The macroeconomic determinants of small-cap stock performance in the USA

Master Thesis

Author: Brian Franch Sabaidini
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

Graduation Committee members:
Dr. P. Khrennikova
Prof. Dr. J. Osterrieder

Keywords
Small-capitalization stocks, Regression Analysis, Macroeconomic Factors, Stock Returns, Small-Cap Premium

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my first supervisor, Dr. P. Khrennikova, for her constant support and guidance, which has been of immense value throughout the writing process of this dissertation. Additionally, I would like to extend my sincere thanks to Dr. J. Osterrieder for his exceptional feedback and assistance, which has been truly invaluable to me.

I am also deeply appreciative of all my friends and colleagues who have been with me during my Bachelor's and Master's degrees. Their contributions to this work, through proofreading, feedback, and, importantly, moral support, have been truly significant.

Lastly, but certainly not least, I want to acknowledge the unwavering support of my family, especially my parents, who have supported me and had faith in me throughout this journey, constantly giving me the motivation to complete it.

Abstract

Small-cap equities represent a smaller percentage of the total market capitalization of the USA, nevertheless it is worth researching on this equity class due to its characteristics and higher expected returns. However, higher volatility and sensibility to economic growth might be expected. Therefore, the relationship between macroeconomic factors and the returns of the Russell 2000 are analyzed. The Russell 2000 tracks the performance of the 2000 smallest stocks in the USA and therefore is representative for the US small cap market. An OLS regression will be conducted with logarithmic returns of the Russell 2000 as dependent variables and several macroeconomic variables as independent variables. Macroeconomic variables and return data between 2003 and 2023 will be used for the regression model. Statistically significant results were found for a small part of the total variables included. Positive statistically significant results were found for retail sales at a 5% significance level, and for consumer confidence at the 1% confidence level. Negative relationships were found for industrial production at the 5% confidence level.

Table of Contents

1. Introduction

Small cap equities, also known as small capitalization stocks, usually refer to the shares of companies with market capitalizations ranging from $300 Million to $2 billion (Chopra et al., 2019). These stocks are considered to be more risky but also have the potential for higher returns than their large cap counterparts. Despite their higher risk, small cap equities are an important component of an investment portfolio as they provide diversification benefits. Moreover, these small-cap stocks offer a higher growth potential than many of the bigger capitalization companies. Research has shown that small capitalization stocks offer a higher rate of return when compared to the S&P500 stock index, while also an increased volatility (E. F. Fama & French, 1993).

We often talk about the small cap premium, meaning that the expected return of these small cap stocks are higher, but also have higher systematic risk. On the other hand, this has been put to the test, as researchers also suggest that other factors are responsible for the higher potential return, such as liquidity, information asymmetry or even default risk (Liu, 2006; Vassalou & Xing, 2004; Zhang, 2006).

However, the performance of small cap equities is affected by various factors such as economic conditions, and several other factors. The objective of this study is to investigate the macroeconomic determinants of performance of small cap equities in the USA between the years 2003 and 2023 using a regression analysis utilizing returns from the small-cap index Russel 2000, since research has shown that these companies with smaller capitalization are more sensititve to the economic business cycle.

In this case macroeconomic variables will be included in the research and will shine light into the importance and effect of different factors on the performance of these stocks. The macroeconomic factors that can be taken into consideration include: GDP, CPI, PPI, Industrial Production, Retail Sales, Housing Starts, Consumer Confidence, Interest Rate, Money Supply, Public Debt, Tax Rate, and an energy price index; in line with previous empirical findings showing that these factors were drivers of financial market return dynamics.

The data will be analyzed using multiple regression analysis. The regression model will be estimated using the ordinary least squares (OLS) method. The model will be estimated using the

full sample as well as sub-samples to examine the robustness of the results. This is interesting since the dataset includes observations from 2003 until 2023, with the financial crisis of 2007/08 included in it.

The findings of this study will be useful for investors, portfolio managers, other market participants and even policymakers in understanding the macroeconomic factors and policies that affect the performance of equities and specially small cap equities. This is specially relevant during these times of turbulence in the macroeconomic environment of 2022-2023 where one has experienced some shifts in the policies of central banks and governments. This policies as interest rate hikes and general tightening of the monetary policy comes as a reaction to rising inflation in many countries of the world (Jones, 2023). Also it is interesting to get more insight into the behavior of these small-cap stocks , including data from the recent 2020 Covid crash and also during 2022 where equity markets have been trending down in general, with the S&P 500 index returning -18% the year 2022[1].

In general too much focus is given on the market as a whole and not much in specific categories of equities, such as small cap stocks. Therefore this paper might be of interest to investors and others to further understand the effects of these economic policies and get some insights into the small-cap equity market.

In order to better understand the structure of this research paper, a short overview of its structure will be described. Initially, a well structured overview of past literature will be provided, offering a comprehensive backdrop to the study. This section is aimed at acquainting readers with the existing knowledge surrounding the topic. Following this, the transition will be made into presenting the research questions, the macroeconomic factors being analysed, and outlining the hypotheses driving this investigation. This step is crucial for understanding the specific focus of the study, as well as the rationale behind the research design.

Moving forward, details on the research methodology will be provided, offering transparency into the approach to data collection and analysis. This section serves the purpose of proving the

[1] https://www.forbes.com/sites/stevevernon/2023/01/04/the-sp-500-lost-more-than-18-in-2022-should-retirees-be-worried/

validity of the research process. Subsequently, the core analysis part starts, forming the main part of the paper, where empirical evidence is scrutinized and meaningful conclusions can be drown.

Finally, the paper will be wrapped up with a comprehensive interpretation of the results and a robust discussion of its implications. Here, findings will be contextualized and insights into the scientific contribution will be presented.

2. Literature Review

2.1 Asset Pricing Models (CAPM, 3-Factor, 5-Factor, APT)

The topic of portfolio choice has been a significant area of study in finance and one of the most influential contributions is Markowitz´s paper on Portfolio Theory. This paper suggests that investors are risk averse in nature and seek investments in portfolios with the highest possible return given a specific level of risk. This is the notion of mean-variance efficient portfolios, which are made out of a set of securities that maximize expected return while minimizing the variance of the portfolio (Markowitz, 1952).

2.1.1 CAPM

Later on the Capital Asset Pricing Model (CAPM) was developed in 1964, and explained the relationship between macroeconomic variables and stock performance. Nevertheless, CAPM was only able to explain around 70% of the stock market returns, but CAPM is still widely used to estimate the expected return on an investment/asset and provides a theory for analyzing portfolio returns (Sharpe, 1964).

Banz found that small cap stocks had higher actual returns than those calculated /expected by the CAPM, highlighting the problems with the model (Banz, 1981). Later on, research conducted also demonstrated the poor performance of CAPM. Aktas and McDaniel found that the cost of equity calculated with CAPM was often negative and impossible in reality. They discovered that the results were either negative, under the marginal cost of debt, or even also under the risk-free rate. In total, 925 out of 8361 companies analyzed in the paper had a negative cost of equity when calculated with CAPM, showing that the theoretical mode does not fit the reality often very well (Aktas & McDaniel, 2009).

Despite some unrealistic assumptions of the CAPM, it is still used and taught in finance courses (E. F. Fama & French, 2004). CAPM has been criticized for the assumptions of perfect information and also the assumption that investors are always rational and risk-averse, which in the reality it is not always the case. Nevertheless it remains a widely recognized tool for estimating return.

However, Cremers also argued that the poor performance of the CAPM is due mostly to measurement problems rather than the model itself. In his opinion, the CAPM is still a valid model (Cremers, 2001). These differing views on the validity of CAPM underscore the necessity for continued research and analysis in the finance field in order to better understand it and develop better fitting models.

2.1.2 Fama-French 3 Factor Model/ 5 Factor Model

A further extension of the CAPM model is the Fama-French 3 Factor Model, which was introduced in 1993 and included more factors than the CAPM model. The factors included in this model were market risk, size and book-to-market ratio, which together explain up to 95% of the stock returns in a diversified portfolio. This paper also proves that value stocks outperform large-cap stocks and also small cap stocks outperform large-caps, due to the fact that they are perceived as riskier investments and therefore should offer a higher expected return (E. F. Fama & French, 1993).

A further development of the 3 Factor model was introduced in 2015 by Fama and French. The 5 Factor model included two additional factors, these being profitability and also investment. The model uses the return spread of the most profitable minus the less profitable firms (RMW) and the return spread of companies that invest conservatively minus those that invest aggressively (CMA), in order to explain the returns (E. F. Fama & French, 2015).

There are studies showing conflicting results regarding the effectivity of the 5factor model compared to the older 3 factor model of Fama French. When comparing both models, the 5 factor model ranked higher in explanatory power of anomalies when compared to competing asset pricing models, thus supporting the proposal of the improved Fama French model (Chiah et al., 2016).

According to the Fama-French "small minus big" (SMB) factor, US small-cap stocks have outperformed large-cap stocks by 3.1%, thus suggesting that investors might be able to generate higher returns by investing in these small-cap equities.

Another interesting research focused on the "quality" of stocks has shown that the higher expected returns might be connected to the quality of the companies in which investments are made. This means that "bad quality" stocks are those of companies with bad profitability, poor or negative growth, high risk and also low payouts to investors. Opposite of this would be the "high quality" stocks of companies showing good growth, profitability and adequate returns to investors. Therefore research has shown that investors might achieve higher returns investing the so called "high quality companies" and avoiding junk stocks (Asness et al., 2019).

A short overview on the CAPM formula and the further developments into the Fama and French models can be seen in the following Figure:

Model	Regression equation
Capital Asset Pricing Model	$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + e_{it}$
Fama – French Three-Factor Model	$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + e_{it}$
Fama – French Five-Factor Model	$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}$
Fama – French Four-Factor (combine $\mathbf{R_M} - \mathbf{R_F}$, SMB, and pairs of HML, RMW, and CMA)	(1) $R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + e_{it}$ (2) $R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + c_i CMA_t + e_{it}$ (3) $R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + r_i RMW_t + c_i CMA_t + e_{it}$

Source: 1 (Nguyen, 2018)

2.1.3 Arbitrage Pricing Theory

The arbitrage pricing model was developed by Stephen Ross in 1976 as an alternative to the CAPM which was developed by Sharpe and Lintner, and also tries to describe the relationship between risk and return. Ross states that a linear relationship holds between the expected returns of the assets and their covariance with other random variables (Ross, 1976). For example, the risk factors can include macroeconomic variables such as inflation rates, interest rates and others.

In a market with investors that act in a rational way, there is the "law of one price" which states that rational players will do arbitrage and the asset price goes back into equilibrium, because investors realized the imbalance and implemented arbitrage. Some of the assumptions of this model include: risk-averse investors, efficient markets with not many arbitrage opportunities, risk factors as indicators of undiversifiable systematic risk (D. Basu & Chawla, 2012).

The formula for the APT of Ross could be described as following:

$$R_i = a_i + \beta_i 1 * F_1 + \beta_i 2 * F_2 + \cdots + \beta_n * F_n + \varepsilon_i$$

where:

R_i: Expected return asset i

β: Sensitivity to systematic factor

F: Systematic factor (such as macroeconomic variable)

ε: Error term (unsystematic risk)

Arbitrage pricing theory is another asset pricing model showing the factors that influence stock returns. Such factors include industrial production, risk premium, yield curves, inflation and other macroeconomic variables (Chen et al., 1986). The Arbitrage Pricing Theory is an alternative to the CAPM which also tries to explain the relationship between expected returns and risks.

In conclusion, the literature on small-cap stocks has evolved significantly over time. Markowitz's portfolio theory and the CAPM were early contributions to the field, but the Fama-French Three Factor Model and Five Factor Model have become widely used frameworks for analyzing stock returns, particularly for small-cap stocks and growth companies. While small-cap stocks may have higher perceived risk and liquidity constraints, they have consistently outperformed large-cap stocks and are likely to continue to do so in the future. Researchers have also highlighted the importance of quality companies for generating higher returns.

2.2 Determinants of stock returns

The determinants of stock returns have been a subject for research for a long time. Several studies have explored the factors that affect the returns of stocks and also they way these factors interact with each other. This literature review summarizes some of the most notable findings in

this field, with a focus on studies that used regression analysis to identify the determinants of stock prices.

Enow and Brijlal examined the determinants of share prices of companies listed in the Johannesburg stock exchange. The study shows that the factors influencing prices were mostly the dividend per share, earnings per share, and also the PE ratio. These factors together accounted for 57.8% of the price movements. This somewhat shows the importance of these financial ratios and KPIs (ENOW & BRIJLAL, 2016).

Another study focusing on the Dhaka stock exchange, one of the two stock exchanges of Bangladesh, used a linear regression model to show the influence of several variables such as liquidity, leverage, profitability, growth, size and dividend rate (Mondal & Imran, 2010). Also some qualitative factors were found that affected share prices, including company goodwill, market sentiment, company announcements and also even political turmoil. The quantitative factors that had an influence on the returns were: dividend, market capital, PE ratio, EPS, Net income, ROI, earnings, demand and supply of the stock, inflation, interest rates and exchange rates (Mondal & Imran, 2010). This shows some resemblance with the earlier research by Enow and Brijlal, who also found dividends, EPS and PE ratio to be relevant factors influencing stock returns (ENOW & BRIJLAL, 2016).

When talking about the importance of the price-earnings ratio (P/E), Basu has shown that in functioning capital markets, stock prices should fully reflect the information in a fast and unbiased way and thus provide estimates of the intrinsic value of the companies. This research of 1977 shows that low P/E firms had higher returns than those with high P/E ratios, being in line with the results of Enow & Brijlal and also Mondal & Imran (S. Basu, 1977).

It is also interesting to research on the effects of foreign exchange rate and interest rate risk on the value of stocks. A study of 2006 suggests that these variables significantly affect the value of common stocks and highlights the importance of monitoring global economic indicators if trying to anticipate changes in stock prices (Lael Joseph & Vezos, 2006). Nevertheless it is important to remember that predicting shifts in the macroeconomic environment is hard and it might be impossible for individuals to accurately predict these changes.

Research on Japanese companies also showed a significant relationship between earnings yield, size, book-to-market ratio and also the cash flow yield (Chan et al., 1991).

It is also worth mentioning that using historical betas estimates from annual returns, rather than monthly returns how it is normally done, yielded better results and produces a stronger relation between the returns and beta (Kothari et al., 1995). Some years later Kothari and Shanken argued that Fama and French tended to ignore positive evidence on historical betas and overemphasize the importance of the price-to-book ratio of companies (Kothari & Shanken, 1997). Another important aspect is the fact that survivorship bias might exaggerate the relationship between returns and P/B ratio (Kothari et al., 1995).

It is also argued that the size and PB effects are only a result of the overreaction of investors to corporate news, either expecting extremely high or low growth in the future (Lakonishok et al., 1994). This leads to the underpricing of "value", and overpricing of "growth". These terms can be described with their capitalization and PB ratio, "value" meaning companies with smaller market capitalization and high PB ratio and "growth" as companies with bigger market capitalization and low PB ratios (Lakonishok et al., 1994).

Pettengil suggests in the same year that the relationship between the beta coefficient and the stock performance is only significant when the analyzed sample is dividend into bull and bear markets (Pettengill et al., 1995).

Overall, these studies show that the relationship between market factors and asset risk-return relationships are complex and multifaceted and require careful consideration of many factors to provide complete understanding of this topic.

2.3 Background on regression for stock prediction

Regression models are a common tool for the investigation of the relationship between stocks and several macroeconomic variables, with many studies using regression analysis and other similar tools (Asprem, 1989). When doing this it is important to remember that the introduction of too many variables which are as well too closely related to each other might harm the explanatory power of the model and the effects of each variables will not be clear (Gujarati & Porter, 2009).

Other possibilities apart from the OLS model are Vector Autoregressive models and also GARCH. These models are also very useful to analyze the relationship of stock performance and several variables. Nevertheless in this paper, the analysis will be carried out with a regular OLS regression model, similar to the one by (Özdemir & Yeşilyurt, 2013).

2.4 Background and empirical evidence on small-cap stock performance

Small cap stocks are usually described as publicly traded companies with a market capitalization between $300 million and $2 billion. Nevertheless it is worth mentioning that most small cap indexes do not only contain stocks in that range of market capitalization. Most small-cap indexes and the ETFs representing those indexes also include stocks with higher market capitalization than the $2 billion.

Examples of that would be the CRSP US Small Cap Index and also the Russel 2000, the latter is going to be used in this analysis. The CRSP US Small Cap Index tracks the bottom 2 – 15% of the investable market capitalization of the US market, and its biggest constituent as of March 2023 is Idex Corporation with a Mkt. Cap of $17.4B. Regarding the Russel 2000 index, the biggest constituent as of May 2023 is Super Micro Computer Inc with a market capitalization of $ 13.3B as it can be found in the constituent list of the index provider for the Russell.2000

Small cap stocks have received attention in academic literature due to their characteristics and potential for higher returns. Research has shown that large firms tend to react faster to new information compared to smaller companies, suggesting that small-cap stocks may be more information opaque than large-cap stocks which can create opportunities for investors wanting to generate excess returns (Baharumshah et al., 2009).

Nevertheless there are also challenges associated with the investment in small-cap equities, involving lack of coverage and also liquidity. The excess returns mentioned above which are also a result of the inefficiency of the market and also the companies themselves might not even be realizable to an investor because of liquidity constraints. Research shows that companies in the cheapest decile in regards to P/B ratio tended to be stocks considered illiquid (Stanhope & Meredith, 2015).

[2] https://research.ftserussell.com/analytics/factsheets/Home/DownloadConstituentsWeights/?indexdetails=US2000

Another interesting factor could be the geographical location/country of the equities, being a possible determinant of performance for these small-cap equities, as proposed by (Switzer & Fan, 2007).

