MASTER
IN FINANCE

MASTER´S FINAL WORK
DISSERTATION

THE EFFECT OF COMPANY FUNDAMENTALS ON MARKET CAPITALISATION – A QUANTITATIVE ANALYSIS OF THE ENERGY SECTOR IN BRAZIL

LUKAS BINDER

SUPERVISION:
LUÍS FILIPE ÁVILA DA SILVEIRA DOS SANTOS

OCTOBER, 2024

LIST OF ABBREVIATIONS

3B – Brasil, Bolsa, Balcão

ADF – Augmented Dickey-Fuller

BRL – Brasilian Real (Reais)

BRLm – Brasilian Reais in milion

CapEx – Capital Expenditures

COGS – Cost of Goods Sold

D&A – Depreciation & Amortisation

EBITDA – Earning Before Interest Depreciation and Amortisation

EVA – Economic Value Added

FCF – Free Cash Flow

GDP – Gross Domestic Product

IRR – Internal Rate of Return

JEL – Journal of Economic Literature

LTM – Last Twelve Months (on a rolling basis)

MFW – Master's Final Work

NPV – Net Present Value

PC1 to PC5 – Principal Component 1 to 5

PCA – Principal Component Analysis

PP&E – Property Plant and Equipment

Q1 to Q4 – Quarter 1 to 4

SE – Standard Error

VIF – Variance Inflation Factor

ABSTRACT, KEYWORDS, ACKNOWLEDGEMENT AND JEL CODES

(EN) This dissertation aims to explain market capitalisation in the energy market in Brazil using a dataset of 36 companies and the corresponding ratios: Asset Turnover, Debt Ratio, Current Ratio, Price Earnings Ratio, Net Profit Margin, and Book-to-Market Ratio. The method used is a regression model based on a principal component analysis to address several problems with the time series dataset, namely multicollinearity and to reduce the complexity of the model. There is also a robustness check and discussion of the model to evaluate the impact and validity of the results. This aims to help investors seeking value maximisation to use the model and its results, as well as company managers, to make the right decisions to optimise the fundamental company structure and maximise value. Key results are that net profit margin has a positive impact if increased on market capitalisation, which is aligned with the expectation that one of the most important objectives of a company is to generate cash. In general, the model is able to explain market capitalisation accurately so, the dissertation is a valuable tool for investors and company managers.

(PT) Esta dissertação tem como objetivo explicar a capitalização bolsista no mercado de energia no Brasil, utilizando um conjunto de dados de 36 empresas e os respetivos rácios: Volume de Negócios do Ativo, Rácio de Endividamento, Rácio Corrente, Rácio Preço/Lucro, Margem de Lucro Líquido e Rácio Cotação/Valor Contabilístico. O método utilizado é um modelo de regressão baseado numa análise de componentes principais para resolver vários problemas com o conjunto de dados de séries temporais, nomeadamente a multicolinearidade e para reduzir a complexidade do modelo. É também efetuada uma verificação da robustez e uma discussão do modelo para avaliar o impacto e a validade dos resultados. O objetivo é ajudar os investidores que procuram maximizar o valor a utilizar o modelo e os seus resultados, bem como os gestores das empresas, a tomar as decisões certas para otimizar a estrutura fundamental da empresa e maximizar o valor. Os principais resultados são que a margem de lucro líquida tem um impacto positivo, se for aumentada, na capitalização bolsista, o que está de acordo com a expetativa de que um dos objetivos mais importantes de uma empresa é gerar dinheiro. Em geral, o modelo é capaz de explicar com exatidão a capitalização bolsista, pelo que a dissertação constitui uma ferramenta valiosa para investidores e gestores de empresas.

ACKNOWLEDGEMENT: I would like to thank my supervisor, Luís Filipe Ávila Da Silveira Dos Santos for his continuous support and suggestions for the model and writing part of the dissertation. Also, thank you to the master´s coordination team for providing guidelines and being available regarding open questions. Finally, I would like to thank my family and friends during the time I was working on the dissertation for their patience, motivation and support.

KEYWORDS: Energy Market Brazil, Listed Companies, Emerging Market, Valuation Model, Market Capitalisation

JEL CODES: G10; G12; G15; G17; G30; G31.

TABLE OF CONTENTS

FIGURES AND TABLES

Figures

Tables

Headnote

This dissertation aims to explain the market capitalisation of the Brazilian Energy Market, an Emerging Market, using data from 36 publicly listed companies and five common financial ratios within the PCA regression model. The motivation is to provide a useful tool for investors and company managers and insights into what impacts market capitalisation within this sector.

1. INTRODUCTION

Brazil's energy market is considered one of the country's most important markets due to Brazil’s wealth in natural resources (Amaro et al. 2008). Amaro et al. (2008) claimed in their study that Brazil offers a great availability of energy resources. The energy sector as a preliminary characterisation includes, according to Ferreira et al. (2022), important commodities such as natural gas, WTI crude oil, NY gasoline, gulf gasoline diesel, heating oil and propane. The objective of this dissertation is to analyse the Brazilian energy sector using publicly listed companies at B3 (São Paulo stock exchange) and determine which of the company fundamentals, namely total asset turnover, total debt ratio, current ratio, price-earnings ratio, net profit margin, and book-to-market ratio are statistically significant to explain the corresponding market capitalisation. In particular, this can be applied to obtain an estimate of the market capitalisation of any company in this sector and determine which fundamental inputs are relevant to consider in a company analysis (Ozlen, 2014). The dissertation aims to explain the energy sector's fundamental impacts under the research question: “What is the effect of company fundamentals on market capitalisation in the energy sector in Brazil?” This is important due to different current literature covering similar fundamentals in different markets.

Furthermore, the results of the dissertation suggest which fundamentals company managers should improve to increase market capitalisation and, therefore, shareholder value. Here, especially the debt ratio analysis is interesting since the debt ratio usually increases the probability of financial distress (Septyanto et al. 2022) while improving the cost of capital to a certain extent because debt is cheaper than equity. Moreover, one needs to differentiate between a pure shareholder maximisation approach and a general welfare maximisation approach. This means that the model focuses on shareholder maximisation (market capitalisation) not implying the best output for all stakeholders (welfare maximisation). Nevertheless, other studies, such as the one of Michalak (2016), linked the main objective for managers of companies to value maximisation. The results of the dissertation suggest what leads to an increase in shareholder value, but company managers should also take care of all company stakeholders. For example, Queen (2015) claimed in his study that no one benefits if shareholder maximisation is pursued without regard to the negative impacts on other stakeholders. The motivation to create this model comes from analysing stocks from Brazil's energy sector and determining which fundamentals are important to maximise shareholder value.

Besides the main research question, the dissertation further investigates the development of the combined market capitalisation in the energy sector in Brazil as well as the impact of other fundamentals such as free cash flow, EBITDA, depreciation and amortisation (D&A), cost of revenue, property plant and equipment (PP&E), income tax, and capital expenditure (CapEx). These variables are used as proxy variables but will be analysed as well.

This approach differs from another common approach in predicting the probability of failure and, therefore, provides different insights into how to evaluate financial ratios. In general, Beaver (1966) claims that if one wants to test financial ratios, this can be only done with a particular purpose. The model is expected to show evidence that in an asset-heavy industry variables such as book-to-market value and CapEx are important. The book-to-market ratio should be especially relevant in determining market capitalisation. Furthermore, previous studies, such as the one of Hovakimian (2006), concluded that if company managers want to make a long-term effect on this ratio, equity transactions and their timing are considered to be irrelevant. However, the book-to-market ratio also includes market capitalisation as an input and cannot be relied on solely. Moreover, it is interesting to analyse the growth rates of companies in this industry, which usually differ from what is considered to be a growth stock, such as technological companies and rely more on fundamental values – value investing principles. Here, one must differentiate between established companies (operating for a longer time period) and new companies with characteristics of start-ups. Lund (2014) found evidence in his study that there is an upper cap of growth rates for companies in the energy sector, using the market circumstance in 2013, of 15-25%. The academic community heavily researched value and growth investments and concluded that value investment strategies (including investment in the energy sector) outperform growth investment strategies (Chan & Lakonishok 2004).

Other expected results are the hypothesis that profit margin and cash-generating ratios, such as free cash flow, are especially important since researchers widely use especially free cash flow to analyse companies. Regarding the findings from the perspective of company managers, this dissertation suggests what is significantly important for running the company efficiently and how sensitive the market capitalisation is reacting to these changes.

For current research in finance, this dissertation provides additional value simply because of the input of the most recent data and covering a different sector. For example, Ozlen (2014) covered in his study similar main inputs in different sectors and the market in Turkey. These results may significantly differ from similar models used in different markets, countries, or a combination of both. This is related to the characteristics of the energy sector and different political frames in countries, including the ability to export goods, differences in taxation, bureaucracy, and general political stability. Also, it is important to mention that other research focuses on a different data basis regarding variables analysed and sectors analysed, amongst others (Ozlen, 2014; Behera, 2020).

The key results of the dissertation are that an increase in the book-to-market value has a negative impact on market capitalisation, similar to an increase in the debt ratio. Other ratios, if increased, positively impact market capitalisation, especially net profit margin. This suggests that the financing structure of the companies is slightly overloaded on debt and that higher profitability, which in the end leads to a higher cash generation, does have the strongest impact. On the other hand, the strong result of book-to-market value is also explainable due to market capitalisation as an input to the ratio itself. The results are discussed in depth in the results section of the dissertation.

The structure of the dissertation followed after this is divided into a literature review and a background analysis of where the companies operate. Thereafter, it continues by describing the dataset with a focus on the source, adaption of the data, and some statistical properties. This is then followed by the introduction of the empirical strategy and why this strategy is used to face issues such as multicollinearity. Thereafter is a discussion part of the results and their validity, followed by limitations implied by the model assessed. As the final part of the dissertation, all the results are concluded, and the most important findings are broken down into a short summary. This also suggests further research ideas on how to continue the work based on this dissertation and an outlook on the sector. In the appendix are additional tables and the R code used to perform the statistical model.

2. INSTITUTIONAL BACKGROUND AND LITERATURE REVIEW

2.1. Definition of the Energy Sector

The energy sector can be characterised by businesses operating in the sectors of electricity generation, electricity transmission, oil and natural gas as well as biofuels (Tolmasquim, 2012). In this thesis also companies with one or several divisions operating in these areas have been considered because they count by the definition above to the energy sector. In addition, academic researchers often differentiate the energy sector between renewable energy and fossil energy e.g. oil and natural gas (Ogulata, 2003). However, for the dissertation, the combined energy sector including the renewable and fossil energy sectors has been taken into account because the renewable energy sector fits into the definition. Also, it would be difficult to address this problem from a corporate finance perspective since lots of large energy conglomerates do have renewable energy and fossil energy business divisions but report consolidated financials.

In the context of the dissertation, data has been extracted from the São Paulo stock exchange B3 related to (i) Oil, Gas and Biofuels, (ii) Basic Materials, and (iii) Utilities. Therefore, the sectors can be further broken down into (i) Exploration, Refining & Distribution, and Equipment & Services, (ii) Metallic Minerals, Steel, Iron & Steel Products, Copper Products, Petrochemicals, Fertilizers, Chemicals – Other, Pulp & Paper, Packaging, and Other Materials, (iii) Electric Utilities, Water Utilities, and Gas Utilities. This aligns with the characterisation of the energy sector of Tolmasquim (2012).

2.2. Market Characteristics and Limitations to Other Markets

Usually, the business model is asset- and CapEx-heavy[1] and dependent on energy needs (Chan & Lakonishok, 2004). This implies that companies need a high amount of capital to operate their vehicle fleet and production facilities with ongoing investments. Moreover, this can lead to pressure on profitability margins and financial pressure due to higher fixed costs for the assets even in economic downturns. He & Lin (2018) described as well that the industrial sector has a pro-cyclical impact on the energy sector meaning

[1] CapEx-heavy: Business models characterised by a high amount of Capital Expenditure.

that in positive shocks also the energy sector is positively impacted. Also, Bourghelle, et al. (2021) highlight that the oil price (one of the goods from the energy sector) reacted strongly to the economic shock of COVID-19.

Nevertheless, the energy sector is differentiated from other sectors, such as technology, finance and more similar industrial sectors (Tolmasquim, 2012). Technically speaking, the industrial manufacturing sector is similar due to the often asset- and CapEx-heavy nature of the business model, but still significantly differs in cyclicality versus the energy sector. Energy sector companies are generally considered more cyclical due to the volatile demand for oil and gas or electrical energy, whereas basic energy suppliers are commonly considered resilient (He & Lin, 2018). Also, Akhtaruzzaman & Sabri Boubaker (2021) provided some evidence that oil suppliers such as oil production and integrated oil and gas companies benefit the most from an increase in the oil price. Hence, they are negatively affected in economic downturns, with a decrease in oil demand and a corresponding drop in price. These companies are a remarkable part of the energy sector, but not limiting it. Also, Petrobras (included in the analysis) is an oil and gas company and contributes with a high market capitalisation to the model. As an example of a recent shock, during the Covid 19 crisis, the price per barrel of crude oil dropped significantly, and demand dropped to very low levels. This decrease in demand led to a price decrease and difficulties for companies in the oil and gas sector to sell their oil profitably. On the other hand, some of the companies related to this sector are considered to be not cyclical due to their basic supply function, such as more local electricity companies. They have a relatively secure demand since private households and most companies still need energy on a basic level to keep their business running even though costs are high.

