram quais são as características específicas de determinado setor que fazem com que o endividamento das empresas que nele atuam seja maior ou menor do que o das demais” (p. 20). O autor sugere o uso de seis variáveis específicas dos setores, calculadas em geral a partir de subcontas do Balanço Patrimonial e da Demonstração do Resultado do Exercício (DRE). Entretanto, os resultados mostraram que apenas três destas variáveis se mostraram importantes para explicar o endividamento de empresas listadas nos países estudados: concentração setorial, ciclo de vida do setor e poder de barganha dos clientes (p. 135).

No tocante à relação entre endividamento e listagem, Den Berg (2021) e Raalte (2021) revisam a literatura e avaliam que o sinal da relação não é clara. Por um lado,

empresas listadas são mais conhecidas, enfrentam custos de juros mais baixos sobre a dívida, além de terem maior poder de barganha em relação a bancos e outras instituições. Assim, esperar-se-ia que fossem mais endividadas. Raciocínio semelhante ao aplicado para a variável “listagem” poderia ser aplicado para a variável “participação no mercado de valores mobiliários”. Uma empresa CVM, listada ou não, tende a enfrentar custos mais baixos de captação, portanto seria natural esperar que elas desejassem um grau maior de endividamento.

Por outro lado, para Den Berg (2021), o custo de emissão de ações/quotas é maior para empresas não listadas, pois é muito difícil vender uma ação de uma empresa não listada. Nesse sentido, empresas listadas seriam menos alavancadas, pois a emissão de ações é mais barata. Raalte (2021) ainda apresenta estudos que concluíram que variáveis como tamanho da empresa, crescimento, risco do negócio e lucro afetam de maneira diferente as empresas listadas e as não listadas. Assim, é possível que, dentre as empresas CVM, as listadas tenham endividamento menor que as não listadas.

2.5 Técnicas econométricas utilizadas em estudos que testam o setor e/ou a listagem como determinantes do endividamento empresarial e dificuldades relacionadas

As dificuldades econométricas envolvidas na identificação dos determinantes do endividamento empresarial são muitas, e já foram bem documentadas por Parsons e Titman (2008), Bastos e Nakamura (2009), Raalte (2021) e Miniaci e Pantegini (2021). Por causa dessas dificuldades, a literatura acadêmica nacional e internacional utiliza técnicas variadas.

O Quadro 2 traz exemplos de trabalhos que utilizam amostra de empresas listadas e não listadas e resume as técnicas utilizadas. Destacam-se as técnicas de regressão em painel, em especial GMM (Generalized Method of Moments, ou Método Generalizado dos Momentos) e efeitos fixos, mas mesmo entre essas técnicas há muito debate, como pode ser verificado em Dang et al (2015) e Barros et al. (2020), respectivamente contra e a favor de modelos GMM com variáveis instrumentais¹⁸.

Quadro 2 – Estudos sobre endividamento empresarial que utilizam amostras que incluem empresas não listadas: técnicas econométricas, variáveis dependentes e critérios de exclusão de dados

¹⁸ Gaibulloev et al (2014 p. 261) afirmam que esse método funciona mal quando o coeficiente da variável defasada se aproxima da unidade, pois as variáveis instrumentais se tornam muito fracas.

Autor	Técnica usada	Variável dependente	Crítérios para exclusão das empresas da regressão/ tratamento de <i>outliers</i>
Sant' Ana (2001)	Regressão linear simples e quadrática, ano a ano		
Andrtzky (2003)	POLS	EG	<ul style="list-style-type: none"> • Endividamento > 100%
Pohlman e Iudícibus (2010)	POLS	EG (1); ECP(1); ELP(1)	<ul style="list-style-type: none"> • Empresas que apresentaram prejuízo médio no período • Empresas para os quais o IRPJ e a CSLL, somados, apresentam valor positivo (suposta "receita" com tributos)
Azevedo (2013)	Regressão em painel e múltipla	EG, ECP e ELP	<ul style="list-style-type: none"> • Empresas que não continham todos os dados de todos os anos (painel balanceado) • Empresas com EBITDA negativo (p. 37).
Correa et al (2013)	Painel balanceado, efeitos fixos + efeitos aleatórios + GMM-Sys	EG	<ul style="list-style-type: none"> • Financeiras
Forte, Barros e Nakamura (2013)	Painel desbalanceado - POLS e GMM-Sys -	EG	<ul style="list-style-type: none"> • Endividamento > 100% • Winsorização
Marinšek (2015)	Painel balanceado- MLM (<i>Multilevel linear modeling</i> , ou Modelo Linear Múltiplo)	EG; EF (p. 35)	<ul style="list-style-type: none"> • Ativo menor que 5 milhões de euros • Endividamento > 100% • Empresas financeiras • 9 critérios para exclusão de outliers • Winsorização em 1%, entre 0 e 86,18%
Nehrebecka e Białek-Jaworska, (2015)	Painel desbalanceado - GMM-Sys	PEG	<ul style="list-style-type: none"> • Endividamento > 100%
Demirgüç-Kunt et al (2020)	GLS (<i>Generalized Least Squares</i> , ou Mínimos quadrados generalizados). Prais-Winsten estimator.	EG; ELP; PLP	<ul style="list-style-type: none"> • Firms com endividamento ou ativo zerado. • Firms com menos que 5 empregados • Microempresas • Firms que não têm pelo menos seis anos consecutivos na base • Financeiras, estatais • Países com menos de 20 empresas em cada ano, empresas de Taiwan, empresas de países que são paraísos fiscais • Empresas com financiamento de longo prazo antes da crise zerado
Raalte (2021)	Painel desbalanceado - POLS	EG, ECP, ELP	<ul style="list-style-type: none"> • Empresas com menos de 5 anos de dados • Financeiras • Winsorização em 5% (retira 2,5 superiores e 2,5 inferiores, exceto <i>dummies</i>) • Considera somente empresas grandes ou muito grandes (receita operacional superior a 10 milhões de euros, ou ativos totais superiores a 20 milhões de euros ou mais que 150 empregados)

			<ul style="list-style-type: none"> • Redução do número de não listadas de maneira que a mediana do ativo de listadas e não listadas seja semelhante.
--	--	--	---

Fonte: Elaborado pelos autores. *POLS = Pooled Ordinary Least Squares (Mínims Quadrados Ordinários Agrupado)* *GMM-Sys = S-GMM é o estimador de Blundell R, Bond S (1998)*. $EG = (\text{passivo total})/(\text{ativo total})$; $ECP = (\text{passivo circulante})/(\text{ativo total})$; $ELP = (\text{passivo exigível a longo prazo})/(\text{ativo total})$; $EF = (\text{Passivo financeiro})/(\text{ativo})$; $PLP = (\text{Passivo exigível a longo prazo})/(\text{passivo total})$. $PEG = (\text{passivo total})/(\text{passivo total-reservas de reavaliação})$. (1) o ativo total foi ajustado pelos efeitos da inflação.

Mergulhar nesta discussão econométrica foge ao escopo desse trabalho. Entretanto, os autores entenderam importante abordar três questões essenciais para estudos econométricos na área de finanças corporativas: (i) o critério de exclusão de dados, (ii) o fato de a variável dependente ser (na maioria dos estudos) uma razão com transformação logarítmica e (iii) endogeneidade.

2.5.1 Critérios de exclusão de dados

Além de detalhar as técnicas econométricas utilizadas por cada trabalho sobre os determinantes do endividamento empresarial, o Quadro 2 também destaca as diferentes escolhas de exclusão de dados utilizadas. Em primeiro lugar, a maioria dos trabalhos exclui empresas financeiras, inclusive as sociedades de participação (*holdings*). Embora essa exclusão faça sentido, é preciso ter em mente que há empresas não financeiras (ENFs) detidas por estas *holdings*, e estas ENFs detidas dificilmente estarão em qualquer análise, pois os balanços individuais das empresas detidas raramente são divulgados. Se estivermos certos, as amostras disponíveis para trabalhos de finanças corporativas contêm menos ENFs detidas por *holdings* do que uma amostra representativa de ENFs recomendaria.

Em segundo lugar, pela dificuldade de lidar com dados zerados, alguns trabalhos excluem empresas sem endividamento (ou seja, que utilizam apenas capital próprio em determinado ano)¹⁹. Em terceiro lugar, pela dificuldade de lidar com dados negativos, muitos trabalhos excluem todos os dados de empresas com endividamento maior que 100% (portanto com patrimônio líquido negativo). Nehrebecka e Białek-Jaworska (2015, p. 9) dizem que a exclusão desses dados é uma das principais limitações de seu estudo.

¹⁹ A exclusão de empresas com passivo zerado não é problema para os trabalhos que utilizam somente dados de empresas que participam do mercado de valores mobiliários, pois praticamente todas têm algum passivo. É provável, inclusive, que a decisão da empresa de participar do mercado de valores mobiliários reflita interesse em aumentar o uso de capital de terceiros.

Este trabalho diverge de parte da literatura por incluir na análise tanto ENFs com endividamento zero quanto ENFs com endividamento superior a 100%.

Em quarto lugar, provavelmente para excluir erros de declaração e outros *outliers*, alguns estudos descartam dados extremos pelo método de winsorização (Dang et al, 2015; Forte, Barros e Nakamura, 2013; Raalte, 2021; Marinšek, 2015, p. 64). Este estudo seguiu essa prática apenas na análise paramétrica (seção 4.3).

Em quinto lugar, além das exclusões acima mencionadas, os estudos que utilizam painéis balanceados forçosamente excluem todas as observações de empresas que não têm dados para todos os anos, gerando assim um “viés de sobrevivência” (Demirgüç-Kunt et al; 2020; Marinšek; 2015). Em outras palavras, painéis balanceados excluem empresas novas (que iniciam suas atividades em momento posterior ao início do período de análise), empresas que pararam de atuar no meio do período de análise (inclusive por excesso de endividamento e falência), e as que, por outros motivos, não têm demonstrativo financeiro em todos os anos na base²⁰. Como esse estudo demonstra na seção 4.3, o uso de painel balanceado pode alterar as conclusões.

2.5.2 Variável dependente como um logaritmo de uma razão

A variável de interesse deste trabalho é uma razão - passivo/ativo -, que assume sempre valores não negativos.

Segundo influente artigo de Firebaugh e Gibbs (1985), a questão da utilização de uma razão como variável dependente de uma regressão econométrica suscitou grandes debates na literatura. Os autores apontam, entretanto, que essa utilização pode ser feita sem maiores ajustes no modelo desde que o intercepto seja zero (aqui, se o passivo for zero quando o ativo for zero, o que é verdade em quase todos os casos pela identidade contábil²¹).

Bartlett e Partnoy (2020), entretanto, sugerem que os trabalhos que utilizem uma razão como variável dependente devem adicionar algumas variáveis adicionais para evitar viés. Assim, suponha que “ x_i ” seja uma variável qualquer que supostamente explique a

²⁰ A exclusão de empresas para formação de um painel balanceado pode ser um problema especialmente grande na análise de empresas com duração curta e pré-determinada, como as sociedades de propósito específico formadas para construção e venda de empreendimentos imobiliários (enquadradas em geral no grupo 681 da classificação ISIC/CNAE – veja IBGE, 2021, para mais informações sobre essa classificação).

²¹ Existe apenas um caso em que o ativo é zero mas o passivo não: quando o patrimônio líquido é negativo e exatamente igual ao passivo com terceiros em módulo, isto é, (passivo circulante + passivo não circulante)=(patrimônio líquido). Entendemos que é um caso tão raro que pode ser ignorado.

razão passivo/ativo. Os autores afirmam que, em vez de utilizar uma equação como a 2, deve-se utilizar equação como a 3, que permite a possibilidade de a variável explicativa “ x_i ” estar relacionada ao passivo, independentemente do ativo.

$$\frac{Passivo_i}{Ativo_i} = \beta_0 + \beta_1 x_i + \epsilon_i \quad (2)$$

$$\frac{Passivo_i}{Ativo_i} = \beta_0 + \beta_1 x_i + \underbrace{\beta_2 \left(\frac{1}{Ativo_i} \right)}_{\substack{\text{desnecessário} \\ \text{se passivo for} \\ \text{0 quando ativo} \\ \text{for 0}}} + \beta_3 \left(\frac{x_i}{Ativo_i} \right) + v_i \quad (3)$$

Um outro ponto importante na literatura de endividamento empresarial é a utilização de transformação logarítmica da variável dependente, que se subdivide em duas questões: (i) se o uso da especificação logarítmica é teoricamente apropriado e (ii) quais são as alterações necessárias na especificação do modelo para acomodar a transformação logarítmica.

Em relação à primeira questão, a transformação logarítmica é normalmente utilizada para lidar com dados com *outliers*, com assimetria positiva e/ou com muita dispersão. Além disso, a transformação permite interpretar os coeficientes como elasticidades ou semi-elasticidades. Entretanto, essa decisão deve ser cuidadosa, pois a transformação logarítmica não é sempre a melhor escolha e traz custos ao pesquisador, pois dificulta a interpretação dos resultados da regressão e os testes de hipóteses (Lee, 2020, e Feng et al, 2014).

Em relação à segunda questão – alterações necessárias na especificação do modelo para acomodar a transformação logarítmica-, Bartlett e Partnoy (2020) afirmam que essa transformação, aplicada a uma razão, evita o problema da variável omitida da equação 2 (pois a equação 4 é matematicamente equivalente à equação 5).

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

$$\ln[Passivo_i] = \beta_0 + \beta_1 x_i + 1 \cdot \ln[Ativo_i] + \epsilon_i \quad (5)$$

Entretanto, continuam os mesmos autores, as equações 4 e 5 assumem uma relação de um-para-um entre $\ln(Passivo_i)$ e $\ln(Ativo_i)$, o que raramente é verdade. Assim, os

autores sugerem a utilização da equação 6 (que inclui um controle, o logaritmo do denominador como variável dependente) em vez da equação 4.

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

Finalmente, como mencionado na seção 2.5.1, muitos autores que modelaram econometricamente os determinantes do endividamento empresarial optaram por excluir da amostra empresas com endividamento zero ou maior que um, inclusive porque a transformação logarítmica exige valores não nulos. Bellégo, Benatia e Pape (2019) afirmam que este procedimento adiciona viés de seleção, ou, no mínimo, muda o escopo do estudo, ou. A literatura aponta alternativas utilizadas para evitar a exclusão de dados, mas nenhuma é mencionada como a ideal:

- adicionar uma constante antes de fazer a transformação logarítmica, de modo que a variável dependente Y_{it} é transformada em $\tilde{Y}_{it} = \ln(Y_{it} + c)$ (Feng et al, 2014)²². Utilizamos esta alternativa nos testes paramétricos, com $c=1$.
- utilizar a transformação IHS, em que a variável dependente Y_{it} é transformada em $\tilde{Y}_{it} = \left\{ \ln \left[\theta Y_{it} + \sqrt{\theta^2 Y_{it}^2 + 1} \right] \right\} \div \theta$ (sendo que normalmente $\theta = 1$) (Bellégo, Benatia e Pape, 2019)²³;
- utilizar um modelo Poisson (Bellégo, Benatia e Pape, 2019).
- utilizar a transformação de Box-Cox, em que $\tilde{Y}_{it} = \begin{cases} \frac{(Y_{it} + \lambda_2)^{\lambda_1 - 1}}{\lambda_1} & \text{se } \lambda_1 \neq 0 \\ \log(Y_{it} + \lambda_2) & \text{se } \lambda_1 = 0 \end{cases}$ (sendo que normalmente $\lambda_2 = 0$) (Atkinson et al, 2021).

2.5.3 Endogeneidade

A endogeneidade, que descreve a situação em que uma variável explicativa está correlacionada ao erro, ocorre geralmente na forma de variáveis omitidas, erro de mensuração, simultaneidade (Wooldridge, 2010, p. 50) ou grau de persistência

²² Veja Lee (2020), Feng et al (2014) para uma discussão detalhada das desvantagens da adição da constante. Bellégo, Benatia e Pape (2019) calculam o viés de endogeneidade trazido por essa escolha e comentam que esse procedimento também diminui os erros padrão.

²³ Veja Aihounon e Henningsen (2019) para uma discussão detalhada do viés trazido pela transformação IHS (*inverse hyperbolic sine*) e sua relevância dependendo da unidade de mensuração.

significativo²⁴ (endogeneidade dinâmica). Barros et al (2020) descrevem como todas essas formas aparecem em painéis de finanças corporativas. Acrescentamos aqui apenas os complicadores associados a bases de dados que contêm empresas fora do mercado de valores mobiliários, que se referem a omissão de variáveis e erros de mensuração.

As bases de dados que contêm apenas empresas que negociam valores mobiliários normalmente contêm mais variáveis que as bases que incluem empresas que não os negociam. Com isso, é natural pensar que o problema de omissão de variáveis seja ainda mais importante em estudos que usam o segundo tipo de base de dados. Por causa desse problema, muitos desses estudos reduzem significativamente a amostra inicial e constroem modelos apenas com as empresas que contêm todas as rubricas desejadas²⁵. Essa estratégia, embora compreensível, cria outro problema (o viés de seleção).

Em relação a erros de mensuração, os autores desse trabalho acreditam que o passivo das empresas não-CVM é normalmente mensurado com mais acurácia que o ativo. Isso ocorreria, também na opinião dos autores, porque as dívidas são recalculadas periodicamente de acordo com contratos, enquanto o ativo, especialmente o ativo realizável a longo prazo, está sujeito à avaliação contábil menos precisa no caso de empresas menores:

- As empresas CVM e as empresas não-CVM de grande porte estão sujeitas a atendimento de regras contábeis que, muitas vezes, exigem avaliação do seu ativo a valor justo ou a valor recuperável.
- Já as empresas não-CVM de menor porte tendem a registrar o ativo com base em métodos de custo amortizado ou custo histórico.

Assumindo um cenário (comum) em que o custo histórico é menor que o preço de mercado do ativo, o endividamento das empresas não-CVM de menor porte *pareceria mais alto do que realmente é*, pois o numerador (passivo) é atualizado periodicamente (com novas dívidas e juros de dívidas anteriores), enquanto parte do denominador (ativo) tem seu valor sempre reduzido pela depreciação. Entretanto, não encontramos referência na literatura sobre esse tema.

²⁴ A maioria dos autores mencionados nas seções 2.2 e 2.3 verificaram que o endividamento do período anterior é uma importante variável para explicar o endividamento atual (o que poderia acontecer, por exemplo, porque a empresa não consegue se livrar das dívidas rapidamente ou porque mantém ao longo do tempo um nível de endividamento que considera adequado). Entretanto, a inserção da variável dependente defasada como variável explicativa viola a hipótese da exogeneidade estrita (Wooldridge, 2010, p. 256).

²⁵ É o caso de Forte, Barros e Nakamura (2013), Marinšek (2015, p. 37) e Raalte (2021).

Barros et al (2020) fazem simulações de Monte Carlo com painéis de empresas para verificar como as diferentes fontes de endogeneidade alteram os coeficientes das regressões. Os autores concluíram que os estimadores OLS, efeitos fixos e efeitos aleatórios podem ser inconsistentes na presença de problemas de endogeneidade que são muito prováveis no contexto de finanças corporativas. Por outro lado, os estimadores GMM (em especial o GMM-Sys) se bem especificados, são capazes de lidar adequadamente com esses problemas.

3. Empresas não financeiras (ENFs) analisadas neste trabalho: representatividade

Conforme já mencionado, os autores deste trabalho construíram uma base de dados - que denominaram Alexandria - sobre entidades formais residentes no Brasil, identificadas por seu número de identificação, o CNPJ. Alexandria contém uma rica combinação de informações cadastrais, contábeis, societárias, creditícias (inclusive em relação à dívida externa) e de fluxos bancários. Os demonstrativos contábeis, que representam a maior novidade da base, foram compilados a partir de diferentes fontes, com longa averiguação manual para diminuir inconsistências. A estrutura da base e suas variáveis, inclusive as que não foram utilizadas no exercício empírico deste trabalho, são apresentadas pela primeira vez no anexo A.

As empresas não financeiras (ENFs) representam a maioria das entidades contidas em Alexandria e são o foco do estudo exploratório deste trabalho. Este capítulo se destina a comparar essas ENFs de Alexandria com a população de ENFs do Brasil para avaliar sua representatividade.

3.1 Escopo do exercício empírico: dados anuais de empresas não financeiras (ENFs) de 2015 a 2021

As ENFs de Alexandria serão divididas em doze grupos, de acordo com o setor e o registro no mercado de valores mobiliários, conforme Quadro 3.

Quadro 3– Empresas não financeiras (ENFs): definição dos setores utilizados neste trabalho e nome das variáveis que identificam os grupos

Código do setor utilizado neste trabalho	Descrição do setor	Seções CNAE/ISIC incluídas	Nome do subgrupo de ENFs	
			Registradas na CVM	Não registradas na CVM

Agr	Agropecuária	A	<i>CVM_Agr</i>	<i>NCVM_Agr</i>
Ind	Indústria (Extrativa, de transformação, eletricidade e gás)	B a E	<i>CVM_IND</i>	<i>NCVM_IND</i>
Constr	Construção	F	<i>CVM_Constr</i>	<i>NCVM_Constr</i>
Com	Comércio e reparação de veículos	G	<i>CVM_Com</i>	<i>NCVM_Com</i>
Imob	Serviços imobiliários	L	<i>CVM_Imob</i>	<i>NCVM_Imob</i>
Serv	Serviços, exceto financeiros, domésticos, imobiliários e administração pública	H a S, exceto K, L e O	<i>CVM_SERV</i>	<i>NCVM_SERV</i>

Fonte: elaborado pelos autores. Para uma descrição detalhada da classificação CNAEs, que é correspondente à classificação internacional ISIC no nível das seções, veja IBGE (2021). Note o leitor que as ENFs consideram apenas CNPJ8 com natureza jurídica iniciada em 2, que identifica as entidades empresariais, o que retira todos os produtores rurais pessoa física. A natureza jurídica 221 também foi retirada, por identificar entidades empresariais não residentes.

Serão considerados apenas demonstrativos contábeis anuais. Essa definição inclui tanto os demonstrativos efetivamente encerrados em dezembro como aqueles que se encerram em período diferente, normalmente para coincidir com o ano safra. Além disso, foram utilizados apenas dados de 2015 a 2021²⁶.

3.2 Quantidade de empresas ENFs no Brasil e em Alexandria

Entre 2015 e 2021, o número de CNPJ8s ativos no Brasil ficou entre 17,3 e 20,8 milhões, a vasta maioria dos quais é microempresa não financeira. Entretanto, um grande número dessas entidades ativas estava, na prática, sem atividade. Por exemplo, das 19,1 milhões de ENFs ativas em 2020,

- apenas 10,7 milhões pagaram algum tributo federal durante 2020 (RFB, 2021b);
- apenas 6,6 milhões registraram algum fluxo de recebimento ao longo de 2020 e/ou tinham algum passivo bancário, externo ou em títulos no fim de 2020²⁷.