These above mentioned inefficiencies and lack of coverage are factors that allow investors to possibly generate a higher alpha when investing in more risky, small capitalization stocks. This implies that the lack of information and research on small-cap stocks makes them riskier investments but also provides an opportunity for higher returns (Moreolo, 2019).

Pandey and Sehgal identified several factors which caused higher risk stocks, supporting the evidence that the company size was a influencing factor. They found that small-cap stocks tend to have much greater volatility than large-cap stocks due to their higher exposure to market risk factors (Pandey & Sehgal, 2016). Therefore the expectancy of higher returns as additional risk premium required for bearing the volatility of the small-cap stocks

Another interesting factor might be the economic conditions, since small-cap stocks tend to outperform large-cap stocks during economic downturns and underperform large-cap stocks during economic expansion times. This shows that macroeconomic factors influence the returns on these stocks and highlights the importance of understanding economic cycles when focusing on small-caps (Norland, 2020). Nevertheless, an outperformance during economic downturns doesn't mean that the small-cap equities are the most interesting asset to invest in during these hard economic times necessarily, but just show a comparative outperformance when comparing to large-cap stocks. Also it is worth mentioning that the volatility in these small-cap stocks is higher, and may lead to higher returns for those periods where the market in general experiences higher volatility, as in these economic downturns. As shown by Norland, the Russell 2000 Index outperformed the S&P500 during the period 1979-82 by 76%, during times of high inflation and interest rate hikes, similar to the situation in 2023. Also in the period of 1990-93 the Russell outperformed the S&P500 by 48%, during the recession caused by the savings and loan crisis and the oil price shock (Norland, 2020).

Another possibility of the higher possible return is the fact that small cap stocks might be more often mispriced by investors. A possible reason for this might be the behavioral biases of

investors, leading them to often under valuate these small companies as a result of their biases (Lakonishok et al., 1994).

Liquidity constraints are also an important consideration when investing in small-cap stocks. These stocks have lower trading volumes and in general are much smaller in market capitalization, which can make it difficult to buy or sell stocks when desired. This applies mostly to institutional investors and bigger orders but can affect also small individual investors that have some equity of these small companies. Despite the potential liquidity constraints, the small cap premium has become a staple in corporate finance theory and the potential for outperformance of large-cap stocks over time has been well documented, with its origins in the research by Banz, where he found that the higher adjusted risk of small capitalization stocks led to a higher expected returns. Nevertheless, it was found that the effect described in the research was stronger for the smallest stocks in the sample (Banz, 1981).

Probably the most well-known small-cap premium research was conducted is the 3 Factor model of Fama and French, with one of the three factors being Size (E. F. Fama & French, 1993).

Nevertheless, some researchers have argued that the Size premium has disappeared and that it does not seem to exist anymore when looking at more recent times. Researchers have explained alternative reasons for the perceived size premium during the 70s and 80s, such as an overvaluation of these small caps due to mispricing and also a general lack of efficiency regarding those smaller, more information opaque stocks (Campbell, 2000; Daniel et al., 2001).

3. Research Questions

After gathering some insights of existing literature on the topic this paper is focused on, the research question can be stated. In this case, the research question is quite broad, giving some freedom in terms of research design. Nevertheless, some more specifically formulated hypotheses will also be included.

The main research question that is aimed to be answered in this research paper is:

- **What is the relationship between macroeconomic factors and small cap equity returns?**

With this project, macroeconomic variables will be analyzed in order to find out which ones have significant effects on the returns of small cap stocks. This being said, several main hypotheses will be tested but the analysis as a whole will be conducted with more variables as well, which will be described shortly in the next section.

3.1 Macroeconomic Factors:

Table 1 shows the different independent variables that are going to be included in this research. The first category is the general macroeconomic variables such as GDP or also the Inflation Rate, which are variables that give insights into the development of the economy as a whole and might be able to explain the returns of the stocks. Other variables that will be used include also monetary and fiscal policy, which give information about topics such as Interest Rates or also Government spending. As an addition to these macroeconomic factors, the Energy Price Index will be included since energy prices might have a big influence on the performance of stocks and also of the underlying companies. The variables included in this paper are based on previous research and include macroeconomic factors which are very relevant to most companies.

Table 1: Macroeconomic variables included in this research

Category	Variable Name	Description
Macro	GDP	Gross Domestic Product
	CPI	Consumer Price Index
	PPI	Producer Price Index
	Industrial Production	Industrial Production Index
	Retail Sales	US Retail Sales
	Housing Starts	New Housing Starts (construction)
	CCI	Consumer Confidence Index
Monetary Policy	Interest Rate	3M T-bill rate (USA)
	Money Supply	M2 Money Supply
	Public Debt	Public Debt
	Tax Rate	Corporate Tax Rate
Energy/Commodities	Energy Index	Energy Price Index

3.2 Hypothesis

Table 2: Relationship between IVs and Small-Cap Stock Returns

Category	Variable Name	Relationship with Small-Cap stock returns
Macro	GDP	+
	CPI	-
	PPI	-
	Industrial Production	+
	Retail Sales	+
	Housing Starts	+
	CCI	+
Monetary Policy	Interest Rate	-
	Money Supply	+
Fiscal Policy	Public Debt	-
	Tax Rate	-
Energy Prices	Energy Index	-

Table 2 shows the variables and the expected relationship between the Small-cap returns and the independent variables. The expected relationship between the variables and small-cap stock returns are based on empirical findings. Further information on previous findings can be found in Part 3.3.

Nevertheless, the main hypotheses of this research paper are focused specifically on the variables: GDP, CPI, Interest Rate, and the Energy Price Index.

- **H1**: There is a positive relationship between the variable "GDP" and the small-cap stock returns
- **H2**: There is a negative relationship between the variable "Consumer Price Index (CPI)" and the small-cap stock returns
- **H3**: There is a negative relationship between the variable "Interest Rate" and the small-cap stock returns

- **H4**: There is a negative relationship between the variable "Energy Index" and the small-cap stock returns

3.3 Motivation for Hypotheses

The following section gives some insight into the reasoning behind the hypothesis, showing previous research and findings regarding each of the variables that are tested in this paper.

3.3.1 GDP

Regarding the relationship between GDP and the stock market return there doesn't seem to be clear consensus in existing literature. One could easily think that the relationship between these two must be of course positive, since that seems to be quite reasonable. Nevertheless, research has shown that it does not seem to be positively correlated, even possibly having some negative relationship. It is indeed worth mentioning that GDP growth has different effects on different industries and company specific factors, thus being hard to generalize the effect of economic growth on companies.

In the earlier studies, a positive relationship between the actual economic growth and the stock market returns was found. In theory the stock returns of a company should reflect the discounted earnings expectations of the company, which in turn are dependent or at least affected by the general economic environment and growth (Shapiro, 1988). A strong positive link between stock prices and economic activity was found in 1984 when analyzing the US economy (Fischer & Merton, 1984).

On the other hand, there have been researchers arguing for a decoupling of the stock market returns from real economic activity. Binswanger found that the evidence of a relationship between these two variables was diminishing and no clear indication for significant correlation was found after the 1980s (Binswanger, 2000). Some years after, a further study focused on the relation between stocks and GDP for the G7 countries found no evidence of a long-term stable relationship. Only for some subperiods before the 1980s a relationship could be found (Binswanger, 2004).

A similar study was conducted in recent years, also focusing on the G7 countries. Here little evidence for a long-term relationship was found, with results for Germany, Italy, Japan, UK and the US being not significant. Only for the countries of Canada and France, Fichtner et al. were

able to find some cointegration relation both for the total observation period, as well as for subperiods (Fichtner & Joebges, 2024).

A study showing that the general assumption of a positive relationship might be wrong is the paper of J. Ritter, who analyzed worldwide data from 1900 util 2002. A negative correlation was found between the per capita GDP growth the real equity returns. It is stated that the increase in GDP is caused by either an increased production or increased inputs, but that the economic growth ultimately has a positive effect on consumers, without translating to higher PV of dividends or earnings for the companies. Additionally, the author asserts that countries exhibiting high economic growth may not necessarily offer favorable investment opportunities in equities, unless the valuations are deemed attractive. (Ritter, 2005).

3.3.2 CPI

CPI stands for the Consumer Price Index, which is the measure used to track the increases in price levels of consumer goods and services. This index includes prices of many different categories of goods such as energy, food and other.

There is a general belief amongst economists and researchers that the stock market had the potential to act as an inflation hedge, but many other famous papers on the topic found the opposite relationship. Bodie found that the real return on equity was actually negatively related to inflation, which is in line with Eugene Fama´s paper where he found that the negative relationship was consistent for both expected and unexpected inflation (Bodie, 1976; E. D. Fama, 1981).

Some years later, a similar research was conducted with alternative results when compared to the paper of Fama. Titman & Warga found a positive correlation between future inflation and stock returns and even argue that it might be possible to use these stock prices to improve inflation prediction models (Titman & Warga, 1989). Nevertheless, they found that the use of REITS for inflation forecasting was more accurate than with the broad market since this type of companies are more sensitive to inflation and interest rate changes (Titman & Warga, 1989).

An alternative view is that of Boudoukh et al., which criticized the short term emphasis of previous research, all focusing mostly on short time frames. They stated that since investors usually hold stocks for very long periods, it was even more interesting to see if the relationship

between stock returns and inflation held as well in the long-term. Using different model designs, similar results were found using ex ante inflation as well as ex post inflation, coming to the conclusion that there was enough evidence to determine that on the long-term, stocks and inflation were positively correlated (Boudoukh & Richardson, 1993). According to Boudoukh, long-term stock holding involves stock holding periods of 5 years or more. It is assumed that this timeframe is long enough for irregularities and possible inefficiencies to become less relevant to the returns. In the real world, it can also be argued that investments with holding periods shorter than 5 years are often of a speculative nature. It is common for perceived market inefficiencies or opportunities to remain unresolved for longer periods, even if one has correctly identified them. Ex ante rate is basically the difference between the nominal inflation rate and the expected inflation, whereas ex post rate is the difference between the nominal interest rate and the actual inflation (Clements, 2014).

An very interesting alternative approach is to analyze the relationship between stock returns and inflation in countries suffering from high inflation rates for long periods of time. In these countries, the population needs to seek for inflation hedges actively, in order to protect their purchasing power. A study focusing on Argentina, Chile, Mexico and Venezuela found that there might be a positive relationship on the long term in some cases. The results held for the periods of the 1980s and also 1990s, and in two of the four countries, Argentina and Chile, stocks acted as perfect hedges against inflation (Choudhry, 2001). Nevertheless, it would be interesting to see if that relationship still holds in these countries with a longer time frame.

All in all, one can see that there is no general consensus on the inflation hedging properties of the stock market. Nevertheless, most research cannot prove any positive relationship between the two variables, with some exceptions, such as periods where they acted as such, or also in countries with prolongued inflation. Moreover, the examination of the correlation between small cap stocks, which are characterized by their high volatility and relatively low correlation with the broader market during specific time periods, presents an intriguing area of study.

3.3.3 Interest Rates

Interest rates and interest rate changes are said to have a negative relationship with stock returns. In general there is consensus on the relationship of interest rates and stock returns, nevertheless

this is challenged by some alternative findings stating that there is an asymmetrical relationship, with mainly falling interest rates having an effect on stock returns (Domian, 1996).

Their research paper first released in 1996 analyzed the relationship between stock returns and interest rates in the USA, including data for the time frame between 1952 and 1992 from CRSP and using the T-bill data for the interest rate information. This analysis comes to the conclusion that there is an asymmetrical relationship between the variables, meaning that drops in the interest rates are followed by a 12-month excess returns in stocks; whereas increases in interest rates have little effects on the stock market (Domian, 1996).

On the other hand, research has shown that interest rates have significant effects on stock returns, specially in developed economies such as the USA or the Eurozone. Using quarterly data from 1999 up to 2013 Asseffa et al. found a significant relationship between the two variables. This paper takes data from all around the world, including interest rates and stock returns from 40 countries. When it comes to developing countries, this stated relationship between the interest rates and stock returns becomes insignificant, thus suggesting only a significant effect of benchmark rates on the stock returns of developed countries (Assefa et al., 2017).

A somewhat more vague relationship is found by Martinez-Moya, with the interest rate sensititivy varying across industries and depending on the time horizon considerated. Although this paper was focused on the Spanish stock market, the findings might be transferable to other developed economies (Martínez-Moya et al., 2013). Similar conclusions are found as well in another paper focused in the USA, where the relation between interest rates, exchange rates and stock prices is analyzed. The results show that the relationship between interest rate and the US stock index is significantly different from zero. Nevertheless, this only is true for longer time horizons, also in line with the results of Martinez-Moya (Hamrita & Trifi, 2011). Therefore one might state theat there is evidence about interest rates leading stock returns in the USA.

3.3.4 Government Spending

When it comes to the predictive power of government spending on stock returns there is no clear consensus. Nevertheless, most studies make a differentiation between the exposure of each company to government spending, since not every company profits the same from an increased spending. A instance highlighting the contrast in investment amounts received by companies can

be observed presently, particularly with the substantial investments directed towards big tech and net-zero initiatives. In a paper analyzing the effect of taxes and government spending on US stocks and bonds, Tavares and Valkanov came to the conclusion that an increase in government spending did have a positive impact on the expected returns of both of these assets. The results are only statistically significant for the asset class of bodns and also only for those of short duration (Tavares & Valkanov, 2003).

As mentioned above, an important implication is the sensititivy of each firm to changes in spending. This was also addressed in another paper, where government spending but also political cycles were analyzed as predictors for stock returns (Belo et al., 2013). All in all it was found that the president partisan cycle had an effect on stock returns and profitability, through the changes in government spending. This means that not only the exposure of each firm was important, but in the case of the USA, the political party and changes between Republican and Democratic governments had an effect on stocks. Nevertheless it was recommended to further investigate this relationship, since it was not clear if the differences in returns came as a result of a mispricing of the stock or a compensation for the government exposure risk of each firm (Belo et al., 2013).

When talking about exposure to government spending in the USA, one can argue that the military sector is one of a very high sensitivity. Therefore, many studies have tried to analyze the relationship between government spending, military spending and also the returns of military-related stocks in the USA. Fisher and Peter tried using stock returns to identify government spending shocks, supposing that companies are forward looking and include expectations of future sales in their planning. Moreover, the USA spending for military constitutes a big part of the total government spending, while this military spending is also distributed amongst a not very big amount of firms (M. J. Fisher & Marshall, 2009). They were also not the only ones focusing their research on the military sector, thus being also several more studies on this. For example in 2009, Berndt developed a “Fama-French Guns” portfolio to try forecasting government spending (Berndt et al., 2009). As stated before, the distribution of this big portion of government spending goes in majority to a very small numbers of firms, which directly profit from the increase in military spending.

3.3.5 Engergy Index

Energy prices are known to have an effect on the general economy and also the profitability of companies, specially in energy intense industries, but also food and energy sectors (Wattanatorn & Kanchanapoom, 2012). However, there is no clear consensus on the effect of the energy prices on stock returns, with many different outcomes for different commodities and countries.
A recent study found that energy prices were not robust predictors for stock returns for the period between 1987 and 2015, whereas until 2008 crude, gas and coal might have had a moderately good predictor effect on stock prices (Kim et al., 2019).

When focusing on the Eurozone, research has shown that the crude oil price did have a negative effect on European utility stocks, while gas did not have any significant effect on these. Oberndorfer also found that coal had a minimal effect compared to oil, even if electricity constitutes a big part of the energy usage in the Eurozone. All in all they found that crude oil was the strongest indicator for energy prices, even if it was not used mainly for electricity production, but rather for production plants and transporting goods (Oberndorfer, 2009).

A similar study done by Acaravci et al. focused purely on the relationship between natural gas prices and stock prices in the EU-15 countries during the period of 1990 until 2008. A long term equilibrium relationship was found in countries such as Austria, Denmark, Finland, Germany and Luxembourg. In the rest of the countries no clear relationship was found. Nevertheless, the results are not very conclusive, since it seems like the relationship is not very direct, possibly indicating that industrial production serves as a mediator between energy prices and stock value. Therefore one could argue that the direct relationship is more that of the industrial production with the stock prices in the above mentioned countries (Acaravci et al., 2012).

Another relevant finding is the fact that the relationship between oil price and stock returns is dependent on the cause of the oil price shock as well. Kilian and Park found that depending on whether the oil price changes were caused by a supply or demand shock, the relationship between the variables changed (Kilian & Park, 2009).

All in all, we have seen that there is often not complete consensus on many of these variables. In most of the cases, there are papers with small differences in assumptions, time frames, and also geographic focus, thus yielding often very different results depending on these assumptions. It is

therefore interesting to test these hypotheses in order to get some more insight into the relationship between the variables included in this paper and the returns of the Russell 2000 Index.

4. Research Method

The method implemented for this paper consists of a multi-factor regression model, since several of the macroeconomic variables will be tested for correlation with the dependent variable, which is the logarithmic return of the Russell2000 index. However, before proceeding with the regression model, it is important to have a clear understanding of the data being analyzed. For this reason, the first step in the research methodology will be to conduct a descriptive analysis of the data. Descriptive analysis involves the exploration and summary of the data to gain insights into the distribution, central tendency, variability and other characteristics. By doing this first, one can ensure that the data is adequate for the later analysis and also one is then able to make informed decisions about changes, adaptations and possible problems in the data sets. This involves the preparation of the data, which will be checked for errors and missing values. After that basic statistics will be presented, such as mean, median, variance and also standard deviations. Some information can be also displayed in a graph in order to have a visual representation of the relationships and characteristics of the data being used, thus creating a better understanding. One example are histograms which shows the number of cases for each of the possible outcome, making it easier to quickly assess if the outcomes look normally distributed (M. J. Fisher & Marshall, 2009).