Regarding the companies, some of them also engage in side operations as part of their consolidated business activity, such as logistics and distribution, among others. These companies are still included in the analysis since their core business activity revolves around energy and or energy-related fields. However, this should not strongly affect the results obtained since the side activities are related to enhancing the core operations.

The energy market is an established market, especially in Brazil, where the total market share of the energy sector is quite large (Tolmasquim, 2012). This means there is competition amongst the market participants to gain further market share, limiting revenue growth. On the other hand, market share is created for innovative start-ups operating in a new business field. It does not already exist, so growth rates significantly deviate and are lower for the energy sector. Nevertheless, risks associated with companies within the energy sector are lower because they usually provide basic inputs used in every other industry.

Brazil is considered to be a global leader in the energy sector (Tolmasquim, 2012). Therefore, it is a natural objective that the energy sector is much more dominant here compared to other countries. Furthermore, Brazil is considered to be an emerging market and, therefore, lacks infrastructure compared to higher-developed countries. There are discussions on how to classify countries as emerging markets, developed or otherwise, but Sinkovics et al. (2014) and Ortas et al. (2012), among other studies, claimed Brazil as an emerging market. Business operations in an emerging market can lead to significant challenges, especially when distributing refined or produced goods (Tolmasquim, 2012). Generally, this results in high operating costs for companies within this sector. Their distribution capabilities strongly rely on carrying, for example, oil and gas from the extraction facility to refineries and or further to harbours or logistic centres from where they are further distributed. Moreover, this is not only limited to the oil and gas industry since a similar logic also applies to electricity distribution and electricity gain from power plants.

Sarkar (2020) mentioned another important point in their study: there is a significant correlation between market capitalisation as a ratio of GDP and GDP per capita growth in emerging markets. So, it is beneficial for these countries to develop their infrastructure further since this commonly increases GDP growth rates. Moreover, energy companies often carry lots of assets, off-balance sheet financing, and trade working capital. If companies are unable to handle these efficiently, they usually rely on a high level of debt financing. This implies that managers could especially improve market capitalisation using an efficient way to finance their inventory, such as a fast billing of their accounts receivable.

Regarding other emerging markets active in energy production and distribution, Brazil faces significant challenges due to its size and sometimes ineffective taxation system. Brazil has an ineffective taxation system with high inequality and poverty (Melo, et al., 2014). Therefore, this ineffective system can lead companies to use less efficient ways optimised for the local economic and political situation. This means that the model and further results are limited to the energy industry in Brazil. On the other hand, since Brazil's regional challenges are similar to those of other countries in South America, the results can probably be partially projected to other countries within this region. As an example, also Pereiro (2006) showed in their study that Argentina uses valuation practices often aligned with those used in the United States of America.

2.3. Further topics regarding the Energy Market

The academic literature also extensively reported ongoing trends in the Brazilian energy market and topics regarding emerging market valuation. For example, Abuaf (2011) reported a statistical dependence on CDS spreads and macroeconomic growth rates and Pereiro (2006) showed that discounted cashflow methods as well as NPV, IRR and payback methods are popular choices used by finance managers. Moreover, the valuation models especially highlight the fact of which inputs are important and which models more effectively describe desired outcomes. Hereby, one needs to limit that these valuation models typically only reflect niche markets or certain sectors within a market under given rules such as economic (i.e. taxes) environment. In contrast, there are also broader studies, such as the one by Behera (2020), where he stated that EVA models could be used to determine economic profits in a market by subtracting market profits from company-specific profits, using a sample of small, mid, and large-cap companies. Other studies, such as Young & Jung-Jin (2003), stated that when applying relative valuation analysis in the Japanese stock market, the price-to-sales ratio delivers the best returns for investment analysis within their dataset. However, these two examples do not limit the finding of results and should be seen as highlights that there is a demand and interest in determining the best valuation model in certain economies.

Recent trends in Brazil show that they have many new possibilities regarding renewable energies and gained lots of media exposure Amaro et al. (2008). This is a relatively common effect since the world is worrying more about climate change and trying to find sustainable solutions to this issue. Moreover, countries like Brazil are rich in renewable energy resources and have lots of raw materials such as oil, natural gas, and other mineral resources that are key inputs within the energy transition. Amaro et al. (2008) highlight that energy and its availability are one of the key driving forces for a country towards economic development which is favourable for countries like Brazil.

3. DATA

Obtaining the right data is the key point to ensure that the results of the model are replicable, and the data is still providing sufficient results. Furthermore, the data needs to be adapted so it can be used in the specific model to describe the variables. Also, dependent on the structure of the model, the dataset has been adjusted using transformations and enhancements through margins.

This analysis includes companies from the sectors: oil, gas and biofuels, basic materials, and utilities. The basic materials sector can be broken down into mining, steel and metal, chemicals, wood and paper, packaging, and diversified materials. Moreover, the utilities sector is divided into electric, water, and gas utilities. In total, the extracted data list includes 108 companies, whereby for the final analysis, 36 companies were used. This is related to the partially limited availability of companies' data and lack of reporting. However, the data has been collected using Bloomberg, which is considered to have the highest standard for financial data according to academic literature, and manually enhanced, if needed, by data from Yahoo Finance as well as average margins (Abuaf 2011). The basic structure of the dataset is quarterly data, which includes an annualisation (LTM) for operating metrics such as revenue, profit, EBITDA, and free cash flow. At the end of this passage, there are summarising statistics to get an overview of the dataset describing key characteristics.

3.1. Source of the Data

First, it is important to mention that the list of all companies from this sector was obtained through the São Paulo stock exchange B3 (“Brasil, Bolsa, Balcão”). The filters (Basic Materials, Oil, Gas and Biofuels, Utilities) were related to all companies operating in the energy sector, as characterised in the background section. In total, the list of companies related to this sector comprises 108 different companies, with 13 operating in the oil, gas and biofuels subsector, 31 operating within the basic materials area, and the remaining 64 operating in the utility space. The figure below (Figure 1) provides an overview of the exact structure. Furthermore, in the appendix is a list with the full names and tickers of each company, as well as an explanation.

Original Segmentation

Sector	#	Subsector	#	Segment	#
Oil, Gas and Biofuels	13	Oil, Gas and Biofuels	13	Exploration, Refining and Distribution	10
				Equipment and Services	3
Basic Materials	31	Mining	7	Metalic Minerals	7
		Steel and Metalurgy	9	Steel	5
				Iron and Steel Products	3
				Copper Products	1
		Chemicals	7	Petrochemicals	2
				Fertilizers	3
				Chemicals - Others	2
		Wood and Paper	6	Wood	2
				Pulp and Paper	4
		Packaging	1	Packaging	1
		Diversified Materials	1	Other Materials	1
Utilities	64	Electric Utilities	53	Electric Utilities	53
		Water Utilities	8	Water Utilities	8
		Gas Utilities	3	Gas Utilities	3
SUM	108		108		108

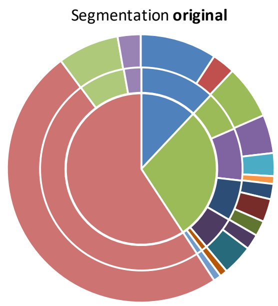

Figure 1: Segmentation Original

Figure 2 shows the number of companies where sufficient data was available. The general objective was to make sure that at least 95% of the data collected for the inputs of these variables were from Bloomberg to meet high-quality standards and the same source. All the Companies without this information have been dropped from the dataset.

Available Segmentation

Sector	#	Subsector	#	Segment	#
Oil, Gas and Biofuels	8	Oil, Gas and Biofuels	8	Exploration, Refining and Distribution	7
				Equipment and Services	1
Basic Materials	10	Mining	3	Metalic Minerals	3
		Steel and Metalurgy	3	Steel	3
				Iron and Steel Products	0
				Copper Products	0
		Chemicals	1	Petrochemicals	1
				Fertilizers	0
				Chemicals - Others	0
		Wood and Paper	3	Wood	1
				Pulp and Paper	2
		Packaging	0	Packaging	0
		Diversified Materials	0	Other Materials	0
Utilities	18	Electric Utilities	15	Electric Utilities	15
		Water Utilities	3	Water Utilities	3
		Gas Utilities	0	Gas Utilities	0
SUM	36		36		36

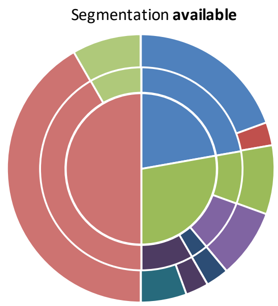

Figure 2: Segmentation available

Using this list, data was collected from all these companies from Q3 2017 to Q3 2023 using Bloomberg and, later, Yahoo Finance to enhance missing values. Missing variables (e.g. FCF) were collected from the last 4 quarters available and there was an average margin based on revenue calculated (a detailed list of which companies data was extended is provided in the appendix). With this approach, it is possible to fill in the missing data points from the past based on the newest observations. It is important to mention here that more than 95% of the data collected, especially all the data from market capitalisation and the ratios explained, was gathered from Bloomberg. This was done to ensure that the quality is as high as possible since Bloomberg offers a very comprehensive database.

Herein, it is also noteworthy that 11 out of these 108 companies were not available on Bloomberg, so they were dropped from the model.

3.2. Data Adjustments

Not all companies had sufficient data available to create the model ensuring good comparability. This is related to using data from Bloomberg and not mixing up sources to ensure comparability. Regarding that, for the final model, there were 36 companies out of the initial 108 extracted from the São Paulo stock exchange. This comes down due to some companies simply do not have a sufficient amount of values available for each variable. This regards companies that did not have values for all variables needed to calculate the ratios used in the model - asset turnover, debt ratio, current ratio, price-earnings ratio, net profit margin, and book-to-market value. Similar main variables were used in the study of Ozlen (2014).

In the further process, for some of the companies (a detailed list in the appendix), the data for proxy variables has been enhanced by the latest quarterly figures from Yahoo Finance for the last quarters. This was made to be able to use more companies for the analysis but still stick to the rule of using Bloomberg as the majority source. Then, the averages of the corresponding revenue figures were calculated and divided by the average of the newly collected data. In this way, it was possible to get the numbers for companies with some missing variables to enhance the whole number of companies analysed in the model. Revenue was used as a base parameter because this is the most unbiased estimate, which is a crucial part of almost every valuation model. The following process is summing up variables to an accumulated variable to then calculate the ratios analysed by the model.

The raw inputs of the variables are as follows (to calculate the ratios and proxy inputs for the model): total revenue; cost of revenue; depletion, depreciation & amortisation (D&A); income tax expense; EBITDA; net income; current assets; property, plant & equipment, net; total assets; current liabilities; total liabilities; total shareholders' equity; capital expenditures (CapEx); free cash flow (FCF); and market capitalisation. These inputs then have been transformed for all operating and cash flow lines to annualise the data on a twelve-month rolling basis (LTM), so they have been summed up using the values of the last four quarters. All the items using that approach are total revenue, cost of revenue, D&A, income tax expense, EBITDA, net income, CapEx, and FCF. This was not necessary for the balance sheet items because they reflect the value of assets or liabilities at a certain point and are not generated over time. The same logic applies to market capitalisation.

In the ongoing process, all the quarterly data from each company has been summed up to obtain an accumulated total value of the companies representing the energy sector. This has been made because the model uses quarterly data as input and the quarterly market volatility is smaller than the annual one. As a demonstration, all revenues from the 36 companies were summed up for each quarter to obtain the total revenue generated by all these companies at this time. Thereafter, the ratios have been calculated: asset turnover as total sales divided by total assets; debt ratio as total debt divided by total assets; current ratio as current assets divided by current liabilities; price-earnings ratio by market cap divided by net profit; net profit margin as net profit divided by total revenue; and book-to-market as book value divided by market capitalisation. Following that, the data of the 36 companies, including the ratios, has been multiplied by the EBITDA except for market capitalisation. This was made to be able to add companies that later published their data without biasing the ratios. Without this process, a ratio would only slightly change compared to market capitalisation if a company is added at a later stage in the time series. So, if the market capitalisation increased in value because one or more companies were added to the list (i.e., newly listed or provided sufficient data), the ratios also changed in value and are still able to explain the market capitalisation respectively.

As the last point, it is important to mention that some minor adjustments[2] have been made to key input variables (to calculate the ratios).

3.3. Quality and Structure of the Dataset

Data quality is one of the key factors regarding all quantitative valuation models, especially when it comes to determining the significance of inputs. The data collected within has been focused predominantly on the source Bloomberg and enhanced by Yahoo Finance. Bloomberg was the first source of the data with approximately 95% of the data and the rest was gathered by Yahoo Finance. Other academic research also relies on data from Bloomberg aiming to keep the quality high. Therefore, the quality of the data is considered to be quite high but also needs to be limited regarding the timeframe. The data

[2] If at least 95% of a data series for one Company is available (e.g. revenue) but for one quarter there is a missing value, the average of the following and previous value has been calculated and inserted to get an approximation of the missing value. This process is necessary to keep the specific company in the dataset.

collected starts in Q3 2017 and ends in Q3 2023. Most of the companies did not report later figures. Therefore, the end of Q3 2023 was at the point where this dissertation was created, the best option.