Mesmo considerando apenas as entidades que tiveram alguma atividade econômica efetiva, uma minoria produz demonstrativos contábeis, como explicado na seção 3.1. Em 2021, a RFB recebeu demonstrativos de 1,19 milhão de entidades contábeis (SPED, 2023). Assumindo que 95% dessas empresas são ENFs, podemos concluir que, das 19,6 milhões de ENFs ativas em 2021, não mais que 1,13 milhão produz demonstrativos contábeis.

²⁶ O período foi escolhido devido ao sensível aumento do tamanho da amostra a partir de 2015 (devido ao Censo quinquenal e às mudanças regulatórias que tornaram o Quadro Societário obrigatório – veja seção A.5.1).

²⁷ Veja o leitor o anexo A, item A.8 e A.7 para a definição das variáveis “fluxo de recebimento” e os diversos tipos de passivos compilados em Alexandria.

Dessas ENFs que produzem demonstrativos financeiros (no máximo 1,13 milhão, segundo estimativa acima), um número pequeno, 27.573, está em Alexandria²⁸. Mesmo assim, trata-se de amostra maior que a utilizada em qualquer outro estudo sobre endividamento empresarial no Brasil, como mostrou a seção 2.3.

3.3 Representatividade de Alexandria em termos de volume: fluxos de recebimentos identificados e crédito.

Por construção, Alexandria capta apenas empresas que produzem demonstrativos contábeis²⁹, e, portanto, a imensa maioria das microempresas e empresas de pequeno porte estão fora desta base. A Tabela 1 mostra que entre 2015 e 2021 (período de interesse deste trabalho) o ativo mediano das ENFs de Alexandria variou entre R\$ 7,83 e R\$ 46,88 milhões (coluna A), o que corresponderia a uma empresa de médio porte (empresas de grande porte são aquelas com ativo superior a R\$ 240 milhões, segundo Brasil, 2007).

Além disso, A Tabela 1 também mostra o ativo mediano apenas do grupo de ENFs para os quais Alexandria dispõe de dados de lucro líquido (coluna B), pois somente essas foram utilizadas na avaliação paramétrica da seção 4.3. Como se percebe pelos dados da coluna B, estas ENFs ainda são em geral de porte médio³⁰, mas são bem maiores que as ENFs para as quais Alexandria não dispõe do dado de lucro líquido (coluna C).

Tabela 1 – Ativo mediano e quantidade de ENFs com dados de ativo em Alexandria, por disponibilidade ou não de dados de lucro em cada data base

Ano	Ativo mediano das ENFs (R\$ milhões)			Quantidade de ENFs		
	(A) Com dados de ativo na data	(B) Das quais, com dado de lucro	(C) Das quais, sem dado de lucro	(D) Com dados de ativo na data	(E) Das quais, com dado de lucro	(F) Das quais, sem dado de lucro
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	5836	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

Fonte: Alexandria. Note o leitor que o ativo mediano de 2015 e 2020 é inferior aos dos demais anos, e a quantidade de ENFs, inclusive com dados de lucro, é maior. Isso ocorre porque nesses anos aconteceu o

²⁸ Note o leitor que o número de entidades em Alexandria é bem maior: 42.143. Destas, 42.068 são entidades empresariais, e destas, 31.233 são ENFs. Destas ENFs, 27.573 estavam ativas no fim de 2020.

²⁹ Veja a esse respeito o anexo, item A.1.

³⁰ Exceto em 2019 e 2021, provavelmente pelo efeito inflação. A definição de empresa de grande porte foi definida em 2007 pela lei em valores nominais (Brasil, 2007).

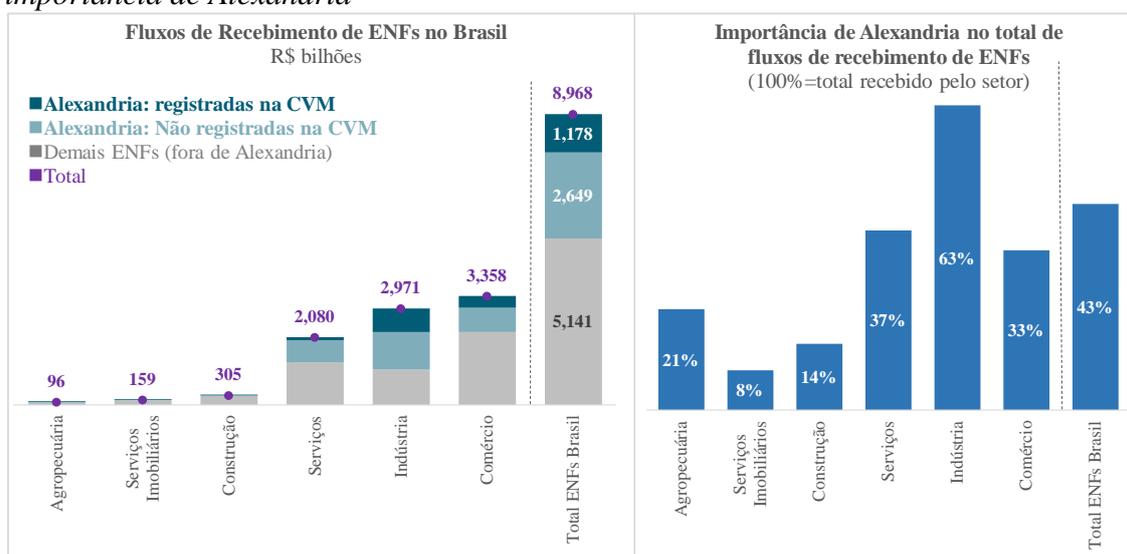
Censo Quinquenal de Capitais Estrangeiros, que têm declarações de todas as empresas detidas por não residentes, e Alexandria utilizou todos esses demonstrativos. Ver seção A.5.1.

Para estimar a importância das ENFs de Alexandria, frente ao total de ENFs ativas em cada data base, esta seção utilizou indicadores para os quais os valores totais (referentes a todas as ENFs) são conhecidos:

- Fluxos de recebimento recebidos pelas ENFs (variável descrita no item A.8 do anexo)
- Crédito ampliado às ENFs (variável descrita na seção A.7)
 - Do qual, crédito do Sistema Financeiro Nacional às ENFs (variável descrita na seção A.7.1)

A Figura 1 mostra que, dos R\$ 8,97 trilhões de fluxos de recebimento que entraram na conta bancária de todas as ENFs ao longo de 2020, 43% foi recebido por ENFs integrantes da Alexandria. A representatividade de Alexandria varia por setor: apenas 8% dos fluxos do setor de serviços imobiliários são recebidos por ENFs de Alexandria, enquanto na indústria esse número é de 63%³¹.

Figura 1 – Fluxos financeiros recebidos por ENFs no Brasil em R\$ bilhões (2020) e importância de Alexandria



Fonte: elaboração dos autores com base em dados do BCB e Alexandria. Nota: As ENFs “Alexandria: não registradas na CVM” são as ENFs ativas em 31 de dezembro de 2020, sem dados na CVM nesta data e

³¹ Na opinião dos autores, essa diferente representatividade pode refletir incentivos diferentes em cada setor. É possível, por exemplo, que o setor industrial tenha significativos ganhos de escala e incentivos para atrair capital estrangeiro e estar presente no mercado de valores mobiliários. Por isso as empresas industriais seriam em geral maiores e mais presentes na base do BCB e na base “Economática”, o que faria com que Alexandria captasse mais esse setor (veja anexo A.5.2). Mais estudos são necessários para corroborar essas hipóteses.

com pelo menos um balanço patrimonial disponível entre 2013T1 e 2021T4 (condição para estar na Alexandria). Ou seja, são consideradas todas as ENFs de Alexandria, mesmo que o balanço patrimonial não esteja disponível especificamente para 2020.

Seguindo raciocínio semelhante, a representatividade das empresas não financeiras (ENFs) de Alexandria foi analisada usando os dados de crédito. As empresas de Alexandria respondiam por 75% do saldo ampliado de crédito às ENFs³² ao final de 2020. Considerando somente os R\$ 1,47 trilhões³³ de crédito devido ao Sistema Financeiro Nacional (SFN) pelas ENFs no fim de 2020, as empresas de Alexandria respondiam por 49%.

Em resumo, embora o número de ENFs de Alexandria seja pequeno perto do número total de ENFs brasileiras, e mesmo do número de ENFs com balanço patrimonial (1,13 milhão, no máximo, em 2021, segundo nossa estimativa na seção 3.2), a representatividade na atividade econômica é bem expressiva.

É interessante notar que a importância das “ENFs CVM” em Alexandria é menor que a importância das “ENFs não CVM”, tanto em termos de fluxos de recebimento quanto em termos de crédito bancário doméstico (“crédito SFN”)³⁴. Sem dúvida, essa será uma das grandes contribuições de Alexandria para a pesquisa contábil e financeira no Brasil.

4. Painel de empresas não financeiras (ENFs): diferenças no índice de endividamento setorial

Esta seção manterá o recorte da seção 3 e continuará a analisar os demonstrativos anuais das empresas não financeiras (ENFs) por setor, no período de 2015 a 2021.

A essência dessa seção é verificar se existem diferenças estatisticamente significantes no endividamento de acordo com setor e participação no mercado de valores

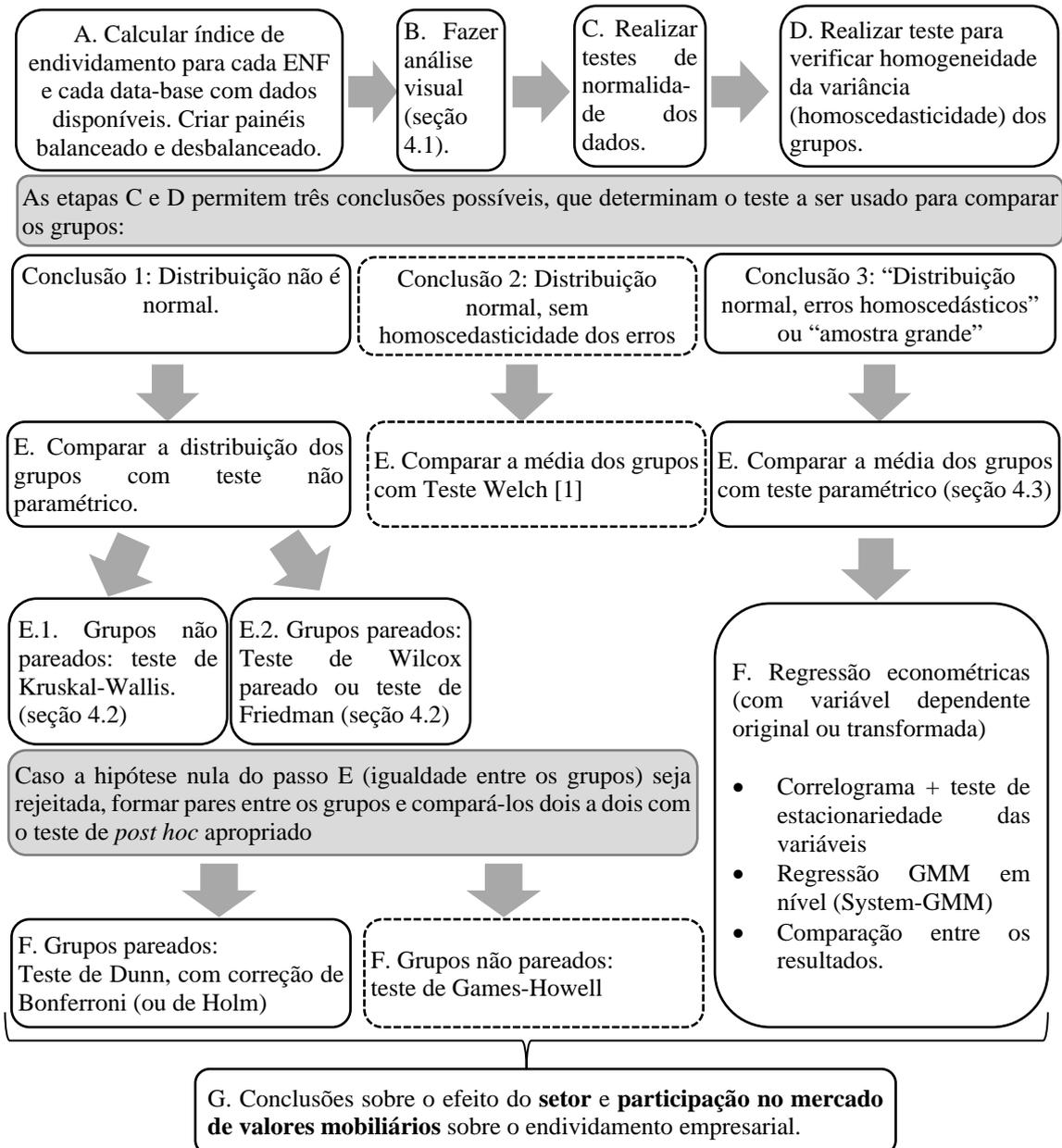
³² Esse valor elevado se explica, em grande parte, porque a amostra de Alexandria contém todas as empresas detidas por não residentes, e essas empresas têm maior facilidade de captação externa. Este resultado também parece apoiar a conclusão de Freire et al. (2021), de que “empresas com maior grau de internacionalização possuem maior endividamento”. Entretanto, estudos posteriores podem ser necessários para verificar a existência de causalidade entre as variáveis.

³³ Esse saldo é muito semelhante ao informado na série SGS 22047 - Saldo das operações de crédito por tipo de cliente - Setor privado - Pessoas jurídicas - R\$ (milhões). A diferença é que essa série inclui crédito concedido a holdings, explicitamente excluídas das ENFs.

³⁴ Entretanto, a maior importância das empresas não CVM de Alexandria não se verifica no conceito “crédito ampliado”: em dezembro de 2020, as ENFs “CVM” de Alexandria deviam R\$ 1,43 tri de crédito ampliado, enquanto as empresas “não CVM” de Alexandria deviam R\$ 1,22 trilhão. As empresas não CVM fora de Alexandria deviam o R\$ 0,87 trilhão restante.

mobiliários. As etapas dessa análise estão resumidas na Figura 2. As etapas que não foram necessárias (pois os testes levaram a outro caminho) estão com linhas tracejadas.

Figura 2 –Empresas não financeiras (ENFs) em Alexandria: etapas da análise do endividamento



Fonte: elaborado pelos autores. [1] Se a assimetria for forte, preferir o teste de Brown-Forsythe ANOVA.

4.1 Análise visual, de normalidade e homoscedasticidade

A primeira coisa notável acerca do endividamento empresarial, segundo a literatura, é que ele muda muito pouco ao longo do tempo, independente do setor. De fato, um conjunto grande de autores observa que o endividamento de um determinado ano é parecido com o do ano anterior (ver, por exemplo, Nehrebecka e Białek-Jaworska, 2015). Como muitas empresas têm dívidas cujo pagamento pode demorar mais que doze meses, essa observação pode simplesmente significar que a empresa ainda não acabou de pagar a dívida que contraiu no ano anterior. Porém, também é possível que essa estabilidade reflita fatores individuais da firma (como a escolha deliberada de manter certo patamar de endividamento, por exemplo). Para tentar separar o efeito contábil entre um ano e outro de características supostamente mais duradouras, usamos um período mais longo. Tomamos o endividamento de cada ENF em 2015 e 2020 e classificamos o endividamento de cada uma em relação ao setor em que opera em “baixo”, “normal” e “alto”. O resultado, apresentado na Tabela 2, mostra que, das 8.206 ENFs para a qual Alexandria dispunha de dados de endividamento tanto em 2015 quanto em 2020, a grande maioria não mudou de classificação entre 2015 e 2020. Em outras palavras, quem tinha endividamento considerado baixo em 2015 continuou a manter esse patamar em 2020, e o mesmo vale para as outras classificações. A mediana da variação é de 0,16 ponto percentual, o que significaria, para quem tinha endividamento de 50% em 2015, aumentar para 50,16% em 2020. O resultado reforça a percepção de que as empresas – independente do setor – tendem a manter seu padrão de endividamento ao longo do tempo, e que as empresas de Alexandria refletem o padrão já visto em outras bases de dados.

Tabela 2 - Índice de endividamento de empresas não financeiras (ENFs) em 2015 e 2020, classificadas em relação ao endividamento de seu setor.

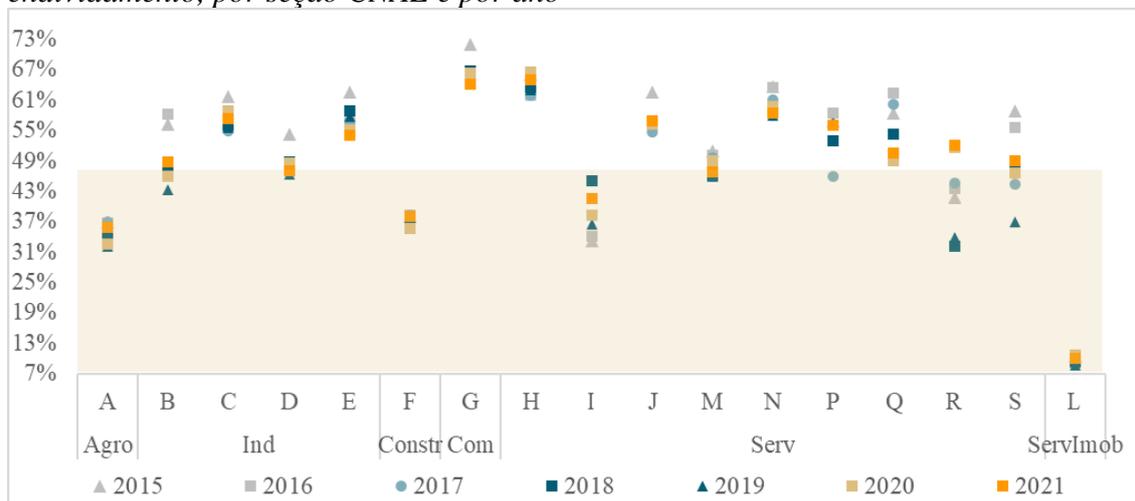
Classificação do endividamento em 2015	Classificação do endividamento em 2020			Total de ENFs
	Baixo para o setor (menor que o primeiro quartil)	Normal para o setor (entre o primeiro e o terceiro quartil)	Alto para o setor (maior que o terceiro quartil)	
Baixo para o setor (menor que o primeiro quartil)	1.154	642	97	1.893
Normal para o setor (entre o primeiro e o terceiro quartil)	696	3.185	646	4.527

Alto para o setor (maior que o terceiro quartil)	125	538	1.123	1.786
Total	1.975	4.365	1.866	8.206

Fonte: Alexandria.

Voltando a atenção agora para a análise estática, a Figura 3 apresenta o endividamento mediano das seções CNAE das ENFs de Alexandria, sem detalhamento por participação no mercado de valores mobiliários. Essa primeira análise visual sugere que o setor é uma variável importante para explicar o nível de endividamento. Em especial, três seções CNAE/ISIC apresentaram endividamento mediano inferior a 40% em todos os anos analisados: “A” (Agropecuária), “F” (Construção) e “L” (Atividades imobiliárias)³⁵. Na outra ponta, duas seções CNAE/ISIC apresentaram endividamento mediano acima de 60% em todos os anos: “G” (Comércio) e “H” (Transporte). Em linhas gerais, os valores medianos são ligeiramente maiores que as medianas da base de dados da empresa Klook (a que os autores tiveram acesso), mas o ordenamento setorial (qual setor é mais - ou menos - endividado) é muito semelhante³⁶. A diferença importante é a seção G (Comércio): na base Klook, o endividamento mediano é muito menor que na base Alexandria, e por isso nesta base privada as seções H e E, e não a G e H, são as mais endividadas.

Figura 3—Empresas não financeiras (ENFs) em Alexandria: mediana do endividamento, por seção CNAE e por ano



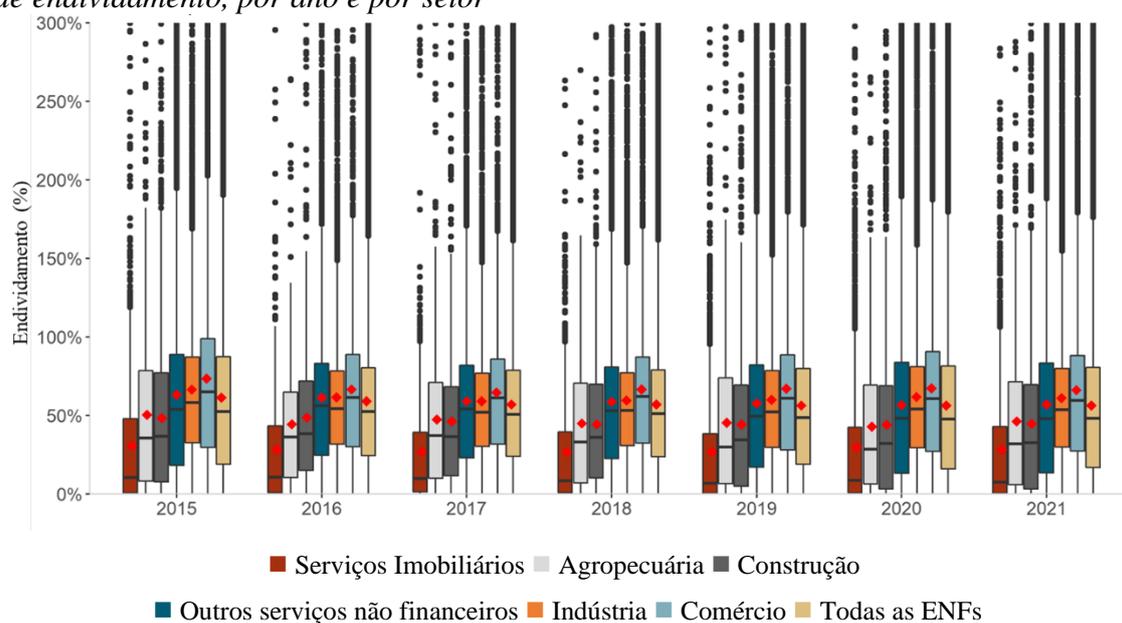
Fonte: Alexandria.

³⁵ Estas seções CNAE/ISIC coincidem com os grupos “Agricultura”, “Construção” e “Serviços Imobiliários” apresentados na seção 3.1.

³⁶ O Anexo B compara os endividamentos medianos das ENFs de Alexandria e da base Klook.

A Figura 4 agrega as seções CNAE em setores (sem separar por participação no mercado de valores mobiliário) e apresenta a distribuição do endividamento setorial no formato *box-plot*. A análise visual sugere que (i) a amostra apresenta expressiva variabilidade no endividamento e grande número de *outliers* (representando empresas muito endividadas)³⁷; (ii) em todos os anos analisados, os menores endividamentos médios e medianos são o do setor de serviços imobiliários, enquanto os maiores endividamentos médios e medianos são os do comércio³⁸.