After the descriptive statistics have been done and the data set is ready to be used, a multi-factor Regression model will be implemented for assessing the relationship between stock performance and the different independent variables. This regression will be run on the program SPSS. A dummy variable for economic crisis will also be included. As a result, the accuracy of estimates of the regression analysis might be improved as the control variable should help isolate the effect of the macroeconomic factors on the small-cap stocks.

In the book on the basics of econometrics by Damodar and Porter the basic assumptions of such linear regression models are described (Gujarati & Porter, 2009). The first assumption for linear regression is the fact that there needs to be a linear relationship between the dependent and

independent variable being analyzed. Secondly, the errors must be normally distributed, which can be tested with a graphical demonstration with a histogram or Q-Q Plot or also in an analytical way with the Shapiro-Wilk test or Kolmogorov test. The third main assumption is the fact that the residuals need to have a constant variance, which we refer to as homoskedasticity. If this was not the case, the residuals would be heteroskedastic and the assumption would not be fulfilled. The heteroskedasticity assumption may be tested with the Breusch-Pagan test (Özdemir & Yeşilyurt, 2013). Very important as well is the assumption of multicollinearity, which will be also relevant in this paper, since some of the variables used might be very closely related to each other. Multicollinearity of variables is when two independent variables in regression are too closely correlated and the measurement of the influence of each variable will be harmed (Gujarati & Porter, 2009).

For analyzing the effect of the macro variables, a small cap index will be used, representing the market for small-cap stocks the USA. The most used index is the Russell 2000 and tracks the performance of the smallest 2000 companies listed in the USA.

After the regression model is run, the results will be examined and interpreted. This means that the coefficients and also the p-values for each independent variable will be taken into account. Moreover it might make sense to separate the sample into different parts, such as bull and bear markets. This allows to better understand how these macroeconomic factors affect the small-cap stocks in the different market conditions as stated in prior literature (Pettengill et al., 1995). Therefore it will be tested as a whole and also separated into pre 2007-2008 financial crisis and after, in order to see how the results differ from each other just because of the separation.

Also the face validity will be interpreted, if the signs and also magnitudes of coefficient make common and economic sense. Also statistical validity of the model has to be tested, for example diagnosing the model for heteroskedasticity and also testing the assumption of multicollinearity. In order for the assumption of heteroskedasticity to be tested, either the White´s test or the Breusch-Pagan test will be used. This will show whether the variance in the errors in the regression are constant (Brooks, 2019). Testing multicollinearity is testing the variables to see whether two variables are too closely linearly related and thus one can be also predicted by the other and must then not be included into the regression model. Two possibilities of doing this are by creating a correlation matrix in SPSS and looking at the correlation values between the

independent variables, and also by looking at the variance inflation factors (VIF) values. These values should be under 10, since values exceeding 10 are considered as indicating multicollinearity (Senaviratna & A. Cooray, 2019).

4.1 Regression model

The model is a regression analysis that seeks to analyze the effect of several independent variables on the stock prices of small-cap stocks. The equation takes form of linear regression model, with the dependent variable R, being Market returns of these small cap indexes. Therefore for each of the analyzed small-cap indexes one regression analysis is done.

Equation 1:

$$R = \alpha + \beta_1 * InterestR + \beta_2 * MoneyS + \beta_3 * RetailS + \beta_4 * IndProduction + \beta_5 * PPI + \beta_6 * HousingS + \beta_7 * ConsumerConf + \beta_8 * EnergyPrice + \beta_9 * GDPgrowth + \beta_{10} * CPI + \beta_{11} * CrisisDummy + e$$

Where:

R : Small-Cap Index Log Returns (Russell 2000)

α: Constant

e: error term

The independent variables on the right include the constant term α, the different coefficients β1, β2,...,βn, representing the change in dependent variable for a one-unit increase in corresponding independent variables when holding all other constant. The error term ε captures the variability in the dependent variable that cant be explained by the independent variables included in the research. Overall this model will be used to test the relationship between the independent variables and market returns. The study will also conduct diagnostic tests to check for the assumptions of the model such as normality, heteroskedasticity and autocorrelation, in order to sort out variables that are too closely correlated or don't have explanatory power.

5. Data

This data is mostly going to be retrieved database of the St. Louis Federal Reserve "FRED". The exceptions will be the return of the Russell 2000 Index, which will be retrieved from Yahoo Finance and also the data for Retail sales which are going to be retrieved from the U.S. Census database. The time frame is 20 years and will consist of monthly observations between 2003 and 2023. This is interesting due to the fact that during these times the pandemic happened in 2020 which includes very interesting data points for this analysis, as the situation was new and unexpected. This time frame would allow for 240 monthly observations which should be a big enough number for the regression analysis. Research suggest that the minimum number of observations needed for a regression analysis should be: $N > 50+ 8*m$; m being the number of independent variables included in the regression analysis (Green, 1991). In the case of this research the minimum number would be $50+ 8*14 = 162$, thus 240 monthly observations being an adequate sample size.

It is worth noting that the variables are going to be transformed into returns/ % change for this regression model. This is due to the fact that the focus in this paper is to see how variations in these variables affect the stock returns. Moreover, it is more interesting to note the effect of the change in a said variable, rather than to see what the absolute level of this variable predicts. For example, we are interested in how a 1% increase in Energy prices have an effect on Small-cap stock returns.

The country in which the research is going to be focused on is the USA as this area has very developed capital markets and a big quantity of publicly listed companies. Other areas might be of interest too, which is an possible future expansion of this research.

5.1 Data Sources

Table 3: Data sources for the variables included

Variable	Description	Source
Small-cap Index	Russell 2000 Index (Small-cap Index)	https://finance.yahoo.com
GDP Growth	GDP Growth (Monthly, Seas. Adj.)	https://fred.stlouisfed.org/
Interest Rate	Effective FED Rate (FRED)	https://fred.stlouisfed.org/
Retail Sales	Monthly Sales For Retail and Food – Seas. Adj.	https://census.gov/data
Industrial Prod.	Industrial Production Total Index	https://fred.stlouisfed.org/
PPI	Producer Price Index: PPI all Commod.	https://fred.stlouisfed.org/
CPI All Urban	CPI All Urban Consumers: All Items	https://fred.stlouisfed.org/
Housing Starts	New Privately Owned Housing Units Started	https://fred.stlouisfed.org/
CCI	Consumer Confidence Index USA	https://fred.stlouisfed.org/
Public Debt	US Public Debt	https://fred.stlouisfed.org/
Corporate Tax	US Corporate Tax	https://tradingeconomics.com/united-states
Crisis Dummy	NBER based Recession Indicator USA	https://fred.stlouisfed.org/
Energy Price Index	CPI All Urban: Energy (Seas. Adjusted)	https://fred.stlouisfed.org/

Table 3 displays short descriptions and also the sources for all the included variables. Most data for the independent variables used will be taken from the official database of the St. Louis Federal Reserve Bank, with exception of the retail sales and corporate tax data. The return data for the Russell 2000 Index will be downloaded from Yahoo Finance.

5.2 Stock Returns

For the representation of the small cap stocks in the USA different indexes can be utilized. One possibility is the Russell 2000, which is composed of the 2000 smallest companies in the Russell 3000 index. These 2000 stocks represent around 10% of the market capitalization of the equities in the USA and can also be considered as a representative proxy for the small cap companies in the USA.

Until 2007, the construction of the Russell 1000 and 2000 index functioned as following: the 1000 largest stocks by market cap in the USA were included in the Russell 1000 and the next 2000 with the biggest market cap were included in the Russell 2000. Since 2007, the methodology was changed and three factors were used in order to determine which stocks will be included in the 2000 Index. Nevertheless the basic method is similar, with the Russell 2000 including the smaller capitalization stocks found in the USA (Appel et al., 2018).

The following graph shows the percentage return of the Russell 2000 index for the exact time frame of the analysis, also showing the S&P500 Index as a benchmark in red color. In the last 20 years we have not seen a too clear outperformance of the Russell 2000, with the index falling in value more during crashes and performing better coming out of them. This leads to both indexes to have a very similar performance at the moment, after having a more clear outperformance in the last year.

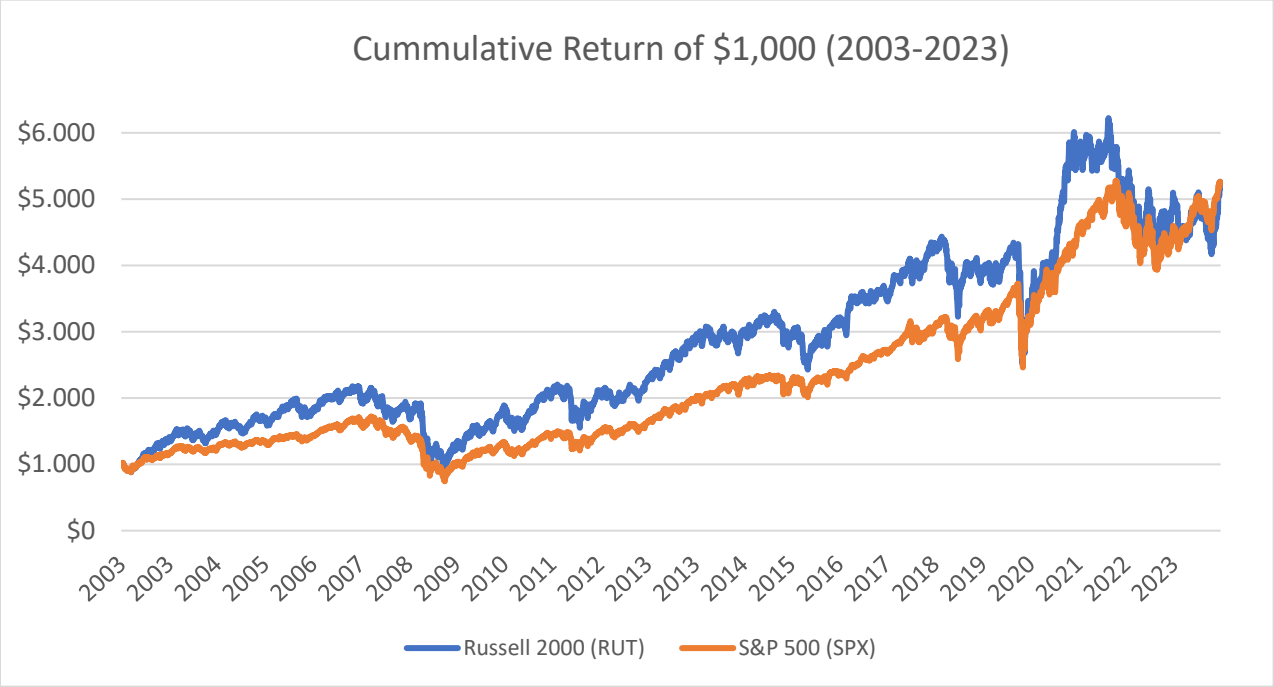


Figure 1: Cummulative Returns Russell 2000 (2003-2023) and S&P500 as benchmark (in orange)

The stock returns that will be analyzed need to converted into logarithmic returns. This means that the difference between the natural logarithm of the price at the end of period and the natural logarithm of the price at the beginning is calculated. This transformation into logarithmic returns is used because of the fact that stock returns are not normally distributed. Research has shown

that when analyzing stock returns over a long period of times, logarithmic returns are the best fitting measure (Hudson & Gregoriou, 2015).

The calculation of logarithmic returns or continuously compounded return functions as following:

Equation 2:

$$R_{log} = \ln\left(\frac{Vf}{Vi}\right)$$

Where

Rlog : Logarithmic Return

Vf: Final value

Vi: Initial Value

Some of the reasons why logarithmic returns should be used in this type of research include the fact that these continuously compounded returns prevent these returns from becoming negatives in some models of returns (Jorion, 2001). Also it is important that the use of these logarithmic returns allow for better comparison of stock returns across assets and over long periods of time (Hudson & Gregoriou, 2015). Research has also shown that for the calculation of future cumulative returns, using the logarithmic returns yields better results and predicts better future median cumulative returns than when using simple returns and compounding them (Hughson et al., 2006).

6. Analysis: Descriptive Statistics

As mentioned before in the methodology section, before proceeding with the actual regression model, the descriptive statistics of the several used variables will be shortly analysed. This helps better understand the data that is used, and make it also possible for one to create a basic understanding of the behavior of each of the variables. This first table contains the descriptive statistics for all the variables included in this paper. This includes the dependent variable "Small-Cap LogReturn" and all the independent variables with which the Returns of the Small-Cap index are going to be predicted.

Almost all these variables are in a return % Change form, due to the fact that the paper focuses on the sensibility of the stock returns to changes in the independent variables, with exception of the Eff. Interest Rate, which is left on the nominal interest rate in %.

Table 4: Descriptive Statistics (2003-2023)

Whole Period (2003-2023)					
Variables	**N**	**Mean**	**Max**	**Min**	**SD**
Small Cap LogReturn	242	0.64%	16.79%	-24.72%	5.74%
Eff. Fed Rate	243	1.36%	5.26%	0.05%	1.62%
Change Money Supply	243	0.53%	6.38%	-1.12%	0.66%
Change Retail Sales	243	0.37%	19.01%	-14.62%	2.09%
Change Industrial Production	243	0.06%	6.50%	-13.38%	1.26%
Change PPI	243	0.27%	3.21%	-5.33%	1.28%
Change CPI	243	0.21%	1.38%	-1.77%	0.33%
Change Housing Starts	243	0.26%	19.01%	-26.99%	8.71%
GDP Growth	243	0.18%	5.78%	-9.87%	1.02%
Change Consumer Confidence	241	0.26%	0.59%	-1.09%	0.25%
Crisis Dummy	243	0.0%	100.00%	0.00%	0.28%
Change Energy Index	243	0.36%	13.46%	-18.02%	3.20%
Corporate Tax	243	31.31%	35.00%	21.00%	6.18%
Change Public Debt	242	0.66%	5.48%	-0.64%	0.84%

Monthly Observations for all the variables (2003-2023)

Variables (Except Eff. Fed Rate) are in % Change/ Returns

Table 5 shows the basic descriptive statistics, but for the two different subgroups of the dataset. This separation into groups is based on the financial crisis of 2007-2008. The first subgroup "Pre-Financial Crisis" goes from the beginning of the observations until the last month before the official financial crisis breaks out. In other words, this first subgroup includes only observations up to the point where the dummy variable "Crisis" is 1, meaning that the financial crisis happens. The other subgroup "Post-Financial Crisis" contains all the rest of the observations that come from the point where the financial crisis starts until the end of the observation period, 2023.

It is worth mentioning that the second subgroup contains data with higher deviations from the mean, also with the mean being quite different than that from the fist subgroup. For example, in our dependent variable one is able to see that the max drawdown in one month is of about -24.7%, which is extremely different than the maximal monthly drawdown of -7.55% in the first subgroup. Also the biggest monthly increases in the Small-cap Index come directly after the crisis, with a 16.7% monthly return. In order to further offer some visualization of the data and its properties, time-series plots for all the variables will be included in the appendix. A histogram showing the distribution of the dependent variable "LogReturn Russell 2000" can also be used to better understand the data one is working with. From the time-series plots one can see that the volatility in % change of the independent variables were very high around the pandemic era of 2020. For most of the variables this is the period where big changes can be seen, with some exceptions as for the interest rates, where the pandemic did not create an extreme spike.

Table 5: *Descriptive Statistics Subgroups*

	Pre-Financial Crisis (2003-2007)					**Post-Financial Crisis (2008-2023)**				
Variables	**N**	**Mean**	**Max**	**Min**	**SD**	**N**	**Max**	**Min**	**Mean**	**SD**
Small Cap LogReturn	58	1.25%	10.09%	-7.55%	4.15%	185	16.79%	-24.72%	0.45%	6.15%
Eff. Fed Rate	58	3.15%	5.26%	0.98%	1.74%	185	4.83%	0.05%	0.79%	1.09%
Change Money Supply	58	0.43%	1.24%	-0.46%	0.30%	185	6.38%	-1.12%	0.56%	0.74%
Change Retail Sales	58	0.44%	2.86%	-1.38%	0.92%	185	19.01%	-14.62%	0.35%	2.34%
Change Ind.l Production	58	0.20%	1.22%	-1.94%	0.51%	185	6.50%	-13.38%	0.01%	1.41%
Change PPI	58	0.49%	2.92%	-3.12%	1.14%	185	3.21%	-5.33%	0.21%	1.32%
Change CPI	58	0.25%	1.38%	-0.50%	0.32%	185	1.19%	-1.77%	0.20%	0.33%
Change Housing Starts	58	-0.50%	14.59%	-15.54%	7.13%	185	23.96%	-26.99%	0.50%	9.16%
GDP Growth	58	0.26%	1.25%	-0.67%	0.48%	185	5.78%	-9.87%	0.14%	1.14%
Change Consumer Conf.	58	-0.01%	0.59%	-0.73%	0.25%	185	0.57%	-1.09%	-0.01%	0.25%
Crisis Dummy	58	0.00%	0.00%	0.00%	0.00%	185	100.00%	0.00%	12.00%	32.00%
Change Energy Index	58	0.96%	13.46%	-7.24%	3.55%	185	9.57%	-18.02%	0.17%	3.07%
Corporate Tax	58	35.00%	35.00%	35.00%	0.00%	185	35.00%	21.00%	30.16%	6.68%
Change Public Debt	58	0.62%	1.71%	-0.42%	0.48%	185	5.48%	-0.64%	0.68%	0.92%

Monthly Observations for all the variables (2003-2023)

Variables (Except Eff. Fed Rate) are in % Change/ Returns

When taking a look at the dependent variable, one can see that the volatility of returns is quite high, as can also be derived from the statistics from table 4 and 5. Specially during the financial crisis in 2008 and the corona pandemic in 2020 one can see extreme returns, both negative and positive. When taking the histogram into account as well, one can see that the returns are not perfectly normally distributed and that the returns are skewed towards the right. However, there is also some outliers in negative performance with observations with monthly returns of more than -20%.