The benefit of using quarterly data is that there are more observations within a smaller timeframe and seasonal fluctuations are included even though they should not be that high. Compared to annual data reflecting for instance only the 31st of December Magnusson, et al., (2005) found in their study comparing financial ratios that result from the previous quarter affect also the following quarter. Referring to that it reflects the more current years with a higher accuracy compared to longer time frames with annual data. Hence, if investors would wish to make long-term forecasts a higher time frame would be beneficial.

The quality of the current dataset can also be limited by the negative impacts such as the one of COVID-19 in 2019 and 2020 (Li, et al., 2022), with a strong recovery thereafter. Moreover, especially companies from the oil and gas sector are strongly affected by this since the price per barrel of oil as well as the gas prices dropped to historic lows, and the demand was deficient. On the other hand, this has a strong compensation effect due to the speedy recovery.

3.4. Empirical Strategy

In the analysis part of creating the appropriate strategy, AI tools were used to assist in addressing problems occurring with the dataset. The analysis of the data is based on linear regression. This aimed to make results simple and more valuable without changing the initial information. The problem is that some models, such as machine learning models, could detect even more effects but would be harder to replicate. Using this approach, additional challenges are that the data is a time series, so there are problems regarding stationarity and autocorrelation combined with potential heteroskedasticity. Furthermore, the variables have a strong multicollinearity, resulting in a high r-squared with high p-values. First, the data of the market capitalisation that represents a time series was tested using a unit root test (Augmented Dickey-Fuller Test) to ensure that the data is stationary. To address then all these problems including achieving stationarity, the model has been adjusted with log transformations (reduce the dispersion on both the dependent and independent variables), and a differencing strategy (achieve stationarity) coupled with a principal component analysis (PCA) to address the multicollinearity issue. In academic research it is common to use the PCA model to address multicollinearity issues. The logarithmic transformation is a common approach in statistics, also described by Kato (2012), to mitigate homoskedasticity among other benefits such as a better overall fit of the model.

More general information about why the PCA has been used can be broken down into 2 key points. First, the PC analysis addresses the multicollinearity issue, which is a crucial point for the model since almost all variables correlate strongly. By decomposing the data matrix via a linear eigenvector decomposition, the PC includes information with respective loadings of the variables from the dataset that are free from correlation. Secondly, the PC analysis reduced the number of regressors drastically and simplified the model. So, this approach fits the model issues perfectly and still enables the determination of the effect on market capitalisation.

3.5. Descriptive Statistics and Noteworthy Observations

The descriptive statistics of the dataset are provided below in Table 1 and Table 2, whereby Table 2 is Table 1 continued.

	Market Capitalization	*Asset Turnover*	*Debt Ratio*	*Current Ratio*	*Price Earning Ratio*	*Net Profit Margin*	*Book to Market*
Mean	1,267,209.3	239,706.5	288,300.5	650,483.0	4,546,301.8	60,353.2	319,781.4
Standard Error	61,083.0	27,156.0	23,610.6	58,103.9	538,466.2	10,517.0	30,305.1
Median	1,320,443.2	148,676.3	239,207.8	507,647.2	3,594,162.5	28,560.4	235,872.4
Standard Deviation	305,415.2	135,780.1	118,053.0	290,519.7	2,692,331.0	52,585.1	151,525.6
Sample Variance	93,278,469,628.6	18,436,231,692.6	13,936,520,892.5	84,401,705,174.1	7,248,645,987,944.9	2765,194,278.9	22,960,003,318.8
Kurtosis	-0.3	-1.5	-1.6	-1.5	6.1	-1.6	-1.4
Skewness	-0.4	0.5	0.2	0.4	2.4	0.5	0.6

Table 1: Descriptive Statistics

	LTM Cost of Revenue	*LTM D&A*	*LTM Income Tax*	*LTM EBITDA*	*Property, Plant & Equipment, Net*	*LTM CapEx*	*LTM FCF*
Mean	841,312.6	100,195.7	55,655.7	439,489.4	1,181,501.0	-114,027.4	212,294.7
Standard Error	44,298.1	4,819.6	10,855.2	37,983.8	28,500.3	8,882.9	18,007.8
Median	734,728.6	100,252.0	36,178.0	339,479.9	1,167,082.2	-93,048.9	208,943.8
Standard Deviation	221,490.4	24,098.2	54,275.9	189,919.2	142,501.5	44,414.3	90,039.0
Sample Variance	4,905,7998,6447	580,723,846.0	2,945,875,016.8	36,069,312,785.6	20,306,673,370.9	1,972,630,586.0	8,107,026,381.3
Kurtosis	-1.0	-0.8	-0.9	-1.6	1.1	0.7	-1.4
Skewness	0.5	0.4	0.2	0.2	-0.6	-1.4	0.2

Table 2: Descriptive Statistics (continued)

The main point here is that all the figures presented diverge from the actual ratios as well as variables and their meaning. This is due to the adjustment process, including the EBITDA multiplication. The adjustment process and EBITDA multiplication were done to address the data set issues before and to be able to add companies in a later stage as well since otherwise ratios do not change in absolute value. The descriptive statistics shown here are the ones from the variables just multiplied by EBITDA (raw inputs for the model) before performing the log transformation and first differences. Furthermore, the figures presented (especially for market capitalisation) are in BRLm. Therefore, the mean of the market capitalisation is roughly BRL 1.5bn with a standard deviation of BRL 569,800m. All the variables have 25 observations that are used in the model; after the first differencing, there are 24 left over. The other variables are all included for completeness. Another essential thing for other researchers that may try to replicate results is that the LTM CapEx presents a cost line on the income statement and is sometimes reported as a positive number, whereas here, it is left in the original as a negative number.
The Graph (Figure 3) below presents the development of the total market capitalisation of the 36 companies. It shows the development of the number of companies considered. Even excluding them during the observed time frame, there was real growth within the energy sector in Brazil. Furthermore, in the ongoing part, the development of the individual companies is provided separately.

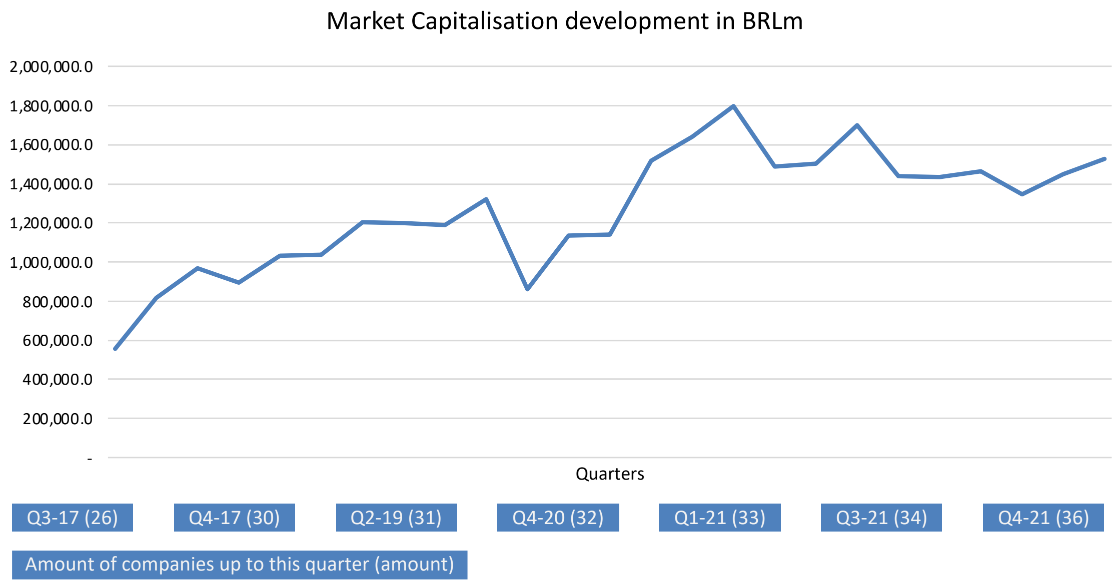

Figure 3: Development of the Market Capitalisation in BRLm

One problem the model addresses is the underlying correlation between the variables since several of them use the same variables as inputs as part of their calculation. For example, in the study of Hovakimian et al. (2006), they found that past profit ratios are an important predictor of observed debt ratios. Hence, one can also think of a correlation between these ratios as well as other ratios. The table below (Table 3) shows an overview of the correlation matrix before transforming the data using logarithms and differences.

Correlation Matrix (original)	MC	AT	DR	CR	PE	PM	BM
MC	1	0.720	0.803	0.800	-0.258	0.760	0.618
AT	0.720	1	0.974	0.970	-0.482	0.975	0.973
DR	0.803	0.974	1	0.975	-0.367	0.952	0.952
CR	0.800	0.970	0.975	1	-0.398	0.975	0.921
PE	-0.258	-0.482	-0.367	-0.398	1	-0.542	-0.434
PM	0.760	0.975	0.952	0.975	-0.542	1	0.913
BM	0.618	0.973	0.952	0.921	-0.434	0.913	1
LTM COGS	0.657	0.944	0.926	0.873	-0.424	0.863	0.970
LTM D&A	0.702	0.830	0.889	0.809	-0.156	0.754	0.879
LTM IT	0.614	0.909	0.821	0.862	-0.741	0.937	0.839
LTM EBITDA	0.801	0.986	0.996	0.983	-0.412	0.969	0.958
PP&E	0.797	0.830	0.877	0.799	-0.379	0.778	0.838
LTM CapEx	-0.484	-0.756	-0.751	-0.652	0.331	-0.647	-0.842
LTM FCF	0.833	0.906	0.954	0.965	-0.248	0.924	0.846

Correlation Matrix (original)	LTM COGS	LTM D&A	LTM IT	LTM EBITDA	PP&E	LTM CapEx	LTM FCF
MC	0.657	0.702	0.614	0.801	0.797	-0.484	0.833
AT	0.944	0.830	0.909	0.986	0.830	-0.756	0.906
DR	0.926	0.889	0.821	0.996	0.877	-0.751	0.954
CR	0.873	0.809	0.862	0.983	0.799	-0.652	0.965
PE	-0.424	-0.156	-0.741	-0.412	-0.379	0.331	-0.248
PM	0.863	0.754	0.937	0.969	0.778	-0.647	0.924
BM	0.970	0.879	0.839	0.958	0.838	-0.842	0.846
LTM COGS	1	0.902	0.807	0.932	0.907	-0.899	0.785
LTM D&A	0.902	1	0.587	0.875	0.882	-0.905	0.781
LTM IT	0.807	0.587	1	0.859	0.674	-0.605	0.746
LTM EBITDA	0.932	0.875	0.859	1	0.870	-0.752	0.947
PP&E	0.907	0.882	0.674	0.870	1	-0.817	0.766
LTM CapEx	-0.899	-0.905	-0.605	-0.752	-0.817	1	-0.538
LTM FCF	0.785	0.781	0.746	0.947	0.766	-0.538	1

Table 3: Correlation Matrix Original

As obviously visible, there is a high correlation indicated by the red colour of the cells. In academic research variables with a correlation coefficient of close to 1 are considered highly correlated. The only variable with a relatively low correlation is the price-earnings ratio. Using these raw data as input, a simple regression model would simply result in a high R squared combined with low significance levels of the estimates. After the transformations (Table 4), the collinearity of the variables improved but still resulted in strong correlations among the variables. Moreover, the VIF values were also calculated. However, VIF values are not the primary reason to transform the data, but the VIF values are a traditional measure applied to detect the presence of collinearity in linear regression models (Fahrmeir, et al., 2021; Salmerón et al. 2018).