Figura 4 – Empresas não financeiras (ENFs) de Alexandria: gráfico *box-plot* do índice de endividamento, por ano e por setor



Fonte: Alexandria. O eixo vertical foi cortado em 300% para facilitar a visualização. Os valores médios e medianos aqui apresentados estão no Anexo C.

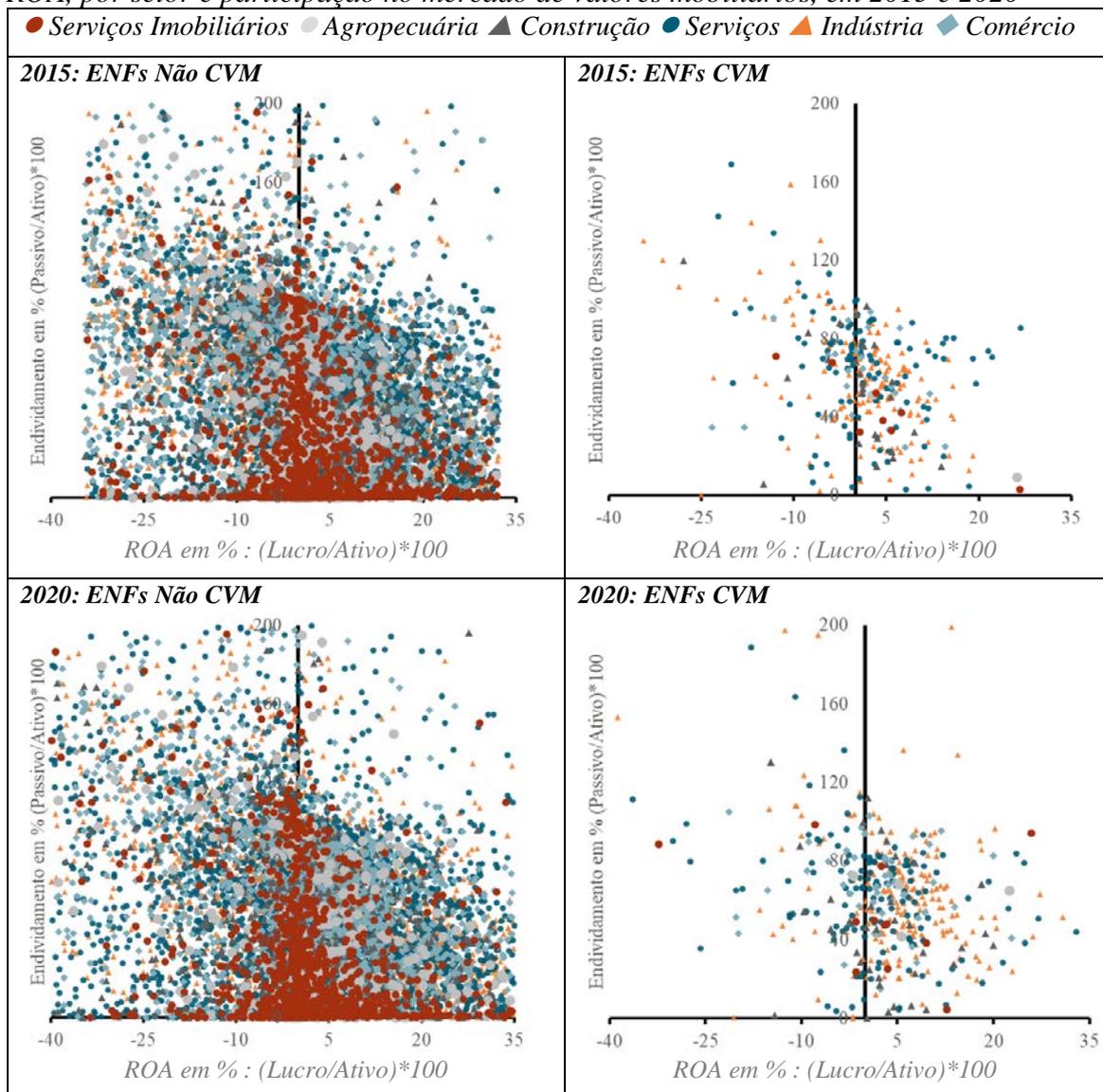
A Figura 5 apresenta os valores de endividamento e de lucratividade (medida em ROA – razão entre lucro líquido e ativo), em 2015 e 2020, por setor e por participação no mercado de valores mobiliários. O resultado sugere uma relação levemente negativa, ou

³⁷ A existência de número expressivo de empresas com endividamento acima de 100%, inclusive muito maior que isso, não é exclusividade de Alexandria. Marinšek (2015, p. 82) cita diversos autores que encontraram enormes diferenças no endividamento entre as firmas de suas amostras. Banco de Portugal (2019b, p. 32-33) informa que em 2017 e 2018 cerca de 26% das empresas não financeiras portuguesas apresentavam capital próprio negativo (portanto apresentavam endividamento superior a 100%) - situação mais comum em microempresas e nos setores “comércio” e “outros serviços”. Urionabarretxea et al (2016, p. 1), em artigo que utilizou a base Amadeus, informam que em 2012 quase 20% das empresas europeias apresentavam capital negativo.

³⁸ O painel balanceado apresenta conclusões semelhantes quanto às medianas setoriais, mas apresenta menos *outliers*.

seja, em geral empresas mais lucrativas apresentam um endividamento menor que empresas com prejuízo – resultado que vai em linha com a teoria do *Pecking Order*, segundo a qual as firmas preferem recursos financeiros internos (Bastos e Nakamura, 2009). Entretanto, é impossível não notar (i) a expressiva dispersão dos dados, que diminui a correlação entre as duas variáveis e (ii) o grande número de empresas do setor de “Serviços Imobiliários” que opera com endividamento igual a zero e fora do mercado de valores mobiliários (talvez, justamente, porque não tenha condição ou interesse em captar capital de terceiros). Esse grupo de empresas não é captado em estudos que excluem empresas sem endividamento, como Demirgüç-Kunt et al (2020).

Figura 5 - Empresas não financeiras (ENFs) de Alexandria: Endividamento versus ROA, por setor e participação no mercado de valores mobiliários, em 2015 e 2020

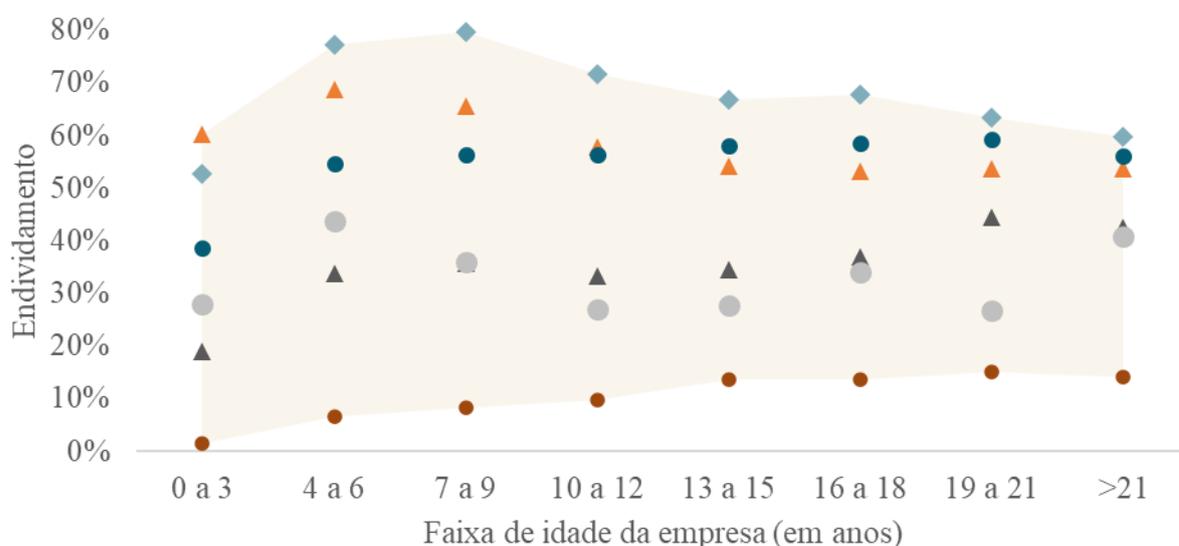


Fonte: Alexandria. O eixo vertical foi cortado em 200% para facilitar a visualização. Note que a amostra de ENFs com dado de lucro é menor que o número total de ENFs com dado de endividamento, por isso este gráfico apresenta uma amostra menor de empresas que a maioria dos gráficos desta seção 4.1.

A Figura 6 apresenta a mediana de endividamento por faixa de idade da empresa. A análise visual sugere que (i) todos os setores aumentam o endividamento a partir do terceiro ano de idade, mas (ii) alguns setores, marcadamente comércio e indústria, reduzem lentamente o endividamento a partir do nono ano de idade.

Figura 6 – Empresas não financeiras (ENFs) de Alexandria: mediana do endividamento por faixa de idade da empresa

● Serviços Imobiliários ● Agropecuária ▲ Construção ● Serviços ▲ Indústria ◆ Comércio



Fonte: Alexandria. (i) Os dados referem-se ao período de 2015 a 2021, portanto a mesma empresa pode aparecer mais de uma vez na mesma faixa de idade.

Passando para a análise por setor e por participação no mercado mobiliário conjuntamente (testes com dois fatores³⁹), as seções CNAE foram agregadas nos dez grupos descritos na seção 3.1. O número de ENFs CVM do setor “Agro” (grupo “CVMAgro”) é muito pequeno (4 empresas em 2021), e por isso foi necessário desconsiderá-lo em algumas análises.

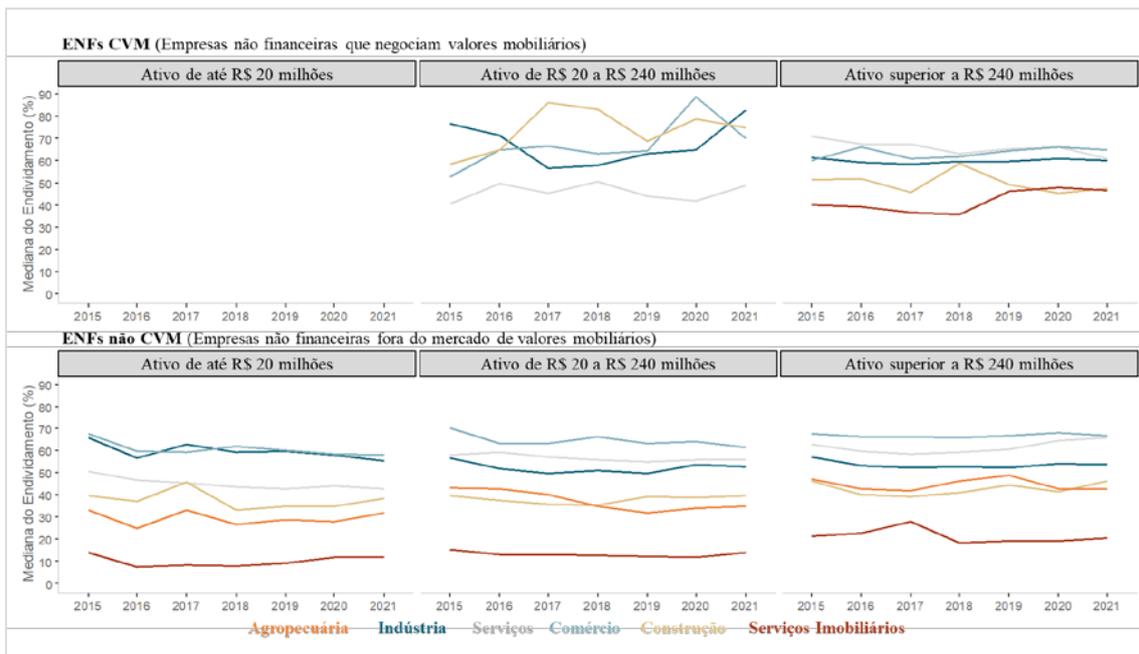
A Figura 7 traz a mediana do índice de endividamento separado por setor, participação no mercado de valores mobiliários e porte. A análise visual sugere que:

³⁹ Testes de dois fatores a serem utilizados (veja seção 5.2): 1. Teste não paramétrico de Mann-Whitney pareando as amostras dos pares de grupos; 2. Teste não paramétrico de Kruskal – Wallis com análise post-hoc de Dunn.

- comparando empresas do mesmo setor e porte, as empresas CVM são em geral mais endividadas que as não-CVM;
- entre as empresas não-CVM,
 - o setor “Serviços Imobiliários” utiliza menos capital de terceiros que todos os demais setores em todos os anos.
 - “Agropecuária Não-CVM” e “Construção Não CVM” têm endividamento mediano parecido (um pouco maior que “Serviços Imobiliários Não-CVM”, mas menor que o dos demais setores).
- o tamanho da empresa (aqui medido pelo ativo total)
 - impacta o endividamento das empresas não-CVM, no sentido de que quanto maior a empresa, maior o nível de endividamento (em conformidade com o encontrado pela maioria dos trabalhos acadêmicos⁴⁰). Somente no setor industrial essa conclusão não parece tão clara.
 - não parece impactar o endividamento das empresas CVM de maneira tão clara (há setores que mostraram aumento do endividamento com o aumento do ativo, mas há setores que mostraram o contrário).

Figura 7– Empresas não financeiras (ENFs): Mediana do endividamento (%) por tamanho do ativo, participação no mercado de valores mobiliários e setor

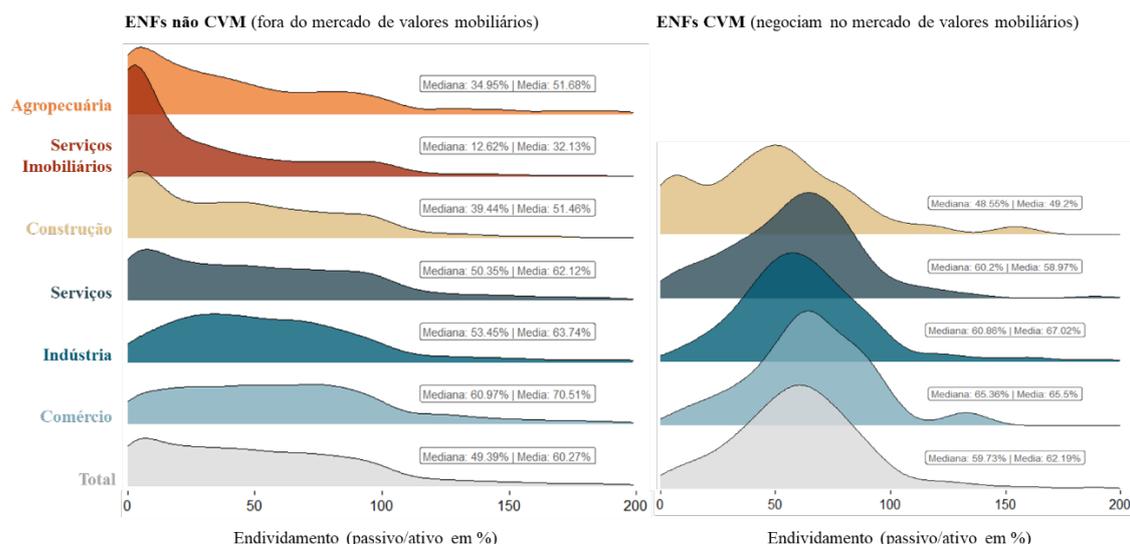
⁴⁰ A seção 5.3 apresenta uma lista de variáveis consideradas significativas pela literatura para explicar o endividamento empresarial.



Fonte: Alexandria. Devido ao pequeno número de empresas, não foram apresentadas: (i) as ENFs CVM da Agricultura (CVMAgr); (ii) as empresas CVM do setor de Serviços Imobiliários com Ativo entre R\$ 20 e R\$ 240 milhões; (iii) as empresas CVM com ativo de até R\$ 20 milhões.

Para permitir uma análise além da mediana, a Figura 8 apresenta a distribuição de frequência do endividamento em 2021 de acordo com os grupos. Em especial para as empresas não-CVM, a análise visual mostra grandes diferenças nas distribuições e não sugere distribuição normal. Em outras palavras, as empresas CVM apresentam uma distribuição mais concentrada em torno da mediana, as empresas não CVM apresentam endividamento mais disperso (maior variância). Em termos contábeis e estatísticos, isso significa que a análise visual não sugere um padrão único ou dominante de estrutura de capital.

Figura 8 - Empresas não financeiras (ENFs): distribuição de frequência do endividamento, por setor e participação no mercado de valores mobiliários (2021)



Fonte: Alexandria. Os eixos horizontais foram cortados em 200% para facilitar a visualização.

A seguir, realizaram-se diversos testes de normalidade e dois testes de homogeneidade das variâncias (passos C e D da Figura 2). Os resultados (detalhados na Tabela 3) mostram que se deve rejeitar a hipótese nula de normalidade, mas não a de homogeneidade das variâncias⁴¹.

Tabela 3 - Índice de endividamento de empresas não financeiras (ENFs), exceto CVMAgr e CVMImob: Testes de normalidade e de homogeneidade das variâncias

	Testes de Normalidade (p-valor) [Distribuição é normal?]						Testes homogeneidade variâncias (p-valor) [Variância é homogênea?]	
	Anderson-Darling	Lilliefors	Shapiro-Wilk	Cramer-von Mises	D'Agostino-Pearson	Jarque Bera	Levene	Bartlett
2015	7208,89 (0,00) [Não]	0,5021 (0,00) [Não]	0,01 (0,00) [Não]	1555,28 (0,00) [Não]	242,31 (0,00) [Não]	41,18 × 10 ⁹ (0,00) [Não]	0,61 (0,77) [Sim]	49943,21 (0,00) [Não]
2016	4223,61 (0,00)	0,5025 (0,00)	0,00 (0,00)	911,13 (0,00)	190,19 (0,00)	39,68 × 10 ⁹ (0,00)	0,14 (1,00)	52977,04 (0,00)

⁴¹ Os resultados da Tabela 3 mostram que os testes de Levene e Bartlett discordaram. O teste de Levene é um procedimento relativamente insensível a desvios de normalidade, o que o torna um teste robusto, pois na ausência de normalidade seu tamanho real é próximo do nível de significância fixado para uma grande variedade de distribuições de probabilidade. Por outro lado, o teste de Bartlett é sensível em relação à hipótese de normalidade dos dados, sendo robusto se as variáveis possuem distribuição aproximadamente normal (ALMEIDA, ELIAN e NOBRE, 2008). Assim, preferimos a conclusão do teste de Levene: não podemos rejeitar a hipótese de homogeneidade das variâncias.

	[Não]	[Não]	[Não]	[Não]	[Não]	[Não]	[Sim]	[Não]
2017	4102,42 (0,00)	0,5019 (0,00)	0,00 (0,00)	885,01 (0,00)	190,30 (0,00)	49,67× 10 ⁹ (0,00)	0,23 (0,99)	98539,02 (0,00)
	[Não]	[Não]	[Não]	[Não]	[Não]	[Não]	[Sim]	[Não]
2018	4400,02 (0,00)	0,5015 (0,00)	0,00 (0,00)	949,24 (0,00)	198,32 (0,00)	61,23× 10 ⁹ (0,00)	0,22 (0,99)	103247,79 (0,00)
	[Não]	[Não]	[Não]	[Não]	[Não]	[Não]	[Sim]	[Não]
2019	6876,52 (0,00)	0,4998 (0,00)	0,00 (0,00)	1483,49 (0,00)	254,98 (0,00)	193,98× 10 ⁹ (0,00)	0,19 (0,99)	77999,69 (0,00)
	[Não]	[Não]	[Não]	[Não]	[Não]	[Não]	[Sim]	[Não]
2020	8105,23 (0,00)	0,5040 (0,00)	0,01 (0,00)	1748,58 (0,00)	248,54 (0,00)	23,99× 10 ⁹ (0,00)	0,42 (0,91)	63283,92 (0,00)
	[Não]	[Não]	[Não]	[Não]	[Não]	[Não]	[Sim]	[Não]
2021	7178,51 (0,00)	0,5013 (0,00)	0,00 (0,00)	1548,74 (0,00)	237,61 (0,00)	29,42× 10 ⁹ (0,00)	0,79 (0,61)	102640,20 (0,00)
	[Não]	[Não]	[Não]	[Não]	[Não]	[Não]	[Sim]	[Não]
Concl.	{ Não }	{ Não }	{ Não }	{ Não }	{ Não }	{ Não }	{ Sim }	{ Não }

Fontes: Ghasemi e Zahediasl (2012) e cálculos dos autores. Notas: (i) Para o teste de Shapiro-Wilk, quando $n > 5.000$, extraiu-se uma amostra aleatória devido à limitação do teste para amostras entre 3 e 5.000 observações. (ii) O teste de Levene robusto é o teste Teste Brown-Forsythe F. (iii) [Sim] indica que não rejeitamos a hipótese nula de normalidade da distribuição ou de homogeneidade das variâncias, com 95% de confiança; [Não] indica que rejeitamos a mesma hipótese. (iv) A última linha é “conclusão” apresenta a conclusão aplicável para a maioria dos anos analisados, ou seja, se na maioria dos anos a hipótese nula foi aceita {Sim} ou rejeitada {Não}; (v) para os testes de homogeneidade de variâncias usamos os dez grupos de setor, o que resultou em 9 graus de liberdade; (vi) retiramos os grupos CVMAGR e CVMIMOB dado o reduzido tamanho da amostra, o que traria um baixo poder dos testes executados.

A rejeição da hipótese de normalidade sugere a utilização de testes não paramétricos, que são apresentados na seção 4.2. Porém, como essa sugestão não é consensual na literatura, realizamos também os testes paramétricos, que são apresentados na seção 4.3.

4.2 Comparação das medianas de endividamento usando testes não paramétricos

Como não se pôde concluir pela normalidade das variâncias dos dados de endividamento, e como os grupos tinham tamanhos diferentes, adotou-se o teste não paramétrico de *Kruskal-Wallis* para amostras independentes (etapa E.1 da

Figura 2), seguindo Marinšek (2015, p. 42). O teste, feito com 11 grupos de ENFs (todos menos CVMAgr), levou à conclusão de que havia pelo menos um grupo com distribuição diferente dos demais, em todos os anos analisados⁴². Por essa razão, realizou-

⁴² Alguns autores afirmam que o teste de Kruskal-Wallis compara a mediana dos grupos, mas sem homoscedasticidade das variâncias, é mais correto afirmar que a hipótese nula deste teste é homogeneidade estocástica (VARGHA e DELANEY, 1998, p. 175). Se adotarmos a conclusão do teste de Levene apresentado na Tabela 3 (que a variância é homogênea entre os grupos), podemos adotar a explicação mais

se a análise de *post hoc* de Dunn, de cada um dos 55 pares formados pelos 11 grupos, com correção de significância de Bonferroni⁴³ para verificar eventuais diferenças entre as medianas (etapa F da Figura 2). Vale ressaltar que esse procedimento é sensível às diferenças de distribuição das amostras.