There are some major differences on the behavior of the variables, with some having bigger movements during the financial crisis and others during the pandemic in 2020. For example, the inflation rate and energy price index were more affected by the global economic crisis, with these having a smaller variation in Covid times. The opposite is the case for Industrial production, with its biggest decrease since 25 years during 2020. Also GDP growth and money supply were affected greatly by the pandemic, with money supply jumping up in a record time historically as well.

7. Variable Selection

The first thing done for selecting relevant variables to include in the regression model is creating a correlation matrix for all the variables. It is important that the independent variables which are used to predict the Small-cap Stock returns. As a rule of thumb I won´t be including any of the variables with correlations higher than 0.7, since that could create collinearity issues, where two variables are too closely related to each other, affecting the effect of each other and creating unsatisfactory results. There are different thresholds for the correlation between variables, with some researchers putting the cutoff at 0.5, and others going up until 0.8 (P. Vatcheva & Lee, 2016). In this case I personally choose to put the threshold at 0.7, since in this case the correlation between some components was to be expected and it is preferrable to leave out some variables and not risking having too much multicollinearity in the model and it was the threshold adviced by (Pallant, 2020) in her SPSS Survival Manual.

Table 6: Pearson Correlation between the Independent Variables

	Variables												
	3M rate	**Money Supply**	**R.Sales**	**Ind. Prod.**	**PPI**	**CPI**	**Housing Starts**	**GDP growth**	**CCI**	**Crisis dummy**	**Energy Index**	**corpo tax**	**Public Debt**
3M rate	**1**	-.123	.132*	.053	.038	.136*	.093	.098	.078	-.044	.107	.008	.123
Money Supply	-.123	**1**	-.074	-.420**	-.145*	-.232**	-.125	-.307**	-.388**	.244**	-.199**	-.097	.411**
R.Sales	.132*	-.074	**1**	.630**	.403**	.309**	.390**	.742**	.197**	-.291**	.336**	-.063	-.065
Ind. Prod.	.053	-.420**	.630**	**1**	.290**	.242**	.361**	.777**	.230**	-.422**	.240**	.005	-.201**
PPI	.038	-.145*	.403**	.290**	**1**	.770**	.169**	.233**	-.107	-.171**	.764**	-.075	-.225**
CPI	.136*	-.232**	.309**	.242**	.770**	**1**	.162*	.188**	-.115	-.170**	.916**	-.195**	-.145*
Hous. Starts	.093	-.125	.390**	.361**	.169**	.162*	**1**	.321**	.130*	-.187**	.179**	-.028	.022
GDP growth	.098	-.307**	.742**	.777**	.233**	.188**	.321**	**1**	.244**	-.306**	.202**	.001	-.189**
CCI	.078	-.388**	.197**	.230**	-.107	-.115	.130*	.244**	**1**	-.201**	-.071	.130*	-.097
Crisis dummy	-.044	.244**	-.291**	-.422**	-.171**	-.170**	-.187**	-.306**	-.201**	**1**	-.155*	.091	.283**
Energy Index	.107	-.199**	.336**	.240**	.764**	.916**	.179**	.202**	-.071	-.155*	**1**	-.022	-.109
corpo tax	.008	-.097	-.063	.005	-.075	-.195**	-.028	.001	.130*	.091	-.022	**1**	-.017
Public Debt	.123	.411**	-.065	-.201**	-.225**	-.145*	.022	-.189**	-.097	.283**	-.109	-.017	**1**

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

The Pearson Correlation of the variables “CPI and “Energy” is of more than 0.9, indicating a huge collinearity in these variables. This is due to the fact that the energy index is a component of the CPI, which measures the price development of many different items and categories, and constitutes the index for what is known as inflation. Moreover, the correlation between “GDP” and “Industrial Production” , as well as that of “GDP” and “Retail Sales” are too high for all to be included in the regression. Because of this, the variables “CPI” and “GDP” will be excluded from the multiple regression model.

In order to still test whether GDP growth and Inflation rate have any explanatory power on the small-cap stock returns, a single regression for each of these variables is going to be run separately. It was complicated to include GDP into th multiple regression model used with all the other variables, since these might all be correlated to GDP or even being drivers of GDP themselves, yielding worse results.

Table 7: Coefficients and Collinearity Statistics

	Unstandardized Coefficients		Standardized Coefficients	Collinearity Statistics	
Model 1	**B**	**Std. Error**	**Beta**	**Tolerance**	**VIF**
(Constant)	0.384	2.416			
Money Supply	0.806	0.746	0.091	0.500	1.999
Interest Rate	-.005	.007	-.048	.913	1.095
R.Sales	0.383	0.290	0.140	0.316	3.169
Industrial production	-1.224	0.491	-0.268	0.306	3.268
PPI	0.372	0.464	0.083	0.330	3.028
CPI	-3.446	3.117	-0.195	0.113	8.811
Housing Starts	-0.074	0.044	-0.113	0.810	1.235
GDP GROWTH	0.493	0.647	0.087	0.270	3.710
CCI	7.192	1.628	0.311	0.717	1.395
Crisis dummy	-4.331	1.373	-0.217	0.747	1.339
Energy Index	0.471	0.302	0.263	0.125	8.013
corpo tax	0.010	0.064	0.011	0.768	1.303
Public Debt	0.464	0.481	0.068	0.724	1.381

a. Dependent Variable: Russel LogR

Another test for multicollinearity with SPSS is analysing the variance inflation factor (VIF), which can be found in table 7. The VIF measures as well the correlation between the independent variables and also shows the strength of this correlation

As a rule of thumb, one can state that values between 1 and 5 indicate moderate correlation, but is usually not a cause for concern. More concerning are values above 5, which indicate a high correlation between variables. There is some discussion in research regarding the acceptable VIF threshold, with some books stating that VIF values of above 10 are problematic (Vittinghoff et al., 2012). Alternatively, reseachers have also argued that VIF > 5 can be a cause for concern in terms of collinearity (Menard, 2002). In this case, the only pair of variables showing high VIF statistics are the variables of "Energy" and "CPI". As mentioned above, this will be avoided by not including the variable "CPI" due to the high correlation with the energy price index.

8. Results

After removing the conflicting predictor variables then the first regression model can be run without the excluded variables of "GDP" and "CPI", which in turn were an important part of the stated hypotheses. This removal was indeed somewhat to expect since the two variables are closely related to other variables included in the regression model. E.g. the CPI and Energy Price Index are closely related since the Energy Price Index is one component of the CPI measure. On the other hand, GDP growth is going to be affected by changes in the other several variables such as consumer confidence or even the interest rates, so including it in the main regression specification would have a negative impact on the explanatory power of the model itself.

8.1 Main Regression Specification

Table 9 shows the model summary for the main regression specification 1, which includes most of the variable, as seen in the model description in the lower part of the table.
Table 8 shows the summary for the Regression specification 1, which includes all the variables apart from GDP growth and CPI, due to the fact that the correlation between these were too high. Also the VIF value for the pair CPI and Energy Price Index shows a potential problem with correlation, which also indicates that the CPI variable might be removed from the model specification 1.

Table 8: Regression Model 1

Model Summary[b]					
Model	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Durbin-Watson**
1	.433[a]	0.187	0.148	5.30733%	2.250

Predictors: (Constant), Housing Starts (% change), Public Debt (% change), corpo tax, 3M rate (% change), CCI (% change), Energy Index (% change), Crisis dummy, R.Sales (% change), Money Supply (% change), Industrial production (% change), PPI (% change)

b. Dependent Variable: Russell LogR

One important consideration is the assumption of autocorrelation of the residuals of the regression model. For this, the Durbin-Watson test is run and the result interpreted. In this case, a Durbin-Watson value of 2.250 is inside an acceptable range of 1 to 3, in which one should not bee too worried about autocorrelation. It is suggested that values under 1 or above 3 are a cause for concern in the model (Field, 2009).

After considering the autocorrelation as not problematic, we might concentrate on the expalatory power of the regression model itself. The R-Squared shows how well the fit of the model ist, and therefore how good it explains the variation in the actual data. In this case, a R-Squared of 0.187 states that the model accounts for 18.7% of the variation in the actual data. That R-squared is very low and typically shows a bad model fit and lack of explanatory power. Nevertheless it is still necessary to look at the individual contribution and significance of each independent variable.

In order to present the actual results, we can take a look at the regression equation for the model specification 1, after having removed the two variables mentioned above:

Equation 3: (Model Specification 1)

$$R = \alpha + \beta_1 * InterestR + \beta_2 * MoneyS + \beta_3 * RetailS + \beta_4 * IndProduction + \beta_5 * PPI + \beta_6 * HousingS + \beta_7 * ConsumerConf + \beta_8 * EnergyPrice + \beta_{10} * CrisisDummy + e$$

Table 9 shows the actual coffficients of the regression for model 1 and includes as well collinearity statistics such as the VIF, we talked about it before testing the assumptions as well Here one can see the coefficients also the P-values of the coefficients of each variable. In this case the threshold for acceptance of significance will be $P<0.05$. Nevertheless, coefficients with $P<0.1$ will be analyzed as well, even though the significance is not at perfect levels.

As seen on the table, out of the 10 independent variables included in this specification, only 4 are statistically significant at P-value < 0.05. These would be the IVs: Retail Sales growth, Industrial Production growth, Consumer Confidence growth and also the Crisis dummy. These variables are the ones with the biggest regression coefficient when compared to other variables without statistical significance.

Retail Sales has a positive significant coefficient of 0.541, which states that there is a positive relationship between the Retail Sales change and the Returns of the Russell 2000 Index. On the other hand, the Industrial production variable has a negative coefficient of -1.065, which is kind of surprising, since it was previously stated in Section 3.2 that actually a positive relationship was to be expected.

The highest statistical significance was found in the variables Consumer Confidence and in the Dummy variable for economic crisis. The dummy variable of crisis was 1 in the times where economic crisis where happening, as stated by the FED. There are different ways of tracking the economic expansion and contraction cycles of the USA, but the FRED provides a dummy variable representing recessions/crisis in the USA, called “NBER Recession Indicator”.
When expanding the acceptance level to P-value < 0.1, we find one more variable with a statistically significant result, this being the Housing Starts growth. Nevertheless, the coefficient of -0.075 shows that the effect is minimal and also could be ignored since it is also not inside the threshold of $P<0.05$ used in this paper. The main result in this case is the strong impact of the CCI and of the crisis dummy variable. CCI being the consumer confidence index has a positive effect on the performance of the Russell 2000.

There might be several reasons for that, including also the fact that consumer sentiment is similar to stock returns, in terms of both representing or being constituted by feelings and sentiment of customers and investors respectively. Alternatively, one can argue that the lack of consumer confidence will lead to decreased spending, which in turn contribute to a decrease in business for the companies that are included in the Russell 2000.

Table 9: Regression Model Specification 1; Results summary

Coefficients[a]							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	-1.118	1.907		-.585	.559		
Interest Rate	-.005	.007	-.048	-.768	.443	.913	1.095
Money Supply	.807	.743	.091	1.087	.278	.506	1.978
R.Sales	.541**	.247	.197	2.190	.030	.437	2.288
Ind. Production	-1.065**	.423	-.233	-2.517	.013	.412	2.424
PPI	.180	.446	.040	.403	.687	.358	2.790
Housing Starts	-.074*	.044	-.113	-1.705	.090	.811	1.233
CCI	7.413***	1.615	.320	4.589	<,001	.729	1.372
Crisis dummy	-4.268***	1.369	-.214	-3.117	.002	.752	1.330
Energy Index	.205	.172	.114	1.190	.235	.385	2.595
corpo tax	.041	.057	.044	.713	.477	.948	1.055
Public Debt	.485	.486	.071	.998	.319	.709	1.410

a. Dependent Variable: Russel LogR

b. Predictors: (Constant), Public Debt (% change), Housing Starts (% change), corpo tax, CCI (% change), Energy Index (% change), Crisis dummy, R.Sales (% change), Money Supply (% change), Industrial production (% change), PPI (% change)

Note: Significance levels → P<0.1 */ P<0.05 **/ P<0.01 ***

It is also important not to neglect the effect of the crisis dummy, since during economic crisis the returns of the stock market, specially in this case the Russell 2000, are affected in a clearly negative manner. This does make logical sense due to the fact that economic crisis are times of uncertainty and also economic stress for companies. Not only companies are affected, but individuals notice economic crisis as well since unemployment usually rises and the economic situation of the average individual worsens.

We can implement a P-P plot to test for the normality of the error terms. In this case, we can see that the observations do not follow perfectly the straight line, nevertheless it seems to be still in an acceptable range. We can therefore argue that the error terms of the regression are normally distributed.

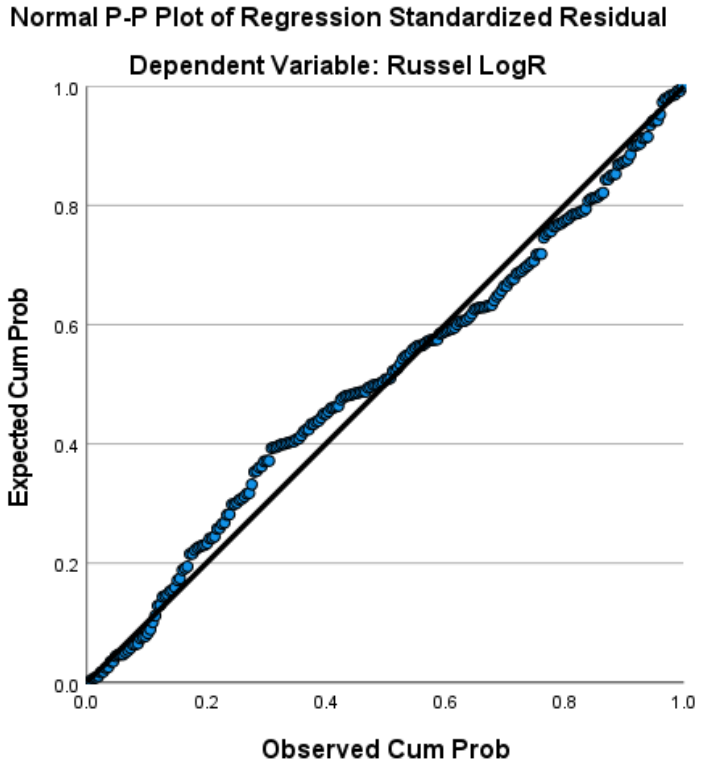


Figure 2: *P-P Plot of Regression Standardized Residual values*

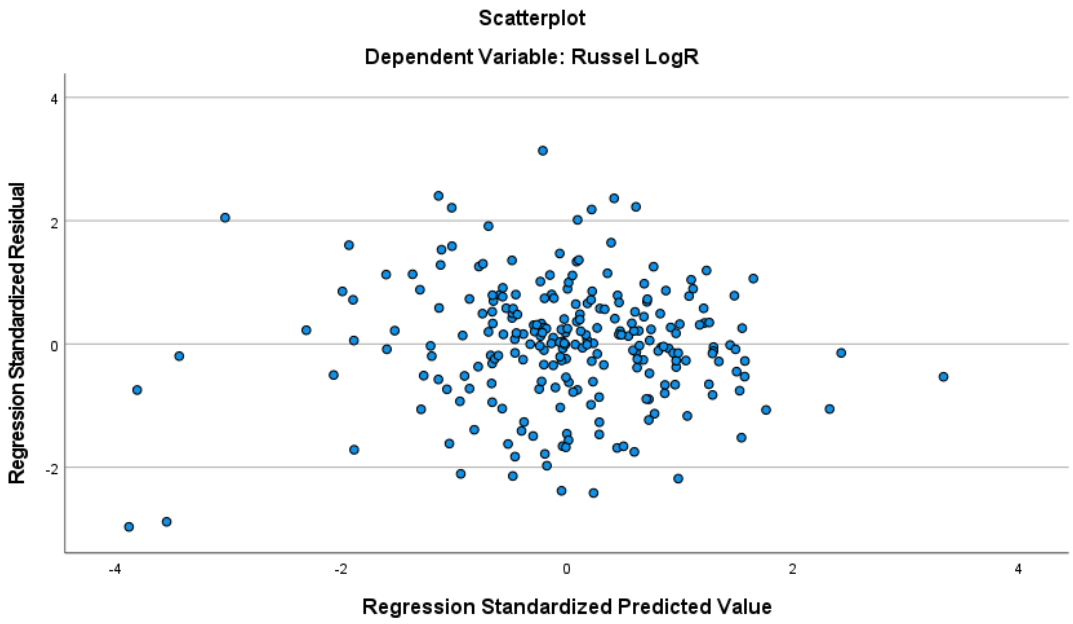


***Figure 3**: Scatterplot for heteroskedasticity check*

One also looks for heteroskedasticity problems in the models, using scatterplots and also running the Breusch-Pagan test as mentioned above. For doing this, when running the regression model, the residuals will be saved as a new variable, and then it will be squared. In order to do the Breusch-Pagan test, a new regression will be run with all the same IVs as before, but this time with the squared residuals as dependent variable. Here, we theoretically make a test of the null hypothesis indicating the lack of heteroskedasticity.