Correlation Matrix (data transformed)	Diff Log(MC)	Diff Log(AT)	Diff Log(DR)	Diff Log(CR)	Diff Log(PE)	Diff Log(PM)	Diff Log(BM)
Diff Log(MC)	1	0.229	0.103	0.205	0.015	0.472	-0.588
Diff Log(AT)	0.229	1	0.847	0.791	-0.635	0.774	0.581
Diff Log(DR)	0.103	0.847	1	0.708	-0.614	0.707	0.655
Diff Log(CR)	0.205	0.791	0.708	1	-0.476	0.635	0.531
Diff Log(PE)	0.015	-0.635	-0.614	-0.476	1	-0.854	-0.567
Diff Log(PM)	0.472	0.774	0.707	0.635	-0.854	1	0.299
Diff Log(BM)	-0.588	0.581	0.655	0.531	-0.567	0.299	1
Diff LTM COGS	0.093	0.755	0.649	0.544	-0.382	0.448	0.546
Diff Log(LTM D&A)	0.172	0.204	0.237	0.180	0.157	0.011	0.048
Diff LTM IT	0.346	0.674	0.372	0.545	-0.555	0.678	0.149
Diff LTM EBITDA	0.165	0.934	0.763	0.777	-0.547	0.674	0.543
Diff PP&E	0.443	0.415	0.581	0.208	-0.422	0.595	0.131
Diff LTM CapEx	0.058	0.070	0.029	0.149	-0.031	0.102	-0.035
Diff LTM FCF	0.047	0.761	0.687	0.660	-0.384	0.496	0.529

Correlation Matrix (data transformed)	Diff LTM COGS	Diff Log(LTM D&A)	Diff LTM IT	Diff LTM EBITDA	Diff PP&E	Diff LTM CapEx	Diff LTM FCF
Diff Log(MC)	0.093	0.172	0.346	0.165	0.443	0.058	0.047
Diff Log(AT)	0.755	0.204	0.674	0.934	0.415	0.070	0.761
Diff Log(DR)	0.649	0.237	0.372	0.763	0.581	0.029	0.687
Diff Log(CR)	0.544	0.180	0.545	0.777	0.208	0.149	0.660
Diff Log(PE)	-0.382	0.157	-0.555	-0.547	-0.422	-0.031	-0.384
Diff Log(PM)	0.448	0.011	0.678	0.674	0.595	0.102	0.496
Diff Log(BM)	0.546	0.048	0.149	0.543	0.131	-0.035	0.529
Diff LTM COGS	1	0.300	0.381	0.647	0.461	-0.304	0.396
Diff Log(LTM D&A)	0.300	1	-0.243	0.120	0.327	-0.524	0.132
Diff LTM IT	0.381	-0.243	1	0.738	0.086	0.302	0.480
Diff LTM EBITDA	0.647	0.120	0.738	1	0.246	0.165	0.850
Diff PP&E	0.461	0.327	0.086	0.246	1	-0.252	0.179
Diff LTM CapEx	-0.304	-0.524	0.302	0.165	-0.252	1	0.442
Diff LTM FCF	0.396	0.132	0.480	0.850	0.179	0.442	1

Table 4: Correlation Matrix Transformed

Another interesting observation is the distribution of the quarterly returns of the total market capitalisation of the energy sector. The distribution in Figure 4 (below) is quite limited due to the number of observations, but it has a similar structure to the theoretical distribution of returns. The distribution seems to be similar to a normal distribution with a slight shift towards positive returns. The return distribution is also interesting regarding the analysis of the chosen time frame since it could be biased towards positive or negative returns depending on the economic cycle from which the data was extracted. Here, it includes the impact of COVID-19 as well as the recovery phase. However, the total returns here are still relatively well-distributed as theoretically expected for normal market cycles (without crisis).

Figure 4: Return Distribution

Even though the returns calculated here include the increasing number of companies over time, the graph does not highly indicate this. However, this comes due to the small impact of the added companies. In other words, the main drivers of market capitalisation are two companies.

The graph below (Figure 5) shows the impact of the market capitalisation development by each company. Here, it is visible that around 50% of the final market capitalisation of this sector is characterised by 2 companies, namely Petrobras (blue) and Vale (red). The rest of the companies are relatively even distributed, contributing to the market capitalisation. On the other hand, it is unclear whether returns are impacted by adding companies over time. So also, with the removal of these special effects, the distribution remains mainly the same.

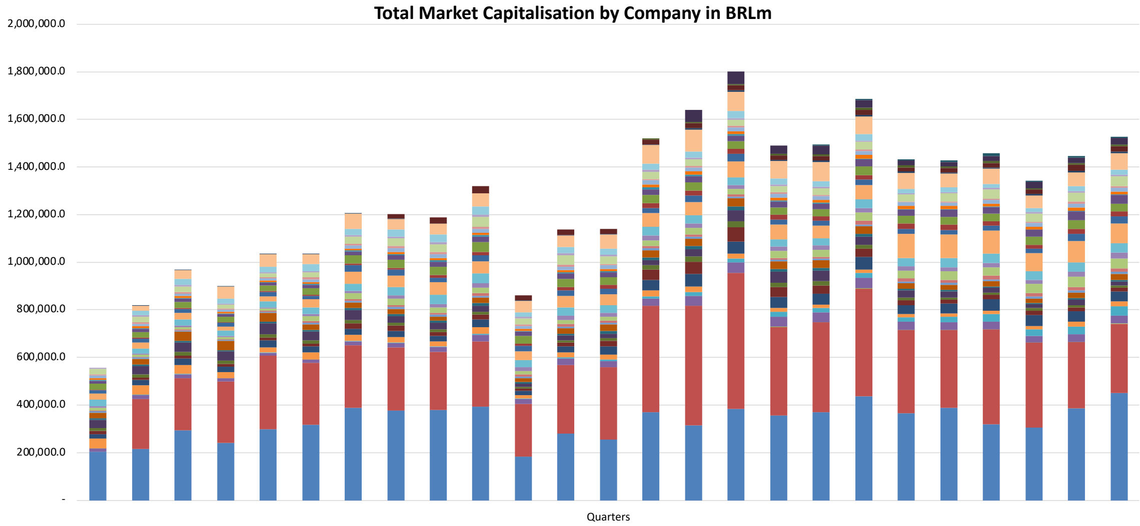

Figure 5: Total Market Capitalisation by Company in BRLm

The graph of total market capitalisation by companies summed up shows the same development as the total market capitalisation graph during the quarters of the observations.

The final transformation for the variables is as follows (log refers to log transformation and diff refers to first difference): diff[3] log[4] market capitalisation; diff log asset turnover; diff log debt ratio; diff log current ratio; diff log price-earnings ratio; diff log net profit margin; diff log book to market ratio; diff cost of revenue; diff log D&A; diff PP&E; diff income tax; diff EBITDA; diff CapEx; and diff FCF. As the last variable included in the PC analysis, there is also the log diff market capitalisation with a lag of one to fix the autocorrelation problem.

The first test is the Augmented Dickey-Fuller Test (Said & Dickey, 1984 cited in Valenzuela, et al., 2021). This consists of the testing of $\rho^* = 0$ against $\rho^* < 0$ in the augmented regression ($\Delta = 1 - B$) as follows:

$$\Delta y_t = \mu_t + \rho^* y_{t-1} + \sum_{j=1}^{k-1} \phi_j^* \Delta y_{t-1} + \varepsilon_t$$

Regarding the ADF stationarity test results, the market capitalisation time series is after the log transformation and using the first difference stationary to the 5% significance level. The H_0 hypothesis is that there is a unit root and, therefore, is not stationary. The

[3] Diff: first difference (used the R code “diff(variable)”)
[4] Log: Logarithm transformation (used the R code: “log(variable)”)

results of the test suggest that to the 5% level (p-value: 0.02375), there is no unit root present for the Diff_Log_Market Capitalisation series, so the H_0 can be rejected. There could also be a higher significance achieved by further differencing the time series, but then it loses the ability to explain market capitalisation adequately.

4. RESULTS

The final model was tested for stationarity regarding the time series market capitalisation using the Augmented Dickey-Fuller Test (Said & Dickey, 1984 cited in Valenzuela, et al., 2021) and a test for checking autocorrelation (first order and up to the 4^{th} order due to the quarterly data). Moreover, also the first difference of the other variables was taken to ensure stationarity and the highest R squared possible. Applying all these fixes, the model consists of 5 principal components explaining more than 90% of the variance. Using the loadings of each principal component, one can then interpret the significance of the underlying variables. So, the whole model is more difficult than a simple regression but ultimately based on this. For the model 5 PCs have been chosen because this is the simplest form of the model that can explain the variance well (adjusted R squared of 0.81), while with 4 PCs, the adj. R squared drops strongly to 0.23 with insignificant estimates So, the model, including 5 PCs, provides the best trade-off of simplicity and explanation.

The final equation of the regression used is as follows:

General:

$$Y_t = \beta_0 + \beta_1 PC_{t1} + \beta_2 PC_{t2} + \beta_3 PC_{t3} + \beta_4 PC_{t4} + \beta_5 PC_{t5} + \beta_6 Y_{t-1} + u_t$$

With the following estimates:

$$\begin{aligned} Diff(\widehat{Log}(MC_t)) \\ &= 0.041 - 0.011PC_{t1} + 0.001PC_{t2} - 0.062PC_{t3} + 0.068PC_{t4} \\ &- 0.128PC_{t5} - 0.032Diff(Log(MC_{t-1})) \end{aligned}$$

The Principal Components are calculated using the linear combinations (Mukherjee, et al., 2018) $PC_1 = l_{11}X_1 + l_{21}X_2 + \cdots + l_{p1}X_p = e_1'X$, $PC_2 = l_{12}X_1 + l_{22}X_2 + \cdots + l_{p2}X_p = e_2'X$, …, $PC_p = l_{1p}X_1 + l_{2p}X_2 + \cdots + l_{pp}X_p = e_p'X$. The result obtained is the $X' = (X_1, X_2, \dots X_s)$ with the covariance matrix $\sum$ with eigenvalue-eigenvector pairs $(\lambda_1, e_1)\dots$ (λ_p, e_p) where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$. Then the $Y_1 = e_1'X, Y_2 =$

$e_2'X, \dots Y_p = e_p'X$ with $var(Y_i) = \lambda_i (i = 1,2, \dots, p)$ and $\sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp} = \sum_1^p var(X_i) = \lambda_1 + \cdots + \lambda_p = \sum_1^p var(PC_i)$. The $Y_1, Y_2, \dots, Y_k$ $(k < p)$ are called the principal components with the first component Y_1 explaining the largest variance.

The model results suggest that Diff_Net_Profit_Margin is especially heavily negatively impacting market capitalisation. All the other variables (diff_log transformation) for the ratios, asset turnover, current ratio, price-to-earnings ratio, and net profit margin positively impact market capitalisation. Also, Fairfield & Yohn (2001), in their study, showed evidence that return on asset, decomposed into asset turnover and net profit margin, provides information about future profitability - impacting market capitalisation. Moreover, in the context of a case study in Indonesia, Septyanto et al. (2022) found that the current ratio positively impacts financial distress using a Z-score model. Even though the study and model used differ from the analysis concluded here, the impact of the current ratio in both cases is positive (market capitalisation as well as financial distress). Therefore, this is relevant in predicting future market capitalisation since it is strongly connected to future profitability. Only the debt ratio has a slight negative impact on market capitalisation if increased. Other studies claim a negative impact of increased debt on market capitalization aligned with the results here (Desai, 2021). The proxy variables tested all have a slight positive impact except for FCF and COGS. However, their effect is closer to zero except for D&A, PP&E, and Income Tax expense, which slightly positively impact market capitalisation. In general, Income Tax Expenses vary a lot in academic research (Thomas & Zhang, 2014).

4.1. Robustness Checks

The model was checked for all the common statistical issues and was adjusted to fix them accordingly. Lu & White (2014) elaborated that it is a common exercise for empirical studies to use robustness checks to check how the model behaves by adding or removing some of the regressors. This work has been done in the background to show the final version providing the best fit. However, the final model was tested using the Breusch-Godfrey test for serial autocorrelation (Godfrey, 1976; Tomas Cipra, 2020) and for heteroskedasticity using the Studentized Breusch Pagan test (Breusch & Pagan, 1979 cited in Fahrmeir, et al., 2021). The motivation for using the Breusch-Godfrey test comes from Uyanto (2020), claiming that this test is usually superior compared to other statistical tests for autocorrelation when also using a lagged version of the dependent variable, which here is the case.

The Breusch Godfrey autocorrelation test (Godfrey, 1976; Cipra, 2020) consists of the calculation of the residuals $\hat{u}_t$ of the model. Then one needs to estimate the auxiliary model with the form: $\hat{u}_t = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_k x_{tk} + \varphi_1 \hat{u}_{t-1} + \varphi_2 \hat{u}_{t-2} + \cdots + \varphi_p \hat{u}_{t-p} + \varepsilon_t$

This is followed by the testing of $H_0: \varphi_1 = \varphi_2 = \cdots = \varphi_p = 0$ against $H_1: \varphi_1 \neq 0 \; or \; \varphi_2 \neq 0 \; or \dots \varphi_p \neq 0$ using the F-test $F = \frac{S_2^2}{S_1^2}$ with the significance level α (the quantile is $F_{1-\alpha}(p, n-k-p)$). The Breusch Pagan Test (Breusch & Pagan, 1979 cited in Fahrmeir, et al., 2021) uses the model for the error variances: $\sigma_i^2 = h(\alpha_0 + \alpha_1 z_{i1} + \cdots + \alpha_q z_{iq})$ with h as a function not depending on the unit index i and z are covariances that may influence the error variance. Then, homoscedasticity is given for $\alpha_1 = \cdots = \alpha_q = 0$ with $H_0: \alpha_1 = \cdots = \alpha_q = 0$.

The results of the tests are shown in Table 5 below.

Test	**p-value**
Augmented Dickey-Fuller Test	0.0238
Breusch Godfrey Test (order 1)	0.4087
Breusch Godfrey Test (order 4)	0.7509
Studentized Breusch Pagan Test	0.3770

Table 5: Econometric Tests

Moreover, the test for autocorrelation, both to the 1st and up to the 4th order, indicates the absence of residual autocorrelation by a p-value of 0.4087 and 0.7509, respectively. The H_0 of this test says that there is no autocorrelation present up to the chosen order. In addition, also the Studentized Breusch Pagan test confirms no heteroskedasticity after the applied transformation to any common significance level (10% or lower). The H_0 of this test says that homoskedasticity is present; here, it is indicated by the high p-value of 0.377. The test suggests H_0 should not be rejected, so the error is homoscedastic. Therefore, the results of the model can be considered valid and analysed in the following subsection.

As a minor point but also worth discussing, the Variance Inflation Values (VIF) go up to 71.26 (before the PCA), which is very high. Although this was not the original motivation for changing the model to a PCA model, it still showcases and quantifies the high correlation between the variables.