Para gerar mais robustez às conclusões, realizou-se também o teste não paramétrico de *Mann-Whitney* para amostras independentes e pareadas (etapa E.2 da Figura 2). Em cada um dos 55 pares, o grupo de tamanho menor determina o tamanho da amostra aleatória a ser criada para o grupo maior. Por exemplo, se o grupo de tamanho menor tem $n = 100$, foram criadas 1.000 amostras aleatórias do grupo maior do par, todas de tamanho $n = 100$, permitindo-se assim 1.000 comparações pareadas. Com isso, é possível analisar em quantas comparações a hipótese nula (medianas são iguais) deve ser aceita.

A Figura 9 mostra, para cada ano e para cada par de grupos, o resultado do teste de Dunn, junto com sua conclusão entre colchetes: “=” ou “≠”. Quando o teste de *Mann-Whitney* (“MH-MC”) discordou da conclusão do teste de Dunn, um sinal (“!”) foi acrescentado. Por exemplo, na primeira coluna (2015), linha “CVMCom-CVMConstr”, comparamos a mediana do índice de endividamento das 28 empresas CVM do comércio e das 30 empresas CVM de construção. A célula apresenta o valor “0,71 [=]”, que deve ser assim interpretado:

- o teste de Dunn não rejeitou a hipótese nula de que as medianas são iguais (por isso o sinal de “=”).
- realizou-se o teste de *Mann-Whitney*, ou seja, foram realizadas 1.000 comparações entre as 28 empresas CVM de comércio e uma amostra aleatória de 28 empresas CVM da construção. Em mais de 50% dos casos, aceitou-se a hipótese de igualdade da mediana.
- portanto, nesse caso, tanto o teste de Dunn quanto o teste de *Mann-Whitney* resultaram na mesma conclusão (medianas de endividamento das

simples: o teste de Kruskal-Wallis mostrou que pelo menos um dos nove grupos de ENFs tem mediana diferente dos demais.

⁴³ A correção de Bonferroni, na sua forma mais simples, significa que em vez de utilizar um valor crítico de 5% para cada teste, utiliza-se o valor de 5% dividido pelo número de grupos testados. Um efeito é que a correção de Bonferroni infla o erro do tipo II, isto é, deixamos de identificar diferenças que podem existir.

empresas CVMComerc e CVMConstr são iguais). Por isso a célula está com a cor verde.

Figura 9- Índice de endividamento de empresas não financeiras (ENFs): comparação da mediana dos pares utilizando o teste de Dunn com correção de Bonferroni e teste de Mann-Whitney

Legenda de cores: resultados sugerem que os grupos têm endividamento mediano ■ igual ■ diferente ■ não foi possível concluir (testes divergem)

Grupos	Par	2015	2016	2017	2018	2019	2020	2021	Conclusão
Somente empresas CVM	CvmCOMERC - CvmCONSTR	0,71 [=]	0,79 [=]	0,11 [=]	-0,08 [=]	1,07 [=]	1,46 [=]	1,64 [=]	{ = }
	CvmCOMERC - CvmIND	-0,2 [=]	0,43 [=]	0,14 [=]	0,28 [=]	0,78 [=]	0,44 [=]	0,05 [=]	{ = }
	CvmCOMERC - 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 [=]	{ = }
	CvmCOMERC - CvmSERVIMOB	0,41 [=]	0,81 [=]	1,05 [=]	0,82 [=]	1,04 [=]	0,55 [=]	0,68 [=]	{ = }
	CvmCONSTR - CvmSERVIMOB	-0,12 [=]	0,23 [=]	0,97 [=]	0,87 [=]	0,27 [=]	-0,49 [=]	-0,51 [=]	{ = }
	CvmIND - CvmSERVIMOB	0,6 [=]	0,65 [=]	1,1 [=]	0,75 [=]	0,68 [=]	0,35 [=]	0,73 [=]	{ = }
	CvmSERV - CvmSERVIMOB	0,39 [=]	0,61 [=]	0,93 [=]	0,61 [=]	0,55 [=]	0,02 [=]	0,11 [=]	{ = }
Mesmo setor	CvmCOMERC - NCvmCOMERC	-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 [=]	{ = }
	CvmSERVIMOB - NCvmSERVIMOB	2,57 [=]	2,63 [=]	2,73 [=]	3,15 [=]	3,57* [=]	3,22 [=]	3,24 [=]	{ = }
	CvmCOMERC - NCvmAGRO	2,34 [=]	3,69* [=]	3,34 [=]	3,63* [=]	4,79** [=]	3,73** [=]	3,6* [=]	{ ≠ }
	CvmCOMERC - NCvmCONSTR	2,72 [=]	3,51* [=]	3,89** [=]	4,12** [=]	4,81** [=]	4** [=]	4** [=]	{ ≠ }
	CvmCOMERC - NCvmIND	0,4 [=]	1,84 [=]	1,67 [=]	1,97 [=]	2,92 [=]	1,56 [=]	2,16 [=]	{ = }
	CvmCOMERC - NCvmSERV	0,82 [=]	1,24 [=]	1,11 [=]	1,46 [=]	2,46 [=]	1,68 [=]	1,45 [=]	{ = }
	CvmCOMERC - NCvmSERVIMOB	4,83** [=]	5,38** [=]	6,11** [=]	6,63** [=]	7,65** [=]	6,74** [=]	6,29** [=]	{ ≠ }
Setores diferentes	CvmCONSTR - NCvmAGRO	1,43 [=]	2,84 [=]	3,27 [=]	3,69** [=]	3,39* [=]	1,65 [=]	1,56 [=]	{ = }
	CvmCONSTR - NCvmCOMERC	-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 - NCvmSERVIMOB	3,99** [=]	4,63** [=]	6,11** [=]	6,66** [=]	6,19** [=]	4,4** [=]	4,1** [=]	{ ≠ }
	CvmIND - NCvmAGRO	5,65** [=]	6,29** [=]	5,7** [=]	5,68** [=]	6,64** [=]	5,82** [=]	5,36** [=]	{ ≠ }
	CvmIND - NCvmCOMERC	-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 - NCvmSERVIMOB	12,15** [=]	9,81** [=]	11,16** [=]	11,31** [=]	11,78** [=]	12,46** [=]	9,91** [=]	{ ≠ }
CVM X Não CVM	CvmSERV - NCvmAGRO	4,15** [=]	5,5** [=]	4,73** [=]	4,9** [=]	5,72** [=]	4,28** [=]	3,71** [=]	{ ≠ }
	CvmSERV - NCvmCOMERC	-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 - NCvmSERVIMOB	9,24** [=]	8,54** [=]	9,47** [=]	9,89** [=]	10,25** [=]	9,65** [=]	7,76** [=]	{ ≠ }
	CvmSERVIMOB - NCvmAGRO	1,02 [=]	1,52 [=]	0,98 [=]	1,35 [=]	1,85 [=]	1,57 [=]	1,6 [=]	{ = }
	CvmSERVIMOB - NCvmCOMERC	-0,66 [=]	-0,46 [=]	-0,86 [=]	-0,59 [=]	-0,4 [=]	-0,28 [=]	-0,5 [=]	{ = }
	CvmSERVIMOB - NCvmCONSTR	1,24 [=]	1,34 [=]	1,26 [=]	1,56 [=]	1,74 [=]	1,67 [=]	1,7 [=]	{ = }
	CvmSERVIMOB - NCvmIND	-0,22 [=]	0,22 [=]	-0,2 [=]	0,21 [=]	0,52 [=]	0,25 [=]	0,49 [=]	{ = }
	CvmSERVIMOB - NCvmSERV	0,04 [=]	-0,16 [=]	-0,54 [=]	-0,08 [=]	0,25 [=]	0,32 [=]	0,07 [=]	{ = }
Somente empresas não CVM	NCvmAGRO - NCvmCOMERC	-10,6** [=]	-7,34** [=]	-6,36** [=]	-6,58** [=]	-7,24** [=]	-8,66** [=]	-5,57** [=]	{ ≠ }
	NCvmAGRO - NCvmCONSTR	1,24 [=]	-0,67 [=]	0,78 [=]	0,56 [=]	-0,41 [=]	0,31 [=]	0,1 [=]	{ = }
	NCvmAGRO - NCvmIND	-7,97** [=]	-5,12** [=]	-4,35** [=]	-4,16** [=]	-4,62** [=]	-6,32** [=]	-3,18 [=]	{ ≠ }
	NCvmAGRO - NCvmSERV	-6,38** [=]	-6,52** [=]	-5,51** [=]	-5,12** [=]	-5,42** [=]	-6,09** [=]	-4,23** [=]	{ ≠ }
	NCvmAGRO - NCvmSERVIMOB	9,09** [=]	3,26 [=]	5,01** [=]	5,06** [=]	4,49** [=]	7,09** [=]	3,47* [=]	{ ≠ }
	NCvmCOMERC - NCvmCONSTR	18,71** [=]	8,72** [=]	9,53** [=]	9,51** [=]	8,83** [=]	13,55** [=]	7,89** [=]	{ ≠ }
	NCvmCOMERC - NCvmIND	5,88** [=]	5,21** [=]	4,67** [=]	5,5** [=]	6,01** [=]	5,4** [=]	6** [=]	{ ≠ }
	NCvmCOMERC - NCvmSERV	10,05** [=]	2,2 [=]	2,14 [=]	3,29 [=]	4,02** [=]	6,34** [=]	3,29 [=]	{ ≠ }
	NCvmCOMERC - NCvmSERVIMOB	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 - NCvmSERVIMOB	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 - NCvmSERVIMOB	29,24** [=]	10,07** [=]	12,34** [=]	12,26** [=]	11,57** [=]	24,05** [=]	8,71** [=]	{ ≠ }	
NCvmSERV - NCvmSERVIMOB	27,52** [=]	11,5** [=]	13,51** [=]	13,19** [=]	12,32** [=]	24,3** [=]	9,83** [=]	{ ≠ }	

Fontes: Alexandria e cálculo dos autores. Notas: (i) a tabela mostra os valores da estatística Z do teste post hoc de Dunn com ajuste de Bonferroni para cada par de amostra; (ii) * 5% de significância; ** 1% de significância. (iii) [=] indica que não podemos rejeitar a hipótese nula de que a mediana do endividamento dos dois grupos, naquele ano, é igual. [≠] indica que rejeitamos, ao nível de significância de 5%, a mesma hipótese nula.; (v) indica que a maioria das simulações de Mann-Whitney divergiram da conclusão do teste de Dunn; (vi) [=] indica que na maioria dos anos analisados não se rejeita a hipótese nula; (vii) [≠] indica que na maioria dos anos analisados rejeita-se a hipótese nula.

Os resultados da Figura 9 em relação às diferenças entre as medianas dos índices de endividamento das ENFs sugerem que, dentro do mesmo ano:

- Não há diferença entre o endividamento mediano dos diferentes setores de empresas registradas na CVM (por isso todas as linhas do grupo “somente empresas CVM” estão em verde).
- A participação no mercado de valores mobiliários afeta o nível de endividamento do setor industrial (pois a mediana do endividamento das empresas CVM desses setores é diferente das não CVM – ver grupo “mesmo setor” na Figura, coluna “Conclusão”).
 - Para os demais setores, ao contrário, a participação no mercado de valores mobiliários não parece afetar o nível mediano de endividamento.
- Entre as empresas fora do mercado de valores mobiliários (grupo “somente empresas não CVM”), a diferença setorial é muito clara: cada setor têm um nível de endividamento diferente de todos os outros, com exceção de agropecuária e construção, cuja mediana é igual em termos estatísticos, e indústria e serviços, par para o qual não foi possível tirar uma conclusão porque os testes divergiram.
- Repare o leitor a aparente contradição entre as conclusões: o endividamento mediano dos setores CVM de serviços e comércio são iguais entre si, e estes são respectivamente iguais aos setores Não-CVM de serviços e comércio. Entretanto, estes dois últimos são diferentes entre si. Essa aparente contradição (que também ocorre com outros setores) pode ocorrer nos testes de Dunn e Mann Whitney e deve ser avaliada com o uso de outras abordagens, como faremos adiante.

4.3 Comparação das médias de índice de endividamento usando teste paramétrico (modelo econométrico)

O raciocínio “se a distribuição não é normal, faça teste não paramétrico”, que levou aos testes realizados na seção 4.2, não é consensual na literatura. Fagerland (2012) alerta que os testes Wilcoxon-Mann-Whitney (incluindo Kruskal-Wallis) produzem p-valores menores que o teste t, na média, o que pode alterar conclusões. Por isso, o autor sugere que, para amostras grandes, o teste t deve ser usado, mesmo que a distribuição seja fortemente assimétrica. Vargha e Delaney (1998, p. 181) afirmam que, quando a variância entre os grupos difere substancialmente, o teste Kruskal-Wallis pode se provar

inadequado, especialmente em caso de grupos com tamanhos diferentes (caso dos grupos analisados na seção 4.2). Por fim, Ghasemi e Zahediasl (2012) mencionam que, para amostras suficientemente grandes, a violação da hipótese de normalidade não causa grandes problemas e testes paramétricos podem ser utilizados, por força do Teorema Central do Limite (passo F da Figura 2).

Seguindo essa literatura, os autores compararam também as médias do índice de endividamento de cada grupo utilizando regressões econométricas em painel.

Embora o objetivo deste trabalho seja testar apenas a significância do setor e da participação no mercado de valores mobiliários para explicar o endividamento empresarial, o modelo foi construído com outras variáveis, recomendadas na literatura (ver seção 2), para ficar mais robusto. O Quadro 4 resume as variáveis testadas.

Em relação às variáveis explicativas, não foi possível testar todas as variáveis comumente usadas em estudos. Em especial, não foi testado o impacto das variáveis (i) tangibilidade dos ativos⁴⁴, (ii) dedutibilidade de impostos⁴⁵ e (iii) razão entre o valor de mercado e seu valor patrimonial (*Market-to-Book ratio*)⁴⁶, frequentemente citados como variáveis relevantes, por falta das rubricas necessárias em Alexandria. Por isso, nosso estudo possivelmente sofre do problema de variável omitida, e é possível que a variável setor capte parte do efeito da tangibilidade.

No nível teórico, acreditamos que a variável “dedutibilidade de impostos”, não é relevante no Brasil, pois apenas empresas no sistema de tributação “lucro real” podem deduzir pagamento de juros do imposto de renda. Estas empresas representam apenas 16% das empresas que produzem balanços (Sped, 2020), mas incluem 100% das registradas na CVM.

Quadro 4 - Variáveis testadas nos modelos econométricos deste trabalho

Variáveis testadas	Alguns estudos anteriores que utilizaram/testaram a mesma variável (a) ou uma transformação dela.
Variável dependente	

⁴⁴ São exemplos de estudos que incluem empresas não listadas e que concluem que a variável “tangibilidade dos ativos” é relevante para explicar o endividamento: Marinšek (2015) e Den Berg (2021).

⁴⁵ Den Berg (2021) faz estudo que inclui empresas não listadas e conclui que a variável “dedutibilidade dos impostos” é relevante para explicar o endividamento. Demirgüç-Kunt et al (2020) chegam à mesma conclusão para Portugal e Polónia.

⁴⁶ Jaworski e Santos (2021, p. 40) usaram dados de empresas não listadas, mas estimaram a variável “market /book ratio” a partir do fluxo de caixa.

$LN_{EG} = \ln[EG + 1]$ sendo $EG = \left(\frac{\text{Passivo circulante} + \text{Passivo não circulante}}{\text{Ativo total}} \right) * 100$ <i>A variável foi winsorizada</i> ⁴⁷ .	Raalte (2021, p. 26); Nehrebecka e Białek-Jaworska (2015); Den Berg (2021); Forte, Barros e Nakamura (2013)
Variáveis independentes	
Endividamento do ano anterior (<i>lag</i> da variável dependente): LAG_LOG_EG	Nehrebecka e Białek-Jaworska (2015); Correa et al (2013); Forte, Barros e Nakamura (2013)
Tamanho da empresa: LOG_ATIVO =ln(ativo)	Marinšek (2015, p. 37); Raalte (2021, p. 25); Forte, Barros e Nakamura (2013); Andritzky (2003); Pohlman e Iudícibus (2010) [b]
Lucratividade no ano base: ROA = $\left(\frac{\text{Lucro Líquido}}{\text{Ativo total}} \right) * 100$ <i>A variável foi winsorizada.</i>	Marinšek (2015, p. 37); Raalte (2021, p. 26); Forte, Barros e Nakamura (2013); Correa et al (2013); Andritzky (2003) [c]
Lucratividade no ano base: ROE = $\left(\frac{\text{Lucro Líquido}}{\text{Patrimônio Líquido}} \right) * 100$	Forte, Barros e Nakamura (2013)
Idade da empresa na data-base, em anos: LOG_IDADE =LOG(idade)	Jaworski e Santos (2021, p. 40); Forte, Barros e Nakamura (2013); Den Berg (2021)
Crescimento: $\left(\frac{\text{Ativototal}[t] - \text{Ativototal}[t-1]}{\text{Ativototal}[t-1]} \right) * 100$	Marinšek (2015, p. 37); Raalte (2021, p. 30); Forte, Barros e Nakamura (2013)
Estatal: <i>dummy</i> que assume o valor 1 se a empresa é detida, em alguma proporção, pela União, Estados ou Municípios brasileiros em 2022.	Marinšek (2015, p. 40); Nehrebecka e Białek-Jaworska (2015); Azevedo (2013), Pohlman e Iudícibus (2010)
PART_NRES: participação de não residentes no capital social na data de referência. Varia de 0 a 1.	Nehrebecka e Białek-Jaworska (2015); Azevedo (2013), Pohlman e Iudícibus (2010); Correa et al (2013)
Listada: <i>dummy</i> que assume o valor 1 se a empresa é listada na bolsa de valores brasileira (B3) no fim do ano base e 0 em caso contrário.	Raalte (2021, p. 25); Den Berg (2021); Azevedo (2013); Demirgüç-Kunt et al (2020) e Marinšek (2015, p. 39)
CVM: <i>dummy</i> que assume o valor 1 se a entidade participa do mercado de valores mobiliários brasileiro e 0 em caso contrário. Inclui todas as listadas.	Não encontramos literatura com essa variável.
Familiar: <i>dummy</i> que assume o valor 1 se a empresa é uma empresa de responsabilidade limitada detida por dois ou mais pessoas de mesmo sobrenome.	Gottardo e Maria Moisello (2014)
Tem mulher na direção: <i>dummy</i> que assume o valor 1 se pelo menos um dos administradores era uma mulher.	García e Herrero (2021) [d]

⁴⁷ A winsorização manteve muitas ENFs com endividamento superior a 100% na amostra, ao contrário do que fizeram muitos trabalhos anteriores (seção 2.5.1). Esta decisão foi tomada pelos autores após análise manual de centenas de balanços publicados em jornal, que permitiu verificar que existe considerável número de empresas que opera em patamares superiores a 100% de endividamento, inclusive por muitos anos. Em outras palavras, em muitos casos não se trata de erro de declaração.

Setor: <i>dummies</i> que assumem o valor 1 se a empresa pertence ao setor e 0 em caso contrário. Os setores considerados foram os seis descritos no Quadro 3.	Marinšek (2015, p. 41); Raalte (2021, p. 25); Forte, Barros e Nakamura (2013); Azevedo (2013); Andritzky (2003)
Ano: <i>dummies</i> que assumem o valor 1 se o dado pertence ao ano em questão e 0 em caso contrário.	Den Berg (2021); Nehrebecka e Białek-Jaworska (2015); Raalte (2021); Forte, Barros e Nakamura (2013)
D_capital: <i>dummy</i> que assume o valor 1 se a sede da empresa está localizada na capital de alguma unidade da federação e 0 em caso contrário.	Não encontramos literatura com essa variável.

Fonte: elaborado pelos autores. [a] Foram mencionados somente estudos cuja amostra incluía empresas não listadas; [b] Correa et al (2013), Azevedo (2013) e Den Berg (2021) utilizam o logaritmo natural da receita líquida como medida do tamanho da empresa, e Demirgüç-Kunt et al (2020) utilizaram uma *dummy* (pequena empresa ou não). Como Alexandria não contém a receita líquida de todas as empresas, optou-se pelo uso do logaritmo natural do ativo total, inclusive pela possibilidade de viés mencionada na seção 2.5.2. Os estudos mencionados utilizaram a mesma métrica. [c] alguns autores citados medem a lucratividade como lucro operacional/ativo, ou seja, EBIT/ativo. Como o lucro operacional não está disponível em Alexandria, utilizamos o ROA.[d] este é o único trabalho citado na tabela que não utiliza empresas não listadas, pois não encontramos trabalhos que investiguem a variável “gênero” utilizando amostra que inclua empresas não listadas.

As estatísticas descritivas das variáveis utilizadas estão no Anexo D. As variáveis mostraram-se estacionárias e o correlograma (Anexo E) contribuiu para selecionar as melhores variáveis.

O correlograma também mostrou a forte persistência do endividamento, vista pela elevada correlação do endividamento e seu valor defasado – resultado que está amplamente em linha com a literatura nacional e internacional (ver Quadro 4 e Correa et al, 2013). Esse fato e o provável problema de variáveis omitidas sugerem endogeneidade (ver seção 2.5.3) e levaram os autores descartarem regressões do tipo POLS, efeitos fixos e efeitos aleatórios.

O método econométrico escolhido, pela sua robustez e pela capacidade de lidar com o problema da endogeneidade, foi a regressão GMM em nível (GMM-Sys), em linha com o sugerido por Barros et al (2020). A equação final foi:

$$\ln\left(\frac{\text{Passivo}_t}{\text{Ativo}_t} + 1\right) = \beta_0 + \beta_1 \ln\left(\frac{\text{Passivo}_{t-1}}{\text{Ativo}_{t-1}} + 1\right) + \beta_{2-5} \text{dummies setoriais} \\ + \beta_6 \text{dados cvm} + \beta_7 \text{estatal} + \beta_7 \text{emp ltda familiar} + \beta_8 \ln(\text{ativo}) \quad (7) \\ + \beta_9 \text{PartNres} + \beta_{10} \text{sede em capital} + \beta_{11} \text{mulher na adm} \\ + \beta_{12} \ln(\text{idade da empresa})$$

Os resultados da regressão estão detalhados na Tabela 4.