Table 10: Test for Heteroskedasticity (Breusch-Pagan Test)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	106153.973	11	9650.361	6.314	<,001[b]
	Residual	349990.425	229	1528.342		
	Total	456144.398	240			

a. Dependent Variable: SQResiduals

b. Predictors: (Constant), Public Debt (% change), Housing Starts (% change), corpo tax, 3M rate (% change), CCI (% change), Energy Index (% change), Crisis dummy, R.Sales (% change), Money Supply (% change), Industrial production (% change), PPI (% change)

In this case, since we have a P-value of <0.001, which is a strong evidence against the null hypothesis. This therefore means that we do have a problem with heteroskedasticity in the data.

It also might be interesting so see if we have extreme outliers, but the out of the normal returns and changes in macro factors happen in crisis times, which are the most volatile, so if I would take any outlier out, this would be changing the reality. This could also be an indictation that the distribution is skewed, with volatility jumps during and shortly after economic recessions. Outside of the 3 SD there is only one monthly observation of November 2020, where the Russell 2000 had a monthly increase in value of more than 16%. This was following the March 2020 crash that the global equities suffered, but then quickly recovered. The most extreme observations are indeed found near to each other since the highest increases come often relatively quick after big falls in price.

Table 10.1: Casewise Diagnostics Outliers

Casewise Diagnostics[a]				
Case Number	**Std. Residual**	**Russel LogR**	**Predicted Value**	**Residual**
214	3.137	16.79%	0.1464%	16.64786%

a. Dependent Variable: Russel LogR

8.1.1 Parameter Estimates with Robust Standard Errors

Because of the presence of heteroskedasticity in the model, it is useful to also run the analysis with robust standard errors. In this case, the HC3 method is used, which is typically the most commonly employed one. It is argued that when the number of observations is < 250, the HC3 method should be the one used (Long & Ervin, 2000).

One can see that the coefficients estimated are exactly the same as in the normal OLS regression model run above in Table 9. Nevertheless, the significance of these results differs between the OLS and this regression with robust standard errors, with only one variable being still significant now.

Table 11: Heteroskedasticity Robust Standard Errors

Parameter Estimates with Robust Standard Errors

Dependent Variable: Russel LogR

Parameter	B	Robust Std. Error[a]	T	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-1.118	2.288	-.489	.626	-5.625	3.390
Interest Rate	-.005	.011	-.467	.641	-.027	.017
Money Supply	.807	1.535	.526	.599	-2.218	3.832
R.Sales	.541	.454	1.192	.234	-.353	1.435
Ind. Production	-1.065	1.055	-1.010	.314	-3.144	1.014
PPI	.180	.558	.322	.748	-.920	1.279
Housing Starts	-.074	.045	-1.637	.103	-.163	.015
CCI	7.413***	1.804	4.109	<,001	3.858	10.968
Crisis dummy	-4.268*	2.443	-1.747	.082	-9.081	.546
Energy Index	.205	.163	1.259	.209	-.116	.525
Corp. Tax	.041	.068	.603	.547	-.093	.174
Public Debt	.485	1.010	.480	.631	-1.505	2.476

a. HC3 method

Note: Significance levels → P<0.1 */ P<0.05 **/ P<0.01 ***

8.2 Results Model GDP

So in order to test this a further model specification will be run, called model 2 GDP and this will be a single regression only including the variable GDP

Equation 4: (Model Specification 2 – GDP)

$$R = \alpha + \beta_1 * GDPgrowth + e$$

In this model specification one can see in Table 12 that the explanatory power of this model is minimal, at 0.7% of the variation in the dependent variable explained by our GDP growth

variable. This is of course very low and might indicate that the relationship between GDP and equity returns might be more complex and not explainable through linear regression models.

Table 12: Model Specification 2 (GDP)

Model Summary[b]

Model	R	R Square	Adj. R Square	Std. Error	Durbin-Watson
2	.083[a]	.007	.003	5.7434%	2.018

a. Predictors: (Constant), GDP GROWTH

b. Dependent Variable: Russel LogR

Table 13 shows the actual coefficients estimated by the model, with GDP having a coefficient of 0.470, but without statistical significance. Therefore, one can argue that no statistically significant relationship can be seen between GDP and the returns of the Russell 2000 Index.

Table 13 : Regression Specification 2 (GDP)

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	T	Sig.	Tolerance	VIF
2	(Constant)	.572	.374		1.527	.128		
	GDP GROWTH	.470	.363	.083	1.294	.197	1.000	1.000

a. Dependent Variable: Russel LogR

8.3 Results Model CPI

CPI was the other variable removed from the 1. Model specification because of collinearity problems with the other variables. In this case, CPI was very closely related to the variable of the energy index, since this energy index is indeed a part of the calculation of the Inflation rate. It would have also been possible to delete the variable of Energy Index, nevertheless, deleting CPI

might be the better option due to the fact that other variables might also be influencing the CPI variable stronger than the energy price index.

Equation 5: (Model Specification 3 - CPI)

$$R = \alpha + \beta_1 * InflationRate + e$$

In this case also a simple linear regression is run and, similar to the regression for the GDP, the results yielded are not good. In this case, R-Square is 0.003, which means that only 0.3% of the variation is explained by the model. Here of course there is also no statistical significance for the coefficients calculated as seen in Table 14.

Table 14: Model Summary Model Specification 3 (CPI)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.055[a]	.003	-.001	5.74509%	1.983

a. Predictors: (Constant), CPI (% change)

b. Dependent Variable: Russel LogR

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.438	.438		1.000	.318		
	CPI % change	.972	1.136	.055	.855	.393	1.000	1.000

a. Dependent Variable: Russel LogR

8.4 Results hypotheses

In this following section the answer to the different main hypotheses are going to be stated in order to create an easy visualization of the results. The hypotheses are going to be stated and then answered. The results for all the 4 hypotheses are all not statistically significant and therefore no hypothesis could be confirmed.

8.4.1 H1 : GDP

- **H1: There is a positive relationship between the variable "GDP" and the small-cap stock returns**

As stated above, a simple regression was run to test H1, without much success of explaining the relationship between GDP and the returns of the Russell 2000. In this case, the B coefficient for GDP is 0.470, which is positive, but nevertheless there is no statistical significance at the 5% P-value threshold. Because of that, one cannot reject the H0, thus providing no evidence of any positive relationship between GDP growth and the Russell 2000 returns. This is in line with previous research of Binswanger, finding no significant correlation between GDP growth and stock returns after the 1980s, suggesting a decoupling of the returns from the real economic growth (Binswanger, 2000). Also similar results were found by Joebges, when analyzing this relationship in the G7 countries, finding only significant results for 2 out of 7 countries (Fichtner & Joebges, 2024).

Figure 4 shows the scatterplot with the line of best fit, which is in turn the regression equation for this model specification. In this case one can see the coefficient for GDP growth is around 0.47, nevertheless still not significant, as stated above. The distribution of the cases that can be seen in Figure 4 gives some insight into the relationship between GDP growth and the returns of the Russell 2000. In this case, one can argue that it is not possible to describe that with a linear regression model, due to the fact that the relationship does not seem to be linear. Most observations are very near to the 0% GDP change, but with big deviations in terms of the index returns. No clear relationship depending on GDP can be seen in this case.

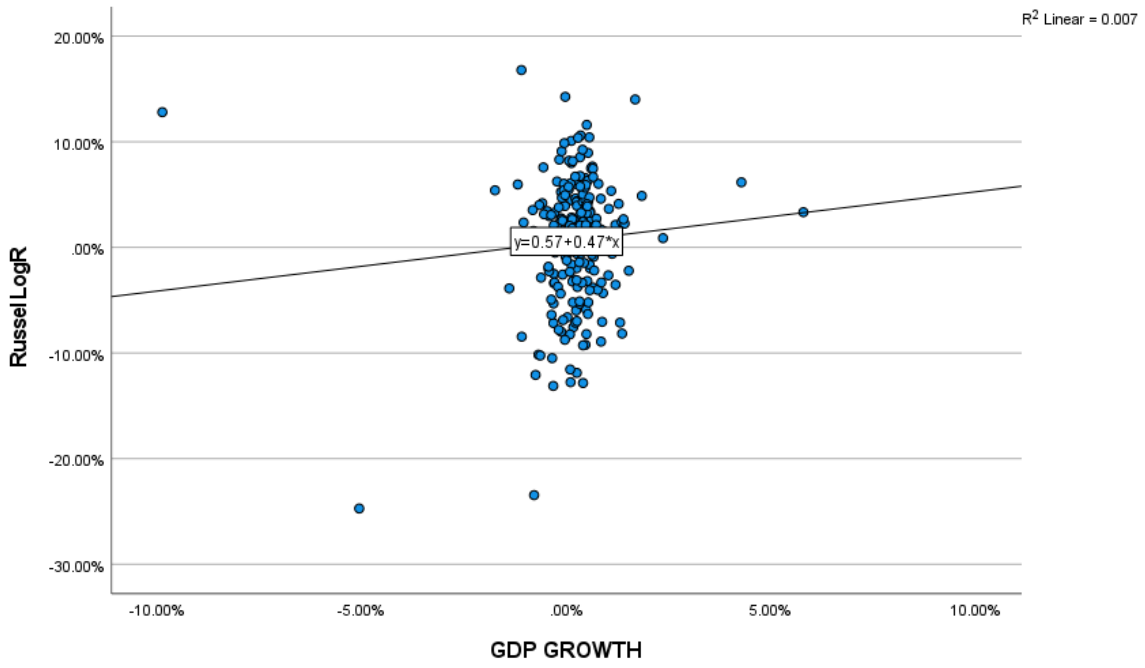


Figure 4: Scatterplot for GDP Regression

8.4.2 H2: Consumer Price Index

- **H2: There is a negative relationship between the variable "Consumer Price Index (CPI)" and the small-cap stock returns**

Similar as for the H1, in order to test this hypothesis a single regression was run only including the variable of CPI Growth, which is the Inflation rate. The results that the regression yielded were also similar, which a very low R-Square of 0.003 and a B coefficient of 0.972, without statistical significance at the 5% level. All in all, we cannot reject the H0 and therefore have no evidence of a positive relationship between the inflation rate and the Russell 2000 returns. Compared to previous research with statistical significance, there is no evidence that can be found in this paper for the relationship between the inflation rate and the stock market. Contrasting to the results found in countries with prolongued high inflation, the USA aims at inflation rates of 2%, thus not obligating the population to protect their purchasing power as much as in countries such as Argentina (Choudhry, 2001).

Figure 5 shows a scatterplot for the regression specification for H2, only including CPI as an independent variable. In this case we can see that there is a slight slope in the regression line,

nevertheless, no clear relationship between the two variables can be found. It does seem like the returns for the Russell 2000 increase as CPI increases, but that might be caused by some outliers rather than because of CPI itself. Similar to Figure 4, Figure 5 shows no indication of a linear relationship between the variables of CPI change and the returns of the Russell 2000.

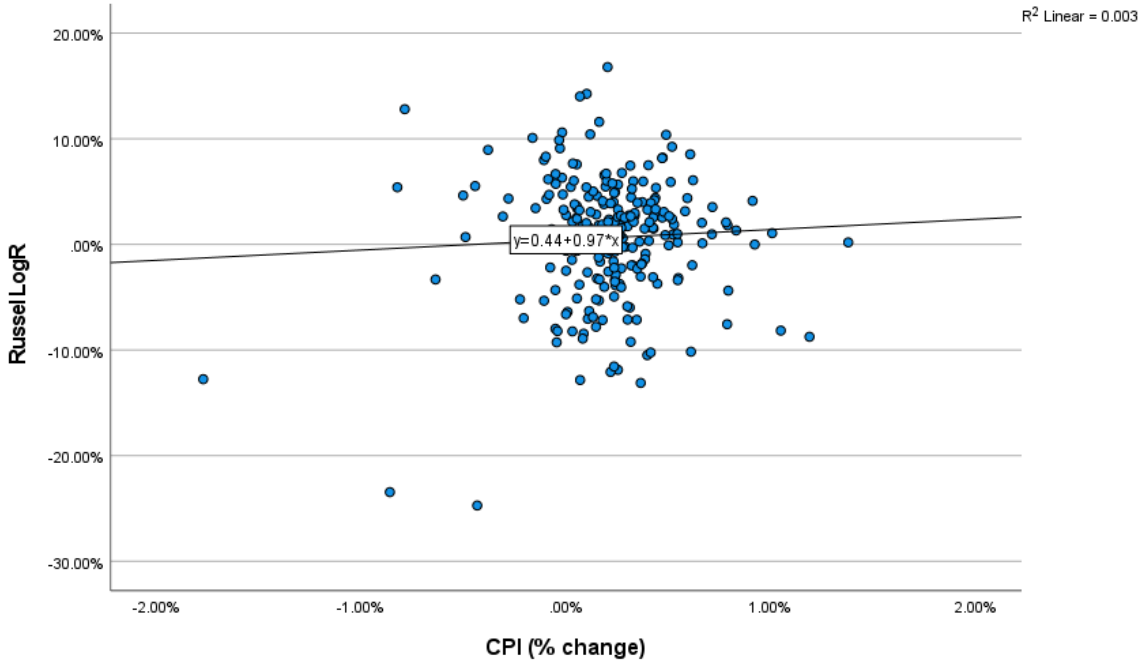


Figure 5: Scatterplot for CPI Regression

8.4.3 H3: Interest Rate

- **H3: There is a negative relationship between the variable "Interest Rate" and the small-cap stock returns**

For the third hypothesis, equation Nr. 3 (regression specification 1) was used. The results for the relationship between the Interest rate changes and small cap returns is also not statistically significant with a P-value of 0.443. The B coefficient is indeed negative as expected by the hypothesis but it is also very close to 0, with -0.005. Therefore, similar to the others, one cannot reject the H0, having no evidence of a negative relationship between the variables.

Figure 6 shows the scatterplot for the relationship between the interest rate change and the returns of the Russell 2000. It is worth noting that the interest rates have experienced some huge

increases on monthly basis during the last years. Here, most observations are centered around a 0% change in interest rates, which outliers with monthly increases of between 300 and 500%. Nevertheless, no linear relationship is found here, indicating that the linear regression implemented might not be able to explain the relationship between these variables.

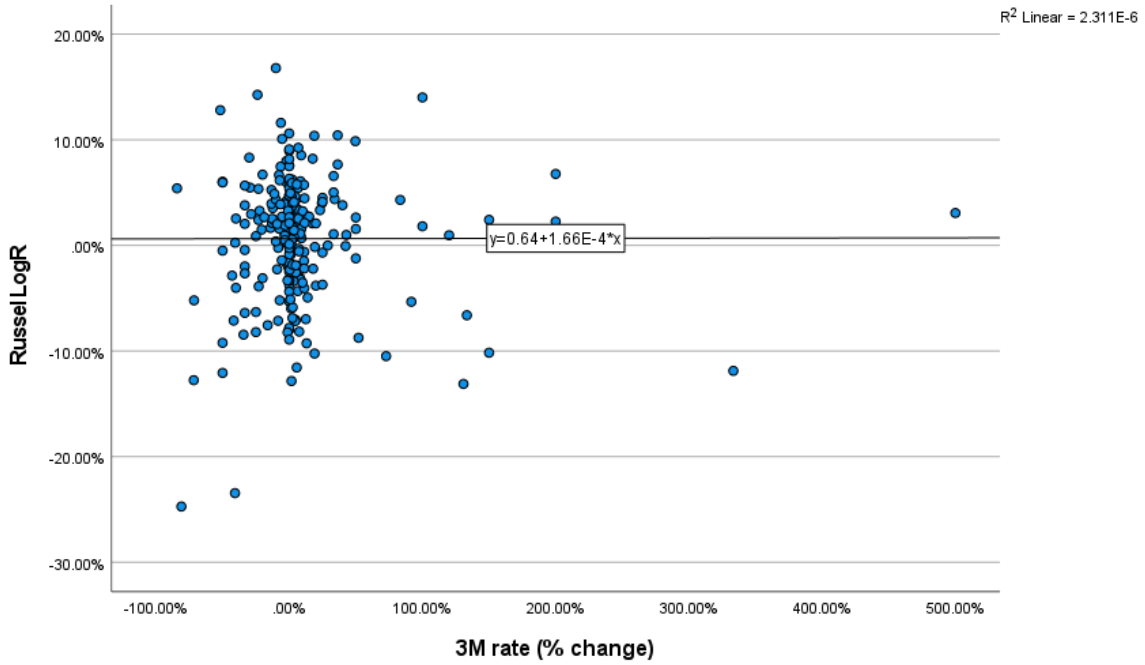


Figure 6: Scatterplot for Interest Rate Hypothesis

8.4.4 H4: Energy Index

- **H4: There is a negative relationship between the variable "Energy Index" and the small-cap stock returns**

Last but not least, the hypothesis regarding the Energy Index, which was also tested together with most variables in model specification 1, yielded similar results when compared to the others. There is still no statistical significance with P-value of 0.235. Nevertheless I was surprised by the positive coefficient, since I expected increases in energy prices to have a negative effect on small-cap returns, specially since smaller companies in theory have less hedging positions against energy prices and also are on average less stable.

Last but not least we have Figure 7, the scatterplot for the hypothesis testing the relationship between the energy index and the returns of the Russell 2000. Here one can also see that the relationship between the variables is not well explained at all by either the multiple regression model run and also the single regression that can be seen with the line of best fit. The observations are not fitting at all, and it becomes clear that the relationship cannot be explained via linear regression models.