4.2. Significance Analysis

Overall, the model's original output is as follows (Table 6):

Standard output

	Estimate	Std. Error	t-value	Pr(>|t|)	
(Intercept)	0.0408	0.0153	2.662	0.0164	*
PC1	-0.0109	0.0060	-1.820	0.0865	.
PC2	0.0010	0.0127	0.077	0.9394	
PC3	-0.0617	0.0133	-4.655	2.27E-04	***
PC4	0.0677	0.0161	4.210	5.89E-04	***
PC5	-0.1283	0.0169	-7.589	7.43E-07	***
lag_Diff_Log_Market_Cap	0.0319	0.1066	0.300	0.7680	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07213	17 degrees of freedom
Multiple R-squared: 0.8628	Adjusted R-squared: 0.8144
F-statistic: 17.82 on 6 and 17 DF	p-value: 1.784e-06

Table 6: Standard Output

The model shows a very good fit with an R-squared of 0.86 and an adjusted R-squared of 0.81. This outcome is favourable since an R-squared too high combined with large p-values for estimated coefficients would indicate strong multicollinearity issues. Moreover, the p-value of the F-statistic is also very low, so one can be confident that the model includes at least one significant variable to explain market capitalisation. Summarising that, the model shows strong adaptability using the input data, and the independent variables significantly influence market capitalisation.

All the tests confirm the classic hypothesis of the regression model; the variables are stationary, and no autocorrelation or heteroskedasticity, at least to the 5% significance level (see section 4.1. Robustness checks). So, the statistical inference shows that PC3, PC4, and PC5 are especially highly statistically significant. The simplified graph below (Table 7) shows that PC3 and PC4 are significant at the 1% level, and PC5 even at the 0.1% level. Therefore, one can be very sure that there is an impact regarding the explainable variables on total market capitalisation. Moreover, PC1 is also considered significant but only to the 10% significance level.

	Estimate	p-value	significance level
PC1	-0.0109	0.0865	10.0%
PC2	0.0010	0.9394	none
PC3	-0.0617	0.0002	1.0%
PC4	0.0677	0.0006	1.0%
PC5	-0.1283	0.0000	0.1%

Table 7: Principal Component Estimates

The estimates of the regression analysis are negative for PC1, PC3, and PC5. This is important because the loadings need to be multiplied by the negative number to see the real impact. The loadings of each PC are summarised in the following table (Table 8) highlighted by a colour code to visualise higher loadings on some variables. The loadings here represent the eigenvectors of the correlation matrix calculated within the PC analysis and provide an overview of how each variable is related to the corresponding PC. The interpretation of the principal components is possible using the loading of each variable multiplying the loading times the estimate of the principal component and summing that up to obtain the overall effect on market capitalisation. This procedure is done in Table 8 (see below).

	PC1	PC2	PC3	PC4	PC5
Diff_Log_Asset_Turnover	-0.372	0.019	0.061	0.018	-0.109
Diff_Log_Debt_Ratio	-0.342	0.164	-0.088	-0.141	0.118
Diff_Log_Current_Ratio	-0.314	-0.025	0.141	-0.219	-0.196
Diff_Log_Price_Earnings_Ratio	0.283	0.247	0.296	-0.094	-0.393
Diff_Log_Net_Profit_Margin	-0.318	-0.172	-0.363	0.193	-0.072
Diff_Log_Book_To_Market	-0.248	0.081	0.206	-0.390	0.618
Diff_LTM_Cost_of_Revenue	-0.282	0.284	0.074	0.188	0.124
Diff_Log_LTM_DA	-0.063	0.673	0.002	-0.026	-0.353
Diff_Property_Plant_Equipment	-0.183	0.309	-0.637	0.095	-0.009
Diff_LTM_Income_Tax	-0.255	-0.443	0.049	0.277	-0.321
Diff_LTM_EBITDA	-0.353	-0.084	0.199	-0.060	-0.218
Diff_LTM_CapEx	0.100	-0.204	-0.418	-0.740	-0.256
Diff_LTM_FCF	-0.295	-0.025	0.281	-0.246	-0.196

Table 8: Eigenvector Loadings of the Principal Components

The highest loading on variables here (not regarding positive or negative and impact on market capitalisation) are Diff_Log_LTM_D&A, Diff_Log_Book_To_Market, Diff_Property_Plant_Equipment and Diff_LTM_CapEx. The loadings here are especially at the higher end for PC2, PC3, PC4, and PC5. However, one still needs to remember that PC2 is not significant and that the explained variance is decreasing over the PCs, meaning that PC1 explains the most variance and PC5 the least variance. The Scree Plot (Figure 6) shows this development.

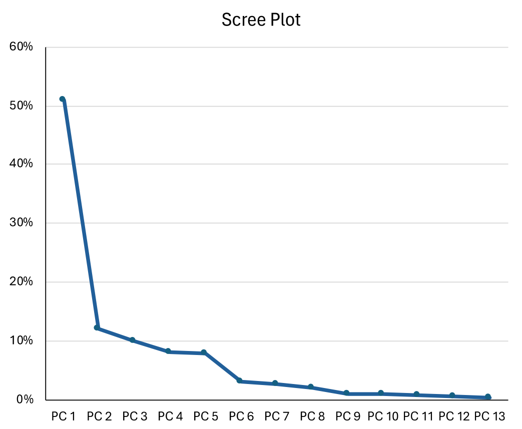

Figure 6: Scree Plot

Putting all this together, there is the following impact obtained (Table 9):

	PC1	PC3	PC4	PC5	Effect		
Estimate	-0.0109	-0.0617	0.0677	-0.1283	All	Ratios	Proxies
Diff_Asset_Turnover	-0.3716	0.0613	0.0176	-0.1090	0.0154	0.0154	-
Diff_Debt_Ratio	-0.3421	-0.0885	-0.1413	0.1177	-0.0155	-0.0155	-
Diff_Current_Ratio	-0.3139	0.1407	-0.2188	-0.1964	0.0051	0.0051	-
Diff_Price_Earnings_Ratio	0.2826	0.2963	-0.0935	-0.3926	0.0227	0.0227	-
Diff_Net_Profit_Margin	-0.3180	-0.3626	0.1927	-0.0716	0.0481	0.0481	-
Diff_Book_To_Market	-0.2484	0.2062	-0.3900	0.6178	-0.1157	-0.1157	-
Diff_LTM_Cost_of_Revenue	-0.2824	0.0736	0.1882	0.1243	-0.0047	-	-0.0047
Diff_LTM_DA	-0.0627	0.0016	-0.0263	-0.3532	0.0441	-	0.0441
Diff_Property_Plant_Equipment	-0.1834	-0.6372	0.0951	-0.0095	0.0490	-	0.04906
Diff_LTM_Income_Tax	-0.2548	0.0490	0.2769	-0.3210	0.0597	-	0.0597
Diff_LTM_EBITDA	-0.3527	0.1987	-0.0600	-0.2181	0.0155	-	0.0155
Diff_LTM_CapEx	0.0997	-0.4179	-0.7403	-0.2561	0.0074	-	0.0074
Diff_LTM_FCF	-0.2952	0.2809	-0.2458	-0.1956	-0.0057	-	-0.0057

(light grey)	no impact	(light grey)	no impact
(green)	positive impact	(blue)	high impact (positive/negative)
(red)	negative impact		

Table 9: Effect of the Input-Variables on Market Capitalisation

This is the main part of the model and shows that multiplying each loading by the respective PC estimator summed up for all five values, resulting in the total effect. The total effect column shows the impact presented by multiplying each loading with the PC estimate summed up for all 4 PCs together. Therefore, it shows the net effect on market capitalisation which is highlighted by a respective colouring with red as negative and green as positive impact. Here, PC2 is excluded due to its lack of significance. So, regarding the model's output, it says that an increase in Diff_Book_To_Market has a negative impact on market capitalisation. This means that if the book-to-market value calculated by book value divided by market capitalisation is increasing, either market capitalisation is decreasing or the book value is increasing, the effect is negative on market capitalisation. This result is aligned with the one of Ozlen (2014), where he found a strong positive impact on stock prices using book value (not book-to-market ratio). The result here was expected because market capitalisation is already input into the ratio changing faster than book value because book value is an accounting value (not changing daily compared to market capitalisation). Nevertheless, as shown by the study of Ozlen

(2014) a large part of the effect goes down to book value directly. Other studies including the one by Chan, et al. (1980) found that an increase in the book-to-market ratio has a positive effect on stock returns in Japan. This slightly deviates from the impact found here but reflects that a higher book-to-market ratio seemingly showcases an undervalued stock implying a buy signal for investors. On the other hand, a significant result of the model is that if debt value is increased (captured by the total debt ratio), the market capitalisation decreases. This means that the energy sector is heavily impacted by debt financing. According to the model, the companies are financed slightly too much by debt because the effect should be positive if only financed by equity and negative if financed by too much debt. Other studies as the one of Smyth & Hsing (1995) were looking for the optimal level and determined the optimal level at 48.9% debt which is well below the level of 66.18% (see Table 10) but linked to economic growth. This is interesting because the study of Smyth & Hsing (1995) took into account the debt-GDP ratio of an economy and therefore allows us to compare the energy sector in Brazil to the overall economy. However, as the colour scale shows, the effect is relatively close to zero, suggesting that the debt financing here should only decrease very slightly.

For the net profit margin, the effect of the PCs is also positive. This aligns with expectations since a higher profit is one of the key figures used by many valuation models, including a DCF model. So, in the end, higher cash-generation abilities or profit-generating abilities result in higher enterprise value and higher market capitalisation. Fairfield & Yohn (2001) found evidence that in forecasting returns asset turnover and net profit margin are not significant but the change in both is significant to forecast returns. Using the Principal component model the results are partially projectable because through the first differencing also the change is used but through the loadings of the principal components it can not be said that it is insignificant or significant to just watch the net profit margin itself. Moreover, it is difficult for companies to improve their net profit margin since this is an overall result of changes regarding other fundamentals, including working capital management and financing structure. A very similar logic applies to the price-earnings ratio, which also has a positive impact on market capitalisation. The price-to-earnings ratio has as input the market capitalisation, so it most likely captures the effect of rising market capitalisation. If the ratio is increasing, which is likely due to an increase in market capitalisation, the result explains itself. On the other hand, if the net income decreases, the price-to-earnings ratio increases as well, but this is normally not considered a positive effect. Hence, the more logical result is to explain the captured effect of rising market capitalisation as an input to the variable. In general, other studies such as the one of Anderson & Brooks (2006) claim that the price-earnings ratio is a useful tool for fund managers and hedge funds.

As the last main explanatory variable, the current ratio, the impact is more on the unclear side. The net effect results in almost zero but still positive. This implies that the current ratio is difficult to use in determining market capitalisation in Brazil's energy sector. This aligns with other research because the current ratio but also the debt-to-assets ratio are often used in measuring the financial distress level (Septyanto, et al., 2022).

Further results of the model are that the proxy variables do not have a strong effect on market capitalisation. In summary, an increase in depreciation and amortisation, property plant and equipment and income tax expense positively affect market capitalisation. Thomas & Zhang (2014) claimed that valuation regressions in prior research show strong variations of the effect of tax expenses from significant positive to significant negative. The reasoning for property plants and equipment sounds logical since the energy industry is very asset-heavy, and these are the most important inputs to produce the end product. On the other hand, the effect of D&A expenses is not directly explainable. If the income tax is increased, there is less cash available for the shareholders. However, the explanation could be the concept that if the company earns more, the total income tax is increasing, and therefore, the model detects a slightly positive effect even though income tax expenses are negative. The explanation for D&A could be that if more revenue is generated, the assets are used more heavily, and therefore, D&A needs to increase. Nevertheless, both effects are assumptions and not key results. LTM EBITDA, LTM CapEx, LTM FCF, and LTM Cost of Revenue do not have an impact and are all very close to 0. However, EBITDA should not be directly considered since it is already part of the ratios since they are all multiplied by EBITDA.

4.3. Discussion

When discussing the results of the model, one can say that they are quite good overall and partially aligned with the similar study of Ozlen (2014). They are mostly in line with expectations except for some of the proxy variable outputs, such as the positive impact on market capitalisation of an increase in income tax expense. However, especially this effect is on the unclear side according to other studies (Thomas & Zhang, 2014). All the variables that were part of the main focus, namely Asset Turnover, Debt Ratio, Current Ratio, Price Earnings Ratio, Net Profit Margin, and Book to Market Value, do have results that align with expectations. For example, Chan, et al. (1980) with their study in Japan reported an increase in book value corresponding to an increase in return. This concept focuses more on outcomes of returns and shows that a higher book-to-market ratio corresponds to a lower valuation (asset multiple). The analysis here focuses directly on market capitalisation which leads to a decreased market capitalisation. The model provides some useful estimates to evaluate any company from the energy sector in Brazil and suggests points that can be improved, similar to the study of Ozlen (2014). It can also be considered good that the influence of the proxy variables is relatively low, so the main drivers for market capitalisation are the ratios mostly unbiased by the development of the proxy variables. However, this is surprising considering that FCF is a proxyvariable of the model and FCF is widely used in valuation models (Beneda, 2003; Gardner, et al., 2012 and Ivanovska, et al., 2014)

Regarding the robustness checks, the model is stationary and free from autocorrelation and heteroskedasticity. This is important because, with non-stationarity, the model's results evaluation would be problematic (Manuca & Savit, 1996). On the other hand, there have been some transformations to achieve stationarity, including taking the first difference, which slightly impacts the capacity to analyse the results. It would be easier to work with the raw inputs because the effect could be described directly in terms of changing the market capitalisation amount. Nevertheless, after the first differencing, the model still contains the main inputs given by the market capitalisation, and therefore, the model offers a high interpretation capacity (Liker, et al., 1985).