Tabela 4 - Estimação dos modelos de endividamento usando o método Sys-GMM em painel balanceado e painel desbalanceado de empresas não financeiras (ENFs) que

participam (“CVM”) ou não (“NCVM”) do mercado de valores mobiliários brasileiro, com dados anuais de 2015 a 2021

	Painel desbalanceado			Painel balanceado		
	(A) ENF's	(B) CVM	(C) N-CVM	(D) ENF's	(E) CVM	(F) N-CVM
lag(log end, 1)	0,7918*** (0,0359)	0,8203*** (0,0583)	0,8665*** (0,0385)	0,9187*** (0,0211)	0,8416*** (0,0450)	0,9065*** (0,0239)
Demais Serviços	0,0038 (0,0026)	-0,0531** (0,0212)	0,0073*** (0,0025)	0,0019 (0,0037)	-0,0111 (0,0106)	0,0050 (0,0046)
Comércio	0,0048 (0,0032)	-0,0008 (0,0102)	0,0043 (0,0030)	0,0051* (0,0028)	-0,0093 (0,0060)	0,0074** (0,0034)
Construção	-0,0137** (0,0056)	-0,0821*** (0,0210)	-0,0078 (0,0051)	-0,0041 (0,0043)	-0,0146 (0,0142)	-0,0048 (0,0048)
Agro	-0,0132* (0,0076)		-0,0082 (0,0055)	-0,0066 (0,0046)		-0,0079 (0,0051)
Serviços imobiliários	-0,0401*** (0,0082)	-0,1648*** (0,0406)	-0,0270*** (0,0073)	-0,0114* (0,0059)	-0,0328 (0,0240)	-0,0151** (0,0067)
ROA	-0,1471*** (0,0132)	-0,1056* (0,0554)	-0,1319*** (0,0334)	-0,1905*** (0,0388)	-0,2384*** (0,0449)	-0,2402*** (0,0395)
Dados na CVM	0,0223*** (0,0057)			-0,0073 (0,0079)		
Estatual	0,0069 (0,0065)	0,0735*** (0,0113)	0,0027 (0,0059)	-0,0138*** (0,0044)	-0,0091 (0,0105)	-0,0184*** (0,0055)
Familiar	0,0021 (0,0058)		0,0064 (0,0052)	0,0032 (0,0048)		0,0048 (0,0054)
log(ativo)	-0,0012 (0,0009)	-0,0822*** (0,0151)	0,0007 (0,0014)	0,0058 (0,0040)	-0,0014 (0,0096)	0,0093* (0,0049)
Participação Nres	0,0020 (0,0031)	0,0895** (0,0368)	-0,0032 (0,0033)	-0,0088* (0,0050)	-0,0045 (0,0242)	-0,0137** (0,0061)
Sede em Capital	-0,0046** (0,0021)	0,0809*** (0,0164)	-0,0058*** (0,0022)	-0,0040 (0,0026)	0,0103 (0,0100)	-0,0068** (0,0029)
Mulher (adm)	-0,0023 (0,0021)	0,0062 (0,0051)	-0,0004 (0,0020)	-0,0012 (0,0012)	0,0010 (0,0046)	-0,0011 (0,0014)
log(idade emp)	0,0017 (0,0015)	0,0015 (0,0088)	0,0003 (0,0014)	0,0035*** (0,0010)	-0,0018 (0,0053)	0,0028** (0,0011)
nobs	53.891	4.162	49.307	20.097	2.151	17.208
Instrumentos	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 [0,2408]	31,3635 [0,1436]	33,4197 [0,0562]	36,0678 [0,0541]	25,1214 [0,1216]	36,0450 [0,0408]
AR1	-11,58 [0,0000]	-2,87 [0,0040]	-11,56 [0,0000]	-14,69 [0,0000]	-4,60 [0,0000]	-13,34 [0,0000]
AR2	-1,58 [0,1149]	1,23 [0,2188]	-1,75 [0,0798]	1,45 [0,1460]	1,50 [0,1349]	1,69 [0,0914]
Wald Coef {GL}	6.610,94 {15} [0,0000]	361,40 {12} [0,0000]	7.702,21 {14} [0,0000]	7.327,56 {15} [0,0000]	1.170,76 {12} [0,0000]	5.676,27 {14} [0,0000]
Wald Time {GL}	17,07 {5} [0,0044]	10,75 {5} [0,0565]	27,74 {4} [0,0000]	28,57 {4} [0,0000]	15,03 {4} [0,0046]	15,85 {4} [0,0032]

Notas: (i) Análise das variáveis anuais da amostra; (ii) Erros robustos à heterocedasticidade em parênteses utilizando clusterização de erros no nível da firma; (iii) * 10% de significância; ** 5% de significância; *** 1% de significância; (iv) AR(1) e AR(2) verificam a presença de correlação seriada de primeira ordem e de segunda ordem nos resíduos de primeira diferença (v) nas regressões B e E, as empresas do grupo CVM agro foram desconsideradas por representarem uma amostra pequena (entre 3 e 4 empresas, dependendo do ano). (vi) Não mostramos as dummies de tempo e o intercepto por concisão. (vii) nas regressões com painel desbalanceado, o número de empresas CVM, somado ao número de não CVM, é inferior ao total porque há empresas que durante um período foram CVM e depois deixaram de ser, ou vice-versa, e, portanto, não puderam ser incluídas nos dois últimos painéis balanceados. (viii) utilizou-se a técnica de winsorização em nível de 1% para as variáveis endividamento e ROA. (ix) variáveis em lag utilizadas como instrumentos: E - ln(endividamento+1) winsorizado; R - ROA winsorizado; A - ln(ativo); valores entre []'s são os lags das variáveis utilizadas como instrumentos. (x)

O Quadro 5 resume as principais conclusões a serem tiradas das regressões. Em resumo, (i) o setor é uma variável significativa para explicar o endividamento em todas as regressões com painel desbalanceado. No painel balanceado, os setores só diferem em endividamento entre as empresas não-CVM; e (ii) tudo o mais constante, a participação no mercado de valores mobiliários aumenta o nível de endividamento das ENFs no painel desbalanceado, mas não no painel balanceado⁴⁸.

Quadro 5 – Principais conclusões sobre as regressões da Tabela 3 (considerando coeficiente de 5% de significância)

	Painel desbalanceado			Painel balanceado		
	(A) ENF's	(B) CVM	(C) NCVM	(D) ENF's	(E) CVM	(F) NCVM
Setor impacta o endividamento médio?	Sim	Sim	Sim	Não (*)	Não	Sim
Se sim, quais são os setores mais endividados? (**)	Indústria, Comércio, Demais Serviços	Indústria, Comércio	Demais serviços	---	---	Comércio
Se sim, qual é o setor menos endividado?	Serviços Imobiliários	Serviços Imobiliários	Serviços Imobiliários	---	---	Serviços Imobiliários
A participação no mercado de valores mobiliários aumenta o endividamento médio?	Sim	---	---	Não	---	---

Obs: () A resposta muda para "Sim" considerando coeficiente de significância de 10% e também nos testes de robustez mostrados no Anexo E. (**) quando mais de um setor é citado, a regressão não mostrou diferença estatística entre o endividamento destes setores, ou seja, eles ficaram empatados como os mais endividados.*

Para testar a robustez dos modelos, estimamos modelos alternativos que contemplam um período menor (2015 a 2019 ou 2017 a 2021). Também estimamos modelos separados para cada setor. Estes testes de robustez, disponíveis respectivamente no Anexo E e F, não alteraram nossas conclusões principais.

Embora não seja o foco deste trabalho analisar as demais variáveis, é interessante mencionar que o endividamento do ano anterior é a variável mais significativa para explicar o endividamento atual da ENF. Em segundo lugar, o lucro das empresas (medido pela ROA⁴⁹) está negativamente correlacionado com o endividamento, resultado em linha com a literatura⁵⁰. Em terceiro lugar, tudo o mais constante, o tamanho da empresa

⁴⁸ Outros estudos serão necessários para verificar se o custo do endividamento difere entre empresas CVM e não CVM.

⁴⁹ As estimativas feitas pelos autores usando o ROE mostraram efeito menos significativo.

⁵⁰ Uma interpretação possível é que, quando a empresa lucra, ela prefira reduzir o endividamento. Observar que a análise visual sugeriu grande heterogeneidade (ver Figura 5).

(medida pelo ativo) não impacta o endividamento. Em quarto lugar, variáveis frequentemente mencionadas como significativas na literatura para explicar o endividamento, como o fato de ser estatal e a idade da empresa, mostraram-se significativos, mas não em todas as regressões e algumas vezes com sinais contraditórios.

Além disso, a participação de não residentes no capital social não se mostrou significativa para explicar o endividamento nas regressões principais (A e D) – e, nas regressões em que foi detectado algum impacto, os sinais foram contraditórios. O resultado não foi diferente quando trocamos o percentual de participação por uma dummy, como fazem alguns trabalhos. Também não se alterou quando trocamos “participação de não residentes” por “empresa de investimento direto”, um subgrupo do primeiro. Como Alexandria contém praticamente todas as empresas diretamente detidas por não residentes no Brasil e a classificação dessas como receptora de investimento direto ou não, este resultado é especialmente forte e contraria literatura anterior (inclusive para o caso brasileiro, como Correa et al, 2013).

Finalmente, é importante mencionar que os autores testaram três variáveis pouco exploradas na literatura brasileira: a presença de mulheres na gestão, o fato de a empresa ser limitada e familiar e o fato de a sede da empresa estar numa capital de estado. Entretanto, somente a última mostrou impacto significativo sobre o endividamento (e com sinal negativo, ao contrário do que prevíamos).

5. Conclusões

5.1 Efeitos do setor e da participação no mercado de valores mobiliários no endividamento de ENFs brasileiras

Alexandria é uma base de dados de valor inestimável para o estudo acadêmico nas áreas de Contabilidade e Economia no Brasil. A base contém uma amostra de empresas cuja expressão na economia nacional é evidente (por exemplo, por representarem cerca de metade de todo o crédito SFN tomado pelas ENFs e pouco menos da metade de todos os fluxos de recebimento). Sua construção abre espaço para debates que antes eram impossíveis e permite verificar se conclusões usando dados de empresas listadas são aplicáveis a empresas não listadas (que participam ou não do mercado de valores mobiliários).

A análise do endividamento empresarial, feita neste trabalho, é um estudo exploratório que exemplifica o potencial de Alexandria.

A maior parte das análises feitas sugerem que o setor influencia o endividamento. Entretanto, há diferenças sutis nas conclusões de cada análise.

As análises visual e não paramétrica sugerem que os setores de “Comércio” e “Serviços” tendem a ter endividamento mediano maior que os demais. Na outra ponta, os setores de “Construção” e “Agropecuária” aparecem com endividamento mediano relativamente baixo, e o setor de “Serviços Imobiliários” tem endividamento menor ainda⁵¹.

A parte das análises paramétricas (econométricas) mostra que as empresas de comércio tendem a ter endividamento médio maior, mas em algumas regressões outros setores aparecem empatados com comércio como os mais endividados. Na outra ponta, o setor de “Serviços Imobiliários” é apontado como tendo o endividamento médio menor.

A conclusão sobre o impacto da participação no mercado de valores mobiliários foi muito clara na avaliação não paramétrica (“a participação aumenta o endividamento mediano”), mas não na avaliação paramétrica (“a participação aumenta o endividamento médio” apenas no painel desbalanceado).

Tanto para a análise do setor quanto para a análise da participação no mercado de valores mobiliários, as análises não paramétricas, que dispõem de uma amostra maior, mostram diferenças muito mais marcantes. Parece-nos, em especial, *que as empresas grandes e as empresas CVM – mais concentradas no painel balanceado - são muito mais parecidas entre si, independente do setor*. As empresas menores são muito mais heterogêneas, e o perfil setorial é mais marcante.

Em linha com a literatura internacional, as análises sugerem que a variável setor capta muitas variáveis não observadas pelo economista, como as diferenças de competitividade, ciclo de negócios e regulação, que muitas vezes estão associadas a cada setor. Não nos parece, entretanto, que a tangibilidade seja um fator tão importante entre

⁵¹ Esses três setores com endividamento menor que os demais têm em comum os ciclos longos de negócios, Entretanto, em especial para a Agropecuária, o resultado (endividamento baixo) pode surpreender dada a oferta de crédito subsidiado para o setor. Nossa amostra é grande (entre 288 observações e 493 observações por ano para este setor), mas mesmo assim é um tema que merece mais estudos – em geral esse setor não aparece nos estudos que encontramos.

empresas não CVM: um dos setores com menos ativos tangíveis, comércio, parece ser o um dos mais endividados.

Conclui-se, assim, que os estudos sobre endividamento empresarial utilizando apenas empresas listadas, e muitas vezes apenas análise paramétrica, devem ser analisados levando em conta as suas limitações, uma vez que os resultados se alteram bastante de acordo com a amostra e a técnica utilizada. De todo modo, a maioria das análises feitas aqui sugerem que: (i) empresas que participam do mercado de valores mobiliários são ligeiramente mais endividadas que as demais, (ii) o endividamento varia significativamente entre setores; (iii) as empresas não alteram muito seu próprio nível de endividamento ao longo do tempo, mas diferem muito entre si, mesmo dentro do próprio setor. Isso reduz o poder explicativo de análises que utilizam somente a média, como fazem alguns trabalhos.

5.2 Limitações do estudo

Evidentemente, este estudo preliminar apresenta limitações, a serem compensadas em estudos posteriores. Em primeiro lugar, a amostra de ENFs de Alexandria *não é aleatória*, e por isso provavelmente não representa adequadamente a população de ENFs que produz demonstrativos financeiros.

Em especial, os resultados para o setor de comércio devem ser olhados com cuidado, pois apesar da amostra grande, a mediana do endividamento de Alexandria se mostrou maior que a mediana de outra base analisada pelos autores (Klooks).

Em segundo lugar, o *erro de mensuração das variáveis contábeis* provavelmente não é uniformemente distribuído na amostra: empresas CVM e não CVM de grande porte (mais concentradas no painel balanceado) têm contabilidade mais precisa e demonstrações contábeis auditadas, mas isso está longe de ser regra entre empresas não-CVM de médio porte. Assim, erros de mensuração podem afetar as conclusões. Acreditamos, em especial, que o ativo das empresas não CVM de médio porte está, em geral, registrado por um valor abaixo do real (pois pelas regras contábeis é muitas vezes medido pelo custo histórico). Se nossa suposição estiver certa, esse grupo *é menos endividado do que os números do balanço patrimonial sugerem*⁵², o que tornaria nossa

⁵² Isso porque o endividamento é calculado a partir da razão entre valores registrados no Balanço patrimonial: $\frac{\text{Passivo registrado}}{\text{Ativo registrado}}$. Se o ativo da empresa não CVM estiver subavaliado mas o passivo não, o endividamento real da empresa é menor que o dado contábil sugere. Se, por exemplo, o valor de mercado

conclusão – de que a participação no mercado de valores mobiliários aumenta o endividamento – mais forte.

Em terceiro lugar, o fato de Alexandria não ter informações de lucro para todas as empresas não financeiras prejudica a comparação entre as análises paramétricas e não paramétricas: enquanto as análises paramétricas (econômicas) utilizaram apenas as empresas com esse dado, as análises não paramétricas utilizaram amostra maior. É possível que essa limitação explique por que o endividamento das empresas de agropecuária e construção parecem muito menor que os demais setores apenas na análise não paramétrica: porque a amostra utilizada nessa análise é maior e inclui mais empresas de médio porte. Sob este ponto de vista, entendemos que a análise não paramétrica é mais acurada para explicar as diferenças de endividamento setorial.

Finalmente, em sua versão 1.0, Alexandria não possui algumas rubricas consideradas importantes por outros autores para explicar o endividamento, em especial a tangibilidade dos ativos e a dedutibilidade de impostos.

6. Referências

AÏHOUNTON, G. e 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. Disponível em: https://www.econstor.eu/bitstream/10419/211087/1/IFRO_WP_2019_10.pdf. Acesso em 11 mai. 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. Revista Colombiana de Estadística, v. 31, p. 241-260, 2008. Disponível em: <http://repositorio.ufc.br/handle/riufc/41465>. Acesso em 22 dez. 2021.

ALMEIDA et al.. Substituição da captação dos dados do CAGED pelo eSocial: implicações para as estatísticas de emprego formal. IPEA, 2020. Disponível em http://repositorio.ipea.gov.br/bitstream/11058/10212/1/BMT_69_substituicao_da_capacidade.pdf. Acesso em 16 jul. 2021.

ANDRTZKY, J. Leverage, Industries, and Countries: Evidence from Seven Countries. Nov. 2003. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=469942. Acesso em 06 mai. 2022.

ANS. Demonstrações Contábeis. 2021. Disponível em: <https://www.gov.br/ans/pt-br/aceso-a-informacao/perfil-do-setor/dados-e-indicadores-do-setor/demonstracoes-contabeis>. Acesso em 16 dez. 2021.

ATKINSON, A. et al. The box-cox transformation: review and extensions. 2021. Statistical Science, 36 (2). 239 - 255. ISSN 0883-4237. Disponível em: <http://eprints.lse.ac.uk/103537/>. Acesso em 18. Mai. 2020.

do ativo seja igual a 120% o ativo registrado, teríamos que endividamento real = $\frac{\text{Passivo registrado}}{1,2 \cdot \text{Ativo registrado}} = \frac{1}{1,2} \cdot \frac{\text{Passivo registrado}}{\text{Ativo registrado}} \cong 0,83 \cdot \text{endividamento medido no Balanço Patrimonial.}$

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). 2019. In: Revista de Gestão em Sistemas de Saúde. Disponível em: <https://periodicos.uninove.br/revistargss/article/view/11537/7948>. Acesso em 10 jan. 2021.

AZEVEDO, Alexandre Bonfim de. Fatores determinantes do endividamento de empresas por setor no Brasil. 2013. Dissertação de Mestrado. Disponível em: <https://repositorio.insper.edu.br/handle/11224/757>. Acesso em 10 jan 2022.

BAJGAR, Matej et al.. Coverage and representativeness of Orbis data. OECD Science, Technology and Industry Working Papers, 2020/06. Disponível em: <https://www.oecd-ilibrary.org/docserver/c7bdaa03-en.pdf?expires=1614629919&id=id&accname=guest&checksum=CA1AFFDD06D086D931025D8966E3DF10>. Acesso em: 1 Mar. 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. Disponível em: <https://www.scielo.br/j/rbgn/a/gPYpsvqrBQKDLgLqH8LHNPz/?lang=en&format=pdf>. Acesso em 16 jun. 2022.

BARTLETT, R. e PARTNOY, F. The Ratio Problem. Mai. 2020. Disponível em <https://ssrn.com/abstract=3605606> e <https://www.law.nyu.edu/sites/default/files/The%20Ratio%20Problem.pdf>. Acesso em 07 mai. 2022.

Banco Central do Brasil (BCB). Crédito ampliado ao setor não financeiro. 2018. In: Relatório de Economia Bancária. P. 56-61. Disponível em: https://www.bcb.gov.br/content/publicacoes/relatorioeconomiabancaria/reb_2018.pdf. Acesso em: 28 dez. 2021.

_____. Relatório de Investimento Direto – 2019. Nov. 2019. Disponível em: https://www.bcb.gov.br/content/publicacoes/relatorioidp/RelatorioID2018/RID_2019.pdf. Acesso em: 22 jul. 2020.

_____. RDE-IED - Manual do Declarante. Maio de 2020. 2020a. Disponível em: https://www.bcb.gov.br/content/estabilidadefinanceira/rde/manuais_RDE/Manual-RDE-IED.pdf. Acesso em 01 jun. 2020.

_____. Censo de Capitais Estrangeiros no País - Manual do Declarante 2020. 2020b. Disponível em: <https://www.bcb.gov.br/content/estabilidadefinanceira/cambiocapitais/Manual2020.pdf>. Acesso em: 07 jul 2020.

_____. Sistema de Pagamentos Instantâneos (SPI). 2021a. Disponível em: <https://www.bcb.gov.br/estabilidadefinanceira/sistemapagamentosinstantaneos>. Acesso em 08 out. 2021.

_____. O que é Pix?. 2021b. Disponível em: <https://www.bcb.gov.br/estabilidadefinanceira/pix>. Acesso em 08 out. 2021.

_____.Relatório de Estabilidade Financeira. Abril 2021c. Disponível em: <https://www.bcb.gov.br/publicacoes/ref/202104>. Acesso em: 30 jun 2023.

_____. Tabelas Especiais - Dívida externa bruta e dívida externa de curto prazo por vencimento residual. 2022. Disponível em: https://www.bcb.gov.br/content/estatisticas/Documents/Tabelas_especiais/DivExt_T.xls. Acesso em: 26 mar 2022.

BANCO DE PORTUGAL. Quadros do setor e quadros da empresa e do setor. Fev. 2019. In: Estudos da Central de Balanços – 36. Disponível em <https://www.bportugal.pt/sites/default/files/flexpapers/2024100/index.html>.

_____. Análise setorial das sociedades não financeiras em Portugal | 2018. 2019b. Estudos da Central de Balanços, v. 40. Disponível em: https://www.bportugal.pt/sites/default/files/anexos/pdf-boletim/estudos_da_cb_40_2019.pdf. Acesso em: 06 jan. 2021.

_____. Central de Balanços. 2021. Disponível em: <https://www.bportugal.pt/page/central-de-balanços?mlid=1045>. Acesso em: 20 out. 2021.

BARROS, L., et al. Endogeneity in panel data regressions: methodological guidance for corporate finance researchers. *Rev. Bras. Gest. Neg. (Online)* 22 (spe), 2020. Disponível em: <https://doi.org/10.7819/rbgn.v22i0.4059>. Acesso em 25 fev. 2022.

BASTOS, D. e 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. Disponível em: <https://www.scielo.br/j/rcf/a/3LpqmV393TZF7wTddnT7NBq/?lang=pt&format=html#>. Acesso em 10 fev. 2022.

BELLÉGO, C.; BENATIA, D. e PAPE, L-D. Dealing with Logs and Zeros in Regression Models. *CREST - Série des Documents de Travail* n° 2019-13. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3444996. Acesso em: 25 fev. 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. Disponível em: <https://www.ucl.ac.uk/~uctp39a/Blundell-Bond-1998.pdf>. Acesso em 07 out. 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. Disponível em: <https://www.sciencedirect.com/science/article/pii/S156601410300061X>. Acesso em: 24 fev. 2022.

BRASIL. Lei n. 11.638, de 28 de dezembro de 2007. Altera e revoga dispositivos da Lei no 6.404, de 15 de dezembro de 1976, e da Lei no 6.385, de 7 de dezembro de 1976, e estende às sociedades de grande porte disposições relativas à elaboração e divulgação de demonstrações financeiras. Disponível em http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/lei/111638.htm. Acesso em 21 mai. 2023.

_____. Lei n. 9.656, de 3 de junho de 1998. Dispõe sobre os planos e seguros privados de assistência à saúde. Disponível em http://www.planalto.gov.br/ccivil_03/leis/19656.htm. Acesso em 16 de. 2021.