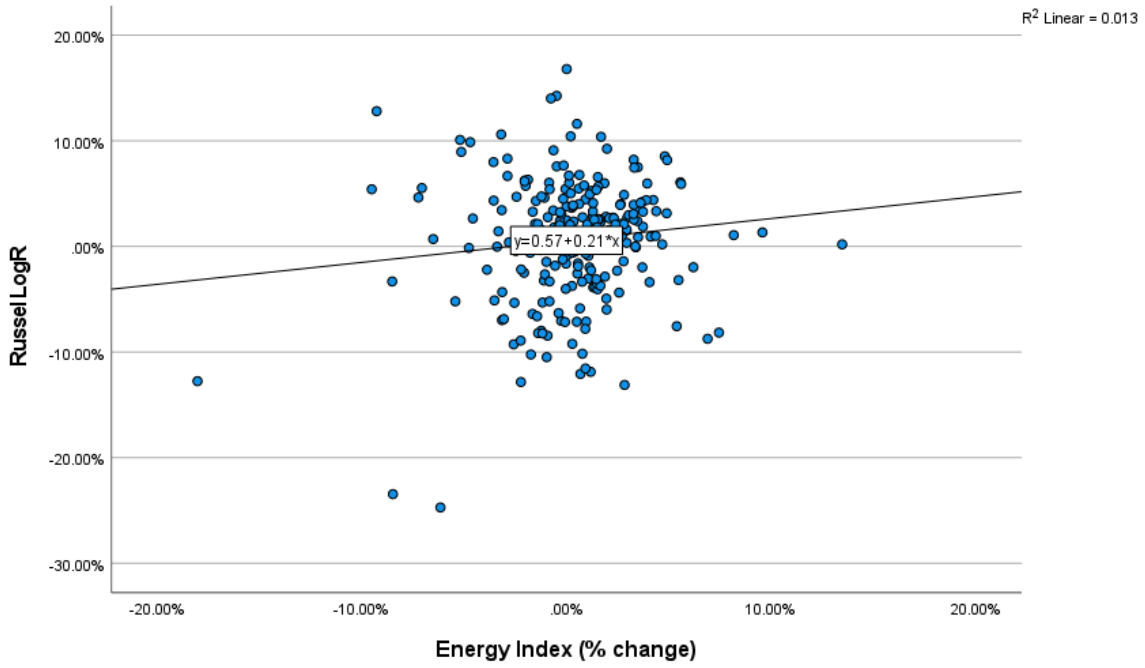


Figure 7: Scatterplot for Energy Index

9. Discussion

Summarizing the results, one can say that the models are not well fitting and all three model specifications have a substantial lack of expalantory power. In order to improve, changes were made to the data and also to the models themselves. First of all, some variables were deleted from the first specification as collinearity problems were present. Moreover, the dependent variable, which in this case is the monthly returns of the Russell 2000 Index were transformed into logarithmic form. Different variable selection methods were also tried, but that did also not

improve the model. Normally all the variables are entered into the regression model as a block, but alternatively one can do it “stepwise”, or even on a “forward selection”. These are methods to choose which IVs to include into the final model. In this case at the end the variables were introduced with the “Enter” procedure, which just includes all IVs as a block, since the selection methods did not improve the models.

It is worth noting that there are not too many statistically significant coefficients in the different specifications. In the model specification 1, which is the main one, 4 variables had statistically significant coefficients. Retail Sales growth had a regression coefficient of 0.541 with a P-value of 0.030, indicating a moderate positive relationship between growth in retail sales and the returns of the Russell 2000 Index. This was in line with the expectations of growth in sales resulting in higher returns for small-cap equities. On the other hand, a result differing greatly from the expectations stated in Table 2 is the variable of Industrial Production growth. In this case, a positive coefficient was expected but at the end a coefficient of -1.065 with P-value of 0.013 came out. This is kind of surprising due to the fact that one might expect increases in industrial production being a good factor for the companies and thus probably for their equities. Previous research has shown that in the past industrial production did have a statistically significant positive relationship with stock returns in the USA (Schwert, 1990). However, consecutive papers have proved that this relationship no longer holds in the USA. It is assumed, that the American economy was undergoing major changes and moving from a manufacturing based economy towards a majority of economic growth coming from the service industry, thus affecting the relationship between stock returns and industrial production (Young, 2006). In this case, a negative significant relationship is found for the periods between 1954 and 1988, but only with the stock reutrns lagged by 12 months, indicating that the effect is not noticeable instantaneously. This result seems in line with the result found in this paper, however in this case, the returns are not transformed into a lagged variable. It is also worth noting that after 1988 that relationship disappears and no statistically significant relationship is found when tested with 6, 9 and 12 months (Young, 2006).

The biggest surprise was the result of the Consumer Confidence Index with a quite strong positive effect on the small-cap returns. Is is indeed normal to think that increases in consumer confidence are going to be noticeable for companies and maybe even more for smaller

companies, as that translates direclty into more private consumption, but it is nevertheless interesting that from all the variables this one is the one with the biggest influence and also with the highest statistical significace. Increases in the CCI not only might translate into more private consumption but also in more private investment into smaller companies, thus triggering improved performance and growth. CCI Growth had a regression coefficient of 7.413 with P-value of <0.001. This result means that a 1% increase in the Consumer Confidence Index leads to a more than 7% returns for equities with small capitalization.It is also worth noting that CCI is the only variable which was also still significant at P-value< 0.05 in the estimation with robust standard errors as seen in Table 11. All of the other significant variables became statistically insignificant after accounting for the violation of the heteroskedasticity assumption. Therefore we can confirm the effect of the Consumer Confidence Index growth on the returns of the Russell 2000. However, this result is conflicting with previous research, as generally the effect of consumer confidence is negative on the stock market. Negative relationship was found by Lemmon & Portniaguina (2006) when analyzing the relationship between these two variables. A main difference with the approach of that paper is the election of a different proxy for consumer confidence, as the Consumer Sentiment Index from the University of Michigan is employed. This proxy for consumer confidence is one of the two main indexes for consumer confidence in the USA, besides the CCI. In this case, no relationship was found before 1977, however, a significant relationship was found between 1978 and 2006, thus suggesting a possible increase in the relevance of consumer sentiment in the USA. But, as mentioned before, the relationship found is negative, a contrasting result compared to this paper (Lemmon & Portniaguina, 2006).

Another similar study suggests also that there is a significant negative relationship between consumer confidence and small-cap returns. It is worth noting that this relationship shows only statistical significance for small-cap stocks and not for the overall market represented by the S&P500 Index (K. L. Fisher & Statman, 2003). This significance only for small cap stocks, which is in line with the previous paper of Lemmon & Portniaguina, can suggest that the small cap equities are more affected by the CCI than the general market would do. The paper argues that increases in consumer confidence lead to more bullish private investors, and simultaneously leads to lower returns in the following month. In this case, the effect of the CCI is lagged by one month, thus suggesting that the effect of changes in consumer confidence are not noticeable inmediately (K. L. Fisher & Statman, 2003).

A recommendation for future research is the introduction of a lagged dependent variable for small cap returns, which might be able to capture the effects of changes in consumer confidence better than as done in this paper. This comes as a result that the effect of changes in such factors might not have an immediate effect on the stock market. Nevertheless, the result is highly significant and does indicate, in line with the results of this paper, that there is a strong relationship between small cap stocks and consumer confidence.

Additionally, the dummy variable representing economic crisis times has also a negative coefficient as it was to expect. During times of crisis and recession it is common for stocks to have a decrease in returns since the overall market as well falls in price. In the case of the Russell 2000, in times of crisis one can therefore expect difference in expected returns of -4.268 %.

In this discussion it is worth talking about the limitations of this paper, starting with the lack of explanatory power of the regression models, thus not giving as much insight as expected into the topic of interest. First possible reason is the fact that many variables were included in the model, which are closely related to each other, since all these macroeconomic variables affect each other as well. Secondly, the model might be just too sensible to outliers if taking into account that the Russell 2000 was quite volatile, specially during crisis times such as 2007-2008 and also for a shorter period in 2020. The technical outliers are for example some months during the Covid crisis, during which the index decreased greatly in price, but was also followed by a quick recovery. Research shows that actually missing the best trading days affects the performance of a portfolio or investment in a great manner and might be even more important to be invested and have exposure to those big positive days occur (Moss, 2022). Additionally, it is very difficult to predict what is going to happen in the macro scale and also in the equity space, meaning that timing the market is really hard. Often the best trading days come shortly after the worst ones, and therefore is patience and time in the market a big factor for success in investing generally. Even though one cannot predict the future, this paper might help to understand what effect these variables can have on the stock market and specifically the small-cap equities in the USA.

Another possibility is that the variables included in the paper just cannot explain what is happening in reality since the relationship between stock returns and real economic growth and output is much more complex. It might be the case that the relationship is just not linear and can thus not be accurately described via an OLS regression. A possible recommendation for future

research is the usage of a different model to test these relationships, even with newer techniques such as machine learning. There are some possibilities to try forecasting and predicting stock returns with models such as neural networks (Abe & Nakayama, 2018). It also seems possible to predict returns on monthly frequencies using the Boosted Regression Trees (BRT) method, which is also another relatively new machine learning methodology (Rossi, 2018). I consider this topic to be of interest because of the potential and novelty and I expect big developments to come in the near future, specially also in my topic of interest which is finance and capital markets.

Even though the relationship might be too compex to understand, volatility in stocks is not necessarily a negative thing, as some investors perceive and also avoid. One has to take into account that volatility goes both ways and of course it means that in crisis times the stocks might fall further, but they will probably rise even further as well. I do therefore consider including small-cap equity investments into each portfolio, and probably even more for younger investors with still a long investment horizon.

One last limitation of this paper is related to the macroeconomic data from the FRED database and its reliability/usability. The data itself that is released on the FRED database gets revised on a regular basis, and thus, it might deviate slightly from the values that were available at other points in time. It is therefore worth noting that it is not possible to download all the data in an unrevised version, in order to run the model again and test for differences. This means that manually creating this historical dataset with each original (non-revised) monthly observation for each of the variables would involve a substantial time investment, and also be prompt to mistakes, because of the manual nature of it.

All in all, I consider the revision of the FRED data to not be a shortfall for this research, since this paper analyses mainly the effects of changes in these macroeconomic variables on the return of the Russell 2000. Meaning that even if some observations get revised, the general trend will broadly stay the same, thus still holding some significance. Also due to the fact that the revisions are often changes in measurement levels and similar adjustments, that also have slight influence on prior observations, one cannot just simply take the original unrevised vintage observation without risking making mistakes.

Some arguments that speak for the use of the latest vintage is the fact that revisions are done often just to correct mistakes, because the data collection methodology has changed or also aiming to improve accuracy. Therefore , it is recommended to take this factor of the revisions of the FRED when wanting to update the model with the latests observations, since previous ones might have changed minimally, as stated in the prior paragraph. It is therefore also recommended that the whole dataset gets updated in order to prevent errors.

10. Conclusion

Coming back to the general research question of: **"What is the relationship between macroeconomic factors and small cap equity returns?"**, one might argue that it is impossible to answer that question with certainty. The relationship between macroecnomic factors and small cap equity returns is too complex to be explained via a regular OLS model .This being said, one result did have interesting outcome, this being the growth in "Consumer Confidence", having a big significant positive relationship with the returns of the Russell 2000.

In the following table the expected relationship of the variables is going to be presented and compared to the actual outcome of the model, in order to visualize the differences in expectations and reality and therefore making it easier to see which relationships were believed to be different.

The biggest surprises were the variables CCI, as mentioned above, and also the Industrial Production, where a different relationship was expected. In theory it was supposed that an increase in Industrial Production would be a positive factor for the small-cap equity returns, but the actual opposite is the case in this study.

Overall, one can argue that the model used for this analysis was probably not the most appropriate one, and alternative models might be better fitting to this type of complex data. The result of the CCI variable seems to be in line with previous research, but it is worth mentioning that the relationship might as well be more complex than it seems. In reality, it seems like rising consumer confidence does come with higher stock returns, but generally a high level of consumer confidence will later be followed by low stock returns (K. L. Fisher & Statman, 2003). That suggests that there is a positive correlation between CCI change and returns, while the general level of CCI is not.

Table 15: Relationship between IVs and Small-Cap Stock Returns

Category	Variable Name	Expected Relationship	Actual Relationship (OLS)
Macro	GDP	+	+
	CPI	-	+
	PPI	-	+
	Industrial Production	+	-**
	Retail Sales	+	+**
	Housing Starts	+	-
	CCI	+	+***
Monetary Policy	Interest Rate	-	-
	Money Supply	+	+
Fiscal Policy	Public Debt	-	+
	Tax Rate	-	+
Energy Prices	Energy Index	-	+

Note: Significance levels → P<0.1 */ P<0.05 **/ P<0.01 ***

This paper adds to the literature on the relationship between macroeconomic factors and stock returns, which is a topic of interest since many years in finance literature.The major differences and coinciding results are discussed in the previous discussion. A part of the findings is realizing that the setting, geography and time differences have a big effect on the results of the analysis, since often conflicting results are found with the same variables, just with a different research design. Since this paper focuses on only smaller capitalization stocks, this also contributes to the research on these small-cap stocks, which are perceived to have a different risk-reward profile than bigger capitalization stocks. Nevertheless, these studies often have contradictory results depending on time frame, lag of variables, research methodology and also depending on the proxy used for each of the variables. Therefore one can expect results to vary also from time to time and also from country to country. These findings on small stock performance and their link to major macroeconomic variables might be of interest to policy makers and investors. Policy makers do consider the financial market into account when designing legislations, knowing that

policy changes are major influences to the stock market. Moreover, it is relevant for investors to understand the effects that these changes in policies and developments in the macroeconomic environment might have on their investments. This is notably important in the case of private equity investors which might consider investing in small-cap equities.

Some suggestions for future research include the introduction of different lags for the dependent variable, as the macroeconomic factors often do not have an instant effect on the stock market. Recommended would be to implement lags for the stock returns of 1, 3, 6 and 12 months in order to capture relationships that might have been ignored in this particular paper due to the research design.

Moreover, the introduction of so many macroeconomic factors, which often are highly correlated or even are moderated by another independent variable might be problematic. It is suggested that a more specific topic of interest within macroeconomic factors is selected, possibly yielding better results.

11. References

1. Abe, M., & Nakayama, H. (2018). *Deep Learning for Forecasting Stock Returns in the Cross-Section*. https://doi.org/10.48550/ARXIV.1801.01777
2. Acaravci, A., Ozturk, I., & Kandir, S. Y. (2012). Natural gas prices and stock prices: Evidence from EU-15 countries. *Economic Modelling*, *29*(5), 1646–1654. https://doi.org/10.1016/j.econmod.2012.05.006
3. Aktas, E., & McDaniel, W. R. (2009). Pragmatic problems in using beta for managerial finance applications. *Applied Financial Economics*, *19*(16), 1345–1354. https://doi.org/10.1080/09603100802570390
4. Appel, I., Gormley, T. A., & Keim, D. B. (2018). Identification Using Russell 1000/2000 Index Assignments: A Discussion of Methodologies. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2641548
5. Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, *24*(1), 34–112. https://doi.org/10.1007/s11142-018-9470-2
6. Asprem, M. (1989). Stock prices, asset portfolios and macroeconomic variables in ten European countries. *Journal of Banking & Finance*, *13*(4–5), 589–612. https://doi.org/10.1016/0378-4266(89)90032-0
7. Assefa, T. A., Esqueda, O. A., & Mollick, A. V. (2017). Stock returns and interest rates around the World: A panel data approach. *Journal of Economics and Business*, *89*, 20–35. https://doi.org/10.1016/j.jeconbus.2016.10.001
8. Baharumshah, A. Z., Mohd, S. H., & Yol, M. A. (2009). Stock prices and demand for money in China: New evidence. *Journal of International Financial Markets, Institutions and Money*, *19*(1), 171–187. https://doi.org/10.1016/j.intfin.2007.11.001

9. Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0
10. Basu, D., & Chawla, D. (2012). An Empirical Test of the Arbitrage Pricing Theory—The Case of Indian Stock Market. *Global Business Review*, *13*(3), 421–432. https://doi.org/10.1177/097215091201300305
11. Basu, S. (1977). INVESTMENT PERFORMANCE OF COMMON STOCKS IN RELATION TO THEIR PRICE-EARNINGS RATIOS: A TEST OF THE EFFICIENT MARKET HYPOTHESIS. *The Journal of Finance*, *32*(3), 663–682. https://doi.org/10.1111/j.1540-6261.1977.tb01979.x
12. Belo, F., Gala, V. D., & Li, J. (2013). Government spending, political cycles, and the cross section of stock returns. *Journal of Financial Economics*, *107*(2), 305–324. https://doi.org/10.1016/j.jfineco.2012.08.016
13. Berndt, A., Lustig, H., & Yeltekin, S. (2009). Does the US Government Hedge against Fiscal Risk? *American Economic Journal*, *Vol 4*(1).
14. Binswanger, M. (2000). Stock returns and real activity: Is there still a connection? *Applied Financial Economics*, *10*(4), 379–387. https://doi.org/10.1080/09603100050031507
15. Binswanger, M. (2004). Stock returns and real activity in the G-7 countries: Did the relationship change during the 1980s? *The Quarterly Review of Economics and Finance*, *44*(2), 237–252. https://doi.org/10.1016/j.qref.2003.07.001
16. Bodie, Z. (1976). Common stocks as a hedge against inflation. *The Journal of Finance*, *31*(2), 459–470. https://doi.org/10.1111/j.1540-6261.1976.tb01899.x
17. Boudoukh, J., & Richardson, M. (1993). Stock Returns and Inflation: A Long-Horizon Perspective. *The American Economic Review*, *Vol. 83*(No. 5), 1346–1355.
18. Brooks, C. (2019). *Introductory econometrics for finance* (Fourth edition). Cambridge University Press.