As a last point here, it is interesting that market capitalisation remained mostly unaffected by the impact of COVID-19 and that the largest change was only visible in the development of the price-earnings ratio. This is a result in itself and can be explained as not all companies in the energy sector are cyclical. Also Harjoto & Rossi, (2023) showed that market capitalisation in emerging markets was significantly negatively affected by COVID-19, especially in the financial sector. Hence the energy sector in Brazil represented by this sample is considered not strongly cyclical.

Summarising this, the model suggests useful information for managers seeking optimisation information regarding the capital structure of energy companies as well as which operating changes lead to higher market capitalisation levels. Also, investors can use models like this to get an overview of whether the shares of a company are overvalued, undervalued or set around the right price level. This model can be seen as further research related to the study of Ozlen (2014) using a different market, time frame and proxy variables as well as a principal component analysis. However, this also needs to be limited because only fundamental values are considered and not company-specific information like the exact business model. Moreover, the timeframe of the model can be seen as a limiting factor and markets can change, so, the loads on variables would change together with their effect on market capitalisation. Nevertheless, the results of the model are satisfying regarding the initial assumptions made about the sector.

4.4. Further Results and Original Output

The graph below (Figure 7) shows an overview of the market capitalisation development and the raw ratios (not multiplied by EBITDA).

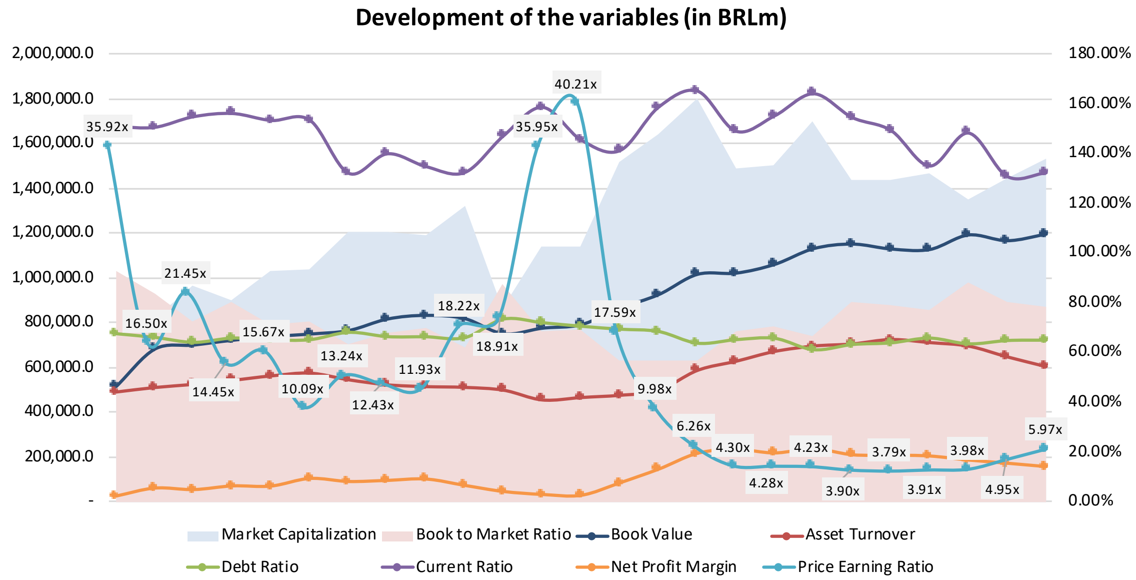

Figure 7: Development of the Variables in BRLm

For the graph, it is important to mention that the ratios remain mostly unchanged since they are relative, and the market capitalisation is influenced positively by the increasing number of companies over time. However, the book value shown here is also increasing with the numbers over time. The graph shows that the book value developed slower over time than the total market capitalisation. Moreover, it is interesting that the market capitalisation is not much influenced by the effect of COVID-19 because the volatility remains mostly the same. On the other hand, the price-earnings ratio increased significantly, most likely due to a short, temporary decrease in earnings. Moreover, it is also visible that the earnings and the price-to-earnings ratio recovered very fast to pre-COVID levels and averaged out at even lower levels, around 4x the annual earnings. All the other ratios were not affected a lot by the impact of COVID-19, which suggests that the Brazilian energy sector can be considered quite resilient.

In Millions of BRL	Average
Market Capitalization	1,267,209
Asset Turnover	51.76%
Debt Ratio	66.18%
Current Ratio	1.48x
Price Earning Ratio	13.52x
Net Profit Margin	11.22%
Book Value	906,042
Book to Market Ratio	72.80%

Table 10: Averages over the whole Time

Table 10 presents the average values of the calculated ratios, which provide an overview of the energy sector during the observed period. It says that the market price is just slightly above the Book Value, which is consistent with the characterisation of a value and not a growth market. Moreover, the debt ratio of around 66% is slightly at the higher end, and the price-earnings ratio of around 13.5x is on the moderate end compared to the broader market. However, in contrast, Smyth & Hsing (1995) concluded that the optimal debt ratio for economic growth is reached with a debt / GDP ratio of around 50%. This is not exactly applicable to the corporate sector, but it implies where optimal levels are expected since the energy sector in Brazil represents an essential part of the country's GDP. In other words, the Brazilian energy sector significantly influences the development of the country. Hence, macroeconomic factors can be applied on a higher level.

5. CONCLUSION

The Brazilian energy sector has a crucial impact on the Brazilian market since it is one of the most important markets due to the country's wealth of natural resources. The dissertation analysed the effect of the variables asset turnover, debt ratio, current ratio, price-to-earnings ratio, net profit margin, and book-to-market value on total market capitalisation. Besides this, the proxy variables LTM Cost of Revenue, LTM D&A, LTM Income Tax expense, LTM EBITDA, Property plant and equipment, LTM CapEx, and LTM FCF have been analysed. The major result of the dissertation is that the ratios are especially important in determining the corresponding market capitalisation. It is interesting for the managers to decide which improvements on fundamentals are crucial to enhance efficiency, resulting in higher market capitalisation as well as for investors looking for the highest returns.

The data for the model consists of 36 listed companies on the São Paulo stock exchange B3, all related to the Energy sector, having in common similar characteristics such as an asset-heavy business model that usually relies on strong debt financing. Moreover, the data consists of 25 quarterly observations from Q3 2017 to Q3 2023.

The model to analyse the dataset is a principal component analysis regression model that addresses the problem of the correlation between the explainable variables. This approach complicates the interpretation of the variables on the input, but on the other hand, the results are more favourable than those of a simple regression model. Using the interpretation tables as shown in the "Results" part, one can clearly see the impact of the variable on market capitalisation. Moreover, using the PCA model, it was possible to drastically reduce the number of variables used in the model which helps to identify structures within the dataset that are not detectable otherwise.

The model's main results have been that an increase in book-to-market ratio results in a decrease in total market capitalisation. This comes mainly from the input of market capitalisation itself to the model since book value is an accounting measure and does not change as fast as market capitalisation. Moreover, another key result is that the companies on average in the dataset have been financed around 66% by debt, which is close to the optimal level just slightly above. Since the estimate for the debt ratio is slightly negative, the companies should slightly reduce their debt levels to around 60%. The other key results are that net profit margin, price-earnings ratio, current ratio and asset turnover positively affect market capitalisation if increased. This is in line with expected results and makes sense for managers and investors to focus on. According to Brooks (2006), the net profit margin is widely used in analysing companies' returns, and he also proposed that the ratio is useful for fund managers and hedge funds, as market players that deploy lots of capital. However, the net profit margin has the highest positive impact on market capitalisation if increased respectively. Therefore, one of the central goals should be an increase in net profit (using the same revenue) and cutting costs as much as possible.

The model also showed that the market capitalisation of the Brazilian energy sector was not heavily affected by COVID-19 and can be considered resilient. This, however, also depends on the specific company analysed since oil and gas companies usually tend to be more volatile, as oil is an important raw input for many companies operating in the industrial sector.

Referring to the initial research question of the beginning: “What is the effect of company fundamentals on market capitalisation in the energy sector in Brazil?” One can answer that the ratios analysed do have a significant effect on market capitalisation, namely Asset Turnover, Debt Ratio, Current Ratio, Price Earnings Ratio, net Profit Margin, and Book Market Ratio. The theoretical and practical implication is that investors and managers can use these variables to calculate respective market capitalisation as well as enhance the company's performance to increase market capitalisation. Also, a practical result is that investors and managers can compare a specific company to the average of the sector.

On the other hand, the proxies sometimes show different effects than expected, such as increased income tax expenses leading to increased market capitalisation. This result is particularly interesting since it matches with the results of Thomas & Zhang (2014) reporting that valuation regressions using tax expense do have substantial variation in the results of the impact of tax expenses ranging from significantly negative to significant positive values. However, the effect captured there was very close to zero. Key results are that companies in the Brazilian energy sector are often already financed favourably, just using slightly too much debt and the average current ratio of around 1.5x is set at a good point. Nevertheless, the model shows high significance overall, coupled with a strong adjusted R-squared of 0.81. Therefore, the model can be used by investors and managers to analyse the energy sector in Brazil.

To limit the objective of the dissertation, it does not aim to provide an exact strategy for implementing these results in optimising companies. This could be part of further research in this field. Other main limitations of the model are especially the limited timeframe that was available to analyse the companies. If the timeframe is increased, effects may significantly differ since practices such as financings of companies change over time, as well as the regulatory environment in which the companies operate.

Moreover, the oil and gas sector probably reacted differently to other shocks in the past, such as the financial crisis in 2008. Besides this, another important limitation to consider is that the model consists of only 36 companies, which could be increased to gain an overview of the broader market in Brazil or narrowed down to analyse specific subsectors. Another important point to mention here is that regression models always use past data as inputs. Therefore, the model estimates effects using past data as input, still leaving questions open regarding the future. This is a common concern, especially in finance, because there is a lot of past data available. Still, many people consider that only future perspectives drive a company's value.

Further research in this field could be relevant to enhance the results via more observations, both added to the past and the future, if available. This would make the model more sophisticated and even provide a more confident estimate for investors. Moreover, this model could be a basis for creating a similar model for other countries in the energy sector, especially in Latin America, due to similar market and geographical characteristics. This would allow us to make a detailed comparison of different fundamental patterns and to show if the results align with the results of the model here.

Concluding the whole dissertation, the model has a good fit and explains changes in market capitalisation with high significance. Also, many of the results are logically explainable and meet expectations while explaining the financing structure of the energy market in Brazil. This is particularly useful because the estimate, in combination with the responding significance level, gives input on how important each ratio is.

REFERENCES

Abuaf, N., 2011. Valuing Emerging Market Equities - The Empirical Evidence. *Journal of Applied Finance,* Volume 2.

Akhtaruzzaman, M. & Sabri Boubaker, M. C. A. Z., 2021. COVID− 19 and oil price risk exposure. *Finance Research Letters,* Volume 42.

Amaro, O. P. J., Jeferson, B. S., Ricardo, G. d. O. & Renato, P. d. Q., 2008. Energy in Brazil: Toward sustainable development?. *Energy Policy,* Volume 36.

Anderson, K. & Brooks, C., 2006. Decomposing the Price-Earnings Ratio. *Journal of Asset Management,* Volume 6.

Beaver, W. H., 1966. Financial Ratios As Predictors of Failure. *Journal of Accounting Research,* Volume 4.

Behera, S., 2020. Does the EVA valuation model explain the market value of equity better under changing required return than constant required return?. *Financial Innovation,* 6(9).

Beneda, N. L., 2003. Estimating free cash flows and valuing a growth company. *Journal of Asset Management,* 4(4).

Birău, F. R., 2012. Econometric approach of heteroskedasticity on financial time series in a general framework. *Economy Series,* Issue 4.

Bourghelle, D., Jawadi, F. & Rozin, P., 2021. Oil price volatility in the context of Covid-19. *International Economics,* Volume 167.

Breusch, T. S. & Pagan, A. R., 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica,* 47(5).

Brooks, C., 2006. Decomposing the Price-Earnings Ratio. *Journal of Asset Management* , Volume 6.

Chan, L. K. & Lakonishok, J., 2004. Value and Growth Investing: Review and Update. *Financial Analysts Journal,* 60(1).

Chan, L. K., Hamao, Y. & Lakonishok, J., 1980. Fundamentals and Stock Returns in Japan. *Working Paper,* Volume 45.

Desai, R., 2021. Nexus Between Debt Financing and Market Value Moderated by Firm Size: Panel Data Evidence from India. *Vision: The Journal of Business Perspective,* Volume 1.

Fahrmeir, L., Kneib, T., Lang, S. & Marx, B. D., 2021. *Regression Models, Methods and Applications.* 2 ed. Berlin: Springer-Verlag.

Fairfield, P. M. & Yohn, T. L., 2001. Using Asset Turnover and Profit Margin to Forecast Changes in Profitability. *Review of Accounting Studies,* Volume 6.

Ferreira, P. et al., 2022. Energy markets e Who are the influencers?. *Energy,* 239(Part A).