_____. LEI Nº 5.172, DE 25 DE OUTUBRO DE 1966 (Código Tributário Nacional). Disponível em http://www.planalto.gov.br/ccivil_03/leis/15172compilado.htm. Acesso em 10 mar. 2021.

_____. Lei Nº 14.286, de 29 de dezembro de 2021. Dispõe sobre o mercado de câmbio brasileiro [...]. Disponível em https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14286.htm. Acesso em 01 jul. 2023.

BRESSAN, Valéria Gama Fully et al. Análise dos determinantes do endividamento das empresas de capital aberto do agronegócio brasileiro. *Revista de Economia e Sociologia Rural* 47(1), Mar. 2009. Disponível em <https://www.scielo.br/j/resr/a/vKR79rd78fjhBqgDrhYwdpr/?lang=pt#>.

BRITO, Paulo; SERRANO, André e FRANCO, Víthor. Determinantes da Estrutura de Capital de Empresas Brasileiras de Capital Aberto em Período de Crise. In: *Revista Ambiente Contábil*, Vol. 10. n. 2, jul./dez. 2018, ISSN 2176-9036. Disponível em <http://www.atena.org.br/revista/ojs-2.2.3-08/index.php/Ambiente/article/view/3347/2587>. Acesso em 10 fev 2022.

CAMPOS, Claudio. *Confronto das Teorias de Pecking Order e Trade-off: Evidências com base nas companhias Brasileiras Abertas*. São Paulo: 2009. Dissertação de mestrado. Disponível em <https://dspace.mackenzie.br/handle/10899/23547>. Acesso em 21 mai. 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. Dissertação de Mestrado. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/16145/000697442.pdf>. Acesso em 16 abr. 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. *Finanças Estratégicas. RAM, Rev. Adm. Mackenzie* 14 (4). Ago 2013. Disponível em: <https://www.scielo.br/j/ram/a/zvYnQf79jh4zDN79nC3vFZn/abstract/?lang=pt>. Acesso em 21 mai. 2023.

CVM. O que é a CVM. 2022. Disponível em: 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. Acesso em: 14 set 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. Disponível em <https://doi.org/10.1016/j.jbankfin.2014.12.009> e <https://www.sciencedirect.com/science/article/pii/S0378426614003902>. Acesso em 18 mai. 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. Disponível em: <https://doi.org/10.1016/j.jcorpfin.2019.101514>. Acesso em 20 dez. 2021.

DEN BERG, Bas. Firm-specific determinants of capital structure: Evidence on UK listed and unlisted firms. Dissertação de mestrado. University of Twente. 2021. Disponível em: http://essay.utwente.nl/85557/1/vandenBerg_MA_BMS.pdf. Acesso em: 20 fev. 2022.

EXAME. Melhores e Maiores – metodologia. Disponível em <https://mm.exame.com/metodologia/>. Acesso em 21 de jul. de 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. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22697476/>. Acesso em 04 jan 2021.

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. Disponível em: <https://www.scielo.br/j/bar/a/XVC8cq8TYXQ9zzPrj79CnHj/?format=pdf&lang=en>. Acesso em 15 fev 2022.

FENG, C. et al. Log-transformation and its implications for data analysis. IN; *Shanghai Arch Psychiatry.* 2014 Apr; 26(2): 105–109. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4120293/>. Acesso em: 10 mai. 2022.

FERNANDES, Pedro Onofre e CAMPOS, Luiz Fernando de Barros. Fundamentos para uma Central de Balanços Brasileira: o uso da XBRL. Nov. 2014. In: *CONTECSI USP - International Conference on Information Systems and Technology Management - ISSN 2448-1041.* Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.517.2033&rep=rep1&type=pdf>. Acesso em 20 out 2021.

FIPECAFI. Melhores e maiores. 2021. Disponível em: <https://fipecafi.org/Fundacao/FundacaoMelhoresMaiores>. Acesso em 04 out 2021.

FIREBAUGH, G. e GIBBS, J. User's Guide to Ratio Variables. In: *Article in American Sociological Review* · October 1985. Disponível em: <https://www.jstor.org/stable/2095384>. Acesso em: 07 mai. 2022.

FREIRE, Janayna; CARVALHO, Eduardo; e CALEGÁRIO, Cristina. Análise da Relação Grau de Internacionalização versus Indicadores Financeiros de Multinacionais Brasileiras. In: *Revista de Negócios*, v. 26, n. 2, p. 39-56, Abril, 2021. Disponível em: <https://bu.furb.br/ojs/index.php/rn/article/view/9710/5123>. Acesso em: 18 out. 2021.

GAIBULLOEV, K. et al. Dynamic Panel Analysis under Cross-Sectional Dependence. *Political Analysis* (2014) 22:258–273. doi:10.1093/pan/mp029. Disponível em: <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>. Acesso em: 19. Mai. 2022.

GARCÍA, J. e 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>. Acesso em 20. Fev. 2022.

GHASEMI, Asghar e 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. Disponível em <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3693611/>. Acesso em 13 jan 2021.

GOTTARDO, Pietro, e MARIA MOISELLO, Anna. The capital structure choices of family firms: Evidence from Italian medium-large unlisted firms. *Managerial Finance* 40.3 (2014): 254-275. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/MF-03-2013-0065/full/html>. Acesso em 29 jun 2023.

IBGE. Bases de dados»metadados»MTE»relação anual de informações sociais – RAIS. 2020?. Disponível em: <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>. Acesso em 16 jul 2021.

_____. Concla – Comissão Nacional de Classificação. 2021. Disponível em: <https://concla.ibge.gov.br/busca-online-cnae.html?view=estrutura>. Acesso em 30 jul. 2021.

IMF. External Debt Statistics – Guide for compilers and users. 2014. Disponível em <http://tffs.org/pdf/edsg/ft2014.pdf>. Acesso em 30 dez 2021.

_____. Balance of Payments and International Investment Position Manual - Sixth Edition (BPM6). 2009. 371 p. ISBN 978-1-58906-812-4. Disponível em: <https://www.imf.org/external/pubs/ft/bop/2007/pdf/bpm6.pdf>. Acesso em: 23 mai. 2020.

JAWORSKI, J. e 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>. Acesso em 19 fev. 2022.

LEE, D. K. Data transformation: a focus on the interpretation. In: *Korean J Anesthesiol*. 2020 Dec; 73(6): 503–508. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7714623/>. Acesso em 10 mai. 2022.

MARINŠEK, Denis. The Impact of Indebtedness on a Firm's Performance: Evidence from European Countries. Tese de doutorado. Ljubljana, jan. 2015. Disponível em <http://www.cek.ef.uni-lj.si/doktor/marinsek.pdf>. Acesso em 4 abr. 2021.

MOTTA, Elaine Cristina de Oliveira. Juros sobre o Capital Próprio – Função e efeitos na tributação da renda. São Paulo: FGV, 2021. Dissertação de Mestrado. Disponível em: https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/30672/Elaine.dissertacao_versao%20final%20%281%29.pdf?sequence=5&isAllowed=y. Acesso em 19 jul 2021.

NAKAMURA, W.; MARTIN, D. e KIMURA, H. Indicadores Contábeis como Determinantes do Endividamento das Empresas Brasileiras. In: Encontro Nacional dos Programas de Pós-Graduação em Administração, 28, 2004, Curitiba, Anais... Curitiba: EnANPAD, 2004. Disponível em: <https://www.anpad.org.br/admin/pdf/enanpad2004-ccg-1177.pdf>. Acesso em: 10 fev. 2022.

NEHREBECKA, Natalia e BIAŁEK-JAWORSKA, Anna. Determinants of Polish Companies' debt Financing Preferences. 2015. Social Sciences 87(1). Disponível em: <https://doi.org/10.5755/j01.ss.87.1.12319>. Acesso em: 07 jan 2021.

NEVES et al.. Qual é o Efeito da Garantia Real na Exigência de Capital para Risco de Crédito no Brasil?. XXXI Encontro da ANPAD – Rio de Janeiro – 22 a 26 de setembro de 2007. Disponível em <http://www.anpad.org.br/admin/pdf/FIN-C720.pdf>. Acesso em 16 de jul. de 2021.

Normas Brasil. Resposta à Consulta nº 8791 DE 11/04/2016. 2016. Disponível em https://www.normasbrasil.com.br/norma/resposta-a-consulta-8791-2016-sp_319569.html. Acesso em 21 out 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. Disponível em: <https://ssrn.com/abstract=2131759>. Acesso em 10 fev. 2022.

POHLMANN, M. e 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. Disponível em: <https://www.scielo.br/j/rcf/a/DS8SPscdbQ8R4CdLZwZtwLv/abstract/?lang=pt>. Acesso em 10 fev. 2022.

PROCIANOY, Jairo Laser e SCHNORRENBARGER, Adalberto. A Influência da Estrutura de Controle nas Decisões de Estrutura de Capital das Companhias Brasileiras. Rio de Janeiro: RBE, 58(1):121-146 jan/mar 2004. Disponível em <https://www.scielo.br/j/rbe/a/PBWn7Bmw85DZsZsRKRCndxv/abstract/?lang=pt>. Acesso em 17 dez. 2021

RAALTE, Floris van. The impact of firm-specific determinants on the capital structure of Dutch listed and non-listed companies. 2021. Tese de mestrado, University of Twente. Disponível em: http://essay.utwente.nl/87321/1/Van%20Raalte_BA_Behavioural%2C%20Management%20and%20Social%20Sciences.pdf. Acesso em 19 fev. 2022.

Receita Federal do Brasil (RFB). Dados Públicos CNPJ. 2021a. Disponível em: <https://www.gov.br/receitafederal/pt-br/assuntos/orientacao-tributaria/cadastros/consultas/dados-publicos-cnpj>. Acesso em 16 jul. 2021.

_____. SIC_Quantidade_de_CNPJs_em_docs_pagos_com_NJ.xlsx. 2021b. Resposta, em 26 abr. 2021 à manifestação 18870.001442/2021-22.

_____. Instrução Normativa RFB nº 2003, de 18 de janeiro de 2021. 2021c. Disponível em <http://normas.receita.fazenda.gov.br/sijut2consulta/link.action?visao=anotado&idAto=114965#2223871>. Acesso em: 21 jul. 2021.

RIBEIRO, Samuel, MENGHINELLO, Stefano e 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. Disponível em: https://www.oecd-ilibrary.org/economics/the-oecd-orbis-database_5kmhds8mzj8w-en. Acesso em: 25 mar. 2021.

SANT'ANA, José Augusto de Almeida. A relação entre endividamento e rentabilidade nas empresas brasileiras de 1996 a 2000. 2001. Dissertação de mestrado. Disponível em: https://repositorio.unb.br/bitstream/10482/15203/1/2001_JoseAugustodeAlmeidaSantAna.pdf. Acesso em 21. dez. 2021.

SILVA, Verônica Favato. Determinantes da estrutura de capital na América Latina e nos Estados Unidos. Universidade Federal de Uberlândia, Uberlândia, 2007. Dissertação de mestrado. Disponível em: <https://repositorio.ufu.br/handle/123456789/12034>. Acesso em 10 fev. 2022.

SILVA, Mailliw. Crédito Direcionado e Desenvolvimento Econômico: uma avaliação da política de crédito brasileira. Dissertação de Mestrado. 2018. Disponível em: <https://www.teses.usp.br/teses/disponiveis/12/12138/tde-12062018-111324/publico/CorrigidaMailliw.pdf>. Acesso em 16 jul. 2021.

SILVA, Marcos Roberto Alves da. Um Painel Setorizado das Variáveis Determinantes da Estrutura de Capital de Empresas Brasileiras. In: *Brazilian Journal of Business*, v. 3, n. 1, 2021. Disponível em: <https://brazilianjournals.com/index.php/BJB/article/view/26314/20866>. Acesso em: 10 fev. 2022.

SPED – Sistema Público de Escrituração Digital. Estatísticas ECD Junta. 2023. Disponível em <https://www.sped.fazenda.gov.br/estatisticascontabil/Default.aspx?ano=2021>. Acesso em 15 abr. 2023.

TERRA, Paulo Renato Soares. Estrutura de capital e fatores macroeconômicos na América Latina. *R. Adm.*, v. 42, p. 192-204, abr/maio/junho 2007. Disponível em: <https://www.revistas.usp.br/rausp/article/view/44435/48055>. Acesso em 17 dez. 2021.

URIONABARRENETXEA, Sara, SAN-JOSE, Leire e RETOLAZA, Jose-Luis. Negative equity companies in Europe: theory and evidence. *Business: Theory and Practice* 17(4) 2016, 307–316. Disponível em: <https://doi.org/10.3846/btp.17.11125>. Acesso em 06 jan. 2022.

VARGHA, A. e DELANEY, H. The Kruskal-Wallis Test and Stochastic Homogeneity. *Journal of Educational and Behavioral Statistics*. Summer 1998, Vol. 23, No. 2, pp. 170-192. Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.899.3263&rep=rep1&type=pdf>. Acesso em: 05 jan 2021.

VALORPRO. Serviço de Informações em Tempo Real. 2021. Disponível em: <https://valorpro.globo.com/#/downloads>. Acesso em: 19 jul 2021.

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

ZONENSCHAIN, Claudia Nessi. Estrutura de Capital das Empresas no Brasil. *Revista do BNDES*, Rio de Janeiro, V. 5, N. 10, P. 63-92, Dez. 1998. Disponível em: <https://web.bndes.gov.br/bib/jspui/handle/1408/13505>. Acesso em: 24 fev. 2002.

Anexo A - Descrição da base Alexandria

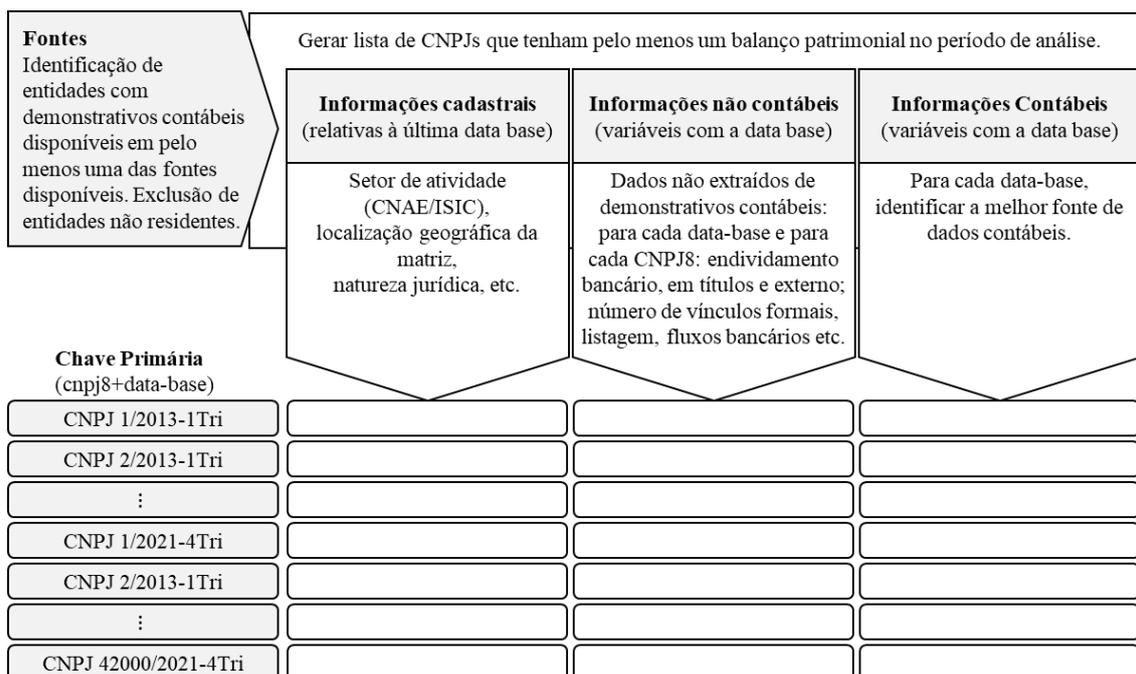
A base de dados Alexandria, construída pelos autores deste trabalho, é uma base de entidades contábeis formais residentes no Brasil – em sua maioria, empresas não financeiras (ENFs) de médio porte. Em sua versão 1.0, contém cerca de 18,0 mil demonstrativos financeiros anuais por ano, em média (dos quais, 13,3 mil de ENFs), e informações que abrangem o período de 2013T1 a 2021T4, sendo ampliada trimestralmente. O período de análise da versão 1.0 precede a lei nº 14.286 de 29 de dezembro de 2021 (Brasil, 2021) - “novo marco cambial” -, que alterou a coleta de dados de capitais estrangeiros no país pelo Banco Central do Brasil (BCB). Portanto, a amostra de Alexandria não sofreu mudanças metodológicas decorrentes dessa lei.

Na visão dos autores deste trabalho, Alexandria é uma contribuição relevante para os estudos acadêmicos brasileiros nas áreas de Contabilidade, Finanças e Economia, pelos seguintes motivos:

- Alexandria é uma base de dados de tamanho expressivo que reúne empresas de médio e grande porte do Brasil. É superior, em tamanho, à principal base utilizada em trabalhos acadêmicos sobre empresas não listadas (a base Maiores e Melhores).
 - Alexandria também inova ao conter base única de mais de 4.300 demonstrativos por trimestre desde 2016.
- Alexandria cruza dados contábeis com dados sigilosos e proprietários do Banco Central do Brasil e de outros órgãos do governo, disponíveis ao BCB via convênio. Nenhuma outra base com dados contábeis no Brasil tem essa característica.

A Figura 10 ilustra os passos da construção de Alexandria, a serem detalhados nas próximas subseções. Em resumo, os autores tomaram todas as fontes disponíveis com demonstrativos contábeis e com isso geraram uma lista de entidades contábeis. Dessas fontes, obtiveram os demonstrativos contábeis. De outras fontes, obtiveram informações cadastrais, de fluxo bancário, de número de empregados etc.

Figura 10 – Representação esquemática da construção de Alexandria



Fonte: elaborado pelos autores.

A.1 Escopo de Alexandria

Alexandria é uma base de dados que reúne informações de entidades contábeis formais residentes no Brasil que tenham pelo menos um balanço patrimonial disponível⁵³ entre 2013 e 2021. Utiliza como chave primária os primeiros 8 dígitos do CNPJ – o “CNPJ8”⁵⁴ - combinado com a data base. Assim, para o CNPJ8 “01234567” e para a data base “2013-1º Tri”, a chave primária seria “01234567/2013-1Tri”.

Os critérios escolhidos para participar de Alexandria (entidade residente e com as principais rubricas do balanço patrimonial) foram propositalmente amplos, para permitir que a Alexandria 1.0 e suas versões posteriores pudessem acolher entidades de diferentes tipos. Como a escolha dos CNPJ8s de Alexandria não foi aleatória (baseou-se na

⁵³ Embora Alexandria traga informações contábeis das principais rubricas do Balanço Patrimonial e da Demonstração do Resultado do Exercício, além de duas rubricas da Demonstração da Mutação do Patrimônio Líquido, nem sempre todas as rubricas estão disponíveis para todas as entidades. Assim, o critério mínimo utilizado foi a disponibilidade de informações de ativo total, passivo total e patrimônio líquido total em pelo menos uma data-base.

⁵⁴ As entidades formais são reconhecidas no Brasil pelo seu número de identificação junto à Receita Federal do Brasil (RFB), aqui chamado “CNPJ14” por ter 14 dígitos. Para os estabelecimentos ou filiais de uma mesma entidade, os primeiros 8 dígitos de um CNPJ - o “CNPJ8” - são iguais. Assim, a utilização do “CNPJ8” significa que a entidade está sendo considerada com todas as suas filiais e estabelecimentos.

disponibilidade de informações contábeis), as análises devem levar em conta esse possível viés.

Como Alexandria é baseada nos demonstrativos contábeis, há certas entidades que dificilmente serão captadas, por não serem obrigadas a produzi-los. São exemplos dessas entidades: (i) microempresas ou empresa de pequeno porte (RFB, 2021c, art. 3); (ii) pessoas jurídicas imunes e isentas (como igrejas e sindicatos) que auferiram receita (num conceito amplo) anual inferior a R\$ 4,8 milhões (RFB, 2021c) e (iii) produtores rurais pessoa física (Normas Brasil, 2016). Se estas entidades decidirem produzir demonstrativo contábil mesmo não sendo obrigadas a isso, elas podem fazer parte da Alexandria, mas esse é um caso raro. Isso significa, na prática, que Alexandria contém em geral empresas de médio e grande porte (simplificadamente, aquelas com receita igual ou superior a R\$ 4,8 milhões).

A.2. Entidades na Alexandria: datas-base válidas e disponibilidade de informação

Alexandria inclui informação do primeiro trimestre de 2013 ao último trimestre de 2021 – um total de 36 trimestres ou datas-base. Entretanto, nem todas as entidades estão ativas em todas as datas-bases. Para cada CNPJ8, tomou-se a “data de abertura” e a “data de fechamento” (se houver) e definiu-se as datas-bases válidas para cada entidade. Quando o CNPJ8 consta como baixado, a data de fechamento é inferida como sendo a última data de alteração do CNPJ.

Para cada CNPJ8 e para cada data-base válida, Alexandria contém uma série de informações sobre a entidade. Entretanto, nem todas as informações estão disponíveis para todas as datas-bases válidas. Em especial, os demonstrativos contábeis estão concentrados na data-base do quarto trimestre e são geralmente informações anuais, pois a maioria das empresas produz demonstrativos trimestrais. Além disso, muitas vezes o balanço simplesmente não está disponível para uma data-base válida (porque não consta em nenhuma das fontes que Alexandria utiliza).

Os 31.233 CNPJ8s de ENFs constantes na Alexandria têm, no total, 936.873 datas-base válidas, o que significa que cada CNPJ8 acompanhado ficou ativo em média 29,99 trimestres. A maioria dos CNPJ8s estava ativo em todas as datas-bases válidas⁵⁵.

⁵⁵ Por questão de espaço e escopo, os comentários desta seção A.2 se restringem a ENFs, apesar de Alexandria conter outras entidades contábeis.

Das 936.873 datas-bases válidas, Alexandria tem demonstrativos financeiros em 166.958 delas - dos quais, 119.839 são demonstrativos anuais. Isso significa que Alexandria tem em média 3,8 demonstrativos financeiros anuais de cada ENF.

Portanto, em termos econométricos, Alexandria é um painel desbalanceado, já que nem todas as entidades estão presentes em todas as datas-base, e algumas variáveis, notadamente demonstrativos financeiros, não estão disponíveis nem mesmo para todas as datas-bases válidas (Wooldridge, 2010, p. 828)⁵⁶. Trata-se, além disso, de um painel curto, em que o número de datas-bases (36) é muito inferior ao número de observações (31.233, considerando apenas ENFs).

A.3 Tipos de informações em Alexandria

Alexandria é uma base de dados que compila e organiza informações de outras bases de dados. As bases compiladas podem ser divididas em dois grandes grupos.