19. Campbell, J. Y. (2000). Asset Pricing at the Millennium. *The Journal of Finance*, *55*(4), 1515–1567. https://doi.org/10.1111/0022-1082.00260
20. Chan, L. K. C., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and Stock Returns in Japan. *The Journal of Finance*, *46*(5), 1739–1764. https://doi.org/10.1111/j.1540-6261.1991.tb04642.x
21. Chen, N.-F., Roll, R., & Ross, S. A. (1986). Economic Forces and the Stock Market. *The Journal of Business*, *59*(3), 383. https://doi.org/10.1086/296344
22. Chiah, M., Chai, D., Zhong, A., & Li, S. (2016). A Better Model? An Empirical Investigation of the Fama–French Five-factor Model in Australia. *International Review of Finance*, *16*(4), 595–638. https://doi.org/10.1111/irfi.12099
23. Chopra, S., Yadav, D., & Chopra, A. N. (2019). Artificial Neural Networks Based Indian Stock Market Price Prediction: Before and After Demonetization. *International Journal of Swarm Intelligence and Evolutionary Computation*, *8*, 1.
24. Choudhry, T. (2001). Inflation and rates of return on stocks: Evidence from high inflation countries. *Journal of International Financial Markets, Institutions and Money*, *11*(1), 75–96. https://doi.org/10.1016/S1042-4431(00)00037-8
25. Clements, M. P. (2014). Forecast Uncertainty— *Ex Ante* and *Ex Post*: U.S. Inflation and Output Growth. *Journal of Business & Economic Statistics*, *32*(2), 206–216. https://doi.org/10.1080/07350015.2013.859618
26. Cremers, M. M. (2001). Reviving Beta? Bayesian Test of Efficiency and the CAPM. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.276177
27. Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (2001). Overconfidence, Arbitrage, and Equilibrium Asset Pricing. *The Journal of Finance*, *56*(3), 921–965. https://doi.org/10.1111/0022-1082.00350

28. Domian, D. L. (1996). Expected inflation, interest rates, and stock returns. *The Financil Review*, *31*(4), 809–830.
29. ENOW, S. T., & BRIJLAL, P. (2016). Determinants of Share Prices: The Case of Listed Firms on Johannesburg Stock Exchange. *Journal of Accounting and Management*, *6*(1), 85–92.
30. Fama, E. D. (1981). Stock Returns, Real Activity, Inflation, and Money. *The American Economic Review*, *71*(4), 545–565.
31. Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5
32. Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, *18*(3), 25–46. https://doi.org/10.1257/0895330042162430
33. Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010
34. Fichtner, F., & Joebges, H. (2024). Stock market returns and GDP growth. *Hans-Böckler-Stiftung, Institut Für Makroökonomie Und Konjunkturforschung (IMK)*, *IMK Study No. 90*. https://www.econstor.eu/bitstream/10419/283065/1/p-imk-study-90-2024.pdf
35. Field, A. P. (2009). *Discovering statistics using SPSS: And sex, drugs and rock "n" roll* (3rd ed). SAGE Publications.
36. Fischer, S., & Merton, R. C. (1984). Macroeconomics and finance: The role of the stock market. *Carnegie-Rochester Conference Series on Public Policy*, *21*, 57–108. https://doi.org/10.1016/0167-2231(84)90005-8
37. Fisher, K. L., & Statman, M. (2003). Consumer Confidence and Stock Returns. *The Journal of Portfolio Management*, *30*(1), 115–127. https://doi.org/10.3905/jpm.2003.319925
38. Fisher, M. J., & Marshall, A. P. (2009). Understanding descriptive statistics. *Australian Critical Care*, *22*(2), 93–97. https://doi.org/10.1016/j.aucc.2008.11.003

39. Green, S. B. (1991). How Many Subjects Does It Take To Do A Regression Analysis. *Multivariate Behavioral Research*, *26*(3), 499–510. https://doi.org/10.1207/s15327906mbr2603_7
40. Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed). McGraw-Hill Irwin.
41. Hamrita, M. E., & Trifi, A. (2011). The Relationship between Interest Rate, Exchange Rate and Stock Price: A Wavelet Analysis. *International Journal of Economics and Financial Issues*, *1*(4), 220–228.
42. Hudson, R. S., & Gregoriou, A. (2015). Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns. *International Review of Financial Analysis*, *38*, 151–162. https://doi.org/10.1016/j.irfa.2014.10.008
43. Hughson, E., Stutzer, M., & Yung, C. (2006). The Misuse of Expected Returns. *Financial Analysts Journal*, *62*(6), 88–96. https://doi.org/10.2469/faj.v62.n6.4356
44. Jones, E. (2023). Managing the risks of quantitative tightening in the euro area. *Spanish and International Economic and Financial Outlook*, *12*(1), 5–13.
45. Jorion, P. (2001). *Value at risk: The new benchmark for managing financial risk* (2. ed). McGraw-Hill.
46. Kilian, L., & Park, C. (2009). THE IMPACT OF OIL PRICE SHOCKS ON THE U.S. STOCK MARKET*. *International Economic Review*, *50*(4), 1267–1287. https://doi.org/10.1111/j.1468-2354.2009.00568.x
47. Kim, J. H., Rahman, M. L., & Shamsuddin, A. (2019). Can energy prices predict stock returns? An extreme bounds analysis. *Energy Economics*, *81*, 822–834. https://doi.org/10.1016/j.eneco.2019.05.029
48. Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics*, *44*(2), 169–203. https://doi.org/10.1016/S0304-405X(97)00002-0

49. Kothari, S. P., Shanken, J., & Sloan, R. G. (1995). Another Look at the Cross-section of Expected Stock Returns. *The Journal of Finance*, *50*(1), 185–224. https://doi.org/10.1111/j.1540-6261.1995.tb05171.x

50. Lael Joseph, N., & Vezos, P. (2006). The sensitivity of US banks' stock returns to interest rate and exchange rate changes. *Managerial Finance*, *32*(2), 182–199. https://doi.org/10.1108/0307435061064193

51. Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, *49*(5), 1541–1578. https://doi.org/10.1111/j.1540-6261.1994.tb04772.x

52. Lemmon, M., & Portniaguina, E. (2006). Consumer Confidence and Asset Prices: Some Empirical Evidence. *Review of Financial Studies*, *19*(4), 1499–1529. https://doi.org/10.1093/rfs/hhj038

53. Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics*, *82*(3), 631–671. https://doi.org/10.1016/j.jfineco.2005.10.001

54. Long, J. S., & Ervin, L. H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician*, *54*(3), 217. https://doi.org/10.2307/2685594

55. Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77. https://doi.org/10.2307/2975974

56. Martínez-Moya, P., Ferrer-Lapeña, R., & Escribano-Sotos, F. (2013). Relationship between interest rate changes and stock returns in Spain: A wavelet-based approach. *Department of Economic and Financial Analysis*, *University of Castilla-La Mancha,*. https://www.academia.edu/download/36405501/2013-4_DT-DAEF.pdf

57. Menard, S. (2002). *Applied Logistic Regression Analysis* (2nd ed.). SAGE Publications, Inc. https://doi.org/10.4135/9781412983433

58. Mondal, S., & Imran, M. S. (2010). Determinants of stock price: A case study on Dhaka stock exchange. *International Journal of Finance*, *2*(3), 1–16.
59. Moss, W. (2022). The Perils Of Market Timing: Missing The Best Days In The Market. *Wessmoss.Com*. https://www.wesmoss.com/news/money/the-perils-of-market-timing-missing-the-best-days-in-the-market/
60. Nguyen, N. M. (2018). *An Empirical Test of the Fama-French Five-Factor Model: Applicability to State-Owned Enterprises in Vietnam* [Aalto University School of Business]. https://core.ac.uk/download/pdf/186477249.pdf
61. Oberndorfer, U. (2009). Energy prices, volatility, and the stock market: Evidence from the Eurozone. *Energy Policy*, *37*(12), 5787–5795. https://doi.org/10.1016/j.enpol.2009.08.043
62. Özdemir, M., & Yeşilyurt, S. (2013). The Relationship between Macroeconomic Variables and ISE 100 Index Returns during the Quantitative Easing Policies of USD Federal Reserve: Evidence from Turkey. *International Research Journal of Finance and Economics*, *107*. https://www.researchgate.net/profile/Ali-Awdeh/publication/257862737_Measuring_the_Degree_of_Competition_in_the_Arab_Banking_Systems/links/00463526258de90448000000/Measuring-the-Degree-of-Competition-in-the-Arab-Banking-Systems.pdf#page=130
63. P. Vatcheva, K., & Lee, M. (2016). Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. *Epidemiology: Open Access*, *06*(02). https://doi.org/10.4172/2161-1165.1000227
64. Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS* (7th ed.). Routledge. https://doi.org/10.4324/9781003117452
65. Pandey, A., & Sehgal, S. (2016). Explaining Size Effect for Indian Stock Market. *Asia-Pacific Financial Markets*, *23*(1), 45–68. https://doi.org/10.1007/s10690-015-9208-0

66. Pettengill, G. N., Sundaram, S., & Mathur, I. (1995). The Conditional Relation between Beta and Returns. *The Journal of Financial and Quantitative Analysis*, *30*(1), 101. https://doi.org/10.2307/2331255

67. Ritter, J. R. (2005). Economic growth and equity returns. *Pacific-Basin Finance Journal*, *13*(5), 489–503. https://doi.org/10.1016/j.pacfin.2005.07.001

68. Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

69. Rossi, A. G. (2018). *Predicting Stock Market Returns with Machine Learning* [University of Maryland]. https://mendoza.nd.edu/wp-content/uploads/2019/07/2018-Alberto-Rossi-Fall-Seminar-Paper-1-Stock-Market-Returns.pdf

70. Schwert, G. W. (1990). Stock Returns and Real Activity: A Century of Evidence. *The Journal of Finance*, *45*(4), 1237–1257. https://doi.org/10.1111/j.1540-6261.1990.tb02434.x

71. Senaviratna, N. A. M. R., & A. Cooray, T. M. J. (2019). Diagnosing Multicollinearity of Logistic Regression Model. *Asian Journal of Probability and Statistics*, 1–9. https://doi.org/10.9734/ajpas/2019/v5i230132

72. Shapiro, M. (1988). *The Stabilization of the U.S. Economy Evidence From the Stock Market* (Working Paper NBER w2645; p. w2645). National Bureau of Economic Research. https://doi.org/10.3386/w2645

73. Sharpe, W. F. (1964). CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

74. Stanhope, E., & Meredith, C. (2015). *Inefficiency Breeds Opportunity in Small Cap Equities*. O'Shaughnessy Asset Management, LLC.

75. Switzer, L. N., & Fan, H. (2007). Spanning Tests for Replicable Small-Cap Indexes as Separate Asset Classes. *The Journal of Portfolio Management*, *33*(4), 102–110. https://doi.org/10.3905/jpm.2007.690611

76. Tavares, J., & Valkanov, R. (2003). Fiscal Policy and Asset Returns. *Society for Economic Dynamics Conference 2003*. https://www.researchgate.net/publication/45579142_Fiscal_Policy_and_Asset_Returns

77. Titman, S., & Warga, A. (1989). Stock Returns as Predictors of Interest Rates and Inflation. *The Journal of Financial and Quantitative Analysis*, *24*(1), 47. https://doi.org/10.2307/2330747

78. Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance*, *59*(2), 831–868. https://doi.org/10.1111/j.1540-6261.2004.00650.x

79. Vittinghoff, E., Glidden, D. V., & McCulloch, C. E. (Eds.). (2012). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models* (2nd ed). Springer.

80. Wattanatorn, W., & Kanchanapoom, T. (2012). Oil Prices and Profitability Performance: Sector Analysis. *Procedia - Social and Behavioral Sciences*, *40*, 763–767. https://doi.org/10.1016/j.sbspro.2012.03.263

81. Young, P. (2006). *Oung, P. (2006). Industrial production and stock returns. Global Asset and Wealth Management MBA Programm, Simon Fraser University, Burnaby CA* [MBA Thesis, Simon Fraser University]. https://core.ac.uk/download/pdf/56366829.pdf

82. Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *The Journal of Finance*, *61*(1), 105–137. https://doi.org/10.1111/j.1540-6261.2006.00831.x

12. Appendix

12.1 Time-series plots for the Independent variables

Figure 8: Time-series plot – Interest Rate

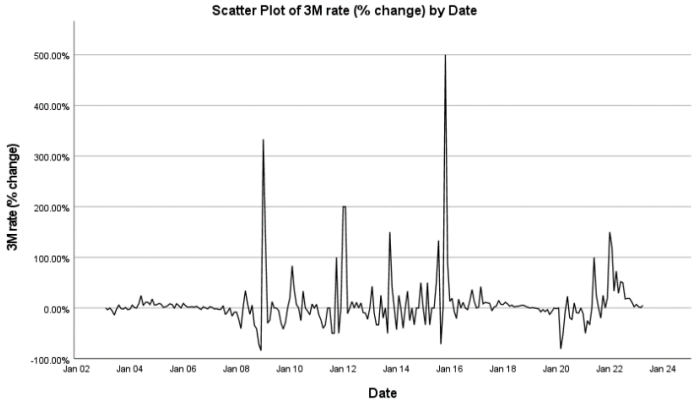


Figure 9: Time-series plot – Money Supply

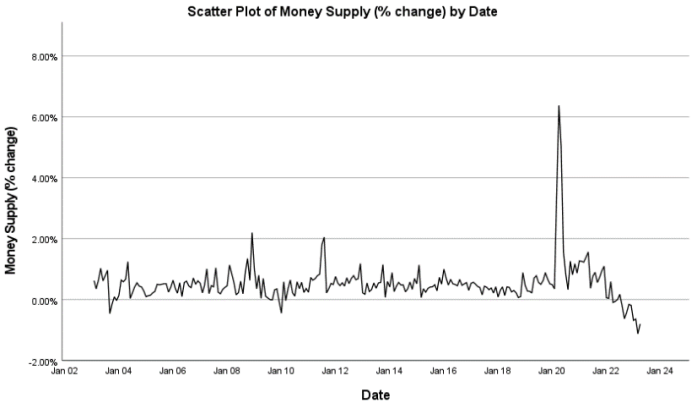


Figure 10: Time-series plot – Retail Sales

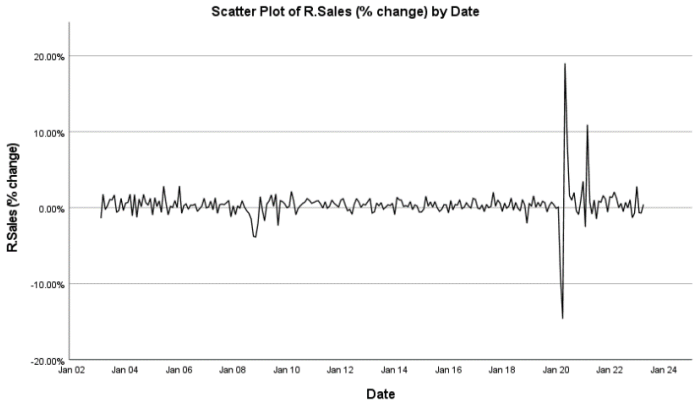


Figure 11: Time-series plot – Ind. Prod.