Gardner, J. C., McGowan, C. B. & Moeller, S. E., 2012. Valuing Coca-Cola Using The Free Cash Flow To Equity Valuation Model. *Journal of Business & Economics Research,* 10(11).

Godfrey, L. G., 1976. Testing for Serial Autocorrelation in dynamic simultaneous Equation Models. *Econometrica,* 44(5).

Halunga, A. G., Orme, C. D. & Yamagata, T., 2017. A heteroskedasticity robust Breusch–Pagan test for Contemporaneous correlation in dynamic panel data models. *Journal of Econometrics ,* Volume 198.

Harjoto, M. A. & Rossi, F., 2023. Market reaction to the COVID-19 pandemic: evidence from emerging markets. *International Journal of Emerging Markets,* 18(1).

He, Y. & Lin, B., 2018. Time-varying effects of cyclical fluctuations in China's energy industry on the macro economy and carbon emissions. *Energy,* Volume 155.

Hovakimian, A., 2006. Are Observed Capital Structures Determined by Equity Market Timing?. *Journal of Financial and Quantitative Analysis,* 41(1).

Hovakimian, A., Opler, T. & Titman, S., 2006. The Debt-Equity Choice. *Journal of Financial and Quantitative Analysis,* 36(1).

Hyun, J.-Y., Mun, H. H., Kim, T.-H. & Jeong, J., 2010. The effect of a variance shift on the Breusch–Godfrey’s LM test. *Applied Economics Letters,* Volume 17.

Ivanovska, N., Ivanovski, Z. & Narasanov, Z., 2014. Fundamental analysis and discounted free cash flow valuation of stocks at Macedonian Stock Exchange. *UTMS Journal of Economics,* 5(1).

Kato, T., 2012. Prediction in the lognormal regression model with spatial error dependence. *Journal of Housing Economics,* Volume 21.

Li, M. C., Lai, C. C. & Xiao, L., 2022. Did COVID-19 increase equity market risk exposure? Evidence from China, the UK, and the US. *Applied Economics Letters,* 29(9).

Liker, J. K., Augustyniak, S. & Duncan, G. J., 1985. Panel Data and Models of Change: A Comparison of First Difference and Conventional Two-Wave Models. *Social Science Research,* Volume 14.

Lu, X. & White, H., 2014. Robustness checks and robustness tests in applied economics. *Journal of Econometrics,* Volume 178.

Lund, P. D., 2014. How fast can businesses in the new energy sector grow? An analysis of critical factors. *Renewable Energy,* Volume 66.

Magnusson, C. et al., 2005. The language of quarterly reports as an indicator of change in the company's financial status. *Information & Management,* Volume 42.

Manuca, R. & Savit, R., 1996. Stationarity and nonstationarity in time series analysis. *Physica D,* Volume 99.

Melo, M. A., Barrientos, A. & Coelho, A. C., 2014. Taxation, redistribution and the social contract in Brazil. *International Research Initiative on Brazil and Africa (IRIBA).*

Michalak, A., 2016. The Cost of Capital in the Effectiveness Assessment of Financial Management in a Company. *Oeconomia Copernicana,* 7(2).

Mukherjee, S. P., Sinha, B. K. & Chattopadhyay, A. K., 2018. *Statistical Methods in Social Science Research.* Singapore: Springer Nature.

Ogulata, R. T., 2003. Energy sector and wind energy potential in Turkey. *Renewable and Sustainable Energy Reviews,* Volume 7.

Ortas, E., Moneva, J. M. & Salvador, M., 2012. Does socially responsible investment equity indexes in emerging markets pay off? Evidence from Brazil. *Emerging Markets Review,* 13(4).

Ozlen, S., 2014. The Effect of Company Fundamentals on Stock Values. *European Researcher,* 71(3-2).

Paparoditis, E. & Politis, D. N., 2018. The asymptotic size and power of the augmented Dickey–Fuller test for a unit root. *Econometric Reviews,* 37(9).

Pereiro, L. E., 2006. The practice of investment valuation in emerging markets: Evidence from Argentina. *Journal of Multinational Financial Management,* Volume 16.

Queen, P. E., 2015. Enlightened Shareholder Maximization: Is this Strategy Achievable?. *Journal of Business Ethics,* Volume 127.

Said, S. E. & Dickey, D. A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika,* 71(3).

Salmerón, R., García, C. B. & García, J., 2018. Variance Inflation Factor and Condition Number in multiple linear regression. *Journal of Statistical Computation and Simulation,* 88(12).

Sarkar, A., 2020. Market capitalization and growth with nominal and real rigidities: the case of emerging economies. *Indian Economic Review,* Volume 55.

Septyanto, D., Welandasari, N. F. & Sayidah, N., 2022. An Empirical Test of The Financial Ratio Effect on Financial Distress in Indonesia (Study in Garment and Textile Industry). *Economics and Business Quarterly Reviews,* 5(2).

Sinkovics, R. R., Yamin, M., Nadvi, K. & Zhang, Y. Z., 2014. Rising powers from emerging markets—The changing face of international business. *International Business Review,* 23(4).

Smyth, D. J. & Hsing, Y., 1995. In search of an optimal debt ratio for economic growth. *Contemporary Economic Policy,* Volume 13.

Thomas, J. & Zhang, F., 2014. Valuation of tax expense. *Review of Accounting Studies,* Volume 19.

Tolmasquim, M. T., 2012. The energy sector in Brazil: Policy and Perspectives. *Astudos Avançados,* 26(74).

Tomas Cipra, 2020. *Time Series in Economics and Finance.* 1 ed. Prague: Springer Cham.

Uyanto, S. S., 2020. Power Comparisons of Five Most Commonly Used Autocorrelation Tests. *Pakistan Journal of Statistics and Operation Research,* 16(1).

Valenzuela, O. et al., 2021. *Theory and Applications of Time Series Analysis and Forecasting.* Granada: Springer.

Young, S. P. & Jung-Jin, L., 2003. An empirical study on the relevance of applying relative valuation models to investment strategies in the Japanese stock market. *Japan and the World Economy,* 15(3).

Yousef, H. H., 2022. The effect of free cash flow on the firm value in a sample of Iraqi commercial companies listed in the Iraqi Stock Exchange. *American Journal of Business Management, Economics and Banking,* Volume 6.

APPENDICES

Code of the Model

```
# Loading of the needed libraries to run the code
library(tidyverse)
library(caret)
library(car)
library(tseries)
library(lmtest)
library(sandwich)
library(forecast)


# Data obtained according to the Excel (Summed Up and the ratios are multiplied by EBITDA except Market Capitalization
) for the underlying companies. The amount of companies is increasing over time, so the EBITDA multiplication is neutr
alising that.


market_cap <- c(557776.836, 818348.266, 968271.890, 897587.156, 1033850.702, 1035595.832, 1204111.846, 1201901.981, 11
88019.999, 1320443.184, 860472.824, 1137088.119, 1139319.769, 1519729.607, 1640519.963, 1800073.995, 1489994.473, 1502
427.179, 1698965.824, 1440397.165, 1435127.909, 1463922.080, 1347893.930, 1448235.926, 1530155.6)

asset_turnover <- c(68374.941, 100266.871, 103933.452, 118389.908, 134168.799, 148676.257, 145561.068, 145211.775, 145
880.493, 144833.814, 145657.851, 131324.081, 142755.285, 164397.640, 204368.520, 302898.887, 354491.523, 409204.121, 4
38450.694, 440768.829, 457088.142, 443591.654, 421418.412, 357955.824, 322994.0)

debt_ratio <- c(104240.349, 143731.927, 140598.997, 159123.716, 172063.964, 186926.220, 201471.079, 204394.504, 208306
.248, 207078.490, 237562.023, 229778.767, 239207.801, 265402.561, 311474.921, 364991.811, 408044.017, 443997.561, 4277
47.719, 438700.410, 447386.239, 455613.189, 426846.831, 397483.715, 385340.5)

current_ratio <- c(233187.139,  328228.991, 340315.576, 378338.620, 406110.264, 440826.870, 391387.444, 431362.544, 42
3901.191, 416326.707, 477145.084, 507647.181, 494206.686, 541118.997, 723412.220, 947275.054, 934758.713, 1051387.586,
1149975.740, 1073986.651, 1046913.611, 936246.891, 998659.174, 803480.554, 785874.4)

price_earning_ratio <- c(5549819.534, 3594162.492, 4710769.704, 3498187.525, 4147879.734, 2897078.384, 3922098.506, 38
31192.147, 3749732.490, 5725335.483, 6125815.626, 11494982.721, 13649604.715, 6737613.434, 4548318.494, 3584982.307, 2
693884.448, 2896656.475, 2960147.868, 2706777.341, 2659617.041, 2709869.029, 2675857.073, 3041527.226, 3545636.3)

net_profit_margin <- c(3507.678, 11907.755, 10665.378, 15325.764, 16570.350, 26762.416, 24039.164, 26355.541, 28560.42
1, 20516.653, 13214.390, 9320.714, 8730.532, 28964.932, 59439.865, 111356.651, 131788.959, 132575.683, 149099.385, 130
958.587, 131025.197, 127179.285, 112547.740, 94853.979, 83563.0)

book_to_market <- c(142780.399, 181662.244, 158893.028, 194828.480, 188960.586, 207979.260, 188350.539, 209470.771, 22
0577.915, 194759.659, 282721.441, 218292.377, 235872.440, 218080.986, 257128.478, 323605.671, 430071.457, 478112.588,
466295.598, 555565.237, 551761.944, 534152.952, 595512.130, 494840.385, 464259.7)

ltm_cost_of_revenue <- c(450779.747, 614366.320, 631131.393, 650122.799, 684497.656, 706721.534, 713712.498, 734840.05
9, 724990.223, 734728.630, 740093.864, 718715.998, 714246.230, 700653.103, 729492.034, 811517.125, 899621.451, 1003900
.170, 1067778.962, 1129593.217, 1180833.727, 1198223.177, 1202842.869, 1168606.724, 1120805.8)

ltm_da <- c(62751.731, 72848.092, 73404.535, 73955.993, 74363.861, 77947.027, 78450.971, 75432.182, 81095.383, 92036.5
10, 95832.835, 101519.487, 104415.312, 99195.944, 100252.049, 108340.647, 109184.270, 116996.541, 117164.305, 112174.3
21, 115438.849, 133809.565, 141364.917, 142981.364, 143936.0)

property_plant_equipment <- c(784344.799, 1020970.731, 1010923.727, 1058347.713, 1062717.961, 1108000.925, 1187012.170
, 1151098.752, 1167082.191, 1162543.918, 1098326.732, 1110718.513, 1103283.511, 1163163.395, 1193573.812, 1185729.523,
1287464.081, 1317070.721, 1277513.123, 1281621.108, 1296915.660, 1339006.720, 1357737.571, 1383630.111, 1428726.4)

ltm_income_tax <- c(19330.948, 24681.898, 25756.117, 31129.933, 34506.694, 39346.927, 32030.719, 36178.020, 35329.823,
18002.545, -14327.9, -29559.9, -36806.6, -4330.4, 44109.710, 91736.083, 102623.070, 117786.576, 150764.665, 132883.265
, 148792.704, 119591.190, 98837.088, 93471.847, 79528.2)

ltm_ebitda <- c(154496.571, 217789.462, 219568.383, 242117.459, 264641.633, 287131.008, 296264.383, 308141.839, 314431
.685, 314243.745, 323924.246, 319734.873, 339479.934, 382977.116, 455539.232, 573058.780, 626905.982, 676722.575, 70024
0.241, 694618.556, 700853.610, 693244.955, 673147.734, 614076.124, 593884.4)

ltm_capex <- c(-66204.7, -83037.6, -81725.1, -82125.6, -88379.7, -89602.6, -87875.6, -86451.7, -82464.2, -86041.0, -91
276.7, -95462.1, -95136.1, -91165.6, -93048.9, -95364.6, -104973.3, -118187.9, -128949.5, -138131.0, -149391.3, -18742
4.4, -199662.4, -209697.6, 218905.5)

ltm_fcf <- c(73821.693, 99931.729, 103109.829, 113118.145, 112280.466, 128194.592, 130196.442, 133191.326, 155832.707,
162355.990, 176631.353, 187091.581, 208943.762, 225004.455, 253906.607, 310064.112, 342607.791, 352082.563, 340830.327
, 336031.323, 317325.674, 295042.160, 278984.239, 243708.113, 227080.6)

quarters <- c(2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25)

q1 <- c(0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0)

q2 <- c(0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0)

q3 <- c(0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1)

q4 <- c(1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0)


# Creation of a data frame saving all variables and differencing the log transformation before (except for dummy varia
```

```
bles)
data <- data.frame(
  Diff_Log_Market_Capitalization = diff(log(market_cap)),
  Diff_Asset_Turnover = diff(log(asset_turnover)),
  Diff_Debt_Ratio = diff(log(debt_ratio)),
  Diff_Current_Ratio = diff(log(current_ratio)),
  Diff_Price_Earnings_Ratio = diff(log(price_earning_ratio)),
  Diff_Net_Profit_Margin = diff(log(net_profit_margin)),
  Diff_Book_To_Market = diff(log(book_to_market)),
  Diff_LTM_Cost_of_Revenue = diff(ltm_cost_of_revenue),
  Diff_Log_LTM_DA = diff(log(ltm_da)),
  Diff_Property_Plant_Equipment = diff(property_plant_equipment),
  Diff_LTM_Income_Tax = diff(ltm_income_tax),
  Diff_LTM_EBITDA = diff(ltm_ebitda),
  Diff_LTM_CapEx = diff(ltm_capex),
  Diff_LTM_FCF = diff(ltm_fcf),
  # Quarters is not in the data set to remove multicollinearity with q1 to q3
  Q1 = q1,
  Q2 = q2,
  Q3 = q3
  # Q4 is not needed, it is already explained by the other 3 quarters (Q1 = 0, Q2 = 0, Q3 = 0 -> Q4 must be 1)
)

data1 <- data.frame(
  Diff_Log_Market_Capitalization = diff(log(market_cap)),
  Diff_Log_Asset_Turnover = diff(log(asset_turnover)),
  Diff_Log_Debt_Ratio = diff(log(debt_ratio)),
  Diff_Log_Current_Ratio = diff(log(current_ratio)),
  Diff_Log_Price_Earnings_Ratio = diff(log(price_earning_ratio)),
  Diff_Log_Net_Profit_Margin = diff(log(net_profit_margin)),
  Diff_Log_Book_To_Market = diff(log(book_to_market)),
  Diff_LTM_Cost_of_Revenue = diff(ltm_cost_of_revenue),
  Diff_Log_LTM_DA = diff(log(ltm_da)),
  Diff_Property_Plant_Equipment = diff(property_plant_equipment),
  Diff_LTM_Income_Tax = diff(ltm_income_tax),
  Diff_LTM_EBITDA = diff(ltm_ebitda),
  Diff_LTM_CapEx = diff(ltm_capex),
  Diff_LTM_FCF = diff(ltm_fcf)
)

#adding log_diff_market_cap as lag variable with lag = 1 to adress the autocorrelation issue
lag_Diff_Log_Market_Cap = c(0, data$Diff_Log_Market_Capitalization[-length(data$Diff_Log_Market_Capitalization)])
lag_Diff_Log_Market_Cap_2 =c(0,lag_Diff_Log_Market_Cap[-length(lag_Diff_Log_Market_Cap)])

# Calculation of the Variance Inflation Factor (VIF) -> using PCA Model to to high Multicollinearity
vif_results_before <- vif(lm(Diff_Log_Market_Capitalization ~ ., data = data))
print(vif_results_before)

# PCA model to fix multicollinearity issue
pca_model <- prcomp(data1[,-1], scale. = TRUE)
print(summary(pca_model))

# Choosing Principal Components based on variance explained
explained_variance <- summary(pca_model)$importance[2,]
num_components <- max(which(cumsum(explained_variance) <= 0.9))

# Putting Principal Components in a Variable (same name)
principal_components <- pca_model$x[, 1:num_components]


# Dataset using Differenced Log Market Capitalization and the Principal Components as inputs
pc_data <- data.frame(Diff_Log_Market_Capitalization = data1$Diff_Log_Market_Capitalization, principal_components,lag_
Diff_Log_Market_Cap)

# Normal linear model on the Principal Components and Difference Log Market Cap
pc_model <- lm(Diff_Log_Market_Capitalization ~ ., data = pc_data)
summary_pc_model <- summary(pc_model)

#scree plot for the coeffficients
# Assuming pca_model is your PCA result
explained_variance <- summary(pca_model)$importance[2,]

# Number of Principal Components
num_components <- length(explained_variance)

# Create a scree plot
plot(explained_variance * 100, type = 'b', xlab = "Principal Component", ylab = "Variance Explained (%)",
     main = "Scree Plot", pch = 19, col = "blue")
abline(h = cumsum(explained_variance) * 100, col = "red", lty = 2) # Optional: Cumulative variance explained

# Adding more descriptive x-axis labels
axis(1, at = 1:num_components, labels = paste("PC", 1:num_components, sep = ""))
```

```
# Newey-West standard errors (heteroskedasticity and autocorrelation robust SE)
# these are optional but not needed since the model is significant to not have autocorrelation nor heteroskedasticity
nw_se <- coeftest(pc_model, vcov. = NeweyWest(pc_model))

# Showing the robust standard errors
print("Summary with Newey-West SE:")

# VIF test after the PCA transformation to ensure that the multicollinearity issue is fixed
# formal calculation -> the model is adressing exactly this problem
#vif_results_after <- vif(pc_model)
#print(vif_results_after)


# Print the summary of the regression model with principal components
print("VIF values of the Principal Components used for the model")

## [1] "VIF values of the Principal Components used for the model"

print(summary_pc_model)

# Eigenvectors (loadings of the variables) for the in the model used Prinicplal Components
loadings <- pca_model$rotation[, 1:5]
print("Eigenvectors (Loadings of the Variables) for PC1 to PC5:")

## [1] "Eigenvectors (Loadings of the Variables) for PC1 to PC5:"

print(loadings)

# PC1 to PC5 Scores calculation
# center and scale the original data
scaled_data <- scale(data1[,-1]) # Exclude the response variable from scaling
# Multiplying by the loadings of the variables to obtain the score
pc_scores <- scaled_data %*% loadings

# Combination of the original data and the responding Principal Component score
data_with_pcs <- cbind(data1, pc_scores)

# Showing the calculations of the PC1 to PC5 and original data
print("Scores PC1 to PC5 and original data (6 observations)")

print(head(data_with_pcs))

# Calculation for the optimal lag used at the ADF test to test for stationarity
# Calculate BIC for lag lengths up to some maximum
max_lag <- 4  # preset maximum lag (short time series so 4 is too much)

aic_values <- sapply(1:max_lag, function(lag) {
  fit <- arima(market_cap, order = c(lag, 0, 0))
  AIC(fit) # R-function for the Akaike Information Criterion
})

# saving the optimal lag accoriding to the Akaike Information Criterion
optimal_lag <- which.min(aic_values)



# Augmented Dickey-Fuller Test (stationarity)
adf_test_result <- adf.test(data1$Diff_Log_Market_Capitalization, k=optimal_lag)

# Results ADF test
print("Augmented Dickey-Fuller Test Result for the 1 time difference Log(Market Capitalization)")

print(adf_test_result)

# Breusch-Godfrey test (Autocorrelation)
bg_test <- bgtest(pc_model, order = 1)  # 1st order autocorrelation (using quarterly data)

# Breusch-Godfrey test results
print("Breusch-Godfrey Test (Autocorrelation):")

print(bg_test)

# Breusch-Godfrey test (Autocorrelation)
bg_test <- bgtest(pc_model, order = 4)  # up to 4th order autocorrelation (using quarterly data)

# Breusch-Godfrey test results
print("Breusch-Godfrey Test (Autocorrelation):")

print(bg_test)

# Breusch-Pagan test (Heteroskedasticity)
bp_test_result <- bptest(pc_model, ~ principal_components + lag_Diff_Log_Market_Cap, data = pc_data)
print(bp_test_result)
```