O primeiro grande grupo corresponde às bases censitárias, ou seja, aquelas que fornecem informações para todos a população de CNPJ8s e – no caso de informações que mudam ao longo do tempo - para todas as datas válidas. É o caso da base de dívida externa (ver seção A.7.3): como praticamente toda dívida externa com recursos ingressados no Brasil deve ser registrada no módulo RDE-ROF⁵⁷ do Banco Central do Brasil, todas as datas válidas contêm alguma informação (mesmo que seja a informação de que a entidade “x”, na data base “y”, tem passivo externo igual a zero).

O segundo grande grupo corresponde às bases amostrais, ou seja, aquelas que fornecem informação para algumas datas válidas e para algum CNPJ8. É o caso das informações de empregados formais (seção A.9) e das informações contábeis (seção A.5).

A.4 Base de dados censitária: cadastro da Receita Federal do Brasil

Alexandria utiliza o cadastro da Receita Federal do Brasil (RFB) para obter dados cadastrais da situação atual da entidade. Estes dados são abertos ao público e representam a totalidade das entidades formais (com CNPJ8). A explicação detalhada de cada variável está em RFB (2021a).

⁵⁶ Wooldridge (2010, p. 809 e 828) comenta as consequências econométricas de um painel não balanceado em que a razão para uma variável ou dado estar faltando pode estar correlacionada com o erro.

⁵⁷ O Sistema RDE-ROF passou a se chamar SCE-Crédito em decorrência do novo arcabouço normativo sobre capitais estrangeiros no país (Brasil, 2021). Aqui utilizamos o nome antigo porque, como já mencionado, captamos as variáveis antes das mudanças trazidas por esta lei.

Os dados apresentam apenas a situação atual da entidade. Assim, por exemplo, se a entidade mudou sua sede de município no período de análise, Alexandria vai considerar somente o último município (que está no cadastro)⁵⁸. Entretanto, no caso de sócios e gestores, existe a informação da data de entrada, o que permite analisar os sócios ao longo do tempo.

Os autores utilizaram as informações dos sócios e administradores para identificar o gênero dos gestores em data base válida. Considerou-se como gestor em primeiro lugar todo adulto (18 a 85 anos) vivo na data base, presente no quadro societário (ou na administração) antes ou na data base que tenha sido formalmente designado como administrador⁵⁹. No caso raro de falta dessa informação, utilizou-se a informação dos adultos vivos e presentes no quadro societário que não fossem formalmente designados como gestores. Na falta dessa segunda informação, todas as pessoas físicas do quadro societário foram consideradas gestoras. Essas informações permitiram a construção de uma variável (mutável ao longo do tempo) que identifica se a entidade tem ou não pelo menos uma mulher na gestão.

Para as empresas limitadas, os autores também criaram um identificador de empresa familiar. Toda empresa detida por duas ou mais pessoas físicas com o mesmo sobrenome foram identificadas como empresas familiares, independente da participação desse grupo no capital social. Além disso, empresas detidas por pessoas jurídicas que são empresas familiares também foram consideradas empresas familiares (por serem detidas indiretamente por famílias).

A.5 Bases de dados amostrais: demonstrativos contábeis (RDE-IED, Censo de Capitais Estrangeiros, Económica, Valor Pro e ANS)

A grande vantagem de Alexandria é a disponibilização das principais rubricas dos demonstrativos contábeis de empresas listadas e não listadas. Essa inovação exigiu um esforço de união de fontes que continham, individualmente e para algumas datas-bases, algumas entidades, mas que reunidas representam um grupo mais expressivo de entidades e de datas-bases.

⁵⁸ Normalmente, todos os CNPJ14 têm a mesma natureza jurídica e desenvolvem a mesma atividade (CNAE/ISIC). Nos poucos casos em que isso não acontece, a Alexandria classifica a entidade de acordo com os dados cadastrais da matriz.

⁵⁹ Considerou-se simplificada que administradores, conselheiros, diretores e presidentes são “gestores”.

Foram coletados apenas dados de balanços individuais, e não consolidados, pela maior disponibilidade de dados.

As fontes dos demonstrativos contábeis coletadas por Alexandria são quatro: BCB (subdividida em RDE-IED e Censo de Capitais Estrangeiros⁶⁰), Economática, ValorPro e ANS. Estas fontes estão descritas nas próximas subseções.

A.5.1 Fonte de dados de demonstrativos contábeis e das informações sobre sócios não residentes: BCB (RDE-IED e Censo de Capitais Estrangeiros)

O BCB, por ser o responsável por coletar estatísticas do setor externo, coleta informações contábeis e não contábeis sobre empresas detidas por não residentes, através de dois sistemas: RDE-IED e Censo de Capitais Estrangeiros. Por serem a principal fonte de balanços da Alexandria e por serem pouco exploradas na literatura, esses sistemas mereceram espaço maior.

Todas as empresas⁶¹ e fundos de investimento residentes⁶² no Brasil que sejam detidos por não residentes em algum percentual devem declarar informações contábeis resumidas ao BCB – seja no Censo de Capitais Estrangeiros, no RDE-IED ou em ambos. A frequência da declaração depende principalmente do porte da entidade. De toda forma, trata-se de uma base de dados censitária⁶³. Por isso, Alexandria é uma base que contém mais empresas detidas por não residentes do que empresas inteiramente detidas por residentes.

A.5.1.1 O Censo de Capitais Estrangeiros no País (Censo)

O Censo de Capitais Estrangeiros no País, em suas versões Anual e Quinquenal, é uma pesquisa eletrônica obrigatória para pessoas jurídicas detidas por não residentes e/ou com saldo devedor total de créditos comerciais de curto prazo acima de determinado

⁶⁰ O Sistema RDE-IED passou a se chamar SCE-IED em decorrência do novo arcabouço normativo sobre capitais estrangeiros no país (Brasil, 2021). Aqui utilizamos o nome antigo porque, como já mencionado, captamos as variáveis antes das mudanças trazidas por esta lei.

⁶¹ Pessoas físicas, incluindo empresários individuais, não podem receber investimento de não residentes. Por isso, aqui não usamos o termo genérico “entidade contábil”, e sim “empresa” (entidade contábil empresarial que é pessoa jurídica).

⁶² O leitor que não estiver familiarizado com o conceito de residência pode consultar IMF (2009, cap. 4, item E).

⁶³ Deixou de ser censitário com a entrada em vigor da lei nº 14.286/2021, principalmente devido a implantação dos pisos declaratórios para a identificação do par entre receptora e investidor não residente, mas não afeta a amostra desse estudo (Brasil, 2021).

montante⁶⁴. A pesquisa capta dados contábeis, societários, de endividamento externo, entre outros (BCB, 2019, p. 48-51).

A.5.1.2 O RDE-IED

O RDE-IED é um módulo de registro obrigatório de investimento de não residentes em empresas residentes no país. Apesar do nome “IED” (que remete à posição de investimento direto segundo o conceito do BPM5 - hoje o nome seria IDP), o RDE-IED não capta apenas investimento direto no país e nem capta todo o investimento direto (BCB, 2020a).

No caso de empresas, o RDE-IED é um sistema de grande abrangência: capta todas as participações de investidores não residentes em empresas não listadas e algumas participações em empresas listadas.

- Como o critério de voto não é levado em conta no registro do RDE-IED, esse sistema capta, para essas empresas, tanto o investimento direto no país (IDP) quanto o investimento de portfólio.
- O RDE-IED capta as participações de investidores não residentes no capital das empresas residentes. Assim, a parcela de investimento externo que chega pelo mercado de capitais não fica registrado no RDE-IED e sim no módulo RDE-Portfólio, pois esses investidores não chegam a integralizar o capital e tampouco integram o quadro societário da empresa receptora.

A abrangência do RDE-IED não se aplica a fundos de investimento, que não são obrigados a fazer declarações nesse módulo. Assim, suas posições de investimento direto e de portfólio só são captadas pelo Censo (BCB, 2020b).

As empresas residentes declarantes do RDE-IED devem preencher informações econômico-financeiras no sistema, informações essas que foram incorporadas à Alexandria.

Sempre que ocorre mudança societária, a empresa declarante deve preencher uma declaração contábil e societária (denominada Quadro Societário, ou QS). Essa declaração

⁶⁴ O saldo devedor de créditos comerciais de curto e de longo prazo passou a ser captada exclusivamente pelo sistema SCE-Crédito (antigo RDE-ROF) a partir da data-base de 31.12.2021 (Brasil, 2021).

contém as seguintes variáveis: ativo total, patrimônio líquido e capital integralizado total (de todos os sócios, residentes ou não). Infelizmente, não contém dados de lucro líquido.

Empresas com ativo e patrimônio líquido inferior a R\$ 250 milhões também devem preencher o QS anualmente, mesmo quando não ocorrer mudança societária. O QS anual, obrigatório desde 2016, é uma declaração periódica e uma fonte de milhares de balanços para Alexandria em cada data-base.

Empresas com ativo ou patrimônio líquido igual ou superior a R\$ 250 milhões devem preencher a Declaração Econômico Financeira (DEF) trimestralmente e não precisam preencher o QS anualmente.

A DEF tem alta qualidade, sendo uma das raras fontes de dados trimestrais de empresas não listadas do país. Além dos dados informados no QS, a DEF contém informações de lucro líquido e distribuição de dividendos.

O Quadro 6 resume a obrigatoriedade de preenchimento de informações contábeis no Censo e no RDE-IED. As informações contábeis e societárias armazenadas nos dois sistemas são em grande parte comparáveis⁶⁵.

Quadro 6 – Obrigatoriedade de preenchimento de informações contábeis no Censo e no RDE-IED de acordo com o atingimento do critério em 31 de dezembro do ano-base

Sistema		Pessoas jurídicas sediadas no país (exceto fundos de investimento), com participação direta de não residentes em seu capital social	Fundos de investimento residentes com cotistas não residentes
Censo de Capitais Estrangeiros	Quinquenal	Todas	Todas
	Anual	Patrimônio líquido igual ou superior a US\$100 milhões	
RDE-IED	DEF (trimestral)	Ativo ou Patrimônio líquido superior a R\$ 250 milhões	----
	QS (anual)	Todos, exceto os que declararam DEF	-----

Fonte: BCB (2019, p. 48-51). Em alguns casos, pessoas jurídicas sediadas no país, com saldo devedor total de créditos comerciais de curto prazo (exigíveis em até 360 dias) concedidos por não residentes também devem preencher o Censo, mas só precisam preencher informações contábeis se cumprirem algum dos requisitos acima.

A.5.2 Fonte de dados de demonstrativos contábeis: Econômica/ CVM

A base de dados Econômica disponibiliza aos assinantes informações contábeis e financeiras de empresas de 45 países – a grande maioria delas listadas. É largamente

⁶⁵ Em especial a partir de 2017, inclusive, os dados contábeis do RDE-IED são de excelente qualidade, pois houve atualização do sistema para *interface web* em janeiro de 2017. Essa atualização manteve os dados anteriores, mas criou uma interface mais amigável para o declarante, além de simplificar diversas regras de negócio. Com isso, facilitou o registro (para o declarante), o monitoramento dos dados (para a equipe do Banco Central do Brasil) e consequentemente a qualidade do dado final.

utilizada em estudos acadêmicos e por analistas de mercado. Alexandria utilizou as informações de todas as 1.152 empresas residentes no Brasil e disponíveis na Economática. Essas empresas estavam registradas junto à CVM para negociar valores mobiliários, e são as empresas que este trabalho denomina “empresas CVM”. Note o leitor que uma empresa pode ser “CVM” num trimestre e deixar de ser no trimestre seguinte, e vice-versa.

A.5.3 Fonte de dados de demonstrativos contábeis: Valor Pro

O ValorPro (2021) é um serviço de informações do grupo econômico Globo. Dispõe de uma base de dados empresarial com 7.360 empresas listadas e não listadas que registraram informações contábeis junto à CVM ou publicaram demonstrativo contábil em diário oficial. Prioriza empresas de grande porte, com receita líquida anual acima de R\$ 100 milhões. Também é uma fonte de boa qualidade e utilizada em literatura anterior (MOTTA, 2021, p. 48).

A.5.4 Fonte de dados de demonstrativos contábeis: ANS

A Agência Nacional de Saúde Suplementar (ANS) fiscaliza as pessoas jurídicas de direito privado que operam planos de assistência à saúde. A agência compila e publica há muitos anos as informações contábeis trimestrais das entidades de pequeno porte (Brasil, 1998 e ANS, 2021), definidas como aquelas com menos de 20 mil beneficiários. Avelar et al (2019) já utilizaram essa base para investigar determinantes (inclusive regulatórios) do endividamento das cooperativas médicas.

A.5.5 União das diferentes fontes de dados contábeis

A Figura 11 apresenta um diagrama de Venn com o número de CNPJ8 com dados contábeis disponíveis em cada fonte (descritas das seções 3.5.1 a 3.5.4). O dado entre colchete informa o número de CNPJ8 que são empresas não financeiras [ENFs]. Uma das conclusões que se pode extrair da Figura 11 é que a fonte “BCB” é a mais importante para a construção da base Alexandria: das mais de 42 mil entidades contábeis de Alexandria, 32.499 têm informações contábeis apenas na fonte “BCB”.

250 milhões. Entretanto, esse pequeno grupo representa 94,3% do ativo total de Alexandria nesta data-base. O número deve ser interpretado com cuidado, pois há alguns casos de entidades de Alexandria que detém outras. Mesmo assim, já dá uma dimensão da importância desse grupo de entidades e justifica a atenção concentrada na qualidade destes dados.

Os autores desse trabalho estabeleceram uma hierarquia de qualidade das fontes de dados contábeis, baseada na presença ou não de dados auditados e em extensa validação manual. Assim, se a empresa possui, em determinada data-base, demonstrativos contábeis cuja fonte é a Economática, empresa reconhecida no mercado e especializada em dados, esse dado prevalece na Alexandria, independente de existir em outra fonte.

A segunda base de maior qualidade é a do Valor Pro. Assim, se os seus dados conflitarem com dados informados em outra fonte (exceto Economática), eles prevalecerão. Em seguida, consideramos como de melhor qualidade, nessa ordem: RDE-IED, ANS e Censo. Como os quadros societários (QS) do RDE-IED não captam informações de lucro, receita bruta e dividendos, estes dados foram buscados no Censo quando disponíveis.

Após a seleção da melhor fonte dos demonstrativos contábeis de cada data-base, para cada data-base e cada CNPJ8, obtém-se as principais rubricas do demonstrativo contábil de melhor qualidade referente àquela data base, conforme seleção explicada acima.

É importante notar que, para o mesmo CNPJ8, podem ocorrer fontes diferentes dependendo da data-base. Por exemplo, uma empresa detida por não residentes pode ter declarado o Censo em dezembro de 2015. Em 2016, passou a ser listada, situação em que a fonte utilizada por Alexandria passa a ser a Economática.

Após a estruturação da base, os dados foram exaustivamente comparados para identificação de *outliers* e erros. Especial atenção foi dada a entidades com ativo total superior a R\$ 10 bilhões.

A.6 Base de dado censitária: participação societária de não residentes nas empresas

Além de dados contábeis, o BCB captura, através das pesquisas Censo de Capitais Estrangeiros e RDE-IED (ver seção A.5.1), as informações do quadro societário das empresas e fundos de investimento detidos total ou parcialmente por não residentes de maneira censitária até 2021.

A combinação das duas fontes, Censo e RDE-IED, gerou informações únicas da participação de investidores nas empresas ao longo dos anos. Essas informações, combinadas a dados societários disponíveis na base da Receita Federal do Brasil, transformaram-se em dados trimestrais na Alexandria.

A.7 Base de dado censitária: crédito ampliado ao setor não financeiro

O saldo de crédito ampliado ao setor não financeiro é uma estatística produzida pelo BCB em conformidade com o padrão metodológico internacional que computa a dívida de entidades não financeiras contra (i) instituições do Sistema Financeiro Nacional (SFN), (ii) detentores de títulos e (iii) entidades e indivíduos não residentes (BCB, 2018).

Alexandria utilizou os microdados que deram origem e essa estatística e captou informações de dívida dos CNPJ8s já constantes em Alexandria, separando o saldo ampliado de cada CNPJ8 por tipo de dívida e por prazo. Esta última separação foi feita para comparar essas informações com o passivo circulante e com o passivo não circulante disponíveis no balanço patrimonial. O Quadro 7 resume os três tipos de dívida captados por essa estatística, o sistema que capta o microdado e as variáveis criadas em Alexandria.

Quadro 7 – Componentes do crédito ampliado emprestado ao setor não financeiro, sistema de coleta do microdado e variáveis geradas para Alexandria

Tipo de dívida registrada no sistema		Empréstimos e financiamentos concedidos por instituições financeiras do SFN	Títulos de dívida públicos e privados	Créditos concedidos por não residentes (dívida externa)
Sistema que coleta o microdado [ver seção]		SCR (BCB) [A.7.1]	Registro da B3 [A.7.2]	RDE-ROF (BCB) [A.7.3]
Variáveis geradas para Alexandria a partir do microdado	Curto prazo (CP- vencimento em até 12 meses)	Passivo SCR curto prazo	Passivo em títulos curto prazo	<ul style="list-style-type: none"> • Intercompanhia declarado CP • Intercompanhia inferido CP, • Outros CP
	Longo prazo (LP - vencimento igual ou superior a 12 meses)	Passivo SCR longo prazo	Passivo em títulos longo prazo	<ul style="list-style-type: none"> • Intercompanhia declarado LP • Intercompanhia inferido LP, • Outros LP

Fontes: Elaborado pelos autores a partir de BCB (2018).

A.7.1 Dívida contra instituições do sistema financeiro nacional: fonte “SCR”

O Sistema de Informações de Crédito (SCR) do Banco Central do Brasil reúne todas as informações individualizadas dos contratos de crédito de clientes a partir de um

certo valor (que em 2021 era igual a R\$ 200). Surgiu em 2002 e seus dados são fonte de inúmeros trabalhos acadêmicos – Neves et al. (2007) e Silva (2018) são apenas alguns exemplos.

As variáveis obtidas do SCR por Alexandria, para cada data-base a partir de 2013T4 (inclusive) e para cada CNPJ8, são as seguintes:

- “Passivo SCR curto prazo”: representa o passivo do CNPJ8 contra todas as instituições do Sistema Financeiro Nacional, a ser pago em até 12 meses.
- “Passivo SCR longo prazo”: representa o passivo do CNPJ8 contra todas as instituições do Sistema Financeiro Nacional, a ser pago em prazo superior a doze meses.

Quando a entidade não tem dívida externa nem emitiu títulos, a soma das duas variáveis acima deveria ser idêntica ao informado na rubrica “Empréstimos e financiamentos” do Passivo no Balanço Patrimonial.

A.7.2 Passivo em títulos emitidos no mercado doméstico: fonte “registro da B3”

Esta base compila títulos emitidos no mercado doméstico por pessoas jurídicas não financeiras e por entidades do setor público, excluídos aqueles detidos por não residentes, que são classificados na dívida externa.

As variáveis obtidas para cada data-base a partir de 2013T4 (inclusive) e para cada CNPJ8 são as seguintes:

- “Passivo em títulos curto prazo”: representa o passivo do CNPJ8 emitido em títulos, a ser pago em até 12 meses. A estrutura de vencimentos é estimada pelo Banco Central do Brasil.
- “Passivo em títulos longo prazo”: representa o passivo restante emitido em títulos (a ser pago em prazo superior a doze meses).

A.7.3 Créditos concedidos por não residentes: fonte “RDE-ROF”

O Banco Central do Brasil compila os dados associados a cada operação de dívida externa, inclusive aquelas ocorridas entre empresas do mesmo grupo econômico no sistema RDE, módulo RDE-ROF. Os dados têm boa qualidade, pois quase todos os pagamentos para credor externo exigem conversão cambial, que é reportada via sistema Câmbio ao BCB. Os únicos passivos externos que não são registráveis no RDE-ROF são:

- Créditos comerciais de curto prazo (como por exemplo financiamento à importação e recebimento antecipado de exportação). Esse passivo é compilado no Censo de Capitais Estrangeiros (seção 3.5.1.1).
 - Entretanto, as operações originalmente contratadas de curto prazo de pagamento e que, ao serem refinanciadas, atinjam prazo de pagamento superior a 360 dias, devem ser registradas no RDE-ROF.
- Passivos externos não ingressados no país, ou seja, que foram tomados no exterior e lá permaneceram (por exemplo, para pagamento de algum compromisso da empresa no exterior). Esse passivo é compilado no Censo de Capitais Estrangeiros (ver seção 3.5.1.1) e é pouco comum, mesmo entre empresas de grande porte⁶⁶.

Para os propósitos desse estudo, foram compilados os saldos devedores de cada entidade da Alexandria, em cada data-base, e os valores foram convertidos para reais à taxa de câmbio da época. O valor compilado é, teoricamente, ligeiramente menor ou igual ao registrado como passivo externo na rubrica “passivo” no balanço patrimonial da entidade devedora (entidade residente). Isso porque o RDE-ROF calcula somente o principal de cada dívida – sem juros acumulados⁶⁷ - e porque existem alguns passivos externos captados apenas pelo Censo.

Os passivos externos registrados no RDE-ROF receberam duas classificações: por relação entre credor e devedor e por prazo de cada parcela.

A classificação de acordo com o relacionamento entre credor e devedor separou os passivos em “intercompanhia” e “demais”. A definição de empréstimo “intercompanhia” de Alexandria é muito semelhante à utilizada em estatísticas do setor externo (IMF, 2009, §6.26), exceto porque estas excluem os empréstimos tomados entre

⁶⁶ Apenas para dar uma ideia aproximada ao leitor, o BCB informa que a dívida externa bruta brasileira do setor privado era de US\$ 465,9 bilhões em dezembro de 2020 (BCB, 2022). Esse valor é compilado pelo sistema RDE-ROF, portanto não inclui a dívida não ingressada. As empresas declarantes do Censo de Capitais Estrangeiros informaram que tinham um passivo “inteiramente não ingressado” de cerca de US\$ 164 bilhões, sendo que duas empresas respondem por 88,9% desse valor. Parte desse valor pode ser devido por uma filial da empresa declarante que seja sediada no exterior, razão pela qual não necessariamente todo esse valor aparecerá no balanço individual da empresa residente.

⁶⁷ A opção por não calcular os juros foi feita para simplificar os cálculos. Os autores entendem que essa simplificação não traz grande prejuízo, pois: (i) as taxas de juros internacionais são muito baixas, e em geral as empresas pagam os juros ao longo do empréstimo (ou seja, há pouco juro acumulado e não pago em cada data-base); em especial no caso de empréstimos intercompanhia, há casos de empréstimo sem juros e casos de juros nunca pagos (um perdão implícito ou explícito dos juros acordados no contrato).

instituições financeiras afiliadas (IMF, 2009, §6.28). Ou seja, na Alexandria, todo passivo entre empresas afiliadas é classificado como intercompanhia em Alexandria, independente do setor do credor ou tomador⁶⁸.