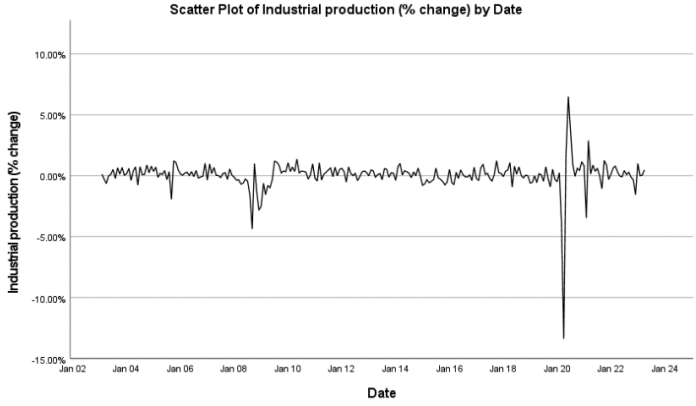


Figure 9: Time-series plot – Inflation Rate

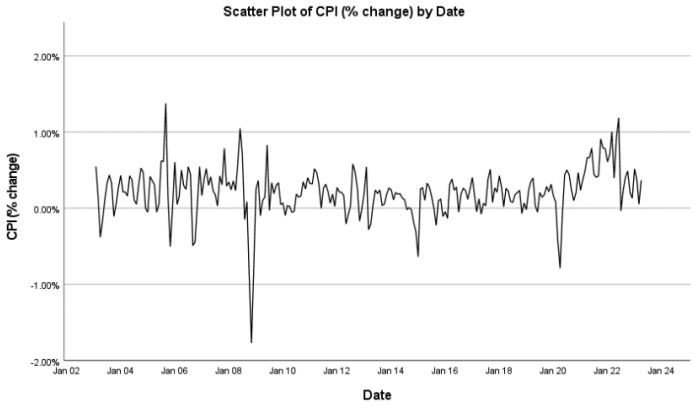


Figure 13: Time-series plot – Housing Starts

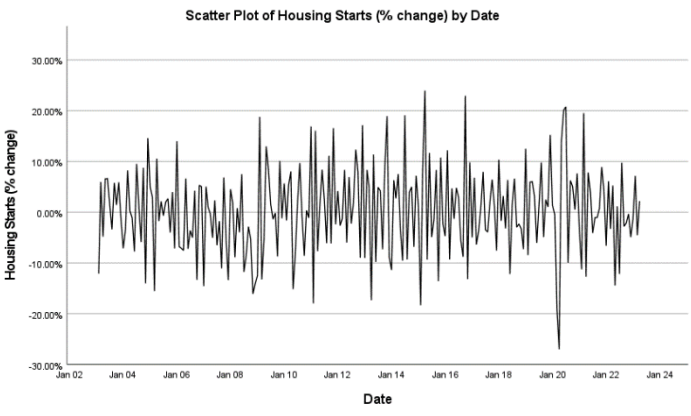


Figure 104: Time-series plot – GDP

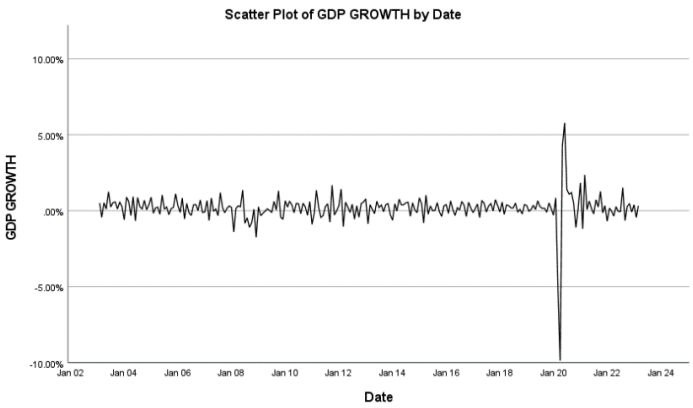


Figure 115: Time-series plot –

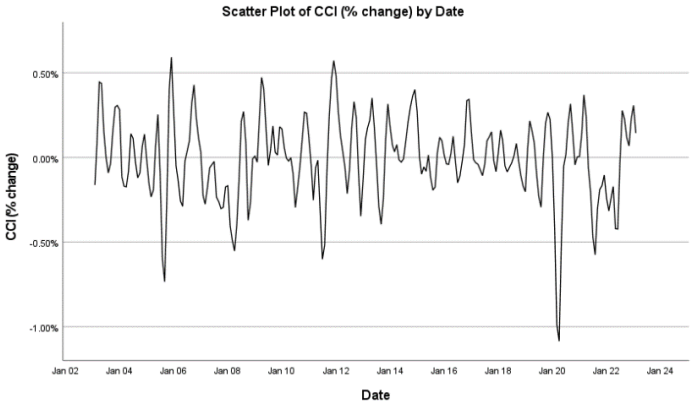


Figure 126: Time-series plot – Energy Index

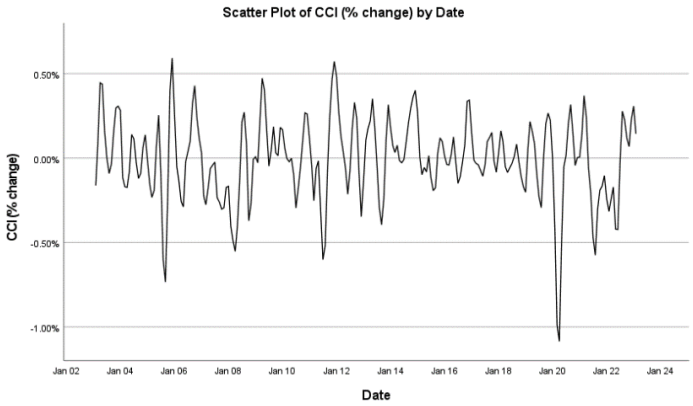


Figure 137: Time-series plot – Public Debt

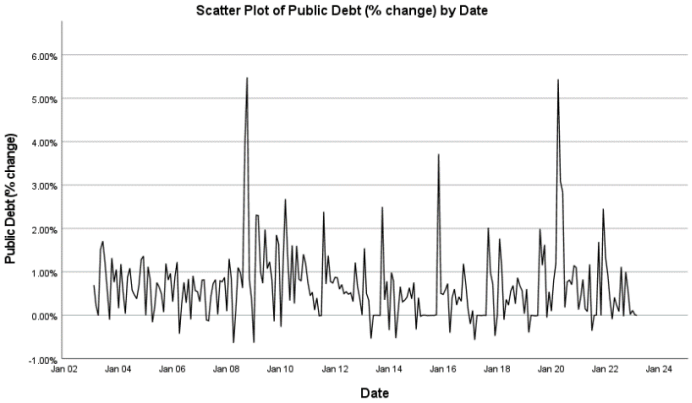


Figure 148: Time-series plot – Corporate Tax

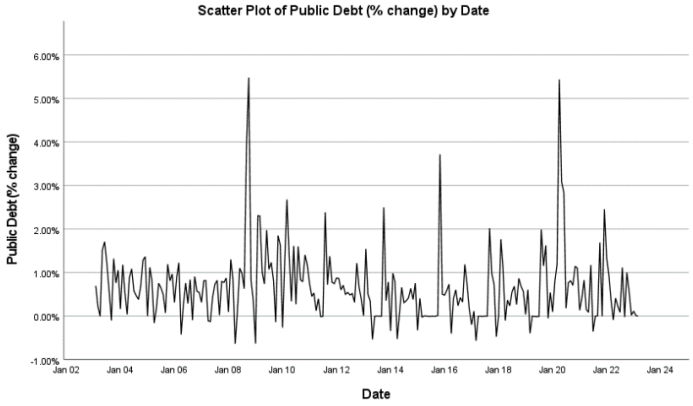
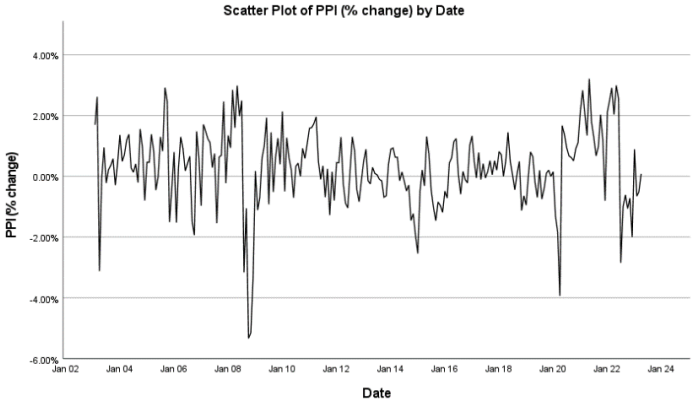


Figure 19: Time-series plot – PPI



12.2 Histograms for the independent variables

Figure 20: Histogram – Interest Rate

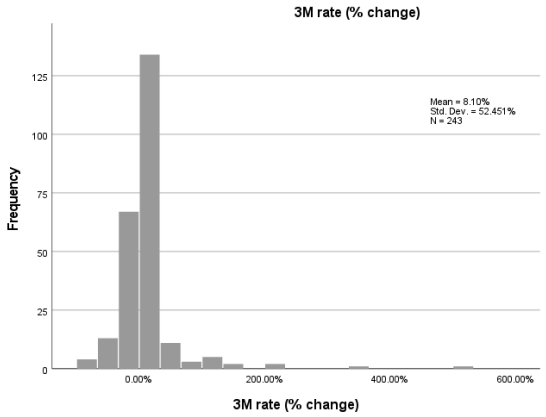


Figure 21: Histogram – Money Supply

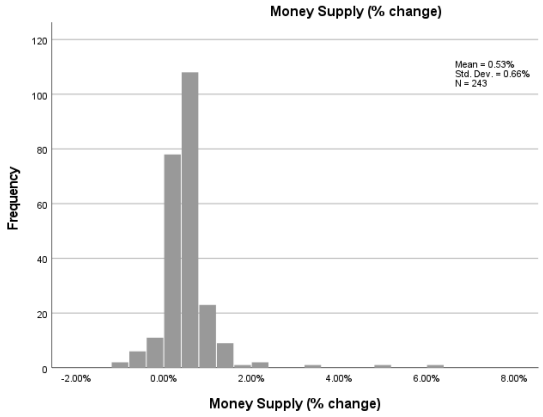


Figure 22: Histogram – Retail Sales

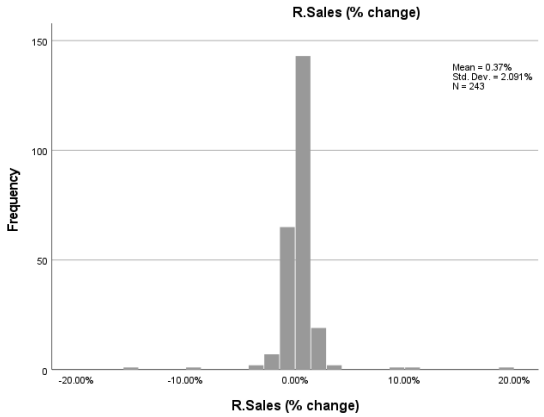


Figure 23: Histogram – Industrial Production

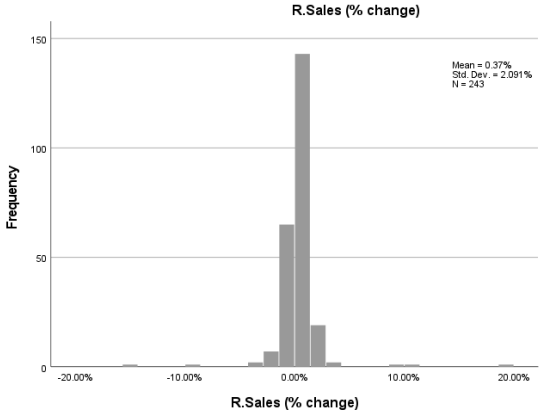


Figure 24: Histogram – Inflation Rate

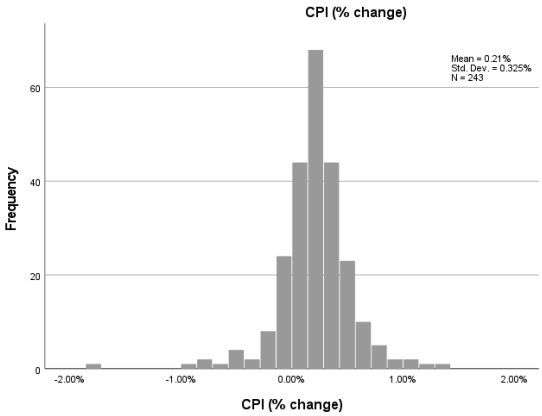


Figure 25: Histogram – Housing Starts

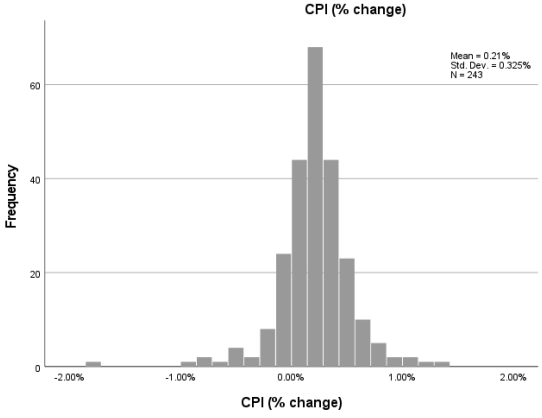


Figure 26: Time-series plot – GDP Growth

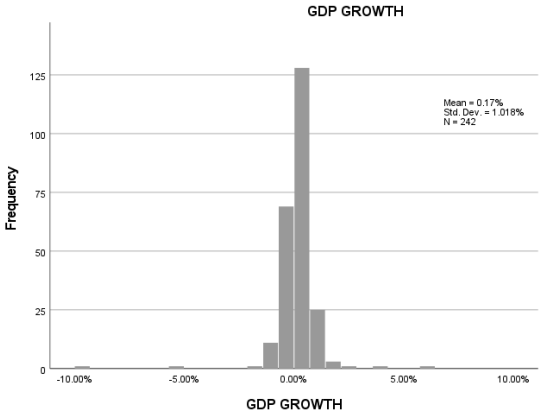


Figure 27: Time-series plot – Cons. Conf.

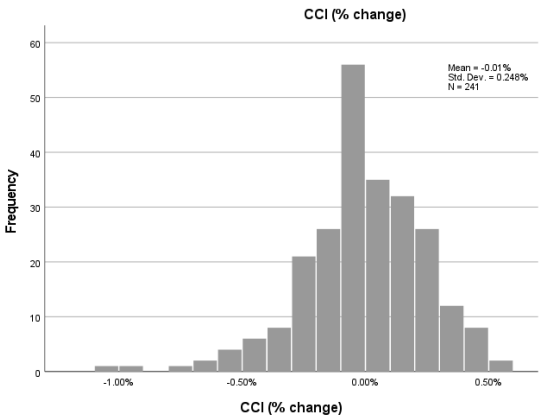


Figure 28: Time-series plot – Energy Index

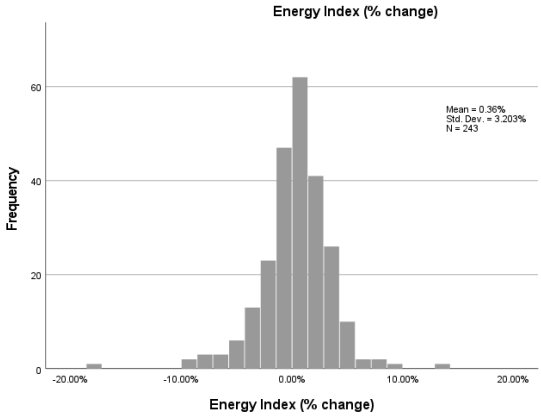


Figure 29: Time-series plot – Public Debt

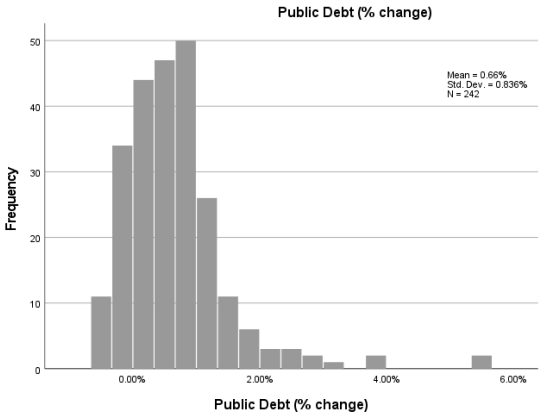


Figure 30: Histogram – Corporate Tax

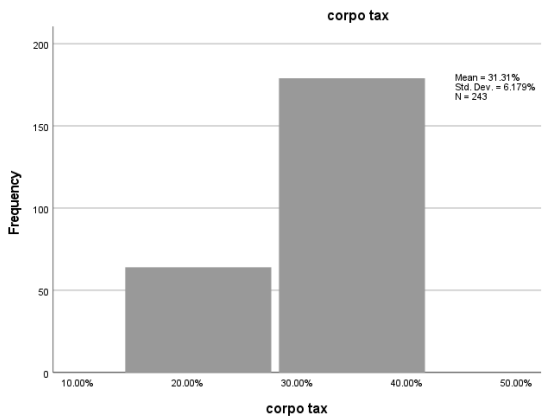
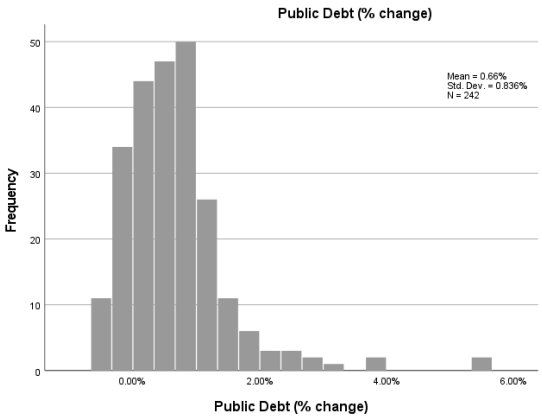


Figure 31: Histogram Producer Price Index



12.3 Scatterplot and Histogram for the dependent variable “LogReturn Russell 2000”

Figure 32: Time-series plot – Russell 2000 Return

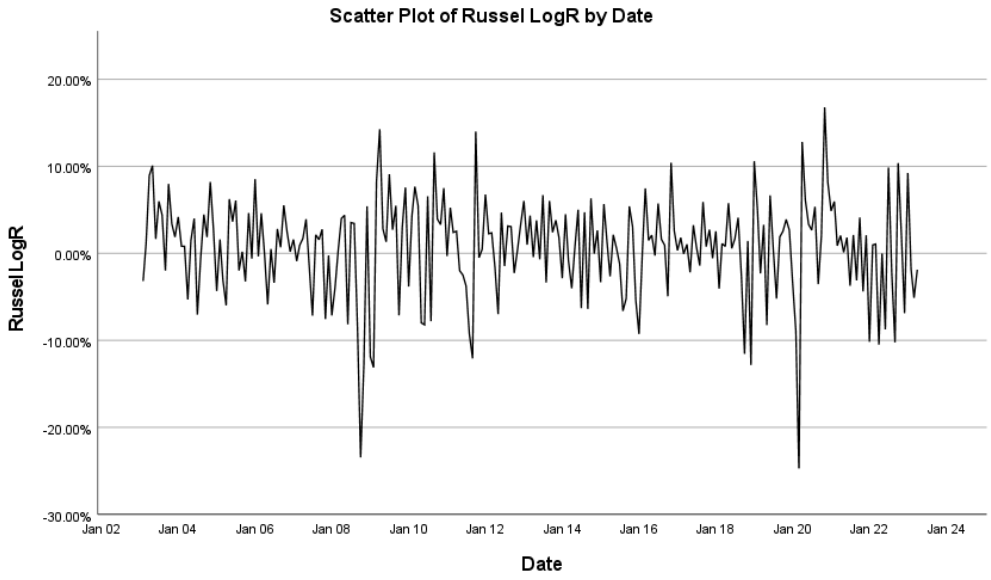


Figure 3315: Histogram Russell 2000 Index Monthly Return