List of the companies included in the dataset

INDUSTRY CLASSIFICATION OF COMPANIES LISTED AT B3

SECTORS	SUBSECTORS	SEGMENTS	LISTING CODE	LISTING SEGMENT
	Oil, Gas and Biofuels	**Exploration, Refining and Distribution**	**10**	**7**
		3R PETROLEUM	RRRP	NM
		COSAN	CSAN	NM
		ENAUTA PART	ENAT	NM
		PET MANGUINH	RPMG	
		PETROBRAS	PETR	N2
		PETRORECSA	RECV	NM
		PETRORIO	PRIO	NM
		RAIZEN	RAIZ	N2
		ULTRAPAR	UGPA	NM
		VIBRA	VBBR	NM
		Equipment and Services	**3**	**1**
		LUPATECH	LUPA	NM
		OCEANPACT	OPCT	NM
		OSX BRASIL	OSXB	NM
SECTORS	SUBSECTORS	SEGMENTS	LISTING CODE	LISTING SEGMENT
Basic Materials	**Mining**	**Metalic Minerals**	**7**	**3**
		AURA 360	AURA	DR3
		BRADESPAR	BRAP	N1
		CBA	CBAV	NM
		CSNMINERACAO	CMIN	N2
		LITEL	LTEL	MB
		LITELA	LTLA	MB
		VALE	VALE	NM
	Steel and Metalurgy	**Steel**	**5**	**3**
		FERBASA	FESA	N1
		GERDAU	GGBR	N1
		GERDAU MET	GOAU	N1
		SID NACIONAL	CSNA	
		USIMINAS	USIM	N1
		Iron and Steel Products	**3**	**0**
		MANGELS INDL	MGEL	
		PANATLANTICA	PATI	
		TEKNO	TKNO	
		Copper Products	**1**	**0**
		PARANAPANEMA	PMAM	NM
	Chemicals	**Petrochemicals**	**2**	**1**
		BRASKEM	BRKM	N1
		DEXXOS PAR	DEXP	N1
		Fertilizers	**3**	**0**
		FER HERINGER	FHER	NM
		NUTRIPLANT	NUTR	MA
		VITTIA	VITT	NM
		Chemicals - Others	**2**	**0**
		CRISTAL	CRPG	
		UNIPAR	UNIP	
	Wood and Paper	**Wood**	**2**	**1**
		DEXCO	DXCO	NM
		EUCATEX	EUCA	N1
		Pulp and Paper	**4**	**2**
		KLABIN S/A	KLBN	N2
		MELHOR SP	MSPA	
		SUZANO HOLD	NEMO	
		SUZANO S.A.	SUZB	NM
	Packaging	**Packaging**	**1**	**0**
		IRANI	RANI	NM
	Diversified Materials	**Other Materials**	**1**	**0**
		SANSUY	SNSY	

SECTORS	SUBSECTORS	SEGMENTS	LISTING CODE	LISTING SEGMENT
Utilities	**Electric Utilities**	**Electric Utilities**	**53**	**15**
		AES BRASIL	AESB	NM
		AES SUL	AESL	
		AESOPERACOES	AESO	
		AFLUENTE T	AFLT	
		ALUPAR	ALUP	N2
		AMPLA ENERG	CBEE	
		AUREN	AURE	NM
		CACHOEIRA	CPTE	MB
		CEB	CEBR	
		CEEE-D	CEED	
		CELESC	CLSC	N2
		CELGPAR	GPAR	
		CEMIG	CMIG	N1
		COELBA	CEEB	
		COELCE	COCE	
		COMERC PAR	COMR	
		COPEL	CPLE	N2
		COSERN	CSRN	
		CPFL ENERGIA	CPFE	NM
		CPFL GERACAO	CPFG	
		CPFL PIRATIN	CPFP	
		CPFL RENOVAV	CPRE	
		EBE	EBEN	
		ELEKTRO	EKTR	
		ELETROBRAS	ELET	N1
		ELETROPAR	LIPR	
		EMAE	EMAE	
		ENERGISA	ENGI	N2
		ENERGISA MT	ENMT	
		ENERSUL	ENER	
		ENEVA	ENEV	NM
		ENGIE BRASIL	EGIE	NM
		EQTL PARA	EQPA	
		EQTLMARANHAO	EQMA	MB
		EQUATORIAL	EQTL	NM
		ESCELSA	ESCE	
		FGENERGIA	FGEN	
		GER PARANAP	GEPA	
		LIGHT	LIGH	
		LIGHT S/A	LIGT	NM
		NEOENERGIA	NEOE	NM
		OMEGAENERGIA	MEGA	NM
		PAUL F LUZ	PALF	
		PROMAN	PRMN	MB
		REDE ENERGIA	REDE	
		RENOVA	RNEW	N2
		STATKRAFT	STKF	
		STO ANTONIO	STEN	
		TAESA	TAEE	N2
		TERM. PE III	TEPE	
		TERMOPE	TMPE	
		TRAN PAULIST	TRPL	N1
		UPTICK	UPKP	MB
	Water Utilities	**Water Utilities**	**8**	**3**
		AMBIPAR	AMBP	NM
		CASAN	CASN	
		COPASA	CSMG	NM
		IGUA SA	IGSN	MA
		ORIZON	ORVR	NM
		SABESP	SBSP	NM
		SANEPAR	SAPR	N2
		SANESALTO	SNST	
	Gas Utilities	**Gas Utilities**	**3**	**0**
		CEG	CEGR	
		COMGAS	CGAS	
		COMPASS	PASS	

DISCLAIMER
This industry classification structure is not a recommendation of investment.
The information received from the public companies are available at our website: www.b3.com.br
Contact your Brokerage Firm for further clarification. It can help you to evaluate the potencial risks and benefits related to securities trading.
This is a material of B3 S.A. - Brasil, Bolsa, Balcão

Green	alll data available on Bloomberg
Light green	Bloomberg + other sources + estimations for missing datapoints using avaerage % of Revenue
Pink	insufficient data avaialable

Disclaimer

I disclose that AI tools were employed during the development of this thesis as follows: AI-based research tools were used to assist in the literature review and data collection. AI-powered software was utilized to assist with data analysis and visualisation. Generative AI tools were consulted for brainstorming and outlining purposes. However, all final writing, synthesis, and critical analysis are my own work. Instances where AI contributions were significant are clearly cited and acknowledged.

Nonetheless, I have ensured that the use of AI tools did not compromise the originality and integrity of my work. All sources of information, whether traditional or AI-assisted, have been appropriately cited in accordance with academic standards. The ethical use of AI in research and writing has been a guiding principle throughout the preparation of this thesis.