O empréstimo é considerado “intercompanhia” em duas situações:

- Declarado: quando o passivo é registrado no RDE-ROF e o declarante informa que o credor detém 10% ou mais do poder de voto.
- Inferido: Quando o devedor consta no RDE-IED (seção 3.2.1.2) como detido pelo credor que tenha capital integralizado maior que zero.

O empréstimo intercompanhia pode, em muitas situações, ser considerado equivalente ao capital próprio. Como este trabalho analisa o endividamento das entidades empresariais em relação a terceiros, seria interessante separá-los dos demais. Entretanto, não dispomos de informações sobre os empréstimos intercompanhia domésticos (isto é, entre empresas residentes do mesmo grupo econômico), somente dos empréstimos intercompanhia externos⁶⁹.

Considerando o saldo de crédito ampliado total de cada ENF, verificamos um dado de boa qualidade: para as 165.754 datas-base para as quais as informação de passivo e crédito ampliado estão disponíveis, em apenas 3.601 casos (2,1%) o valor supera o passivo total (indicando, portanto, um erro, que pode ser (i) falta de atualização do registro por parte do declarante em alguma das informações ou (ii) declaração errada do passivo).

A.8 Base de dado censitária a partir de 2019T1: fluxos de recebimento (DOC, TED, PIX, cartão de débito e exportação)

Os fluxos de recebimentos representam os montantes recebidos pelo CNPJ8 nas seguintes modalidades (BCB 2021c): (i) boletos, (ii) Transferências Eletrônicas Disponíveis (TEDs), (iii) cartões de débito, (iv) cartões de crédito, (v) Exportações e (vi) Transferências via Pix (BCB, 2021b) feitas pelo sistema de pagamentos instantâneos (SPI), que identifica as partes (BCB, 2021a).

⁶⁸ O pesquisador que desejar buscar uma comparabilidade entre o empréstimo intercompanhia das estatísticas externas e o dado de Alexandria pode simplesmente excluir os saldos de empréstimo intercompanhia em que ambos, devedor e credor, são instituições financeiras. Para ENFs, foco deste artigo, a definição coincide com o padrão metodológico do BPM6 (IMF, 2009, §6.26).

⁶⁹ Para comentários metodológicos sobre a mensuração dos empréstimos intercompanhia na dívida externa, consulte IMF (2014, §3.20 e §4.3).

Alexandria toma a variável “fluxo de recebimentos” e associa a cada trimestre o fluxo de recebimentos de cada entidade. Os dados estão disponíveis somente a partir de 2019 (inclusive) e somente para entidades não financeiras⁷⁰.

Os fluxos de recebimento têm sido utilizados por BCB (2021c) como uma *proxy* do faturamento da entidade. Trata-se de análise a ser confirmada (e estudos futuros utilizando Alexandria podem ajudar nisso). De toda forma, é importante salientar que os recebimentos não representam, necessariamente, faturamento, e muito menos indicam a totalidade do faturamento, uma vez que as empresas podem receber por outros meios, não captados por essa variável: (i) dinheiro, (ii) cheque, (iii) transferências intra-bancárias (exceto via TED e Pix identificado), (iv) Pix fora do sistema SPI, (v) DOCs, (vi) TECs e (vii) faturas de contas de água, energia, gás, TV a cabo e outras, cujo pagamento não é feito por nenhuma das formas incluídas nos fluxos (e descritas acima).

Assim, existem entidades empresariais ativas que não registram nenhum fluxo de recebimento durante longos períodos. Além de poder receber por outros meios, como mencionado acima, há casos em que o empresário utiliza sua conta pessoal para movimentar o caixa da empresa (comum no caso de empresários individuais). Entretanto, os autores entendem que isso dificilmente ocorrerá com uma entidade empresarial de médio ou grande porte.

Mesmo com estas observações sobre a variável, é inegável afirmar que se trata de informação preciosa, que permite, no mínimo, investigar o nível de atividade e inclusão financeira da empresa, entre outras possibilidades.

A.9 Base de dado amostral: empregos formais (RAIS, CAGED e eSocial)

A RAIS, o Caged e o sucessor de ambos, o Esocial, são repositórios estatais de informações sobre os vínculos empregatícios formais das entidades contábeis residentes (IBGE, 2020?; Almeida et al., 2020).

Apesar de as declarações RAIS, Caged e eSocial serem obrigatórias para as entidades contábeis, na prática muitas não declaram, em especial aquelas sem empregados formais – por isso foi considerada uma base amostral. Entretanto, IBGE (2020?) estima que a RAIS tenha uma cobertura de “97% do total do setor organizado da economia”.

⁷⁰ Ou seja, esta base não contém os fluxos de recebimento de empresas com CNAE 64 (exceto 64.62, para o qual há dados), 65 (exceto 655) ou 66.

As variáveis de emprego captadas por Alexandria, para cada CNPJ8 e para cada data-base, referem-se (i) ao número de empregados e (ii) à remuneração média, ambos com detalhamento por gênero e escolaridade.

A.10 Resumo das fontes

A Figura 12 ilustra as principais informações que Alexandria têm sobre cada entidade empresarial, em cada data-base, e as classifica de acordo com a fonte utilizada. Trata-se de um resumo, pois Alexandria tem mais de 120 variáveis.

O resumo deixa claro que, apesar de Alexandria coletar apenas algumas rubricas do passivo das entidades, foi possível completar a informação com dados internos do BCB. Assim, o passivo pôde ser detalhado em passivo contra credores externos, contra credores do SFN e em títulos. O restante do passivo foi obtido por resíduo e inclui passivo contra fornecedores, com folha de pagamento, contra o governo, entre outros.

Figura 12 – Informações sobre cada entidade constante na Alexandria, em cada data-base.

CNPJ XXXXXX, Data-base: AAAA/T	
Balanco patrimonial	
1) Ativo total [A]	2) Passivo Total [A] 2.1) Circulante [A] 2.1.1) Credores externos [B] – inclui intercompanhia 2.1.2) Credores do SFN [C] 2.1.3) Títulos detidos por residentes [D] 2.1.4) Demais passivos circulantes (=2.1-2.1.1-2.1.2-2.1.3) 2.2) Não circulante [A] 2.2.1) Credores externos [B] – inclui intercompanhia 2.2.2) Credores do SFN [C] 2.2.3) Títulos detidos por residentes [D] 2.1.4) Demais passivos não circulantes (=2.2-2.2.1-2.2.2-2.2.3)
	2.3) Patrimônio líquido [A] 2.3.1) Capital social [A]
Demonstração do Resultado do Exercício (DRE)	
3) Receita líquida [A]	
4) Lucro líquido [A]	
Outras informações	
5) Empregados Formais, por gênero e escolaridade [E]	
6) Fluxos de recebimento [F]	
7) Entidades detidas [G]	
8) Gênero do administrador [G]	
10) Dados societários: empresa familiar (G); participação de não residentes no capital social e no poder de voto [H]	
Fontes utilizadas para obter cada informação [Sim=disponível em todas as datas-base trimestrais] : [A] Censo (BCB), RDE-IED (BCB), Valor Pro ou Econômica (de acordo com a melhor fonte); [B] RDE-ROF (BCB) [Sim]; [C] SCR (BCB) [Sim a partir de 2013T4]; [D] Cetip/B3 [Sim, a partir de 2013T4]; [E] RAIS/CAGED/Esocial; [F] BCB [Sim, a partir de 2019]; [G] Base de dados de CNPJ (Receita Federal do Brasil) [H] Censo (BCB) e/ou RDE-IED (BCB) [Sim].	

Fonte: elaborado pelos autores.

Anexo B – Endividamento mediano das ENFs da base Alexandria versus base Klooks (sem retirar outliers; em %)

Seção CNAE	Base Alexandria									Base Klooks		
	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
ENFs	58,4	58,3	57,0	55,2	52,8	53,4	51,5	51,6	51,4	39,7	43,2	48,9

Fonte: base Alexandria e base Klooks. As seções K, O, T e U não foram apresentadas por não pertencerem às ENFs, conforme definição da seção 3.1.

Anexo C – Endividamento médio e mediano das ENFs, por setor (em %)

	Média (ignora valores maiores que 300)						Mediana					
	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

Anexo D – Estatísticas descritivas das variáveis utilizadas na seção 4.3

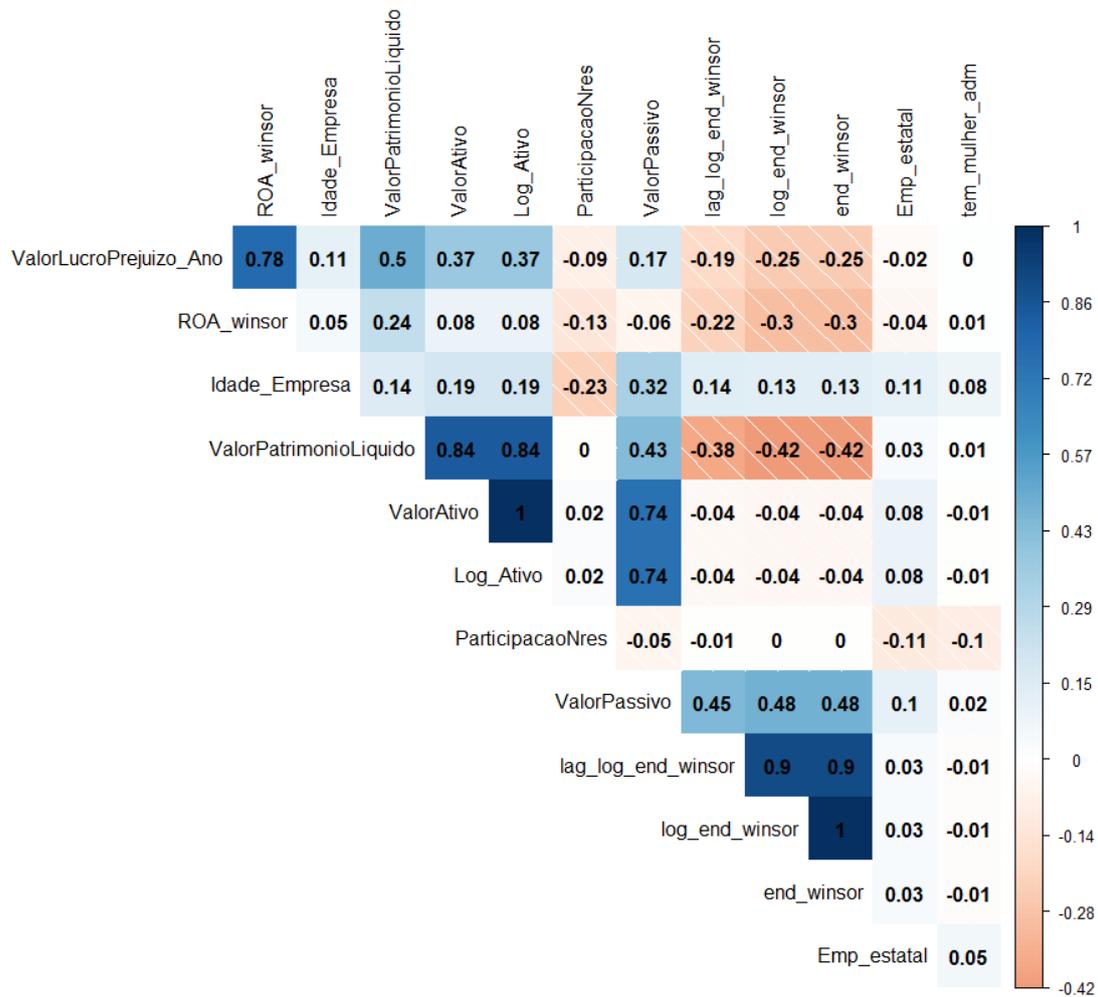
Número de observações utilizadas entre 2015 e 2021, inclusive: 100.944

Variável	Média	Desvio padrão	Valor mínimo	Perc. 25	Mediana	Perc. 75	Valor máximo
Ativo (R\$ milhões)	1.648,58	26.959,45	0	7,074	71,496	403,291	1.988.646,00
Passivo (R\$ milhões)	1.133,179	24.107,94	-12,548	1,022	19,211	142,915	1.854.420,00
Patrimônio Líquido (R\$ milhões)	515,404	5.419,95	-64.485,81	0,996	24,136	183,567	892.952,50
Lucro/Prejuízo Ano (R\$ milhões)	44,892	1.294,914	-44.212,19	-0,506	0,392	13,252	290.820,00
log(endividamento+1) (%) (a)	45,291	49,79	0,00	14,524	39,24	58,667	345,51
endividamento (%) (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(Ativo) (R\$ milhões)	3,737	3,451	-18,644	1,956	4,27	6	14,503
Participação Nres (%)	42,548	47,125	0	0	0,4	100	100

Nota: (a) Estatísticas descritivas utilizando a técnica de winsorização em nível de 1%.

Anexo E – Correlograma das variáveis mencionadas no Quadro 4



Anexo F – Testes de robustez temporal

Tabela 5—Estimação dos modelos de endividamento usando o método Sys-GMM em painel balanceado e painel desbalanceado de empresas não financeiras (ENFs), em diferentes períodos.

	Painel desbalanceado				Painel balanceado			
	(A) ENF's 2015 a 2021 (1)	(A.1) ENF's 2015 a 2019 (2)	(A.2) ENF's 2017 a 2021 (3)	Δ	(D) ENF's 2015 a 2021 (4)	ENF's 2015 a 2019 (5)	ENF's 2017 a 2021 (6)	Δ
lag(log end, 1)	0,7913*** (0,0359)	0,7771*** (0,0497)	0,8339*** (0,0390)	▲	0,9187*** (0,0211)	0,8563*** (0,0247)	0,9086*** (0,0322)	▼
Demais Serviços	0,0038 (0,0026)	0,0058* (0,0031)	0,0014 (0,0025)	▲	0,0019 (0,0037)	0,0122** (0,0050)	0,0028 (0,0035)	▲
Comércio	0,0048 (0,0032)	0,0083** (0,0040)	0,0031 (0,0033)	▲	0,0051* (0,0028)	0,0125*** (0,0036)	0,0066** (0,0033)	▲
Construção	-0,0137** (0,0056)	-0,0126* (0,0069)	-0,0095* (0,0056)	▲	-0,0041 (0,0043)	-0,0031 (0,0057)	-0,0023 (0,0046)	
Agro	-0,0132* (0,0076)	-0,0140 (0,0095)	-0,0087 (0,0067)	▲	-0,0066 (0,0046)	-0,0071 (0,0063)	-0,0083* (0,0051)	▲
Serviços Imob.	-0,0401*** (0,0082)	-0,0382*** (0,0101)	-0,0384*** (0,0084)	▲	-0,0114* (0,0059)	-0,0138* (0,0077)	-0,0147** (0,0071)	▼
ROA	-0,1471*** (0,0132)	-0,1465*** (0,0164)	-0,1335*** (0,0146)	▲	-0,1905*** (0,0388)	-0,2038*** (0,0735)	-0,0840 (0,0650)	▼
Dados na CVM	0,0223*** (0,0057)	0,0295*** (0,0076)	0,0151** (0,0061)	▼	-0,0073 (0,0079)	-0,0139 (0,0113)	0,0016 (0,0065)	
Estatal	0,0069 (0,0065)	0,0111 (0,0080)	0,0006 (0,0061)		-0,0138*** (0,0044)	-0,0179*** (0,0065)	-0,0109** (0,0051)	
Familiar	0,0021 (0,0058)	-0,0005 (0,0069)	0,0101 (0,0065)		0,0032 (0,0048)	0,0151** (0,0075)	0,0014 (0,0049)	
Log(ativo)	-0,0012 (0,0009)	-0,0020* (0,0011)	0,0010 (0,0009)		0,0058 (0,0040)	0,0114* (0,0059)	0,0027 (0,0035)	
Participação Nres	0,0020 (0,0031)	-0,0005 (0,0041)	-0,0003 (0,0030)		-0,0088* (0,0050)	-0,0215*** (0,0068)	-0,0037 (0,0050)	
Sede em Capital	-0,0046** (0,0021)	-0,0047* (0,0026)	-0,0045** (0,0021)		-0,0040 (0,0026)	-0,0068* (0,0038)	-0,0023 (0,0026)	
Mulher (adm)	-0,0023 (0,0021)	-0,0039 (0,0025)	-0,0010 (0,0022)		-0,0012 (0,0012)	-0,0012 (0,0015)	-0,0002 (0,0014)	
log(idade emp)	0,0017 (0,0015)	-0,0005 (0,0017)	0,0022 (0,0016)		0,0035*** (0,0010)	0,0028** (0,0013)	0,0034*** (0,0013)	
nobs	53.891	35.132	34.427		20.097	15.631	15.631	
Instrumentos	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-valor]	[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-valor]	[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-valor]	[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	5.592,19 {15}	
[p-valor]	[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-valor]	[0,0044]	[0,0490]	[0,4415]		[0,0000]	[0,0000]	[0,3421]	

Notas: (i) Análise das variáveis anuais da amostra; (ii) Erros robustos à heterocedasticidade em parênteses utilizando clusterização de erros no nível da firma; (iii) * 10% de significância; ** 5% de significância; *** 1% de significância; (iv) AR(1) e AR(2) verificam a presença de correlação seriada de primeira ordem e de segunda ordem nos resíduos de primeira diferença (v) Não mostramos as dummies de tempo e o intercepto por concisão. (vi) utilizou-se a técnica de winsorização em nível de 1% para as variáveis endividamento e ROA, (vii) variáveis em lag utilizadas como instrumentos: E - log(endividamento) winsorizado; R - ROA winsorizado; I - log(idade); valores entre [] são os lags das variáveis utilizadas como instrumentos.

Anexo G – Testes de robustez setorial

Tabela 6– Análise de robustez dos modelos de endividamento usando o método Sys-GMM em painel desbalanceado de empresas não financeiras (ENFs) com quebra por setor.

	Indústria	Serviços	Comércio	Construção	Agro	Serv. Imob.
lag(log end, 1)	0,8987 *** (0,0461)	0,6824 *** (0,0691)	0,9528 *** (0,0541)	0,8423 *** (0,0955)	0,9070 *** (0,0565)	0,8384 *** (0,0833)
ROA	-0,1130 *** (0,0251)	-0,1599 *** (0,0187)	-0,1319 *** (0,0332)	-0,1123 ** (0,0437)	-0,1842 *** (0,0465)	-0,0150 (0,0116)
Dados na CVM	0,0108 (0,0076)	0,0094 (0,0098)	-0,0108 (0,0078)	0,0331 (0,0202)	0,0327 (0,0497)	0,0861 ** (0,0401)
Estatal	0,0014 (0,0076)	0,0026 (0,0132)	-0,0104 (0,0165)	0,0044 (0,0424)		0,0425 (0,0335)
Familiar	0,0055 (0,0071)	-0,0002 (0,0152)	0,0068 (0,0071)	-0,0067 (0,0102)	-0,0045 (0,0128)	-0,0034 (0,0158)
Log(ativo)	0,0004 (0,0012)	-0,0018 (0,0016)	0,0023 * (0,0014)	-0,0043 (0,0037)	-0,0005 (0,0030)	0,0028 (0,0031)
Participação Nres	-0,0046 (0,0030)	0,0058 (0,0075)	-0,0051 (0,0053)	-0,0076 (0,0141)	-0,0118 (0,0144)	0,0081 (0,0150)
Sede em Capital	-0,0031 (0,0029)	-0,0049 (0,0050)	-0,0027 (0,0038)	-0,0070 (0,0090)	0,0045 (0,0110)	-0,0081 (0,0110)
Mulher (adm)	-0,0014 (0,0024)	0,0000 (0,0046)	-0,0013 (0,0044)	-0,0064 (0,0111)	0,0075 (0,0089)	0,0079 (0,0089)
log(idade emp)	0,0014 (0,0018)	0,0004 (0,0035)	0,0017 (0,0041)	0,0063 (0,0056)	-0,0028 (0,0084)	0,0012 (0,0060)
Intercepto	0,0261 (0,0361)	0,1919 *** (0,0541)	-0,0222 (0,0458)	0,1225 * (0,0709)	0,0571 (0,0672)	-0,0105 (0,0627)
2017	0,0106 *** (0,0026)	-0,0083 (0,0055)		0,0013 (0,0104)		-0,0139 (0,0094)
2018	0,0123 *** (0,0029)	-0,0149 ** (0,0060)	0,0063 (0,0053)	0,0155 (0,0102)		-0,0191 * (0,0102)
2019	0,0144 *** (0,0027)	-0,0057 (0,0057)	0,0050 (0,0056)	0,0116 (0,0110)	0,0014 (0,0091)	-0,0144 (0,0108)
2020	0,0202 *** (0,0028)	-0,0036 (0,0061)	-0,0047 (0,0060)	0,0057 (0,0112)	0,0085 (0,0083)	-0,0218 * (0,0117)
2021	0,0149 *** (0,0036)	-0,0072 (0,0064)	-0,0021 (0,0073)	0,0212 (0,0129)	0,0091 (0,0135)	-0,0089 (0,0122)
nobs	23.525	15.807	6.325	3.446	931	1.852
Instrumentos	E [2:99]	E [2:99]	E [3:99]	E, R [2:99]	E [4:99]	E, R [2:99]
Hansen-Sargan	173,1740 [p-valor] [0,2397]	182,0540 [p-valor] [0,1976]	168,2210 [p-valor] [0,2076]	166,6630 [p-valor] [0,5461]	118,6160 [p-valor] [0,4569]	168,4380 [p-valor] [0,5339]
AR1	-9,39 [p-valor] [0,0000]	-5,79 [p-valor] [0,0040]	-5,60 [p-valor] [0,0000]	-3,04 [p-valor] [0,0023]	-1,68 [p-valor] [0,0921]	-2,51 [p-valor] [0,0121]
AR2	-1,6 [p-valor] [0,1087]	-1,2 [p-valor] [0,2308]	-2,06 [p-valor] [0,0396]	1,15 [p-valor] [0,2492]	1,96 [p-valor] [0,0495]	0,34 [p-valor] [0,7307]
Wald Coef {GL}	1870.25 {10} [p-valor] [0,0000]	959.89 {10} [p-valor] [0,0000]	1673.2 {10} [p-valor] [0,0000]	298.58 {10} [p-valor] [0,0000]	648.54 {9} [p-valor] [0,0000]	873.45 {10} [p-valor] [0,0000]
Wald Time {GL}	55,00 {5} [p-valor] [0,0000]	9,06 {5} [p-valor] [0,1067]	5,23 {4} [p-valor] [0,2641]	8,37 {5} [p-valor] [0,1372]	1,18 {3} [p-valor] [0,7590]	4,29 {5} [p-valor] [0,5080]

Notas: as mesmas da Tabela 5.